Relativized Exhaustivity: Mention-Some and Uniqueness

Yimei Xiang, Rutgers University

(Version: August 3, 2021)

Abstract   *Wh*-questions with the modal verb *can* admit both mention-some (MS) and mention-all (MA) answers. This paper argues to treat MS as a grammatical phenomenon, primarily determined by the grammar of a *wh*-interrogative. I assume that MS and MA answers come from a uniform 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 contradicts the widely adopted analysis of uniqueness effects in questions: according to Dayal 1996, the uniqueness effects of singular *which*-phrases come from an exhaustivity presupposition that a question must have a unique exhaustive true answer. To solve this dilemma, I propose that question interpretations presuppose 'Relativized Exhaustivity', which has an effect of evaluating exhaustivity relative to the accessible worlds as opposed to the anchor world. Relativized Exhaustivity preserves the merits of Dayal's exhaustivity presupposition while permitting MS; moreover, it explains the local-uniqueness effects in modalized singular *wh*-questions.

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

Contents

Relativized Exhaustivity: Mention-Some and Uniqueness

1.   Introduction

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

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

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

(2)   (A's belief: John went to the party. Unclear who else went to the party.)

      Q: 'Who went to the party?'

      A: a.   'John did .../'
                 L H*   L–H%

           b.   # 'John did.\'   ⤳ *Only John went to the party.*
                  H*   L–L%

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), one may simply 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 exclusive inference. Following Groenendijk and Stokhof (1984), I call these answers '*mention-some* (MS) answers'. Relatedly, interpretations in which a question seeks a MS answer are called 'MS interpretations', and questions that admit MS interpretations are called 'MS questions'. *Can*-questions also admit '*mention-all* (MA) interpretations': in (3), one may address the question by listing all of the accessible coffee places. Hence, we say that *can*-questions

---

[1]There is no clear consensus on what a RFR contour 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, in the answer in (i), the RFR contour is used to indicate that the addressee 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. Unsure whether he is a math professor.)
      Q: 'Which math professor went to the party?' A: 'John did .../'

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

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

exhibit a 'MS/MA ambiguity'. MA answers to *can*-questions can be stated either as conjunctions as in (3b), or more naturally as disjunctions as in (3c) (Dayal 2017). The 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, including 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. For the direction of treating MS answers as partial answers, it is puzzling that *can*-questions systematically tolerate incompleteness. A common view 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 goals of the conversation. This view is shared by the 'pragmatic approaches', which consider MS as a simple pragmatic phenomenon (Sect. 3.1), as well as most of the 'semantic approaches', which consider MS as an independent interpretation but attribute the licensing of MS to pragmatic factors (Sect. 3.2.1). These approaches typically pursue a joint analysis for MS answers to *can*-questions and non-exhaustive answers to non-*can*-questions.

There is no doubt that pragmatics plays an important role in distributing MS. The problem with novelty, however, is whether pragmatics is 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 are 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 answer and MA answers are derived based on a uniform answerhood which doesn't require global exhaustivity (after Fox 2013). 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 delivers MS when it applies to certain forms of *can*-questions, and moreover, the yielded MS answers are 'mention-one', locally exhaustively, 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*. In particular, origin (ii) bridges the distribution of disjunctive-MA with the modal obviation effect in licensing universal free-choice items.

However, allowing complete answers to be non-exhaustive contradicts an influential exhaustivity presupposition from Dayal 1996: a question is defined only if it has a unique exhaustive true answer. This presupposition, henceforth called '*Dayal's exhaustivity presupposition*', nicely explains the unique-ness effects of singular *which*-phrases. To solve this dilemma, I propose to replace Dayal's exhaustivity presupposition with '*Relativized Exhaustivity*', which has an effect of evaluating exhaustivity relative to the accessible worlds as opposed to the anchor world. Relativized Exhaustivity permits MS and avoids over-generating MS. 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 in

matrix questions. Section 3 reviews the antecedent approaches to the MS phenomenon and presents arguments for 'nucleus-dependent approaches'. Section 4 derives the various interpretations of *can*-questions compositionally, including first-order/higher-order MS, conjunctive-MA, and disjunctive-MA. Section 5 delves into the dilemma between uniqueness and MS and reviews two recent analyses by Fox (2018, 2020) and Hirsch and Schwarz (2020). Section 6 proposes Relativized Exhaustivity, solves the dilemma, and accounts for local uniqueness. Section 7 concludes. The appendices discuss choice questions and explore a variable-free treatment of modal bases.

2. Distributional factors of MS

2.1. Modal flavor and modal force

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

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

(4) Q: 'Where can we get coffee?' A: 'Starbucks.\'

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

(5) (A's belief: There are two coffee places near John's home, namely, Starbucks and Peet's. John could go to either place.)

Q: 'John left home for coffee 15 mins ago. Where could he go?' A: ...
   a.   ? 'Starbucks.\'
   b.   'Starbucks or Peet's.\'

Why can't epistemic modals license MS? Intuitively, in (5) what the questioner is really interested in knowing is where John *actually* went, not where he *possibly* went. However, since the questioner doesn't expect the addressee to know the answer, she chose to ask a *could*-question to gather some possibilities. To maximize the chance that the answer covers the place that John actually went to, the addressee would list out all of the places he possibly went to.[2]

Dayal (2017) has excluded epistemic modals and narrowed down the options to what Portner (2009) classifies as 'priority modals'. In Portner's classification, priority modals include not only the

---

[2]Which possibilities should be mentioned depends on multiple contextual factors. On the one hand, the more possibilities be mentioned, the more likely that the answer would cover the real situation. If an answer specifies n-many likely possibilities $o_1, o_2, ..., o_n$, the probability for this answer to cover the real situation is the sum of the probability of each possibility, namely $\sum_{i=1}^{n} \text{PROB}(o_i)$. On the other hand, an answer that lists too many possibilities is under-informative. For example, to answer 'Where could the criminal be hiding?', saying 'Anywhere' is not helpful; instead, the addressee should highlight or only mention the places that deserve a search.

aforementioned teleological and bouletic modals but also deontic modals, of which the ordering source is related to obligations and permissions. However, MS interpretations are difficult in questions with a deontic modal. In example (6), the modal verb *can* in the question is ambiguous between teleological and deontic. If Bob intends to provide an easy path to complete the work, the MS answer (6a) is sufficient; however, if Bob intends to inform Alice about the regulations, especially for Alice's future reference, then the exhaustive answer (6b) is required. The deontic reading is more salient if *can* is stressed or is replaced with *be allowed to*.[3]

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

Alice: 'Who can I ask to sign this document?' Bob: ...

  a. 'Your mentor.'

  b. 'Your mentor, the GPD, or the chair.'

Answers to a deontic *can*-question are exhaustive because the addressee must ensure that at least one option is feasible to the agents of action regardless of the situations they are in. For example, answer (6b) can be understood as follows: 'You can ask your mentor to sign this document; however, if you can't find your mentor or don't want to interact with your mentor, you can ask the GPD or the chair to sign this document.' This meaning is similar to what I call 'family of conditional MS' interpretation of goal-oriented *can*-questions, which is found in cases where the feasibility of an option varies by the situation where the agent of action is in (see Sect. 2.3).

Modal force concerns the force of quantification, which can be existential, universal, or something in between. Only existential modals may license MS. In (7) and (8), the modal verb *should* is goal-oriented but has a universal modal force. In (7), clearly, the addressee should specify all the individuals who should be invited.

(7) Q: 'Who should we invite to form a discussion panel?'

    A: 'The department chair, the GPD, and the two graduate representatives.'

Example (8) illustrates a multiple-choice scenario. If the choices are comparable, the addressee should provide a free-choice answer as in (8b). The single-choice answer (8b) is infelicitous or false in this context unless the addressee has additional reasons to prefer Starbucks over Peet's (e.g., the Starbucks location is new). If the quantification domain of *should* is restricted to the worlds that best satisfy such preferences, (8a) becomes the only true answer.

(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.' $\qquad\qquad\qquad\qquad\qquad\qquad$ $\Box[\phi_{starbucks} \lor \phi_{peet's}]$

In sum, only teleological and bouletic modals with an existential force can license MS interpretations.[4] In English, such modality is realized as *can* or infinitives. In this paper, unless specified, '*can*-questions' refer to *wh*-questions with an existential teleological/bouletic modal.

---

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

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

2.2. Conversation goals

Conversation goals also play an important role in distributing MS and MA. On the one hand, questions without the teleological/bouletic *can* admit MS answers in and only in goal-oriented contexts (Dayal 2017). (9) and (10) exemplify the case with an epistemic modal *might*. In both cases, the answer *Bill might be in* says: 'There is a world w compatible with the available evidence such that Bill is in the office in w.' However, in (10) it also implies a goal-oriented resolvedness inference as follows: 'In my belief, in the current circumstance, that Bill is in the office leads to a possible satisfaction of your goal to find someone to help you.'[5] The same idea applies to the non-modalized question in (11).

(9)   a. I see a light on in the office. Who might be in at this time?
      b. Bill might be in.\
         ⤳ *The available evidence only suggests that BILL might be in the office.*

(10) a. I need help. Who might be in the office at this time?
      b. Bill might be in.\ He could help you.
         ⤳̸ *The available evidence only suggests that BILL might be in the office.*

(11) a. I need a ride to the party tonight. Who's driving?
      b. Bill is.\ He could give you a ride.
         ⤳̸ *Only Bill is driving.*   ((9)–(11) are modified from Dayal 2017: pp. 77)

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

(12) (After the long-list interview, the committee decided to prioritize candidates who could teach Experimental Semantics or Field Methods.)
      Q: 'Who can teach Experimental Semantics?'
      A: 'Judy can.\'
         ⤳ *Among the candidates on the long list, only Judy can teach Experimental Semantics.*

There are two directions to analyze the MS-licensing effect of goal-oriented contexts. One direction is to think of MS-licensing as a pragmatics-driven phenomenon: MS is primarily licensed by a non-exhaustive conversation goal, independent from the presence of a modal expression. In this view, the reason why the modal verb *can* licenses MS is that it may come with the goal-oriented modality. This 'modal-to-context reduction' analysis, however, cannot explain the distribution of MS in embeddings and the 'mention-one-only' constraint (Sect. 3.3).

Another direction, which I assume, is to take the modal verb *can* as the primary licensing factor; goal-oriented contexts may provide a constraint to 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 underlining parts come from the context.

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

With this 'context-to-modal reduction' strategy, whatever explains the MS-licensing effect of the modal verb *can* also explains the MS-licensing effect of goal-oriented contexts. Moreover, it predicts

---

[5]Note that in (10), the literal meaning of the modalized answer, namely that Bill *might* be in the office, doesn't ensure a possibility of satisfying the questioner's goal. More precisely, this answer says: 'Bill might be in. If he *is* in, he could help you.' Hence, no semantic answer to *Who might be in the office?* ensures a possible satisfaction of the speaker's goal.

that, if a *wh*-construction cannot obtain a covert *can*-restriction from the context, it admits MS only in the presence of the modal verb *can*. This prediction is consistent with the observations with question embeddings and non-interrogative *wh*-constructions (see Sect. 3.3.1).

2.3. Agent-dependency and conditional MS

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

(14) (Context: There are four stores in town where one can ship packages: USPS, UPS, Fedex, and DHL. USPS has a cheaper rate, while the others provide express delivery.)
    Q: 'Where can I go to ship this package?'
    A: 'USPS if you want a cheaper rate, or UPS if you need express delivery.'

(15) (Context: The city has locations of CVS, Walgreen, and Rite Aid, distributed in different regions. All these stores provide flu vaccines.)
    Q: 'Where can I go to get a flu vaccine?'
    A: 'CVS, Walgreen, or Rite Aid.'

In the above two cases, the multiple-option answer addresses a family of conditional MS questions. For instance in (15), the question can be thought of as a family of sub-questions as follows: 'If I live in x, where can I get a flu vaccine?'. In each such sub-question, the modal base of *can* is restricted to a set of worlds where the questioner lives in x.[6] This meaning is formalized as below, where M is a circumstantial modal base, and $\mathcal{C}$ is a set of propositional descriptions of relevant conditions. In each sub-question, the modal verb *can* is interpreted relative to a modal base M' which maps any anchor world w to a subset of M_w consisting of only the worlds where a condition in $\mathcal{C}$ is true. A complete answer to this family of conditional *can*-questions should address all these question, which therefore is the conjunction of a set of MS answers.

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

This paper will not delve into the 'family of conditional MS' interpretations any further. However, for later discussions on the nature of the MS/MA distinction, it is crucial to distinguish between a single *can*-question with a MA interpretation and a family of conditional *can*-questions with a MS interpretation. To avoid confounding factors from the latter, the following discussion on the grammatical constraints on the MS/MA distinction will make the goals explicit and consider only cases where the options are agent-independent.

---

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

3. Directions of analysis

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

	Pragmatic	Semantic	
		Nucleus-independent	Nucleus-dependent
Issue A	Yes	No	No
Issue B	Yes	Yes	No

Table 1: (A) whether or not MS is an independent interpretation; (B) whether or not the availability of MS is primarily determined the question nucleus.

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

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

3.1. Pragmatic approaches

MS was firstly perceived as a pragmatic phenomenon. Earlier works took this direction to maintain a core assumption in question semantics, namely, that question interpretations must be exhaustive (Groenendijk and Stokhof 1984).[7] MS answers, which are non-exhaustive, were therefore treated as partial answers permitted for pragmatic reasons, such as that they are sufficient relative to the conversation goals. For instance, for the question *Where can we get coffee around here?*, if the goal of the questioner is just to get a coffee, the addressee only needs to specify one coffee place suited for the questioner's needs; in contrast, if the goal is to investigate the local coffee market, the addressee needs

---

[7]Groenendijk and Stokhof (1984) assume that questions must be interpreted as *strongly exhaustive*. In this view, for example, the question *who came* denotes a function that maps a world w to the exhaustified proposition 'only x came' that is true in w. Hence for Groenendijk and Stokhof, the puzzle of MS is why certain questions admit interpretations that are not strongly exhaustive. Adopting this view, George (2011: Chap. 2) analyzes the MS/MA distinction as a semantic contrast between MS and strongly exhaustive (reviewed in Sect. 3.2.1).

In contrast to Groenendijk and Stokhof 1984, most recent works take the *weakly exhaustive* meaning as the base meaning of an interrogative. For example, the base meaning of *who came* is a function that maps a world w to the proposition 'x came' such that it is true in w that only x came. Strong exhaustivity, in contrast, is derived from weak exhaustivity via a separate answerhood (Heim 1994; Dayal 1996; Beck and Rullmann 1999) or a strengthening operation (Klinedinst and Rothschild 2011; Nicolae 2013, 2015; Uegaki 2015). Hence for these analyses, the puzzle of MS is why certain questions admit interpretations that are not even weakly exhaustive.

to list all the coffee places in the considered area. This direction of analysis was first sketched in partition semantics (Groenendijk and Stokhof 1984) and has remained popular in various frameworks of question semantics.

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

(17) Q: 'Who can give us a ride?'
     A: 'Bill can.\'

(18) Q: 'I need a ride to the party tonight. Who's driving?'
     A: 'Bill is.\ He could give you a ride.'

van Rooij (2003, 2004) relates question-answering to *decision problems* and provides a statistical characterization for the *utility* of a non-exhaustive answer. In this theory, the fact that *can*-questions admit MS answers has no essential difference from the fact that the two non-*can*-questions below admit non-exhaustive answers:

(19) a. Who has got a light?   (with an existential goal)
     b. Who, for example, came to the party?   (with a partiality marker *for example*)

A commonly raised challenge to pragmatic approaches, as pointed out by Groenendijk and Stokhof (1984) themselves and reiterated by George (2011), is that pragmatics cannot explain the availability of MS interpretations in question embeddings. As seen in (20b), knowing a MS question entails knowing a true MS answer to this question.

(20) a. Jack knows [who arrived].
        ⤳ *For every relevant individual x, if x arrived, Jack knows that x arrived.*
     b. Jack knows [who we can address this question].
        ⤳ *For one relevant individual x such that x can address this question, Jack knows that x can address this question.*

Pragmatic approaches address this challenge by making the semantic denotations of questions or answers context-dependent. For instance, as seen above, Ginzburg (1995) and van Rooij (2003) argue that the resolvedness of an answer is context-dependent, which can affect the truth conditions of some question embeddings. Lahiri (2002) proposes that the interpretation of a question embedding involves picking a sub-question of the embedded question, whose size is determined by the goal of the speaker. An unpublished version of Dayal 2017 analyzes the semantic denotation of a question context-independent but defines an answerhood for MS that is sensitive to the conversation goal.

With these options, the fact that MS is available in embeddings isn't a knockdown argument against pragmatic approaches. However, it remains puzzling to pragmatic approaches as well as to some of the semantic approaches why MS is difficult in embeddings of non-*can*-questions (see Sect. 3.3.1).

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

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

Most existing semantic approaches are 'nucleus-independent approaches': they attribute the MS/MA ambiguity to operations *outside* the question nucleus (e.g., the selection of answerhood). In contrast, nucleus-dependent approaches attribute this ambiguity to structural ambiguities *within* the question nucleus. Below I will review some representatives of both categories.
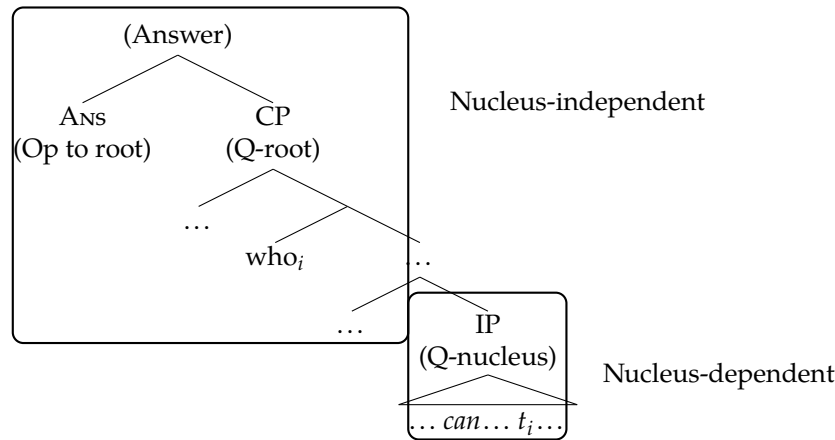
Figure 1: Nucleus-independent versus nucleus-dependent

3.2.1. Nucleus-independent approaches

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

MS as an existential answerhood Beck and Rullmann (1999) attribute the MS/MA ambiguity to the selection of answerhood. They define the root of a question unambiguously as a Hamblin-Karttunen intension (i.e., a function that maps a world to the set of true propositional answers to the

question in this world) but assume multiple answerhood (Ans-)operators, including:

(21) a. $[\![\text{Ans}_{BR1}]\!] = \lambda w_s \lambda Q_{\langle s, \langle st,t \rangle \rangle}.\bigcap \{p \mid Q(w)(p) \wedge p(w)\}$

  b. $[\![\text{Ans}_{BR3}]\!] = \lambda w_s \lambda Q_{\langle s, \langle st,t \rangle \rangle} \lambda P_{\langle s,stt \rangle}.\exists p[P(w)(p) \wedge Q(w)(p) \wedge p(w)]$

The above two Ans-operators differ in exhaustivity. In (22), for the embedding sentence 'x knows Q', applying Ans$_{BR1}$ to the embedded question returns the conjunction of all the true propositional answers to this question, yielding a weakly exhaustive/ MA interpretation. In contrast, as in (23), Ans$_{BR3}$ shifts the embedded question into an existential quantifier over questions. Interpreting this quantifier over the embedding predicate yields an existential/MS interpretation.

(22) a. LF for MA: [x knows [Ans$_{BR1}(w)(Q)$]]

  b. Meaning: $\mathsf{know}_w(x, \bigcap\{p \mid Q(w)(p) \wedge p(w)\})$

(23) a. LF for MS: [Ans$_{BR3}(w)(Q)$ λp_{st} λw_s [x knows$_w$ p]]

  b. Meaning: $\exists p[\mathsf{know}_w(x,p) \wedge Q(w)(p) \wedge p(w)]$

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

(24) Who came?

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

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

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

  d. Without X: MS/ weakly exhaustive
   $[\![\text{Q}(\text{Abs})]\!] = \lambda p_{\langle s,t \rangle}.\exists \beta_e [p = \lambda w_s.\mathsf{came}_w(\beta)]$
     $= \{\lambda w.\mathsf{came}_w(\beta) \mid \beta \in D_e\}$

  e. With X: strongly exhaustive
   $[\![\text{Q}(\text{X}(\text{Abs}))]\!] = \lambda p_{\langle s,t \rangle}.\exists \beta_{\langle e,t \rangle}[p = \lambda w_s[(\lambda x_e.\mathsf{came}_w(x)) = \beta]]$
     $= \{\lambda w_s.[(\lambda x_e.\mathsf{came}_w(x)) = \beta] \mid \beta \in D_{\langle e,t \rangle}\}$

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

3.2.2. Nucleus-dependent approaches

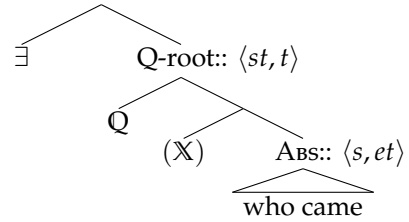Nucleus-dependent approaches attribute the MS/MA ambiguity to structural variations within the question nucleus. Below I review the accounts of George (2011: Chap. 6) and Fox (2013), which analyze the MS/MA ambiguity in *can*-questions as the scope ambiguity of exhaustivity or distributivity relative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

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

(25) Who can chair the committee?
   a. $[\![\text{ABS}]\!] = \lambda x_s.\text{chair}_{w'}(x)$
   b. $[\![\mathbb{X}(\text{ABS})]\!] = \lambda \delta_{et}[\delta = \lambda x_e.\text{chair}_{w'}(x)]$
   c. $[\![can]\!] = \lambda q_{st}.\exists w' \in M_w[q(w')]$
   d. $[\![can(\mathbb{X}(\text{ABS}))]\!] = \lambda \delta_{et}.\exists w' \in M_w[\delta = \lambda x_e.\text{chair}_{w'}(x)]$
   e. $[\![Q(can(\mathbb{X}(\text{ABS})))]\!]$
      $= \lambda p_{st}.\exists \beta_{et}[p = \lambda w_s.\exists w' \in M_w[(\lambda x_e.\text{chair}_{w'}(x)) = \beta]]$
      $= \{\lambda w_s.\exists w' \in M_w[(\lambda x_e.\text{chair}_{w'}(x)) = \beta] \mid \beta \in D_{\langle e,t \rangle}\}$
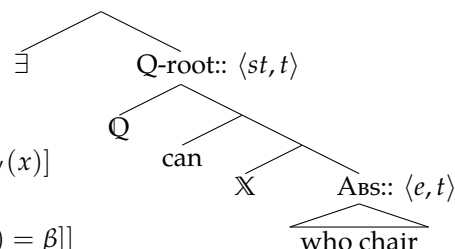
This account predicts that the MS interpretation is only available in questions with an existential expression. It also nicely accounts for the local-exhaustivity effect of MS answers, an observation made in Xiang 2016: Chap. 2. Compare the answers in (26): although it is true that Andy can serve on the committee, (26a) is a bad answer, in contrast to (26b,c) each of which specifies the full composition of a possible committee. This contrast argues that MS answers are subject to local exhaustivity; for example, (26b) is read as: 'It can be the case that *only* Andy and Billy serve on the committee.' In George's account, the $\mathbb{X}$-operator applied under *can* captures this intuition.

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

   Who can serve on the committee?
   a. # Andy.\
   b. Andy and Billy.\
   c. Andy, Billy, and Cindy.\

This account also faces a few problems. First, the composition between *can* and $\mathbb{X}(\text{ABS})$ doesn't follow the conventional rules: *can* selects for a proposition, while $\mathbb{X}(\text{ABS})$ denotes a set of predicates. Second, although the local application of $\mathbb{X}$ nicely predicts the local-exhaustivity effect of MS answers, the global application of $\mathbb{X}$ forces a strongly exhaustive interpretation, which is too strong (see f.n. 7). Third, this account predicts that any expressions with an existential force, including indefinites, can license MS. However, a number of empirical distinctions between *can*-questions and questions with an existential indefinite argue that the non-exhaustive interpretations of these two types of questions have different origins. For details, see Appendix A and references therein.

MS as a scope effect of distributivity Fox (2013) assumes that MS and MA answers are obtained

by a uniform answerhood for complete true answers which doesn't demand global exhaustivity. As defined in (27), an true answer is complete as long as it is not asymmetrically entailed by any of the true answers. Here Q denotes the Hamblin set of the question, also called 'answer space'.

(27) $\text{ANS}_{Fox}(w)(Q) = \{p \mid w \in p \in Q \land \forall q[w \in q \in Q \to q \not\subset p]\}$ (Fox 2013)

Fox's answerhood is stronger than simple existential but weaker than exhaustive. In comparison, Dayal's (1996) answerhood returns the exhaustive true answer, which is the unique true answer that entails all the true answers.[8]

(28) $\text{ANS}_{Dayal}(w)(Q) = \exists p[w \in p \in Q \land \forall q[w \in q \in Q \to p \subseteq q]].$
$\iota p[w \in p \in Q \land \forall q[w \in q \in Q \to p \subseteq q]]$ (Dayal 1996)

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

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

Who can chair the committee?
a.   [$_{CP}$ who λX ... [$_{IP}$ can [[X EACH] λx [x chair the committee]]]]   (*can* $\gg$ EACH: MS)
    i.   True answers: $\{\Diamond\text{chair}(a), \Diamond\text{chair}(b)\}$
    ii. $\text{ANS}_{Fox}(w)(Q) = \{\Diamond\text{chair}(a), \Diamond\text{chair}(b)\}$
b.   [$_{CP}$ who λX ... [$_{IP}$ [X EACH] λx [can [x chair the committee]]]]   (EACH $\gg$ *can*: MA)
    i.   True answers: $\{\Diamond\text{chair}(a) \land \Diamond\text{chair}(b), \Diamond\text{chair}(a), \Diamond\text{chair}(b)\}$
    ii. $\text{ANS}_{Fox}(w)(Q) = \{\Diamond\text{chair}(a) \land \Diamond\text{chair}(b)\}$

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

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

Who can serve on the committee?

---

[8]The following illustrates the difference between Dayal's and Fox's answerhood-operators:

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

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

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

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

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

3.3. Arguments for nucleus-dependent approaches

Pragmatic approaches and nucleus-independent approaches predict that the MS interpretations of *can*-questions are primarily licensed by pragmatic factors, especially whether the question has a non-exhaustive conversation goal. In this view, the existential priority modal *can* may license MS just because it naturally comes with a non-exhaustive conversation goal. This idea is compatible with the data of matrix questions, as seen in (9)–(11). Further, as for why non-exhaustive conversation goals license MS, pragmatic approaches attribute it to the resolvedness/utility of an answer — when the goal is non-exhaustive, a non-exhaustive answer suffices for resolving the question. To this extend, as predicted by the utility theory of van Rooij (2003, 2004), the non-exhaustive interpretations of *can*-questions and non-*can*-questions (see (19)) have the same nature.

In contrast, nucleus-dependent approaches predict that the modal verb *can* (or another equivalent modal expression), which appears within the question nucleus, is the primary source for licensing MS. Existing nucleus-dependent approaches haven't said much about the MS-licensing effect of conversation goals. However, they are compatible with the context-to-modal reduction strategy argued in Sect. 2.2: conversation goals may license MS for non-*can*-questions just because they can restrict the question nucleus and turn a non-*can*-question into a goal-oriented *can*-question. Other pragmatic factors may serve as a blocker for MS answers or a rescuer for incomplete answers.

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

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

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

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

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

13

This background specifies a non-exhaustive goal of the attitude holder (viz., to find two children to co-lead the dance), and it ensures that whatever addresses the embedded *can*-question in (30a) also addresses the embedded non-modalized question in (30b), and vice versa. Next, the participants saw a set of pictures illustrating how children were dressed and two sentences describing Mary's memory. Figure 2 presents a MS trial: Mary remembers only one of the two pairs of children who have an accessary in common and can co-lead the dance. Finally, the participants were asked to judge the truth value of a question embedding sentence. This form of the embedded question varied by two conditions: (i) whether the *wh*-subject is *who* or *which children*, and (ii) whether the embedded question is a *can*-question, as in (30a), or is not modalized, as in (30b). Regardless of the form of the *wh*-subject, fitting data of the MS trials with a logistic mixed effect modal reported a significant effect of [±Modal] ($p < .001$).

How children are dressed:

Ann   Bill   Chloe   Diana

Mary's memory
Bill and Chloe wear the same bowtie, Chloe wears a hat.
Therefore, Bill and Chloe can lead the dance.
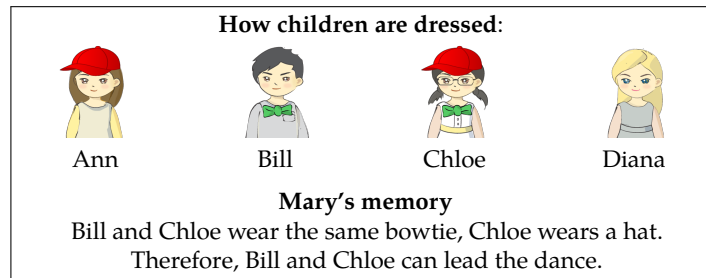
Figure 2: MS trial in Xiang and Cremers 2017

(30) a. Mary remembers {who, which children} can lead the dance. ([+Modal])

b. Mary remembers {who, which children} have an accessary in common. ([–Modal])

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

(31) a. Jack ate [what Mary cooked for him].
⤳ *John ate everything that Mary cooked for him.*

b. Jack went to [where he could go to get help].
⤳ *John went to one of the places where he could go to get help.*

c. Jack went to [where he ought to go to get help].
⤳ *John went to all of the places where he ought to go to get help.*

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

(32) a. Ni   qu-guo nar,   wo jiu qu nar.
        you go-EXP where, I   JIU go where
        Intended: 'I will go to every place where you have been to.'

     b. Nar   neng mai-dao   jiu,   wo jiu qu nar.
        where can   buy-reach liquor, I   JIU go where
        Intended: 'I will go to one of the places where I can buy liquor.'

In sum, despite that non-exhaustive conversation goals sufficiently license non-exhaustive answers to matrix questions, in question embeddings, *wh-* free relatives, and *wh*-conditionals, MS/existential interpretations are more readily available in the presence of the modal verb *can*. This contrast argues that the contextual support for licensing MS and other non-exhaustive interpretations is not accessible to embedded *wh*-constructions, and that the goal-oriented modal *can* is a sufficient MS-licenser.[9]

3.3.2. The 'mention-one-only' constraint of MS answers

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

First, embeddings of *can*-questions allow for mention-one and mention-all interpretations but not mention-few interpretations, even in cases where mention-few fits with the conversation goal better than mention-one/all. In (33), the conversation goal calls for a 'mention-three answer' (viz., an answer that specifies three possible venues). If MS interpretations were primarily licensed by the conversation goal, (33) should have the mention-three interpretation (33b), contrary to fact.

(33) (Context: The committee needs to find out three possible venues for the upcoming conference. They expect about 300 people to attend the conference.)

---

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

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

I argue that the interrogative subject in (i) is semantically ambiguous between MS and MA, but this ambiguity collapses under the selectional requirement of the matrix predicate. The subject of *depend on* must be strongly exhaustive. For example in (ii), if the subject interrogative were weakly/intermediately exhaustive, speaker B's objection would be infelicitous.

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

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

(iii) The strong exhaustivity inference of a question Q in the world w: $\lambda w'[\text{Ans}(w)(\llbracket Q \rrbracket) = \text{Ans}(w')(\llbracket Q \rrbracket)]$

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

Jill knows where in the campus can hold a conference with 300 participants.

a. ✓ 'Jill knows one/all of the places in the campus that can hold a conference with 300 participants.' (Available: mention-one/ mention-all)

b. ✗ 'Jill knows three of the places in the campus that can hold a conference with 300 participants.' (Unavailable: mention-three)

Note that it is difficult to exclude mention-few interpretations based on the truth-conditions of (33). For the sentence 'x knows Q' where Q is a MS question, it is true if and only if (i) x knows a true mention-one answer to Q, and (ii) x has no false belief relevant to Q. However, since knowing a true mention-three answer or knowing three mention-one answers entails condition (i), the mention-one interpretation is true in a scenario where Jill knows three possible venues. To examine the availability of a mention-three interpretation empirically, consider the following conversation. In (34), despite the conversation goal is 'mention-three', to answer a polar question concerning a MS-embedding, the addressee Bob cannot reply with a denial while admitting that he knows one possible venue.

(34) Alice: 'We are looking for a venue for the upcoming conference. We need to find out three options. Do you know where in the campus can hold a conference with 300 participants?' Bob: 'Yes, but I only know one.'/'#No, I only know one.'

Second, in discourse, in contrast to mention-one answers, mention-few answers are not read non-exhaustively in the absence of an ignorance-mark. The following is a comprehension task concerning a typical MS question, presented in English and in Mandarin. Speakers were asked to judge whether the mention-one answer (i) and the mention-two answer (ii) gave rise to an exclusive inference. Notably, among those who judged (i) non-exhaustively, a large majority judged (ii) exhaustively.[10]

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

a. Bob: Which blade can I use to cut carrots? Alice: ...

    i. The green one.

    ii. The green one or the black one.

b. Bob: Na-ge daopian keyi yonglai qie huluobo? Alice: ...
        Which-CL blade can be-used-for cut carrots?

---

[10]The design of example (35) avoids some confounding factors. First, the domain of the *wh*-phrase is small and salient, which avoids confounds from implicit domain restrictions. However, a small domain makes the exhaustive interpretations more appealing, which makes it harder to gather judgments related to the MS interpretation. Second, this example avoids the 'family of conditional MS' interpretation (see Sect. 2.3), which is otherwise hard to be distinguished from mention-few/all.

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

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

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

'Which blade can be used for cutting carrots?'

   i. Lüse-de   zhe-ge keyi.
      green-MOD this-CL can

      'The green one can.'

  ii. Lüse-de   zhe-ge huozhe heise-de   zhe-ge dou keyi.
      green-MOD this-CL or   black-MOD this-CL DOU can

      Intended (free choice): 'The green one and the black one both can.'

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

For answers to questions with a partiality marker (e.g., *for example*, *for instance*, *give me an example*), there is no contrast in exhaustivity between mention-one and mention-few. In (36) and (37), regardless of the presence of *can*, neither mention-one nor mention-few answers imply exclusivity.

(36) Who is on your committee, for example?

    a. Andy.\   ↛ *Only Andy is in my committee.*

    b. Andy and Billy.\   ↛ *Only Andy and Billy are in my committee.*

(37) Who can chair the committee alone, for example?

    a. Andy.\   ↛ *Only Andy can chair alone.*

    b. Andy and Billy (each can).\   ↛ *Only Andy and Billy can chair alone.*

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

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

I would finally clarify a possible confusion: 'mention-one' refers to mentioning one option/possibility, not one atomic individual/entity. In contrast to the above examples where each option is constituted by one single individual/location/entity, there are also cases where an option is made up of the sum or boolean coordination of multiple individuals/locations/entities. For example, to answer (26) *Who can serve on the committee?*, a mention-one answer should specify the sum of a group of individuals who can simultaneously serve on the committee. In (38), the two conjunctive answers are also mention-one answers, where each such answer specifies an option to group teams and assign leaders. Each such answer names a boolean conjunction.

(38)   (The players can be grouped into 2~3 teams. Each team needs one or two leaders.)

Who can we ask to lead a team?

   a. We can ask Andy to lead a team and Billy to lead a team.   (A two-teams option)

   b. We can ask Andy to lead a team, Billy to lead a team, and Cindy and Danny together to lead a team.   (A three-teams option)

4.   Composing MS questions

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

In particular, the treatment of MS will account for a number of grammatical effects of MS answers, including local exhaustivity, mutual independence, and 'mention-one-only'. The treatment of MA relates disjunctive-MA to the derivation and the modal obviation effect of universal free-choice.

4.1.   General assumptions on questions and answers

4.1.1.   Questions as topical properties

In line with categorial approaches, I assume that the semantic denotations of questions are *topical properties* (after Chierchia and Caponigro 2013). As exemplified in (39a,b), the topical property of a *wh*-question is a function that maps an individual in the *wh*-domain to a proposition in the answer space. In other words, as formalized in (39c), the answer space of a question is the *image* of the topical property of this question. I henceforth write topical property as '$[\![Q]\!]$' and answer space as 'Q'.

(39)   Q: 'Which books did John read?' A: 'The Harry Potter books.'

   a. $[\![Q]\!] = \lambda x_e : x \in \mathsf{books}_@.\lambda w[\mathsf{read}_w(\mathsf{j}, x)]$

   b. $[\![Q]\!]([\![HP]\!]) = \mathsf{hp} \in \mathsf{books}_@.\lambda w[\mathsf{read}_w(\mathsf{j}, \mathsf{hp})]$

   c. $Q = \{[\![Q]\!](x) \mid x \in \mathrm{Dom}([\![Q]\!])\}$
           $= \{\lambda w.\mathsf{read}_w(\mathsf{j}, x) \mid x \in \mathsf{books}_@\}$

In contrast to defining questions as partitions or sets of propositions, defining questions as topical properties makes it easy to track short answers and makes the analysis of MS applicable to the existential readings of *wh*- free relatives and *wh*-conditionals. Moreover, this assumption is crucial for my solution to the dilemma between uniqueness and MS: I will argue that question interpretation is subject to a condition called 'Relativized Exhaustivity', the definition of which requires to extract short answers from question denotations. The main reason for this choice is that, in a modalized question, the semantics of a propositional answer varies by the modal base, while that of a short answer doesn't (see f.n. 30 in Sect. 6). However, also see Appendix B for a variable-free analysis of modal base which avoids this issue.[11]

---

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

4.1.2. Answerhood

I adopt Fox's (2013) answerhood for complete true answers: a true answer is complete if and only if it is not asymmetrically entailed by any of the true answers. Such an answer is henceforth called a '*maximally (max-)informative true answer*', in contrast to what Dayal's (1996) answerhood defines as the unique '*strongest/exhaustive true answer*'.

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

(40) Answerhood-operators (to be modified)

     a. For complete true short answers
         $\text{Ans}^S(w)([\![Q]\!]) = \{\alpha \mid \alpha \in [\![Q]\!]_w \wedge \forall \beta[\beta \in [\![Q]\!]_w \rightarrow [\![Q]\!](\beta) \not\subset [\![Q]\!](\alpha)]\}$

     b. For complete true propositional answers
         $\text{Ans}^P(w)([\![Q]\!]) = \{[\![Q]\!](\alpha) \mid \alpha \in \text{Ans}^S(w)([\![Q]\!])\}$

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

4.1.3. First-order versus higher-order interpretations

Wh-questions have first-order and higher-order interpretations (Spector 2007, 2008; Xiang 2021b). For example, the following question with a universal modal can be completely addressed by specifying some particular books, as in (41a), or a generalized quantifier (GQ) over a set of books, as in (41b).

(41) Which books does John have to read?

     a. First-order interpretation: 'Which books x is such that John has to read x?'
         The French novels.

     b. Higher-order interpretation: 'Which GQ π over books is such that John has to read π?'
         The French novels or the Russian novels. (The choice is up to him.)

I assume the LFs and topical properties as follows for the above two interpretations:

(42) First-order interpretation

     a. $[_{CP}$ which-books λx_e $[_{IP}$ have-to $[_{VP}$ John read x $]]]$
     b. $[\![Q]\!] = \lambda x_e : x \in \text{books}_@.\Box\lambda w[\text{read}_w(j, x)]$

(43) Higher-order interpretation ($\Box \gg or$)

     a. $[_{CP}$ which-books $\lambda \pi_{\langle et,t \rangle}$ $[_{IP}$ have-to $[\ \pi\ \lambda x_e$ $[_{VP}$ John read x $]]]]$
     b. $[\![Q]\!] = \lambda \pi_{\langle et,t \rangle} : \pi \in {}^{\text{H}}\text{books}_@.\Box\lambda w[\pi(\lambda x_e.\text{read}_w(j, x))]$

In LF (43a), the fronted *wh*-phrase binds a higher-order trace π across the modal (after Spector 2007, 2008). The interactions between the higher-order *wh*-trace and other scopal expressions inside the question nucleus offer scope ambiguities for free; therefore for a scope-based account of MS, there is no need to assume an additional scopal expression inside the question, in contrast to the analyses of George (2011) and Fox (2013) which analyze MS as a scope effect of exhaustivity or distributivity. In denotation (43b), '${}^{\text{H}}\text{books}_@$' stands for a set of GQs ranging over a set of entities that are books in the

actual world @. Although not all GQs can serve as semantic answers to *wh*-questions (Spector 2007, 2008; Xiang 2021b), for this paper, it only matters that the domain of a higher-order topical property includes Montagovian individuals and their boolean coordinations.

4.2. Deriving MS interpretations

4.2.1. Local exhaustification and first-order MS

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

(44)   Who can serve on the committee? (First-order MS)
       [$_{CP}$ who λx_e [$_{IP}$ can [O_C [$_{VP}$ $x_{[+v]}$ serve on the committee]]]]

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

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

       Who can serve on the committee?
       a.   # Andy.\
       b.   Andy and Billy.\
       c.   Andy, Billy, and Cindy.\

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

(46) $\llbracket O_C \rrbracket = \lambda p \lambda w.p(w) = 1 \wedge \forall q \in C[p \nsubseteq q \rightarrow q(w) = 0]$

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

The LF (44) is computed as follows. Here the O-operator is associated with an *e*-type *wh*-trace x. The [+v] feature of the trace activates a set of 'variable-alternatives', defined like focus-alternatives as in (48).[12] The domain variable C carried by the O-operator denotes a contextually determined subset of the variable-alternatives of the VP.[13] Composing this LF yields the topical property (47b). In the

---

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

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

context described in (45), applying Fox's answerhood to this topical property returns a set consisting of two max-informative true answers as in (47c), each of which is a MS answer.

(47) [$_{\text{CP}}$ who λx_e [$_{\text{IP}}$ can [O_C [$_{\text{VP}}$ $x_{[+\text{v}]}$ serve on the committee]]]]   (=(44))

     a. v-ALT(VP) = $\{\lambda w.\textsf{serve-on-the-comm}_w(x) \mid x \in D_e\}$ (Abbreviated as: $\{\phi_x \mid x \in D_e\}$)

     b. $[\![Q]\!] = \lambda x_e : x \in \textsf{hmn}_@.\Diamond O_C\phi_x$, where $C \subseteq \{\phi_x \mid x \in \textsf{hmn}_@\}$

     c. $\textsc{Ans}^P(w)([\![Q]\!]) = \{\Diamond O_C\phi_{a\oplus b}, \Diamond O_C\phi_{a\oplus b\oplus c}\}$

(48) Variable-alternatives: for any trace and pronoun α,

$$\text{v-ALT}(\alpha) = \begin{cases} D_{\text{type}([\![\alpha]\!])} & \text{if } \alpha \text{ carries a [+v] feature} \\ \{[\![\alpha]\!]\} & \text{otherwise} \end{cases}$$

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

(49) Who can serve on the committee?

     a. Andy and Billy ... maybe also Cindy.

     b. Andy and Billy. I don't know whether we should add a third person.

4.2.2. More on the higher-order MS interpretation

Example (50) illustrates the derivation of a higher-order MS interpretation. Compared with the derivation of a first-order MS interpretation, 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 the modal *can* and the local O-operator. Composing this LF yields the higher-order topical property (50b), where '$^{\text{H}}\textsf{hmn}_@$' stands for a set of GQs over human individuals.

(50) Who can chair the committee? (Higher-order MS)

     [$_{\text{CP}}$ who $\lambda\pi_{\langle et,t\rangle}$ [$_{\text{IP}}$ can [π λx_e [O_C [$_{\text{VP}}$ $x_{[+\text{v}]}$ chair the committee]]]]]

     a. v-ALT(VP) = $\{\lambda w.\textsf{chair-the-comm}_w(x) \mid x \in D_e\}$ (Abbreviated as: $\{\phi_x \mid x \in D_e\}$)

     b. $[\![Q]\!] = \lambda\pi_{\langle et,t\rangle} : \pi \in {}^{\text{H}}\textsf{hmn}_@.\Diamond\pi(\lambda x_e.O_C\phi_x)$, where $C \subseteq \{\phi_x \mid x \in \textsf{hmn}_@\}$

Figure 3 illustrates the answer space of (50). It covers four answers related to the two individuals Andy (*a*) and Billy (*b*), derived by applying the topical property (50b) to the Montagovian individuals $a^{\Uparrow}$ and $b^{\Uparrow}$, their boolean conjunction $a^{\Uparrow} \cap b^{\Uparrow}$, and their boolean disjunction $a^{\Uparrow} \cup b^{\Uparrow}$.[14] 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 relevant current circumstances s.t. only a chairs the committee in w'.

---

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

(Context: Only Andy and Billy can chair the committee. Only single-chairing is allowed.)
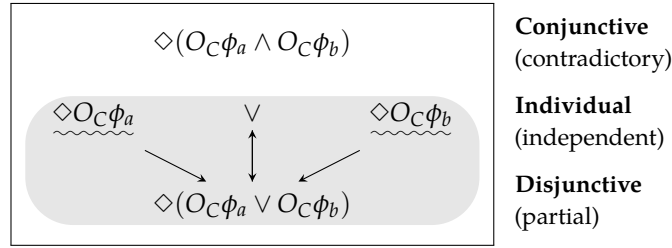

Figure 3: The answer space of the higher-order MS interpretation of *Who can chair the committee?* [Arrows indicate entailment relations, shading marks the true answers, and underwaving marks the max-informative true answers. The bi-implication between '∨' in the middle and the disjunctive answer at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

The following considers the semantics of the three types of answers:

- *Conjunctive answers*: Due to the local O-operator, conjunctive answers are contradictory. In Figure 3, $\Diamond(O_C\phi_a \wedge O_C\phi_b)$ is read as '#There is a world w compatible with the current circumstances s.t. only a chairs the committee in w and only b chairs the committee in w'.

- *Individual answers*: The individual answers can be true and are logically independent from each other. Moreover, due to the non-monotonic O-operator, mutual independence also applies to plural answers (e.g., $\Diamond O_C\phi_{a\oplus b}$), no matter whether the predicate *chair the committee* is interpreted as distributive (i.e., 'chair the committee separately') or collective (i.e., 'chair the committee together'). Hence, any true individual answer is a max-informative true answer.

- *Disjunctive answers*: The disjunctive answer is logically equivalent to the disjunction of the two individual answers. Due to this equivalence, whenever the disjunctive answer $\Diamond(O_C\phi_a \vee O_C\phi_b)$ is true, one of the individual answers ($\Diamond O_C\phi_a$ or $\Diamond O_C\phi_b$) is true and asymmetrically entails the disjunctive answer. Hence, disjunctive answers can never be max-informative and is always partial.

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

To sum up, by adopting Fox's answerhood and assuming an exhaustification operator under the existential modal *can*, this analysis explains three grammatical properties of MS answers, including local exhaustivity, mutual independence, and 'mention-one-only'.

4.2.3. Predictions on the distribution of MS

With Fox's answerhood, a question has a MS interpretation if and only if there is a world in which the answer space of this question has multiple max-informative true propositions. This requirement is met only if the two conditions below are met:

(A) the semantically independent answers are not mutually exclusive;

(B) the answer space is not closed under conjunction.

In a *can*-question, the presence of *can* allows to satisfy 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, while other conditions being equal, if the existential modal is dropped as in (51b) or replaced with a universal modal as in (51c), the individual answers would be mutually exclusive.

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

    a. $\Diamond O_C \phi_a \wedge \Diamond O_C \phi_b \neq \bot$

    b. $O_C \phi_a \wedge O_C \phi_b = \bot$

    c. $\Box O_C \phi_a \wedge \Box O_C \phi_b = \bot$

For condition (B), the answer space of a *can*-question with a higher-order interpretation is not closed under conjunction if the higher-order *wh*-trace π takes scope under the existential modal. In Figure 3, the conjunction of the two individual answers is not in the answer space. 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 4a–c, the conjunctive answer is semantically equivalent to the conjunction of the two individual answers.

a. *Who came?*   b. *Who must come?* ($\Box \gg \pi$)   c. *Who must come?* ($\pi \gg \Box$)
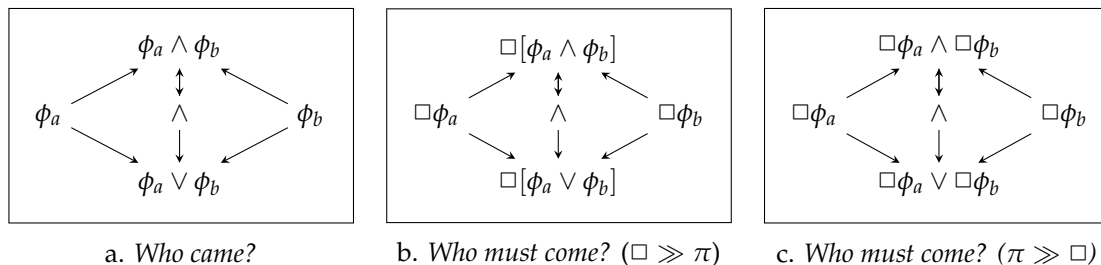
Figure 4: Answer spaces of non-modalized questions and *should*-questions

However, the assumptions made in this section cannot fully explained why MS interpretations are *only* available in *can*-questions: it's possible to satisfy conditions (A) and (B) even if the question doesn't contain an existential modal. For a concrete example, consider the non-modalized question in (52). Here the predicate *formed a team* is 'stubbornly collective' (Xiang 2021b): in contrast to *formed teams*, *formed a team* doesn't admit a covered/(semi-)distributive interpretation. Although the answer space derived in the higher-order interpretation is closed under conjunction, the true answers allowed in the first-order interpretation are logically independent, as seen in (52a). Hence, without reasons to rule out cases like (52a), the assumptions made in this section would predict a MS interpretation. (Q_w abbreviates the set of propositional answers true in w.)

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

    a. First-order interpretation: 'Which plurality of children x is such that x formed a team?'
       $Q_w = \{\phi_{a \oplus b}, \phi_{c \oplus d}\}$ (ϕ_x abbreviates 'x formed a team')

    b. Higher-order interpretation: 'Which GQ π over children is such that π formed a team?'
       $Q_w = \{\phi_{a \oplus b}, \phi_{c \oplus d}, \phi_{a \oplus b} \wedge \phi_{c \oplus d}, \phi_{a \oplus b} \vee \phi_{c \oplus d}\}$

There are two ways to solve the over-generation problem in (52). One way is to enrich the answer space of the first-order interpretation so that it is closed under conjunction. Fox (2018, 2020) proposes that *wh*-phrases may quantify over higher-order pluralities. Accordingly, the conjunctive answer '$a + b$ formed a team, and $c + d$ formed a team' is derived based on the higher-order plurality $\{\{a, b\}, \{c, d\}\}$ (for more details, see f.n. 25 in Sect. 5.2.1). In this proposal, if higher-order plurality answers are available in first-order interpretations, the answer space for (52a) is closed under conjunction.

The other way is to rule out (52a) by an independent constraint. As I will argue in Sect. 5.1.2, the over-generation problem arises not only in questions with a collective predicate, but also in questions with a uniqueness effect (e.g., *Which boy came?*) and questions with an existential indefinite (e.g., *Which movie(s) did one of the boys watch?*). It is more appealing to rule out these over-generating cases uniformly. More generally, this problem is a penalty for abandoning Dayal's exhaustivity presupposition: if question interpretations had to be exhaustive, any interpretation that does not render the existence of an exhaustive true answer would be deviant. To predict the distribution of MS, I will argue for a condition called 'Relativized Exhaustivity' which works like Dayal's exhaustivity presupposition except in modalized questions. According to this condition, only *can*-questions can have an interpretation that allows for multiple max-informative true answers (Sect. 6.2).

4.3.   Deriving conjunctive-MA interpretations

Recall that the MA answer to a *can*-question can be expressed either as a conjunction or as a disjunction. I argue that the two forms of MA answers correspond to two distinct MA interpretations.

(53)   Who can chair the committee?

    a.   Andy can.   (MS)

    b.   Andy can, and Billy can.   (Conjunctive-MA)

    c.   Andy or Billy.   (Disjunctive-MA)

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

(54)   Who can chair the committee?

    a. $\Diamond \gg \pi$: MS   b. $\pi \gg \Diamond$: Conjunctive-MA
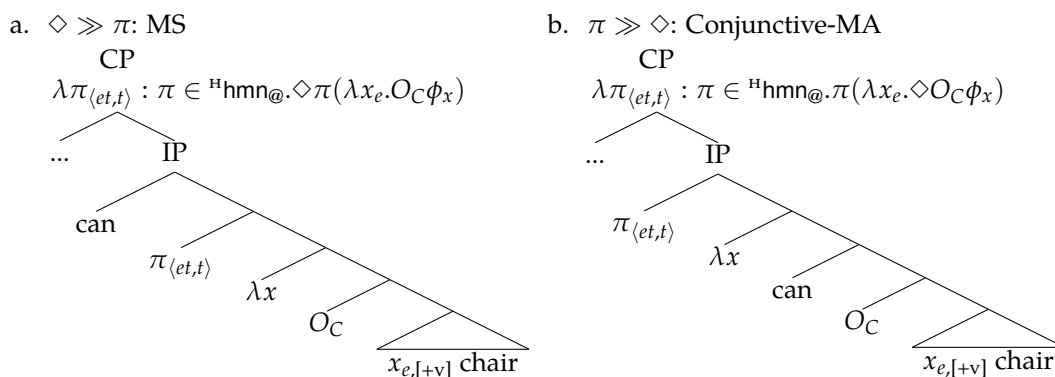
The answer space yielded from (54a/b) is illustrated in Figure 5a/b. (Figure 5a is identical to Figure 3.) Legends and abbreviations are the same as in Figure 3.

24

(Context: Only Andy and Billy can chair the committee. Only single-chairing is allowed.)

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

$$\diamond O_C\phi_a \qquad \vee \qquad \diamond O_C\phi_b$$

$$\diamond(O_C\phi_a \vee O_C\phi_b)$$

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

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

$$\diamond O_C\phi_a \qquad \wedge \qquad \diamond O_C\phi_b$$

$$\diamond O_C\phi_a \vee \diamond O_C\phi_b$$

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

Figure 5: Answer spaces yielded from (54a,b)

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

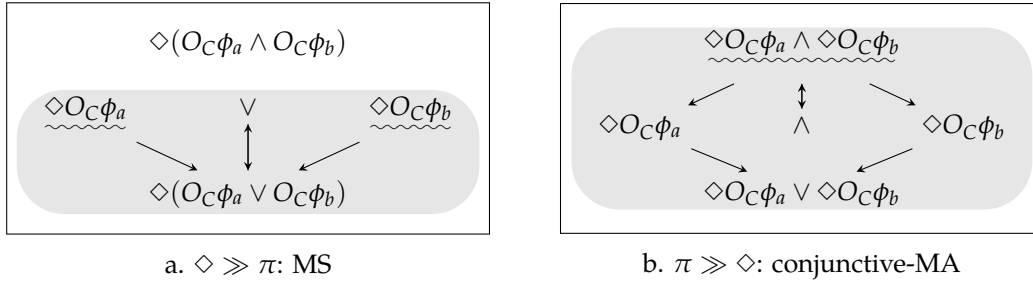### 4.4. Deriving disjunctive-MA interpretations

MA answers to *can*-questions are more naturally expressed as disjunctions. As exemplified in (55), in responding to a *can*-question, an elided disjunction may convey either epistemic ignorance or universal free-choice (FC), used as a partial answer and a mention-few/all answer, respectively.

(55)   Who can teach Intro Chinese?

    a.   Andy or Billy.... but I don't know who.   (Ignorance: partial)
        ⤳ *Either Andy or Billy can teach Intro Chinese, but I don't know who.*   ($\diamond\phi_a \vee \diamond\phi_b$)

    b.   Andy or Billy.   (FC: mention-few/all)
        ⤳ *Andy can teach Intro Chinese, and Billy can teach Intro Chinese, too.*   ($\diamond\phi_a \wedge \diamond\phi_b$)

I argue that the MS/MA ambiguity in *can*-questions has the same origin as the ignorance/FC-ambiguity in *can*-disjunctions: in a *can*-question, a disjunctive-MA interpretation arises if and only if the disjunctive answers are turned into universal FC statements.

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

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

The Mandarin particle *dou* has various uses. These uses can be disambiguated by the form of the occurring environment and the meaning of the associated expression. In particular, two uses in *can*-questions and in *can*-disjunctions exhibit an interesting parallel.

25

In a *can*-question, as seen in (56), associating *dou* with the *wh*-phrase across the modal *keyi* 'can' blocks the MS interpretation. (Underling marks the expression associated with *dou*.)

(56) a. (**Dou**) <u>shui</u> keyi jiao jichu hanyu?
          (DOU) who can teach Intro Chinese

          Without *dou*: 'Who can teach Intro Chinese?'   (✓MS, ✓MA)
          With *dou*: 'Who all can teach Intro Chinese?'   (✗MS, ✓MA)

     b. Mali (*dou*) keyi zai <u>nali</u>   mai dao kafei?
          Mary (DOU) can at where buy get coffee

          Without *dou*: 'Where can Mary get coffee?'   (✓MS, ✓MA)
          With *dou*: 'Where all can Mary get coffee?'   (✗MS, ✓MA)

I call *dou* in this use descriptively an 'exhaustivity marker' (after Beck and Rullmann 1999) and assume the LFs in (57) for the above two sentences.[15] In these LFs, *dou* appears within the IP and is associated with the *wh*-trace in its c-commanding domain.[16]

(57) a. $[_{CP}$ who$_i$ C^0 $[_{IP}$ *dou*$_C$ $[_{VP}$ $t_{i,[+v]}$ can teach Intro Chinese $]]]$
     b. $[_{CP}$ where$_i$ C^0 $[_{IP}$ Mary$_j$ *dou*$_C$ $[_{VP}$ t_j can get coffee $t_{i,[+v]}$ $]]]$

In *can*-declaratives, associating *dou* with a pre-verbal disjunction evokes a conjunctive FC inference, as exemplified in (58). Note that these two sentences do not give rise to an exclusive inference; for

---

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

(i) a. Yuehan dou keyi qu <u>nali</u>   mai kafei?
       John   DOU can go where buy coffee?

       'Where all can John buy coffee?" (MA)

     b. Yuehan <u>mei-ci</u>   dou keyi qu nali   mai kafei?
       John   each-time DOU can go where buy coffee?

       'In each case, where can John can buy coffee?' (Pair-list MS)
       'John always can buy coffee from where?' (Individual MS)

[16]The following explains why the exhaustivity marker *dou* has this syntax. First, as seen in (56b,b'), *dou* must appear on the right side of the subject unless the subject is an interrogative *wh*-item. Since a non-interrogative subject stays within IP, this fact argues that *dou* is also within IP.

(56) b'. (*Dou) Mali keyi zai <u>nali</u>   mai dao kafei?
          (*DOU) Mary can nat where buy get coffee

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

(i) (Context: John can give Mary either all the apples or some (but not all) of the cookies; the choice is up to him.)

     a. Yuehan dou keyi ba <u>shenme</u> gei Mali?   b. Yuehan keyi ba <u>shenme</u> dou gei Mali?
       John   DOU can BA what   give Mary   John   can BA what   DOU give Mary

       'What all is John allowed to give to Mary?"   'What x is s.t. John can give all of x to Mary?'
       Proper reply: 'The apples or some of the cookies.'   Proper reply: 'The apples.'

example, (58a) doesn't imply that no one else can teach Intro Chinese. This fact argues that FC is independent from exclusivity.

(58) a. <u>Yuehan huozhe Mali</u> (*dou*) keyi jiao   jichu hanyu.
          John   or   Mary (ᴅᴏᴜ) can teach Intro Chinese

   Without *dou*: 'Either John or Mary can teach Intro Chinese.'   (Ignorance)
   With *dou*: 'Both John and Mary can teach Intro Chinese.'   (Universal FC)

   b. Mali zai <u>Xingbake huozhe Maidanglao</u> (*dou*) keyi mai dao kafei.
       Mary at Starbucks or   McDonalds (ᴅᴏᴜ) can buy get coffee

   Without *dou*: 'From either SB or MD, Mary can get coffee.'   (Ignorance)
   With *dou*: 'From both SB and MD, Mary can get coffee.'   (Universal FC)

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

(59) a. <u>Yuehan huozhe Mali</u>   dou keyi/*bixu jiao   jichu hanyu.
          John   or   Mary ᴅᴏᴜ can/*must teach intro Chinese

   b. <u>Yuehan huozhe Mali</u>   (*dou) jiao   -guo jichu hanyu.
       John   or   Mary (*ᴅᴏᴜ) teach -ᴇxᴘ intro Chinese

(60) a. Anyone can/*must teach Intro Chinese.

   b. * Anyone taught Intro Chinese.

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

4.4.2. Derive the FC-trigger use: *Dou* as an anti-exhaustification operator

Kratzer and Shimoyama (2002) derive the FC inference as a result of 'anti-exhaustification'. Their idea is as follows. The speaker prefers $\diamondsuit(\phi \vee \psi)$ to the two stronger alternatives $\diamondsuit\phi$ and $\diamondsuit\psi$ because she is unhappy with the strengthened meaning of these alternatives. Yet, the reason cannot be that the speaker is unhappy with the non-strengthened meaning of these alternatives, because negating both $\diamondsuit\phi$ and $\diamondsuit\psi$ yields an inference that contradicts $\diamondsuit(\phi \vee \psi)$. Hence, uttering $\diamondsuit(\phi \vee \psi)$ implicates that the strengthened meaning of each stronger alternative is false, which is therefore called 'anti-exhaustification'.

Xiang (2020) analyzes the Mandarin particle *dou* as an anti-exhaustification operator. Roughly, as schematized in (61), *dou* affirms the prejacent and negates the (innocent-exclusion (IE-)based) exhaustification of each 'non-innocently-excludable alternative' except the prejacent itself.[17]

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

---

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

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

(62) 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
$O_C^{\text{IE}} = \lambda q \lambda w. q(w) = 1 \wedge \forall r \in \text{IExcl}(q, C)[r(w) = 0]$

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

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

For the narrow-scope $\Diamond$-disjunction $\Diamond(\phi \vee \psi)$, the subdomain-alternatives $\Diamond\phi$ and $\Diamond\psi$ are not I-excludable because $\{\neg\Diamond\phi, \neg\Diamond\psi\} \cup \{\Diamond(\phi \vee \psi)\}$ is inconsistent (or say, $\Diamond(\phi \vee \psi) \wedge \neg\Diamond\phi \wedge \neg\Diamond\psi = \bot$). In contrast, for the narrow-scope $\Box$-disjunction $\Box(\phi \vee \psi)$, the subdomain-alternatives $\Box\phi$ and $\Box\psi$ are I-excludable because $\Box(\phi \vee \psi) \wedge \neg\Box\phi \wedge \neg\Box\psi \neq \bot$. A summary is given in Table 2, where 'Neither' means neither weaker nor I-excludable. The application of *dou* is semantically vacuous unless the prejacent sentence has alternatives that are neither weaker nor I-excludable.

	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 $\Box\phi, \Box\psi$ $\Diamond\phi, \Diamond\psi$		
$\phi \vee \psi$ $\Diamond\phi \vee \Diamond\psi, \Diamond(\phi \vee \psi)$ $\Box\phi \vee \Box\psi$		ϕ, ψ $\Diamond\phi, \Diamond\psi$ $\Box\phi, \Box\psi$	
$\Box(\phi \vee \psi)$			$\Box\phi, \Box\psi$

Table 2: Subdomain-alternatives of conjunctions and disjunctions

I henceforth call the alternatives that participate in anti-exhaustification 'anti-excludable alternatives' (abbreviation: AntiExcl).[18,19] The semantics in (61) is re-written as follows:

(64) $\llbracket dou_C \rrbracket = \lambda p \lambda w : p(w) = 1 \wedge \forall q \in \text{AntiExcl}(p, C)[O_C^{\text{IE}}(q)(w) = 0]$
$\text{where AntiExcl}(p, C) = (C - \text{IExcl}(p, C)) - \{p\}$

The FC inference of (58a) is computed as in (65). Associated with *or*, *dou* quantifies over the set of the domain-alternatives of its prejacent, which includes two anti-excludable alternatives (viz., the two

---

[18] In Xiang 2020, the alternatives that participate in anti-exhaustification are called 'sub-alternatives'. The function of *dou* varies by what alternatives count as sub-alternatives. Briefly, sub-alternatives are primarily alternatives logically weaker than the prejacent, but in particular syntactic and prosodic environments, the ordering source in defining sub-alternatives gets shifted from logical strength to innocent excludability or likelihood, yielding different functions of *dou*. In particular, the FC-trigger use of *dou* arises when sub-alternatives are defined in terms of innocent excludability, as the definition in (61).

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

disjuncts). Employing *dou* affirms the prejacent disjunctive sentence and negates the exhaustification of each disjunct, yielding a conjunctive-FC inference.

(65) [*dou$_C$* [$_s$ John or$_{[+D]}$ Mary can teach Intro Chinese]]

    a. $[\![S]\!] = \Diamond\phi_j \vee \Diamond\phi_m$ (ϕ_x abbreviates 'x teach Intro Chinese')

    b. $C = \text{D-Alt}(S) = \{\Diamond\phi_j, \Diamond\phi_m, \Diamond\phi_j \vee \Diamond\phi_m\}$

    c. $\text{AntiExcl}([\![S]\!], C) = \{\Diamond\phi_j, \Diamond\phi_m\}$

    d. $[\![dou_C(S)]\!] \Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge \neg O_C^{\text{IE}}\Diamond\phi_j \wedge \neg O_C^{\text{IE}}\Diamond\phi_m$

                    $\Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge [\Diamond\phi_j \rightarrow \Diamond\phi_m] \wedge [\Diamond\phi_m \rightarrow \Diamond\phi_j]$

                    $\Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge [\Diamond\phi_j \leftrightarrow \Diamond\phi_m]$

                    $\Leftrightarrow \Diamond\phi_j \wedge \Diamond\phi_m$

Due to separate syntactic constraints, *dou* cannot be associated with a narrow-scope disjunction in overt syntax. For example, it cannot grammatically occur in the Mandarin translation of 'I (*dou) can ask John or Mary to teach Intro Chinese' (*can* ≫ *or*). Yet, nothing prevents *dou* from being associated with a narrow-scope expression at LF. In light of Matthewson's (2001) '*No Variation Hypothesis*' that semantics has no cross-linguistic variation, I assume that anti-exhaustification is a cross-linguistic source for universal FC. I also assume that anti-exhaustification is realized by the application of the overt *dou* or the covert ᴅᴏᴜ.[20]

It's worth discussing how the proposed analysis of FC differs from the analyses of Fox 2007 and Chierchia 2006, 2013, which also adapt anti-exhaustification to the grammatical approach of implicatures. Fox (2007) analyzes anti-exhaustivity as a consequence of applying the IE-based exhaustification recursively. This analysis predicts that FC and exclusivity are derived at the same time. In contrast, drawing on evidence from (58), I argue to derive these two inferences separately: FC comes from the application of *dou*/ᴅᴏᴜ, while exclusivity comes from the application of a separate IE-based exhaustification operator. This distinction also affects question interpretation: applying recursive exhaustification to the question nucleus inevitably yields strong exhaustivity, while applying ᴅᴏᴜ allows for weak exhaustivity.

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

4.4.3. Modal obviation of FC: Relativized Exclusivity

If anti-exhaustification has 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

---

[20]Mingming Liu (pers. comm.) argues that *dou* itself doesn't assert anti-exhaustivity; instead, by defining *dou* as an *even*-like operator (Liu 2016a), Liu argues that the presupposition of *dou*/*even* — that the prejacent has a more-likely alternative — forces to apply a separate anti-exhaustification operator to strengthen the prejacent disjunction into FC. This idea is reviewed in Xiang 2020: Appendix B. For the interests of this paper, it only matters that associating *dou* with a disjunction forces anti-exhaustification.

including $\Box\phi \vee \Box\psi$. For example, affirming $\phi \vee \psi$ and anti-exhaustifying the two alternatives ϕ and ψ return $\phi \wedge \psi$. This prediction is obviously wrong. Hence, there must be constraints on when anti-exhaustification is available, which cause the modal obviation effect that universal FC is only possible in sentences with an existential modal (i.e., $\Diamond(\phi \vee \psi)$ and $\Diamond\phi \vee \Diamond\psi$).

Explanations to the modal obviation effect fall into two groups. One attributes the unavailability of FC in $\phi \vee \psi$ to the contradiction between the FC inference $\phi \wedge \psi$ and the scalar implicature $\neg(\phi \wedge \psi)$ (Chierchia 2013; Bar-Lev and Fox 2020). The other concerns the (in-)compatibility of interpreting the subdomain-alternatives with local exhaustification (Menéndez-Benito 2010; Dayal 2013; Xiang 2020): for example, $O\phi$ and $O\psi$ are mutually exclusive, while $\Diamond O\phi$ and $\Diamond O\psi$ are not. Following the latter strategy, I propose a novel definedness condition for anti-exhaustification. For a review of an analysis by Fox (2018, 2020) who uses the former strategy, see Sect. 5.2.1.

To account for modal obviation, I propose that the anti-exhaustification operator *dou*/ᴅᴏᴜ has a '*Relativized Exclusivity (RelExcl)*' presupposition (cf. the *Viability Constraint* of Dayal (2013)): for every anti-excludable alternative ϕ stronger than the prejacent, every minimal set of accessible worlds that verifies ϕ also verifies the (IE-based) exhaustification of ϕ. This presupposition is schematized as follows, where M is a modal base, and C/C' denotes a subset of the alternatives of $[\![S]\!]^M/[\![S]\!]^{M'}$:[21,22,23]

(66) *Relativized Exclusivity*
   $[\![dou_C(S)]\!]^M/[\![\text{ᴅᴏᴜ}_C(S)]\!]^M$ is defined in w only if
   $\forall\phi[[\![\phi]\!]^M \in \text{AɴᴛɪExᴄʟ}([\![S]\!]^M, C) \wedge [\![\phi]\!]^M \subset [\![S]\!]^M$
   $\quad \to \forall M'_{\langle s,st\rangle}[M'_w$ is a minimal subset of M_w s.t. $[\![\phi]\!]^{M'}(w) = 1 \to O^{\text{ɪᴇ}}_{C'}([\![\phi]\!]^{M'})(w) = 1]]$
   (For any sentence ϕ such that $[\![\phi]\!]^M$ is an anti-excludable alternative of S stronger than $[\![S]\!]^M$, 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.)

---

[21]For 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 such that 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'.

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

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

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

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

(i) $\forall\phi[[\![\phi]\!]^M \in (\text{AɴᴛɪExᴄʟ}([\![S]\!]^M, C) \wedge [\![\phi]\!]^M \subset [\![S]\!]^M \to \exists M'[M'_w \subseteq M_w \wedge O^{\text{ɪᴇ}}_{C'}([\![\phi]\!]^{M'})(w) = 1]]$
   (For any sentence ϕ s.t. $[\![\phi]\!]^M$ is an anti-excludable alternative of S stronger than $[\![S]\!]^M$, there is a modal base M' such that M'_w is a subset of w-accessible worlds that verifies the exhaustification of ϕ in w.)

The main difference between RelExcl and Viability is the following: in (69b) where the $\Diamond$-disjunction is parsed without local exhaustification, the Viability constraint 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, 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.

The following explains how the RelExcl presupposition accounts for modal obviation. For simplicity, $O_C^{IE}\phi$ is abbreviated as $O\phi$.

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

(67)　* DOU$_C$ [John or Mary teach Intro Chinese]
　　　DOU$_C[\overline{\phi_j \vee \phi_m}]$ inevitably violates RelExcl

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

(68)　* DOU$_C$ [John or Mary must teach Intro Chinese]
　　　DOU$_C[\overline{\Box\phi_j \vee \Box\phi_m}]$ inevitably violates RelExcl

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

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

(69)　DOU$_C$ [John or Mary can teach Intro Chinese]

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

As in (69b), if the sentence is parsed without local exhaustification, RelExcl is satisfied only in w. The minimal set of w-accessible worlds that verifies $\Diamond\phi_j$, namely $\{w_1\}$, also verifies its exhaustification $O\Diamond\phi_j$; likewise for $\Diamond\phi_m$, $\{w_2\}$ verifies both $\Diamond\phi_m$ and its 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 $O\Diamond\phi_m$.

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

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

c.　Given M, RelExcl is satisfied in w and w' but violated in w''. More generally:
　　DOU$_C[\Diamond O\phi_j \vee \Diamond O\phi_m] = \Diamond O\phi_j \wedge \Diamond O\phi_m$, defined only if $\Diamond O\phi_j \wedge \Diamond O\phi_m$

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

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

   b. $\text{DOU}_C[\Diamond(O\phi_j \lor O\phi_m)] = \Diamond O\phi_j \land \Diamond O\phi_m$, defined only if $\Diamond O\phi_j \land \Diamond O\phi_m$

To sum up, anti-exhaustification is a cross-linguistic source for universal FC. This operation is lexicalized as *dou*/DOU. When associated with a disjunction, *dou*/DOU affirms the prejacent and negates the IE-based exhaustification of each anti-excludable alternative, giving rise to a universal FC inference. Moreover, the anti-exhaustification operator *dou*/DOU presupposes Relativized Exclusivity, which can be satisfied only in the presence of an existential modal.

4.4.4. Derive disjunctive-MA

Drawing on the parallel uses of *dou* in *can*-disjunctions and *can*-questions, I argue that the disjunctive-MA interpretations of *can*-questions are derived by applying anti-exhaustification within the question nucleus. The LFs are structured as in (71). The δ-nodes in (71a) and (71b) are simply the IP nodes in the LF of higher-order MS and in the LF of conjunctive-MA, respectively. These two δ-nodes only differ in the scope of the higher-order *wh*-trace π relative to the modal verb *can*. In both LFs, a covert anti-exhaustification operator DOU is applied to the δ-node and is associated with the trace π. The variable C' carried by DOU denotes a set of variable-alternatives of δ, which is equivalent to the answer space derived in the absence of DOU.

(71) Who can chair the committee? (Disjunctive-MA)
      (Context: Only Andy and Billy can chair, and co-chairing is disallowed.)
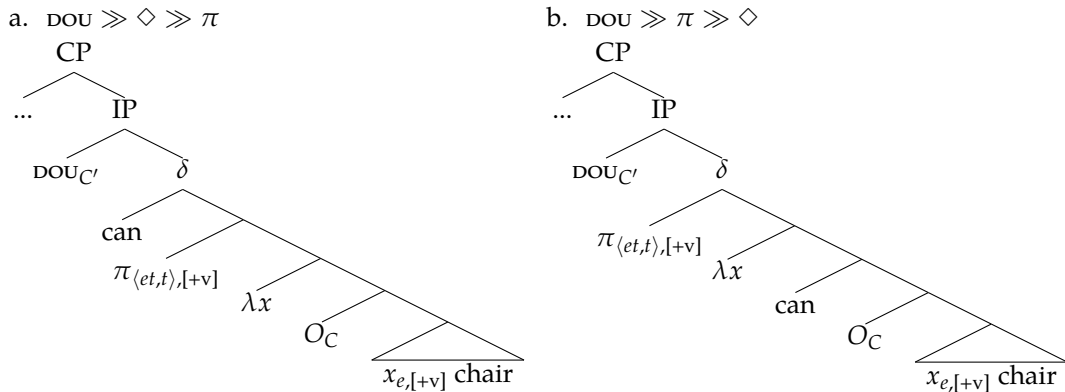
The answer space yielded from the LF (71a/b) is illustrated in Figure 7a/b. The corresponding answer space yielded in the absence of DOU is repeated in Figure 6a/b.
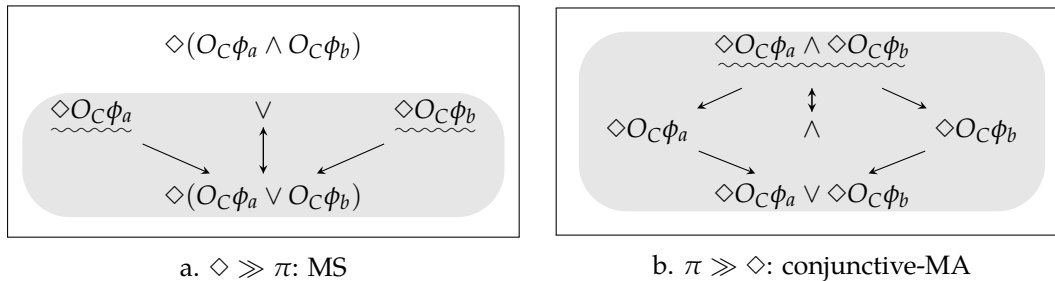

Figure 6: Answer spaces of (71a,b) yielded in the absence of DOU

32

$$\text{DOU}_{C'}\Diamond(O_C\phi_a \land O_C\phi_b)$$

$$\text{DOU}_{C'}\Diamond O_C\phi_a \qquad \land \qquad \text{DOU}_{C'}\Diamond O_C\phi_b$$

$$\underline{\text{DOU}_{C'}\Diamond(O_C\phi_a \lor O_C\phi_b)}$$

$$\text{DOU}_{C'}[\Diamond O_C\phi_a \land \Diamond O_C\phi_b]$$

$$\text{DOU}_{C'}\Diamond O_C\phi_a \qquad \land \qquad \text{DOU}_{C'}\Diamond O_C\phi_b$$

$$\underline{\text{DOU}_{C'}[\Diamond O_C\phi_a \lor \Diamond O_C\phi_b]}$$

a. $\Diamond \gg \pi$: disjunctive-MA

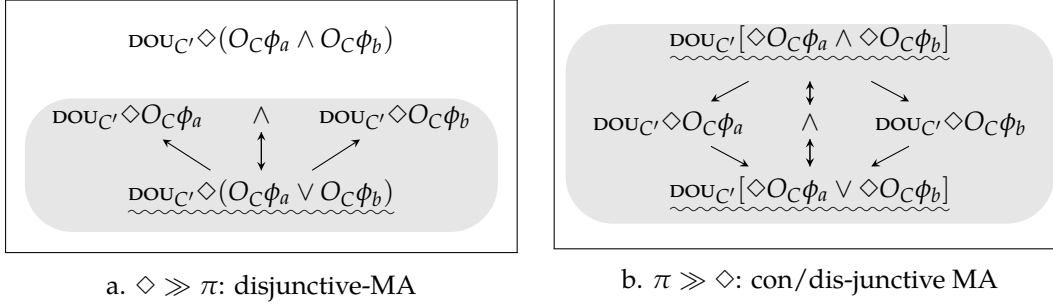b. $\pi \gg \Diamond$: con/dis-junctive MA

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

In both Figures 7a and 7b, the application of DOU in the individual answers and the conjunctive answers are semantically vacuous: the individual answers have no anti-excludable alternative; the conjunctive answers have anti-excludable alternatives, but the anti-exhaustification inference (underlined) is entailed by the prejacent. However, as computed in (72c) and (73c), in both cases, DOU strengthens the disjunctive answer into a universal-FC answer that is logically equivalent to the conjunction of the two individual answers, making the answer space closed under conjunction.

(72) For DOU $\gg \Diamond \gg \pi$ (answer space in Figure 7a)

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

    b. Compute the conjunctive answer

        i. $\textsc{AntiExcl}(\Diamond(O_C\phi_a \land O_C\phi_b), C') = C' - \{\bot\}$

        ii. $\text{DOU}_{C'}\Diamond(O_C\phi_a \land O_C\phi_b)$
$$\Leftrightarrow \Diamond(O_C\phi_a \land O_C\phi_b) \land \underline{\forall\phi \in (C' - \{\bot\})[\neg O_C^{\text{IE}}\phi]}$$
$$\Leftrightarrow \Diamond(O_C\phi_a \land O_C\phi_b)$$
$$\Leftrightarrow \bot$$

    c. Compute the disjunctive answer

        i. $\textsc{AntiExcl}(\Diamond(O_C\phi_a \lor O_C\phi_b), C') = \{\Diamond O_C\phi_a, \Diamond O_C\phi_b\}$

        ii. $\text{DOU}_{C'}\Diamond(O_C\phi_a \lor O_C\phi_b)$
$$\Leftrightarrow \Diamond(O_C\phi_a \lor O_C\phi_b) \land \neg O_{C'}^{\text{IE}}\Diamond O_C\phi_a \land \neg O_{C'}^{\text{IE}}\Diamond O_C\phi_b$$
$$\Leftrightarrow \Diamond(O_C\phi_a \lor O_C\phi_b) \land [\Diamond O_C\phi_a \to \Diamond O_C\phi_b] \land [\Diamond O_C\phi_b \to \Diamond O_C\phi_a]$$
$$\Leftrightarrow \Diamond(O_C\phi_a \lor O_C\phi_b) \land [\Diamond O_C\phi_a \leftrightarrow \Diamond O_C\phi_b]$$
$$\Leftrightarrow \Diamond O_C\phi_a \land \Diamond O_C\phi_b$$

(73) For DOU $\gg \pi \gg \Diamond$ (answer space in Figure 7b)

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

    b. Compute the conjunctive answer

        i. $\textsc{AntiExcl}(\Diamond O_C\phi_a \land \Diamond O_C\phi_b, C') = \{\Diamond O_C\phi_a, \Diamond O_C\phi_b\}$

        ii. $\text{DOU}_{C'}[\Diamond O_C\phi_a \land \Diamond O_C\phi_b]$
$$\Leftrightarrow [\Diamond O_C\phi_a \land \Diamond O_C\phi_b] \land \underline{\neg O_{C'}^{\text{IE}}\Diamond O_C\phi_a \land \neg O_{C'}^{\text{IE}}\Diamond O_C\phi_b}$$
$$\Leftrightarrow \Diamond O_C\phi_a \land \Diamond O_C\phi_b$$

    c. Compute the disjunctive answer

        i. $\textsc{AntiExcl}(\Diamond O_C\phi_a \lor O_C\phi_b), C') = \{\Diamond O_C\phi_a, \Diamond O_C\phi_b\}$

        ii. $\text{DOU}_{C'}\Diamond(O_C\phi_a \lor O_C\phi_b)$
$$\Leftrightarrow \Diamond(O_C\phi_a \lor O_C\phi_b) \land \neg O_{C'}^{\text{IE}}\Diamond O_C\phi_a \land \neg O_{C'}^{\text{IE}}\Diamond O_C\phi_b$$

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

This analysis straightforwardly explains why only *can*-questions admit universal FC-disjunctive answers. In the above *can*-question, the RelExcl presupposition of ᴅᴏᴜ yields a definedness condition equivalent to the FC meaning of the disjunctive answer. However, for other *wh*-questions, their disjunctive answers cannot be strengthened into FC statements via anti-exhaustification. For example, for the $\Box$-question in (74), a disjunctive answer may express ignorance ($\Box\phi_f \vee \Box\phi_r$) or existential FC ($\Box[\phi_f \vee \phi_r]$), but not universal FC ($\Box\phi_f \wedge \Box\phi_r$): applying ᴅᴏᴜ to the disjunctive answer either causes an inevitable violation of RelExcl or is semantically vacuous.

(74)   Q: 'What does John have to read?' A: 'The French novels or the Russian novels.'

   a. If $\pi \gg \Box$: Ignorance

   ᴅᴏᴜ$_C[\Box\phi_f \vee \Box\phi_r]$ isn't a possible answer, because it inevitably violates RelExcl.

   b. If $\Box \gg \pi$: Existential FC

   ᴅᴏᴜ$_C[\Box(\phi_f \vee \phi_r)] = \Box(\phi_f \vee \phi_r)$, because $\Box(\phi_f \vee \phi_r)$ has no anti-excludable alternative.

4.5.   Interim summary

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

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

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

5.   A dilemma

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

Hirsch and Schwarz (2020) novelly observe that the uniqueness inference triggered by a singular *which*-phrase can be interpreted locally. This observation challenges Dayal's exhaustivity presuppo-

sition as well as a modified exhaustivity presupposition offered by Fox (2018, 2020). They argue to abandon Dayal's exhaustivity presupposition and propose that the uniqueness effects come from the lexicon of the determiner *which*. This section discusses the dilemma and reviews the two alternative accounts of uniqueness by Fox (2018, 2020) and Hirsch and Schwarz (2020).

5.1. Dayal's exhaustivity presupposition: merits and challenges

5.1.1. Uniqueness and singular *wh*-questions

Wh-questions with a singular *which*-phrase (called 'singular *wh*-questions') are subject to uniqueness. For example, question (75) can be felicitously uttered only if the speaker believes that this question has only one true answer. This uniqueness effect is standardly explained by 'Dayal's exhaustivity presupposition', which says that a question is defined only if it has an exhaustive true answer.

(75) Which child came? ⤳ *Only one of the children came.*

Dayal (1996) encodes this requirement to question semantics as a presupposition of the answerhood-operator, defined as follows:

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

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 including also sums (Sharvy 1980; Link 1983). Incorporating this distinction into *wh*-questions, Dayal argues that the answer space of a plural *wh*-question includes sum-based propositions while the answer space of a singular *wh*-question doesn't. For example, in a context where two children came, the plural *wh*-question (77a) has an exhaustive true answer derived based on the sum of two children (viz., $a \oplus b$). However, in the same context, the singular *wh*-question (77b) doesn't have an exhaustive true answer and violates the exhaustivity presupposition.

(77) (Among the considered children, only Andy and Billy came. The questioner knows that multiple children came, but she doesn't know who they are.)
   a.   Which children came? $\{\lambda w.\text{came}_w(a), \lambda w.\text{came}_w(b), \lambda w.\text{came}_w(a \oplus b)\}$
   b.   # Which child came? $\{\lambda w.\text{came}_w(a), \lambda w.\text{came}_w(b)\}$

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

Xiang (2021b) observes a disjunction–conjunction asymmetry in higher-order *wh*-quantification: singular *which*-phrases may quantify over boolean disjunctions, despite not boolean conjunctions. For example in (78), the disjunctive answer to a singular □-question may be interpreted under the scope of the universal modal. The narrow-scope interpretation of the disjunctive answer argues that the quantification domain of *which textbook* contains boolean disjunctions over atomic books.

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

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

(79) What can I use for this class?

    a.   *Heim & Kratzer* or *Meaning & Grammar*.

    b.   *Heim & Kratzer* and *Meaning & Grammar*.

(80) Which textbook can I use for this class?

    a.   *Heim & Kratzer* or *Meaning & Grammar*.

    b.   # *Heim & Kratzer* and *Meaning & Grammar*.

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

In addition to singular *which*-phrases, numeral-modified *which*-phrases also trigger a uniqueness presupposition, as exemplified in (81). The same analysis applies: *which two children* may quantify over the 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. In (81b) for example, if *which two children* may quantify over boolean conjunctions, the question would admit a conjunctive answer expressing the composition of two teams, formed out of $(a \oplus b)^\Uparrow \cap (c \oplus d)^\Uparrow$.

(81) a. Which two children came?   ⤳ *Only two of the children came.*

       b. Which two children formed a team?   ⤳ *Only one pair of the children formed any team.*

5.1.2. Consequences of abandoning Dayal's exhaustivity presupposition

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

Despite our interests in MS, it is too costly to abandon Dayal's exhaustivity presupposition. In addition to explaining the uniqueness effects in questions, this exhaustivity presupposition has played an important role in accounting for a variety of linguistic phenomena, such as the negative island effects in degree constructions (Fox and Hackl 2007; Spector and Abrusán 2011; Abrusán 2014) and the maximality effects of definites (von Fintel et al. 2014). Moreover, it can rule out many unwanted question interpretations; without further restrictions, abandoning this presupposition and applying Fox's answerhood would let these unwanted interpretations be rebranded as acceptable MS interpretations. The following discusses three such problematic cases.

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

(82)   *Which child came? / Which two children formed a team? / Did you invite Andy, Billy, or Cindy?*

      a. With Dayal's exhaustivity presupposition: MA with uniqueness ✓

      b. Without Dayal's exhaustivity presupposition: MS without uniqueness ✗

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

(83) *Which children formed a team?*

      a. With Dayal's exhaustivity presupposition: MA without uniqueness ✓

      b. Without Dayal's exhaustivity presupposition: Ambiguous between MS and MA ✗

Case 3: *wh*-questions with an existential indefinite. *Wh*-questions with an existential indefinite have choice interpretations and individual interpretations. In (84), the choice interpretation requests to specify a boy-movie(s) pair, while the individual interpretation requests to specify some movies but does not request to specify the names of the boys who watched these movies.

(84) (Among the relevant boys, 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, 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 such that one of the boys watched y?' (Individual)

        i.   One (of the boys) watched *Ironman* and *Spiderman*, and one watched *Hulk*.

        ii. # One (of the boys) watched *Hulk*.

Existing literature is focused on the choice interpretation. A few works on MS (George 2011; Fox 2013; Nicolae 2013; a.o.) treat the choice interpretation as a MS interpretation and claim that existential indefinites can license MS. However, a number of empirical distinctions argue that choice interpretations and MS interpretations should be treated separately (Appendix A). Here, let's switch our focus to the individual interpretation. In this interpretation, the indefinite is interpreted within the local IP. Here arises a problem similar to what we've seen in questions with a collective predicate. In principle, in analogous to *can*-questions, questions with an existential indefinite could have an LF involving local exhaustification, as in (85a). In the context given above, the true answers yielded from this LF are as in (85b), both of which are max-informative. However, as seen in (84b-ii), such non-exhaustive individual answers are deviant. This problem also applies to the LF in (86a) for the higher-order *wh*-quantification.

(85)   a. $[_{CP}$ which-movie(s) λy_e $[_{IP}$ one-boy λx_e $[$ O_C $[_{VP}$ x watched $y_{[+v]}$ $]]]]$

       b. $\{\lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{i} \oplus \mathsf{s})]], \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{h})]]\}$

(86)   a. $[_{CP}$ which-movie(s) $\lambda \pi_{\langle et,t \rangle}$ $[_{IP}$ one-boy λx_e $[$ π λy_e $[$ O_C $[_{VP}$ x watched $y_{[+v]}$ $]]]]]$

b. $\left\{ \begin{array}{c} \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{i} \oplus \mathsf{s})]], \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{h})]] \\ \lambda w.\exists x[\mathsf{boy}_w(x) \wedge [O_C[\mathsf{watch}_w(x, \mathsf{i} \oplus \mathsf{s})] \vee O_C[\mathsf{watch}_w(x, \mathsf{h})]]] \end{array} \right\}$

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

(87)   a. $[_{\mathrm{CP}}$ which-movie(s) $\lambda\pi_{\langle et,t\rangle}$ $[$ π λy_e $[_{\mathrm{IP}}$ one-boy λx_e $[$ O_C $[_{\mathrm{VP}}$ x watched $y_{[+\mathrm{v}]}$ $]]]]]$

     b. $\left\{ \begin{array}{c} \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{i} \oplus \mathsf{s})]] \wedge \exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{h})]] \\ \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{i} \oplus \mathsf{s})]] , \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{h})]] \\ \lambda w.\exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{i} \oplus \mathsf{s})]] \vee \exists x[\mathsf{boy}_w(x) \wedge O_C[\mathsf{watch}_w(x, \mathsf{h})]] \end{array} \right\}$

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

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

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

     b.   Which letter could be missing in *fo_m*?
        A or *r*. (Intended: 'The unique letter missing in *fo_m* could be *a* and could be *r*.')

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

(89)   Q: Which chapter should we assign to the students?
        $\rightsquigarrow$ *We shouldn't assign more than one chapter to the students.*
     A: Chapter 1 or chapter 2, either is good. $(\Box(\phi_{c_1} \vee \phi_{c_2}))$

Dayal's exhaustivity presupposition cannot account for the local uniqueness inference in (89). To see why, recall that the singular *which*-phrase only ranges 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 as in (90a). Next, assume the scenario in (90b) where local uniqueness is violated: among the three w-accessible worlds w_1, w_2, w_3, we assign only c_1 in w_1, only c_2 in w_2, but both c_1 and c_2 in w_3. In this scenario, the answer space has only one true member $\Box(\phi_{c_1} \vee \phi_{c_2})$, and hence Dayal's exhaustivity presupposition is not violated.

(90)   Which chapter should we assign to the students?
     $[_{\mathrm{CP}}$ which-chapter $\lambda\pi_{\langle et,t\rangle}$ $[$ should $[$ π λx_e $[_{\mathrm{IP}}$ we assign x to the students $]]]]$

     a. $Q = \{\Box\phi_{c_1}, \Box\phi_{c_2}, \Box(\phi_{c_1} \vee \phi_{c_2})\}$ (ϕ_x abbreviates 'we assign x to the students')
     b.   Let $M_w = \{w_1, w_2, w_3\}$ and $\mathsf{assign} = [\; w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\}, w_3 \rightarrow \{c_1, c_2\}, ... \;]$

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. This work covers a number of issues, including also the negative island effects of higher-order *wh*-quantification. The following will focus on the part that is directly relevant to the dilemma between MS and uniqueness.

Fox first re-writes Dayal's answerhood in terms of exhaustification as in (91): for a Hamblin set Q and a world w, applying Ans_{Dayal} returns the unique proposition p in Q such that the exhaustification of p relative to Q is true in w, 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 the prejacent.

(91) $\text{Ans}_{Dayal}(w)(Q) = \exists p \in Q[O_Q(p)(w) = 1].\iota p \in Q[O_Q(p)(w) = 1]$
$\qquad\qquad$ where $O_Q(p) = \lambda w.p(w) = 1 \wedge \forall q \in Q[p \nsubseteq q \rightarrow q(w) = 0]$

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

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

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

Let's see how this this operator derives FC and predicts the distribution of FC. In (93), applying O^{IE+II} to the modalized disjunction $\Diamond(\phi_a \vee \phi_b)$ yields a scalar implicature $\neg\Diamond(\phi_a \wedge \phi_b)$ by innocent exclusion as well as a FC inference $\Diamond\phi_a \wedge \Diamond\phi_b$ by innocent inclusion. In contrast, in (94), applying O^{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^{IE+II} may yield a FC inference for a disjunctive sentence only if the alternative set of this sentence is not closed under conjunction, such as in cases where the disjunction takes scope below an existential modal.

(93) Let $p = \Diamond(\phi_a \vee \phi_b)$ and $C = \{\Diamond\phi_a, \Diamond\phi_b, \Diamond(\phi_a \vee \phi_b), \Diamond(\phi_a \wedge \phi_b)\}$, then we have:
    a. $\text{IExcl}(p,C) = \{\Diamond(\phi_a \wedge \phi_b)\}$ and $\text{IIncl}(p,C) = \{\Diamond\phi_a, \Diamond\phi_b, \Diamond(\phi_a \vee \phi_b)\}$
    b. $O_C^{IE+II}(p) = \underbrace{\neg\Diamond(\phi_a \wedge \phi_b)}_{\text{IE}} \wedge \underbrace{[\Diamond\phi_a \wedge \Diamond\phi_b \wedge \Diamond(\phi_a \vee \phi_b)]}_{\text{II}} = \neg\Diamond(\phi_a \wedge \phi_b) \wedge \Diamond\phi_a \wedge \Diamond\phi_b$

(94) Let $p = \phi_a \vee \phi_b$ and $C = \{\phi_a, \phi_b, \phi_a \vee \phi_b, \phi_a \wedge \phi_b\}$, then we have:
    a. $\text{IExcl}(p,C) = \{\phi_a \wedge \phi_b\}$ and $\text{IIncl}(p,C) = \{\phi_a \vee \phi_b\}$
    b. $O_C^{IE+II}(p) = \underbrace{\neg[\phi_a \wedge \phi_b]}_{\text{IE}} \wedge \underbrace{[\phi_a \vee \phi_b]}_{\text{II}}$

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

(95) *Modified exhaustivity presupposition* (after Fox 2020)
    For any question with a Hamblin set Q, it 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.

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

To see how this analysis allows for MS, consider the *can*-question in (96). With three relevant individuals a,b,c, the answer space is as in (96a).[24] Applying $O_Q^{\text{IE+II}}$ point-wise to this answer space returns a set of mutually exclusive propositions as in (96b), referred to as 'the partition induced by Q'. The modified exhaustivity presupposition is satisfied in any world that is in the union of this partition, including worlds with multiple possible chairs.

(96) Who can chair the committee alone?

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

   (ϕ_x abbreviates 'x chairs the committee alone'.)

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

$= \left\{ \begin{array}{c} \Diamond\phi_a \wedge \neg\Diamond\phi_b \wedge \neg\Diamond\phi_c, \Diamond\phi_b \wedge \neg\Diamond\phi_a \wedge \neg\Diamond\phi_c, \Diamond\phi_c \wedge \neg\Diamond\phi_a \wedge \neg\Diamond\phi_b \\ \Diamond\phi_a \wedge \Diamond\phi_b \wedge \neg\Diamond\phi_c, \Diamond\phi_b \wedge \Diamond\phi_c \wedge \neg\Diamond\phi_a, \Diamond\phi_a \wedge \Diamond\phi_c \wedge \neg\Diamond\phi_b \\ \Diamond\phi_a \wedge \Diamond\phi_b \wedge \Diamond\phi_c \end{array} \right\}$

This partition-by-exhaustification analysis also avoids over-generating MS interpretations for questions with a stubbornly collective predicate. For example, in a multiple-team scenario, question (97) satisfies the exhaustivity presupposition only if its answer space is closed under conjunction. Such an answer space can be created if the *wh*-phrase ranges over higher-order pluralities (as assumed in Fox 2018, 2020, discussed in Sect. 4.2.3) or boolean conjunctions (as assumed in my proposal).[25]

(97) Which children formed a team?
    Andy and Billy formed a team, and Cindy and Danny formed a team.

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

---

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

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

(i) *Non-vacuity Principle* (Fox 2018, 2020)
    Every proposition p in Q is such that the exhaustification of p is identical to a cell in the partition induced by Q.

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

For illustration of uniqueness, consider the singular *wh*-question in (98). With two boys *a* and *b* in the discourse domain, the answer space is Q_{FO} in the first-order interpretation and Q_{HO} in the higher-order interpretation.[26] In (98b), applying point-wise exhaustification to Q_{FO} induces a partition that covers only the worlds where exactly one of the boys came, which does nicely predict uniqueness. However, in the higher-order interpretation, exhaustifying the non-modalized disjunctive answer $\phi_a \vee \phi_b$ with the O^{IE+II}-operator would strengthen it into a FC statement: distinct from the alternative set *C* in (94), here Q_{HO} doesn't contain the conjunction $\phi_a \wedge \phi_b$, and thus the individual answers ϕ_a and ϕ_b are I-includable to the disjunctive answer $\phi_a \vee \phi_b$. Hence in (98c), the partition induced by Q_{HO} covers also the worlds where both boys came, which therefore predicts no uniqueness effect. In short, uniqueness effect is not explained because the O^{IE+II}-operator over-generates FC answers for singular *wh*-questions with a higher-order interpretation.

(98) Which boy came?

    a. $Q_{FO} = \{\phi_a, \phi_b\}$ and $Q_{HO} = \{\phi_a, \phi_b, \phi_a \vee \phi_b\}$ (ϕ_x abbreviates '*x* came'.)

    b. $\{O^{IE+II}_{Q_{FO}}(p) \mid p \in Q_{FO}\} = \{\phi_a \wedge \neg\phi_b, \neg\phi_a \wedge \phi_b\}$ (Prediction: with uniqueness ✓)

    c. $\{O^{IE+II}_{Q_{HO}}(p) \mid p \in Q_{HO}\} = \{\phi_a \wedge \neg\phi_b, \neg\phi_a \wedge \phi_b, \phi_a \wedge \phi_b\}$ (Prediction: no uniqueness ✗)

        [N.B.: $\mathrm{IExcl}(\phi_a \vee \phi_b, Q_{HO}) = \varnothing$ and $\mathrm{IIncl}(\phi_a \vee \phi_b, Q_{HO}) = \{\phi_a \vee \phi_b, \phi_a, \phi_b\}$;

            therefore: $O^{IE+II}_{Q_{HO}}(\phi_a \vee \phi_b) = \phi_a \wedge \phi_b$.]

The O^{IE+II}-operator also over-predicts FC for disjunctions embedded under an existential indefinite, which in turn over-predicts a MS interpretation for questions with an existential indefinite. As argued in Sect. 5.1.2, the answer space of (99) couldn't be like (99a), since it would allow for a non-exhaustive individual interpretation.[27] Applying O^{IE+II}_Q point-wise to this answer space yields the partition in (99b), which includes worlds where two movies each was watched by a different boy.

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

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

                                         ($\phi_{\langle x,m \rangle}$ abbreviates 'boy *x* watched *m*')

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

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

      [N.B.: $\mathrm{IExcl}(\exists x[\phi_{\langle x,m_1 \rangle} \vee \phi_{\langle x,m_2 \rangle}], Q) = \varnothing$ and $\mathrm{IIncl}(\exists x[\phi_{\langle x,m_1 \rangle} \vee \phi_{\langle x,m_2 \rangle}], Q) = Q$;

            therefore: $O^{IE+II}_Q(\exists x[\phi_{\langle x,m_1 \rangle} \vee \phi_{\langle x,m_2 \rangle}]) = \exists x.\phi_{\langle x,m_1 \rangle} \wedge \exists x.\phi_{\langle x,m_2 \rangle}$]

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

To sum up, the partition-by-exhaustification analysis of Fox (2018, 2020) allows for MS; however, due to the particular choice of the O^{IE+II}-operator, this account remains insufficient in predicting uniqueness effects and in avoiding over-generating MS. One way to fix these problems would be to use a different method to derive partition and FC. For example, the partition can be induced by applying both the IE-based exhaustification operator O^{IE} and the FC-trigger DOU to each propositional answer. The O^{IE+II}-operator and the DOU-operator predict different distributions of FC. As discussed

---

[26]Recall that singular *which*-phrases range over boolean disjunctions but not boolean conjunctions or higher-order pluralities.

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

above, exhaustifying a disjunction-sentence by $O^{\text{IE+II}}$ yields FC if and only if the alternative set of this sentence isn't closed under conjunction. The $O^{\text{IE+II}}$-based analysis predicts FC for $\Diamond(\phi \vee \psi)$ as well as $\exists x[\phi_{\langle x,a \rangle} \vee \phi_{\langle x,b \rangle}]$, but not for wide-scope disjunctions including $\Diamond\phi \vee \Diamond\psi$. It also predicts FC for the plain disjunction $\phi \vee \psi$ if the conjunctive alternative $\phi \wedge \psi$ is pruned. In contrast, the DOU-based analysis, which accounts for the distribution of universal FC based on the RelExcl presupposition, predicts universal FC only for $\Diamond$-disjunctions, including $\Diamond(\phi \vee \psi)$ and $\Diamond\phi \vee \Diamond\psi$.

Fox's account doesn't deal with local-uniqueness interpretations. For a recent account of local uniqueness based on partition and local exhaustification, see Kobayashi and Rouillard 2021.

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

To account for local-uniqueness effects in $\Diamond$-questions, in line with Rullmann and Beck (1998), Hirsch and Schwarz (2020) propose that the uniqueness presupposition of a singular *wh*-question comes from the lexical meaning of the determiner *which*. They define *which* as in (100) and assume that *which* is interpreted within the question nucleus.[28]

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

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

(101) Which letter could we add to *fo_m*?
    a. *which* $\gg$ *could*: global uniqueness
        i. [$_{\text{CP}}$? $\lambda1$ [$_{\text{IP}}$ [$_{\text{DP}}$ which t_1 letter] $\lambda2$ [could [$_{\text{VP}}$ we add t_2 to *fo_m*]]]]
        ii. $\{\lambda w : \underline{\exists! y[\Diamond_w[\lambda w'.\text{letter}_{w'}(y) \wedge \text{add}_{w'}(y)]]}.\Diamond_w[\lambda w'.\text{letter}_{w'}(a) \wedge \text{add}_{w'}(a)] \mid x \in D_e\}$
    b. *could* $\gg$ *which*: local uniqueness
        i. [$_{\text{CP}}$? $\lambda1$ [$_{\text{IP}}$ could [[$_{\text{DP}}$ which t_1 letter] $\lambda2$ [$_{\text{VP}}$ we add t_2 to *fo_m*]]]]
        ii. $\{\lambda w : \underline{\Diamond_w[\lambda w'.\exists! y[\text{letter}_{w'}(y) \wedge \text{add}_{w'}(y)]]}.\Diamond_w[\lambda w'.\text{letter}_{w'}(a) \wedge \text{add}_{w'}(a)] \mid x \in D_e\}$

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

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

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

---

[28]The main texts of Uegaki 2018, 2021 also define *which* with a uniqueness presupposition. The main goal of this analysis is to account for the projection behavior of uniqueness in question embeddings. However, for Uegaki's core idea, it only matters that the uniqueness presupposition is carried by each answer (as opposed to by the answer set as a whole), and it does not matter to him how this presupposition is compositionally derived. In Uegaki 2021: Appendix A, he sketches out an alternative analysis which assumes that the uniqueness presupposition is assigned to each of the propositional answers after the application of the ANS-operator. My proposal in Sect. 6 is compatible with this analysis (see also f.n. 31).

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

Despite of these advantages, this account faces many problems. First, this account only deals with uniqueness effects. It cannot avoid over-generating MS interpretations for *wh*-questions with a collective predicate or with an existential quantifier.

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

(103)   (Context: The relevant book has three chapters. The speaker knows that we can assign either chapter 1, or chapter 2, or both chapters 2–3 to the students next week.)

     a.   # I know which chapter we can assign to the students... chapter 3.

Third, the local-uniqueness inference predicted by this account is too weak. Hirsch and Schwarz predict this inference to be existential, read as: "We are allowed to assign one single chapter to the students", not "We are *only* allowed to assign one single chapter to the students". However, (103b) is marginal, despite that uniqueness is satisfied in worlds where we assign chapter 1. It is more natural to express the intended meaning using a number-neutral question as in (103c), or with the stressed modifier *SINgle* as in (103d). (For an explanation of the contrast between (103b) and (103d), see Sect. 6.3.3.)

(103′)   (Continue from above)

     b.   ? I know which chapter we can assign to the students... chapter 1.

     c.   I know what we can assign to the students... chapter 1.

     d.   I know which SINgle chapter we can assign to the students... chapter 1.

Contrary to Hirsch and Schwarz, I argue that the uniqueness requirement in a singular modalized question needs to be satisfied in *every* accessible world that verifies a true answer. Dialogue (104) illustrates this requirement. In a survey through personal communication, 9 out of 11 native speakers reported that TA1's utterance sounded unnatural or sly — it seems that TA1 was unintentionally assuming or intentionally suggesting that they should assign at most one paper next week, conflicting with the instructor's request. In comparison, dialogue (105) is perfectly natural. TA2 clearly didn't imply that the reading for next week must be a journal article; what she asked is simply a sub-question of the more general question 'Which journal article or book chapter should we assign to the students next week? If we choose to assign a journal article, what could it be?'

(104)   Instructor: "We should assign one or two papers to the students each week."
         TA1: "Got it. Which paper could we assign to the students next week?"   [Unnatural]

(105)   Instructor: "We should assign a journal article or a book chapter to the students each week."
         TA2: "Got it. Which journal article could we assign to the students next week?"   [Natural]

Forth, the presuppositional-*which* account under-predicts uniqueness effects for numeral-modified *wh*-questions with a non-divisive predicate.[29] Like singular *which*-phrases, numeral-modified *which*-phrases also trigger uniqueness (Xiang 2021b; see also Sect. 5.1.1). The presuppositional semantics of *which* in (102) does not predict the sensitivity to uniqueness in (106b): in the described scenario, the intersection between the *wh*-domain and the extension of the predicate is closed under sum.

(106) (The students solved three problems in total: $a + b$ together solved one, $b + c$ together solved one, and $a + b + c$ together solved one.)

    a. Which two students solved a problem together?
       ↝ # *Only two of the students solved a problem together.*
       $[\![two\ students]\!]^w \cap [\![s.a.p.t.]\!]^w = \{a \oplus b, b \oplus c\}$ (prediction: with uniqueness ✓)

    b. Which two or three students solved a problem together?
       ↝ # *Only two or three of the students solved a problem together.*
       $[\![two\ or\ three\ students]\!]^w \cap [\![s.a.p.t.]\!]^w = \{a \oplus b, b \oplus c, a \oplus b \oplus c\}$
                                                                (prediction: no uniqueness ✗)

More generally, Dayal's exhaustivity presupposition is concerned with the entailment relation of the propositional answers, while the presupposition of *which* assumed in (102) is concerned with the 'part-of' relation of the short answers. In questions with a non-divisive predicate, the satisfaction of the part-of relation does not ensure the satisfaction of the entailment relation. The uniqueness effects in cases like (106c) argue that the uniqueness effects in questions come from a constraint on propositions, not a constraint on individuals.

Last, in questions with multiple singular *which*-phrases, the uniqueness presupposition assumed for the higher/subject *which*-phrase is too strong to allow for a pair-list interpretation. In the pair-list interpretation (107b), there is a point-wise uniqueness requirement relative to the movies, but no uniqueness requirement w.r.t. the boys. To allow for non-uniqueness w.r.t. the subject domain, it is inevitable to assume a non-presuppositional semantics for the subject *which boy*, which clearly conflicts with Hirsch and Schwarz's explanation of uniqueness.

(107) Which boy watched which movie?

    a. Single-pair interpretation: 'Which unique boy-*x*-movie-*y* pair is such that *x* watched *y*?'
       'Andy watched *Spiderman*.'

    b. Pair-list interpretation: '[Each boy watched at most one movie, tell me:] which boy-*x*-movie-*y* pairs are such that *x* watched *y*?'
       'Andy watched *Ironman*, Billy watched *Spiderman*, Clark watched *Hulk*.'

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

6. Solving the dilemma: Relativized Exhaustivity

Section 5 has presented a dilemma: Dayal's exhaustivity presupposition is incompatible with MS, but abandoning this condition would leave the uniqueness effects unexplained and would over-generate

---

[29] A predicate P is divisive if and only if whenever P holds of something x, it also holds of every subpart of x defined for P. Formally: $\forall x[P(x) \rightarrow \forall y \leq x[y \in \text{Dom}(P) \rightarrow P(y)]]$. For example, the collective predicate *solved a problem together* is not divisive: '*a,b,c* solved a problem together' does not entail '*a,b* solved a problem together'.

MS interpretations for a variety types of questions. Further, the observations on local uniqueness have argue that Dayal's exhaustivity presupposition has a general problem in tackling modalized questions. This problem is independent from the considerations on MS: local uniqueness is observed in various modalized questions, regardless of the modal flavor and the modal force. Hence, we need an alternative of Dayal's exhaustivity presupposition that correctly predicts the distribution of MS and uniqueness.

6.1.   Relativized Exhaustivity

Intuitively, in a modalized singular *wh*-question, local uniqueness arises if uniqueness is evaluated relative to the accessible worlds, as opposed to the anchor world. For example, for the modalized questions in (108), local uniqueness says that we assign at most one chapter to the students in every accessible world. This inference is more accurately stated in (109), where w is the anchor world, and M is the contextually determined modal base of *can/should* which maps the anchor world to a set of accessible worlds.

(108)   a.   Which chapter can we assign to the students?

b.   Which chapter should we assign to the students?

(109)   'For every world w' in M_w s.t. we assign any chapters to the students in w', we assign only one chapter to the students in w'.'

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

I first redefine Dayal's exhaustivity presupposition as in (110), where the question denotation is sensitive to the modal base. Here $\llbracket Q \rrbracket_w^M$ $(:= \{\alpha \mid \alpha \in \mathrm{Dom}(\llbracket Q \rrbracket^M) \wedge \llbracket Q \rrbracket^M(\alpha)(w) = 1\})$ abbreviates the set of short answers to Q that are true in w given the modal base M.

(110)   *Dayal's exhaustivity presupposition* (adapted from Dayal 1996)
Given modal base M, a question Q is defined in w only if
$\exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \forall \beta [\beta \in \llbracket Q \rrbracket_w^M \rightarrow \llbracket Q \rrbracket^M(\beta) \subseteq \llbracket Q \rrbracket^M(\alpha)]]$ [abbreviated as $\boldsymbol{DEP}(w, M, \llbracket Q \rrbracket)$]

Next, I propose that questions are subject to a *Relativized Exhaustivity* (RelExh) condition, defined as in (111). This condition requires Dayal's exhaustivity presupposition to be satisfied relative to every modal base that introduces a singleton set of accessible worlds that verifies a true answer. (See Sect. 4.4.3 on Relativized Exclusivity (RelExcl) for definitions of relevant concepts.)[30]

---

[30]Here RelExh is defined based on short answers, not sentential answers, because the interpretations of sentential answers vary by the modal base. Consider the definition in (i), schematized in parallel to (111). Here Q^M stands for the Hamblin set of Q given the modal base M.

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

Definition (i) is problematic because the two underlined parts are conflicting. For example, assume that the non-modalized question $?\phi_x$ has a true sentential answer ϕ_a in w_1 and let $M_w = \{w_1, w_2\}$ and $M_w = \{w_1\}$. Then relative to both M and M', $\diamond \phi_a$ is a true sentential answer to the modalized question $?\diamond\phi_x$ in w. However, $\llbracket \diamond \phi_a \rrbracket^M \neq \llbracket \diamond \phi_a \rrbracket^{M'}$. In contrast, definition (111) avoids this conflict because the interpretation of a short answer is independent from the modal base: $\llbracket a \rrbracket^M = \llbracket a \rrbracket^{M'}$.

An anonymous reviewer of *NLS* points out that definition (111) of RelExh requires access to the modal base M, which is difficult from a compositional perspective. In the Kratzerian theory of modality, the modal base M is an argument variable of the modal verb. It is unclear how M can be retrieved when evaluating the interpretation of a question, especially if the modal verb is embedded under another scopal expression (e.g., the anti-exhaustification operator DOU). In Appendix B, I will explore a 'variable-free' analysis of modal bases, which defines modalized sentences as functions from modal bases to propositions.

(111) *Relativized Exhaustivity*

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

$\forall M'_{\langle s,st \rangle}[|M'_w| = 1 \wedge M'_w \subseteq M_w \wedge \exists \alpha [\alpha \in [\![Q]\!]^M_w \wedge \alpha \in [\![Q]\!]^{M'}_w] \rightarrow \textbf{DEP}(w, M', [\![Q]\!])]$

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

The next two subsections present the applications and predictions of this condition. The RelExh 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 for modalized questions (Sect. 6.3).

6.2. Predictions of RelExh on the distribution of MS

6.2.1. Permitting MS

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

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

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

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

Assume that the modal base M maps the anchor 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 (113a).

(113) a. Let chair-the-committee $= \begin{bmatrix} w_1 \rightarrow \{a\} \\ w_2 \rightarrow \{b\} \\ ... \end{bmatrix}$ and $\begin{pmatrix} M_w = \{w_1, w_2\} \\ M^1_w = \{w_1\} \\ M^2_w = \{w_2\} \end{pmatrix}$

Given the assumed M, the first-order MS interpretation (112a) 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. It doesn't yield an exhaustive true answer and thus violates Dayal's exhaustivity presupposition. However, if this MS interpretation is evaluated relative to M^1, which maps w to the singleton set $\{w_1\}$, $\Diamond O_C \phi_a$ would be the unique exhaustive true answer in w. Likewise, if this MS interpretation is evaluated relative to M^2, which maps w to $\{w_2\}$, $\Diamond O_C \phi_b$ would be the unique exhaustive true answer in w. Hence, given modal base M, the first-order MS interpretation of the question satisfies RelExh in w.

b. for the first-order MS interpretation (112a), we have:

$[\![Q_{\text{FO}}]\!]^M_w = \{a, b\}$ ETA: not exist

$[\![Q_{\text{FO}}]\!]^{M^1}_w = \{a\}$ ETA: $\Diamond O_C \phi_a$ (ETA abbreviates 'exhaustive true answer')

$[\![Q_{\text{FO}}]\!]^{M^2}_w = \{b\}$ ETA: $\Diamond O_C \phi_b$

This analysis also applies to the higher-order MS interpretation (112b). This interpretation allows for one more true answer, which is formed out of the boolean disjunction $a^\Uparrow \cup b^\Uparrow$. This answer doesn't affect exhaustivity since it remains partial regardless of the modal base.

This analysis overcomes the technical difficulties in retrieving the modal base and it also allows to define RelExh based on sentential answers.

c. for the higher-order MS interpretation (112b), we have:

$$[\![Q_{\text{HO}}]\!]_w^M = \{a^{\Uparrow}, b^{\Uparrow}, a^{\Uparrow} \cup b^{\Uparrow}\} \quad \text{ETA: not exist}$$
$$[\![Q_{\text{HO}}]\!]_w^{M^1} = \{a^{\Uparrow}, a^{\Uparrow} \cup b^{\Uparrow}\} \quad \text{ETA: } \Diamond O_C \phi_a$$
$$[\![Q_{\text{HO}}]\!]_w^{M^2} = \{b^{\Uparrow}, a^{\Uparrow} \cup b^{\Uparrow}\} \quad \text{ETA: } \Diamond O_C \phi_b$$

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

(114) *Generalization of RelExh for MS interpretations*
The MS interpretations of ⌜*Wh-A can P?*⌝ satisfy RelExh if and only if the interpretations of the non-modalized question ⌜*Wh-A P?*⌝ satisfy Dayal's exhaustivity presupposition in every accessible world where *P* holds for an element of *A*.

I revise the definitions of the answerhood-operators to the following. The RelExh presupposition and the max-informativity condition are abbreviated as $\textbf{\textit{REP}}(w, M, [\![Q]\!])$ and $\textbf{\textit{MaxI}}(\alpha, w, M, [\![Q]\!])$, respectively. Applying the answerhood-operators returns a set of max-informative true answers and triggers a RelExh presupposition.[31]

(115) $\textbf{\textit{MaxI}}(\alpha, w, M, [\![Q]\!]) = 1$ if and only if $\alpha \in [\![Q]\!]_w^M$ and $\forall \beta \in [\![Q]\!]_w^M [[\![Q]\!]^M(\beta) \not\subset [\![Q]\!]^M(\alpha)]$.

(116) Answerhood-operators (modified from (40))

a. For complete true short answers
$\text{Ans}^S(w)(M)([\![Q]\!]) = \textbf{\textit{REP}}(w, M, [\![Q]\!]).\{\alpha \mid \textbf{\textit{MaxI}}(\alpha, w, M, [\![Q]\!])\}$

b. For complete true propositional answers
$\text{Ans}^P(w)(M)([\![Q]\!]^M) = \textbf{\textit{REP}}(w, M, [\![Q]\!]).\{[\![Q]\!]^M(\alpha) \mid \textbf{\textit{MaxI}}(\alpha, w, M, [\![Q]\!])\}$

6.2.2. Avoiding over-predicting MS

As first raised in Sect. 4.2.3 and reiterated in Sect. 5.1.2, for the questions in (117) with a stubbornly collective predicate, the answer space derived in a first-order interpretation is not closed under conjunction. In such cases, without further constraints, applying Fox's answerhood over-predicts a MS interpretation.

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

The RelExh presupposition avoids over-predicting MS interpretations for these non-*can*-questions: in a multiple-team scenario, the first-order interpretations of (117a,b) violate RelExh, just like how they violate Dayal's exhaustivity presupposition.

Moreover, RelExh explains why existential indefinites do not license MS: RelExh allows the evaluation of exhaustivity to be relativized w.r.t. a smaller modal base, but not w.r.t. a smaller discourse domain. Hence, RelExh has the same prediction as Dayal's exhaustivity presupposition for non-modalized questions such as (85) *Which movie or movies did one of the boys watch?*.

---

[31] For readers who are familiar with Uegaki 2018, 2021: to account for the projection of the uniqueness presupposition in embeddings, it's better to assume that the RelExh presupposition is carried by each max-informative true answer, as opposed to by the answer set as a whole (see also f.n. 28). With this assumption, the answerhood operators are defined as follows:

(i) a. $\text{Ans}^S(w)(M)([\![Q]\!]^M) = \{\textbf{\textit{REP}}(w, M, [\![Q]\!]).\alpha \mid \textbf{\textit{MaxI}}(\alpha, w, M, [\![Q]\!])\}$

b. $\text{Ans}^P(w)(M)([\![Q]\!]^M) = \{\textbf{\textit{REP}}(w, M, [\![Q]\!]).[\![Q]\!]^M(\alpha) \mid \textbf{\textit{MaxI}}(\alpha, w, M, [\![Q]\!])\}$

6.3. Predictions of RelExh on uniqueness

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

A complete explanation of uniqueness should ensure that every interpretation, if not ruled out independently, delivers a desired global or local uniqueness inference. It should also ensure that each of these inferences can be derived in one or more of the interpretations. Examples of uniqueness inferences and the types of interpretations are gathered in the following:[32]

(118) For singular *should*-questions

    a. Example: "Which chapter should we assign to the students?"

        i. Global uniqueness: *There is only one chapter that we should assign.*

        ii. Local uniqueness: *We should assign at most one chapter.*

    b. First-order (G); wide-scope higher-order (G); narrow-scope higher-order (L)

(119) For singular *can*-questions without local exhaustification

    a. Example: "Which chapter can we assign to the students?"

        i. Global uniqueness: *There is only one chapter that we can assign.*

        ii. Universal local uniqueness: *We can assign one chapter, but not more.*

    b. First-order/higher-order MS (L); wide/narrow-scope disjunctive-MA (L)

(120) For singular *can*-questions with local exhaustification

    a. Example: "Which SINgle chapter can we assign to the students?"

        i. Global uniqueness: *There is only one chapter that we can assign.*

        ii. 'Existential' local uniqueness: *We can assign exactly one chapter, may or may not more.*

    b. First-order/higher-order MS (L); wide/narrow-scope disjunctive-MA (L)

6.3.1. Uniqueness effects in non-modalized questions

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

(82′)   a. Which child came?   (Singular *wh*-question)

       b. Which two children formed a team?   (Numeral-modified *wh*-question)

       c. Did you invite Andy, Billy, or Cindy?   (Alternative question)

Moreover, the RelExh-based analysis of uniqueness extends to cases that are challenging to the presuppositional-*which* account of Hirsch and Schwarz (2020). RelExh can account for the uniqueness effect of (106c) and allows for the point-wise uniqueness effect of (107).

---

[32]For example in (118b), 'first-order (G)' means that the first-order interpretation yields global (G) uniqueness, and 'narrow-scope higher-order (L)' means that the narrow scope higher-order interpretation yields local (L) uniqueness.

(106′c) Which two or three students solved a problem together?

(Numeral-modified *wh*-question with a non-divisive predicate)

(107′) Which boy watched which movie?   (Pair-list multiple-*wh* question)

6.3.2. Uniqueness effects in *should*-question

For *should*-questions (or any $\square$-questions) with a first-order interpretation, RelExh yields the same prediction as Dayal's exhaustivity presupposition. For example in (121), in the first-order interpretation, the question calls for an answer naming an atomic chapter and presupposes that there is only one chapter that we should assign to the students. Dayal's exhaustivity presupposition is violated if there are multiple chapters such that we assign them to the students in every accessible world. Since uniqueness is violated in all of the accessible worlds, the violation of Dayal's exhaustivity presupposition cannot be salvaged by evaluating exhaustivity relative to a smaller modal base. Hence, RelExh yields global uniqueness. The same applies if this question has a wide-scope higher-order interpretation.

(121) Q: Which chapter should we assign to the students?
        ⤳ *There is a unique chapter that we should assign to the students.*   (Global uniqueness)
       A: Chapter 1.

However, when the same singular *should*-question has a narrow-scope higher-order interpretation, RelExh and Dayal's exhaustivity presupposition have different predictions. In (122), the question is exhaustively addressed by a narrow-scope disjunction, read as 'There is no particular chapter that we should assign to the students — we just need to chose between chapter 1 and chapter 2.' (Spector 2007, 2008; Xiang 2021b; see Sect. 4.1.3). In this interpretation, the question implies local uniqueness.

(122) Q: Which chapter should we assign to the students?   (=(89))
        ⤳ *We shouldn't assign more than one chapter to the students.*   (Local uniqueness)
       A: Chapter 1 or chapter 2, either is good. ($\square[\phi_{c_1} \vee \phi_{c_2}]$)

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

(123) assign $= \begin{bmatrix} w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\}, \\ w_3 \rightarrow \{c_1, c_2\}, ... \end{bmatrix}$, $M = \begin{bmatrix} w \ \rightarrow \{w_1, w_2, w_3\} \ \text{(with uniq.-violation)} \\ w' \rightarrow \{w_1, w_2\} \ \ \text{(without uniq.-violation)} \end{bmatrix}$

In contrast, the RelExh presupposition is satisfied in w' but violated in w. The true answer $\square(\phi_{c_1} \vee \phi_{c_2})$ can be verified by three singleton sets of accessible worlds, namely $\{w_1\}, \{w_2\}, \{w_3\}$. RelExh is satisfied in w if and only if this question has an exhaustive true answer in w when interpreted relative to any modal base M' such that $M'_w = \{w_1\}/\{w_2\}/\{w_3\}$. This requirement cannot be satisfied: given any M' such that $M'_w = \{w_3\}$, question (122) has two true answers in w (viz., $\square\phi_{c_1}$ and $\square\phi_{c_2}$) but no exhaustive true answer in w.

In sum, for a singular *should*-question, both RelExh and Dayal's exhaustivity presupposition can predict a global-uniqueness inference, derived in the first-order interpretation or the wide-scope higher-order interpretation. However, only RelExh can predict a local-uniqueness effect, derived in a narrow-scope higher-order interpretation. This finding argues that RelExh is a mandatory condition for question interpretation, not just a salvaging strategy for Dayal's exhaustivity presupposition.

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

According to the generalization in (114), the MS interpretations of the *can*-question (124b) satisfy RelExh if and only if the non-modalized question (124a) has a unique true answer in every accessible world where we assign any chapter(s) to the students. This condition is the desired universal local uniqueness inference.

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

For a concrete illustration, let's interpret (124b) relative to the modal base M assumed in (123) above. In both w and w', (124b) has the same set of true MS answers $\{\Diamond\phi_{c_1}, \Diamond\phi_{c_2}\}$. The MS interpretations of (124b) violate RelExh in w: since $\Diamond\phi_{c_2}$ is a true answer in w and can be verified by $\{w_3\}$, RelExh requires that the question has an exhaustive true answer if interpreted relative to a modal base M' such that $M'_w = \{w_3\}$; however, when interpreted relative to such an M', this question has two true answers in w (viz., $\Diamond\phi_{c_1}$ and $\Diamond\phi_{c_2}$) but no exhaustive true answer in w. In contrast, the MS interpretations of (124b) satisfy RelExh in w' since w' has no uniqueness-violating accessible world like w_3.

In question (125) (see also (103d)), where the *wh*-complement is modified by *SINgle*, the local-uniqueness inference appears to be existential: the question requests to name one chapter which can be the unique chapter that we assign to the students, but it does not rule out the possibilities to assign more than one chapter.

(125) Which SINgle chapter can we assign to the students?   ('Existential' local uniqueness)
      ⤳ *We can assign exactly one chapter to the students, may or may not more.*

Why the local-uniqueness inference in (125) appears to be existential? Due to the modifier *SINgle*, the question nucleus is parsed with local exhaustification, read as: 'Which chapter x is such that we can assign only x to the students?' The set of the true MS answers to this question is $\{\Diamond O\phi_{c_1}, \Diamond O\phi_{c_2}\}$. The RelExh presupposition predicts the following condition: the non-modalized exhaustified question *Which chapter x is such that we assign only x to the students* has a unique true answer in every accessible world where there is a chapter x such that we only assign x to the students. This condition is fairly weak, since it only considers the accessible worlds where uniqueness is satisfied. For example, since none of the locally exhaustified true answers (e.g., $\Diamond O_C\phi_{c_2}$) can be verified by $\{w_3\}$, the violation of uniqueness in w_3 doesn't affect RelExh. In sum, in a singular *can*-question, the local-uniqueness inference appears existential if the question is parsed with local exhaustification.

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

Singular *can*-questions admit only disjunctive-MA, not conjunctive-MA. In (126), the MA answer can be expressed by an elided disjunction, as in (126a), but not by an elided conjunction or a plurality, as in (126b,c). This contrast argues that singular *which*-phrases may quantify over boolean disjunctions, but not boolean conjunctions or pluralities (Xiang 2021b, see also Sect. 5.1.1).

(126) Which chapter can we assign to the students?
      ⤳ *We can assign exactly one chapter to the students, but not more.*
      a.   Chapter 1 or chapter 2.

b. # Chapter 1 and chapter 2.

c. # Chapters 1 and 2.

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

Let's return to local uniqueness. Section 6.3.3 has argued that RelExh can explain the universal local-uniqueness effect of the singular *can*-question in (127) if this question has a MS interpretation.

(127)   Q: Which chapter can we assign to the students?   (Universal local uniqueness)
        ↝ *We are allowed to assign exactly one chapter to the students, but not more.*
        A: Chapter 1 or chapter 2.   (Cf. (124b))

Does the RelExh-based analysis apply if the question has a disjunctive-MA interpretation? To check the predictions, consider the following local-uniqueness-violating context, repeated from (123)/(90b):

(128)   Let assign $= \begin{bmatrix} w_1 \to \{c_1\}, w_2 \to \{c_2\} \\ w_3 \to \{c_1, c_2\}, ... \end{bmatrix}$ and $M_w = \{w_1, w_2, w_3\}$

In this context, the answer space of the question in (127) can be like Figure 8a or Figure 8b, depending on the scope of the higher-order *wh*-trace relative to the modal verb *can*. In both cases, the disjunctive answer at the bottom is the unique exhaustive true answer.

a. DOU $\gg \diamond \gg \pi$: disjunctive-MA   b. DOU $\gg \pi \gg \diamond$: disjunctive-MA
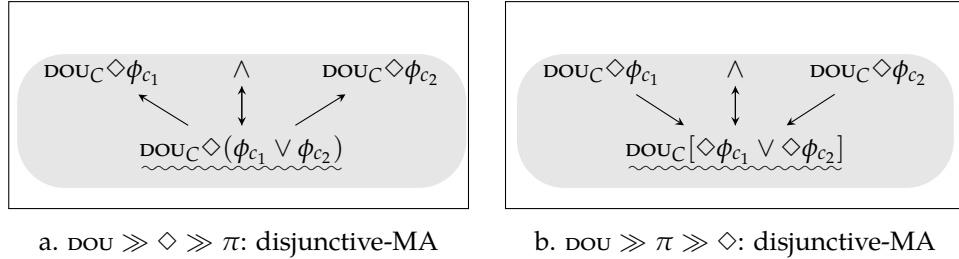
Figure 8: Answer space of (127) in a disjunctive-MA interpretation

In contrast to the MS interpretations in (124b), here the RelExh presupposition alone cannot predict the universal local-uniqueness inference in (127): in a disjunctive-MA interpretation, the violation of local uniqueness does not lead to a violation of RelExh. The true answer $\text{DOU}_C \diamond \phi_{c_2}$ ($= \diamond \phi_{c_2}$) can be verified by $\{w_3\}$, and therefore RelExh requires that the question has an exhaustive true answer if evaluated relative to a modal base M' such that $M'_w = \{w_3\}$. Previously for (124b), I argued that the MS interpretations of this question violate RelExh in w: the MS interpretations of this question do not satisfy exhaustivity if evaluated relative to an M' such that $M'_w = \{w_3\}$. However, here the answer space derived in a disjunctive-MA interpretation includes also the FC-disjunctive answer $\text{DOU}_C[\diamond \phi_{c_1} \vee \diamond \phi_{c_2}]$ / $\text{DOU}_C \diamond (\phi_{c_1} \vee \phi_{c_2})$ ($= \diamond \phi_{c_1} \wedge \diamond \phi_{c_2}$). Given any M' such that $M'_w = \{w_3\}$, this FC answer is an exhaustive true answer to this *can*-question in w, and thus RelExh is not violated.

How does the presented account avoid under-generating uniqueness? I argue that, before RelExh applies, the FC-disjunctive answers $\text{DOU}_C[\diamond \phi_{c_1} \vee \diamond \phi_{c_2}]$ and $\text{DOU}_C \diamond (\phi_{c_1} \vee \phi_{c_2})$ have been ruled out due to violations of Relativized Exclusivity (RelExcl), a condition independently motivated to account

for the modal obviation effect in the licensing of universal FC. As argued in Sect. 4.4.3, the anti-exhaustification operator DOU carries a RelExcl presupposition: for every anti-excludable alternative ϕ that is not entailed by the prejacent, every minimal subset of the accessible worlds that verifies ϕ also verifies the exhaustification of ϕ (for formal definitions, see (66) and (139)). When DOU applies to a $\diamond$-disjunction, no matter whether the disjunction takes scope below or above the existential modal, the RelExcl presupposition yields a definedness condition as follows:

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

In (129a), if the prejacent sentence is parsed without local exhaustification, RelExcl yields a condition that the two disjuncts cannot be simultaneously true. For the same reason, for the singular *can*-question (127), if $\diamond(\phi_{c_1} \wedge \phi_{c_2})$ is true, the disjunctive answers DOU$_C\diamond(\phi_{c_1} \vee \phi_{c_2})$ and DOU$_C[\diamond\phi_{c_1} \vee \diamond\phi_{c_2}]$ are undefined. Once the FC-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 how it does for the MS interpretations in (124b).

This analysis also applies to (130). As previously argued in (125), due to the focused modifier *SINgle*, this question is parsed with local exhaustification, read as: 'Which chapter x is such that we can assign *only x* to the students?' Just like (129b), when parsed with local exhaustification, the disjunctive answers DOU$_C\diamond O(\phi_{c_1} \vee \phi_{c_2})$ and 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 (130), RelExh predicts a seemingly non-universal local-uniqueness effect for the disjunctive-MA interpretations in the same say as how it does for the MS interpretations in (125).

(130) Q: Which SINgle chapter can we assign to the students?   ('Existential' local uniqueness)
   ⤳ *We can assign exactly one chapter to the students, may or may not more.*
   A: Chapter 1 or chapter 2.   (Cf. (125))

In sum, singular *can*-questions admit FC-disjunctive answers formed out of boolean disjunctions. With these answers, RelExh alone predicts only an existential local-uniqueness effect, not a universal one. I argue that whether a FC-disjunctive answer is available is independently restricted by RelExcl, a presupposition carried by the anti-exhaustification operator. Once the FC-disjunctive answers that violate RelExcl are removed, the RelExh presupposition can account for local-uniqueness effects for disjunctive-MA interpretations in the same way as how it does for MS interpretations.

6.4. Interim summary

To predict the distribution of MS interpretations and uniqueness effects in questions, this section has proposed a 'Relativized Exhaustivity (RelExh)' presupposition. This presupposition requires Dayal's exhaustivity presupposition to be satisfied relative to every modal base that introduces a singleton set of accessible worlds that verifies a true answer.

The RelExh presupposition has many advantages: it allows for MS interpretations, carries forward the merits of Dayal's exhaustivity presupposition in deriving uniqueness effects, and avoids over-generating MS interpretations for non-*can*-questions. Moreover, in contrast to Dayal's exhaustivity

presupposition, the RelExh presupposition can account for the local-uniqueness effects in modalized questions. Predictions of Dayal's exhaustivity presupposition and RelExh are summarized as follows:

Modal-type	Reading-type		Dayal's	RelExh
No modal	$\pm$ uniq.		✓	✓
□-modal	$-$ uniq.		✓	✓
	global uniq.		✓	✓
	local uniq.		✗	✓
◇-modal	$-$ uniq.	MS	✗	✓
	$-$ uniq.	MA	✓	✓
	global uniq.		✓	✗
	local uniq.	MS	✗	✓
	local uniq.	MA	✗	✓

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

Note that the mandatorily applied RelExh presupposition predicts only local-uniqueness effects for ◇-questions. To allow for global uniqueness, we can either assume that Dayal's exhaustivity presupposition is applied optionally, or assume that questions have an optional uniqueness condition as follows: there is only one complete true answer to the question.

7. Conclusions

This paper made three contributions to the study of MS interpretations and uniqueness effects in questions. First, observing that MS answers are subject to a 'mention-one-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 structure-ambiguities within the question nucleus. For cases where MS is licensed by a non-exhaustive conversation goal, I argued that the context may provide covert restrictions to the question nucleus and turn the question into a *can*-question. This idea was supported by the unavailability of MS in embeddings of non-*can*-questions.

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

However, allowing non-exhaustive answers to be complete would cause a concerning 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 the most, I proposed that question interpretations can violate Dayal's exhaustivity presupposition but mandatorily presuppose 'Relativized Exhaustivity'. This condition solves the dilemma

between uniqueness and MS, avoids over-generating MS interpretations for non-*can*-questions, and is advantageous in deriving local-uniqueness effects in modalized questions.

A.   Choice questions versus MS questions

As seen in (84), repeated below, questions with an existential indefinite are ambiguous between a choice interpretation and an individual interpretations. In contrast to the individual interpretation discussed in Sect. 5.1.2, the choice interpretation is non-exhaustive: as seen in (131a), a choice answer only needs to specify one boy-movie(s) pair. Should this non-exhaustive choice interpretation be treated as a variant of the MS interpretation?

(131) (Among the relevant boys, 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*, 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 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*.

Dayal (2017: Sect. 3.2.1) has discussed two reasons for keeping choice interpretations separate from MS interpretations. She calls questions with a choice interpretation 'choice questions'. This appendix presents three additional arguments.

First, choice questions admit only 'choose-some' interpretations, not 'choose-all' interpretations. For example, while (132-a) admits a MA interpretation, (132-b) can never be read as that the consultant knows every store's opening time. This contrast argues that the non-exhaustivity of *can*-questions and the (non-)exhaustivity of choice questions have different origins.

(132) a. The consultant knows where we can get coffee.
       b. The consultant knows when one of the stores open.

Second, disjunctive answers to choice questions do not have a FC interpretation. In contrast to *can*-questions, in (133) the disjunctive answer only has an epistemic ignorance reading. This contrast argues that the machinery that makes a *can*-question congruent with FC-disjunctive answers is not available to choice questions. As argued in Sect. 4.4, the disjunctive-MA interpretation of the a *can*-question is derived by applying an anti-exhaustification operator to the local IP. Nevertheless, the existential indefinite in a choice question occupies a fairly high position (Groenendijk and Stokhof 1984; Chierchia 1993; Szabolcsi 1997b; Krifka 2001b; Dayal 2017; Xiang 2019, 2021c; a.o.), and therefore it cannot interact with an anti-exhaustification operator. Hence, it is expected by the presented proposal that choice questions do not have FC-disjunctive answers.

(133) a. Which movie did one of the two boys watch?
       b. Andy watched *Hulk*, or Billy watched *Ironman*.
           (Available: 'Either Andy watched *Hulk*, or Billy watched *Ironman*, I don't know which.')
           (Unavailable: 'Andy watched *Hulk*, and Billy watched *Ironman*.')

Third, the two types of questions behave quite differently w.r.t. local uniqueness. In a *can*-question with a singular *which*-phrase, the uniqueness effect triggered by the singular *which*-phrase applies to every accessible world that verifies a true answer (see (103)–(105) and (124)/(127)). In choice questions, however, the uniqueness presupposition is existential. For example, the choice question (133a) implies that one of the boys watched exactly one movie, not that each boy watched at most one movie. This contrast argues that the local-uniqueness interpretations have different derivations in these two types of questions. See Xiang 2019, 2021c for an analysis that derives this existential uniqueness effect.

B. A variable-free treatment of modal bases

The definitions for Relativized Exclusivity (RelExcl) and Relativized Exhaustivity (RelExh) face some technical challenges.

Consider the RelExcl condition first. As discussed in f.n. 22, the definition for RelExcl in (66), repeated below, is syncategorematic: in this way, we cannot this condition as part of the lexical meaning of *dou*/DOU. Moreover, we cannot give a precise value for C', which is sensitive to the value of M'.

(134) *Relativized Exclusivity* (repeated from (66))
$[\![dou_C(S)]\!]^M/[\![\text{DOU}_C(S)]\!]^M$ is defined in w only if
$\forall \phi [[\![\phi]\!]^M \in \text{AntiExcl}([\![S]\!]^M, C) \wedge [\![\phi]\!]^M \subset [\![S]\!]^M$
$\to \forall M'_{\langle s,st \rangle}[M'_w$ is a minimal subset of M_w s.t. $[\![\phi]\!]^{M'}(w) = 1 \to O_{C'}([\![\phi]\!]^{M'})(w) = 1]]$.

Moreover, in the Kratzerian theory of modality (Kratzer 1977, 1991), the modal base is treated a free argument variable of the modal verb and is interpreted via an assignment function, as formalized in (135). With this treatment, it is unclear how the interpretation of a modalized sentence can make reference to the modal base, especially in cases where the modal verb is embedded.

(135) a. $[\![can_M \phi]\!]^g = \lambda w_s.\exists w' \in g(M)(w)[[\![\phi]\!]^g(w) = 1]$
       b. $[\![should_M \phi]\!]^g = \lambda w_s.\forall w' \in g(M)(w)[[\![\phi]\!]^g(w) = 1]$

To address these technical problems, here I explore a 'variable-free' treatment of modal base. This treatment is inspired by the variable-free analysis of pronouns (Jacobson 1999, 2014). In the Heim-and-Kratzer tradition, pronouns are translated as free variables (e.g., $[\![it_i]\!]^g = g(i)$). In contrast, the variable-free analysis of Jacobson defines a pronoun as an $\langle e, e \rangle$-type function (e.g., $[\![it]\!] = \lambda x_e.x$). Further, any expression that contains a free pronoun has an *e*-type abstraction passed up from this pronoun (e.g., $[\![it \ arrived]\!] = \lambda x_e.\text{arrive}(x)$). Extending this treatment to modal bases, I assume that modalized sentences denote functions from modal bases to propositions, as defined in (136).

(136) a. $[\![can \ \phi]\!] = \lambda M_{\langle s,st \rangle} \lambda w_s.\exists w' \in M_w[[\![\phi]\!](w) = 1]$
       b. $[\![should \ \phi]\!] = \lambda M_{\langle s,st \rangle} \lambda w_s.\forall w' \in M_w[[\![\phi]\!](w) = 1]$

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

(137) For any sentential expression F of type $\langle st, \sigma \rangle$, $\mathcal{M}(F)$ is an expression of type $\langle \langle sst, st \rangle, \langle sst, \sigma \rangle \rangle$ such that $[\![\mathcal{M}(F)]\!] = \lambda \alpha_{\langle sst,st \rangle} \lambda M_{\langle s,st \rangle}.[\![F]\!](\alpha(M))$.

For example, with the application of the $\mathcal{MB}$-operation, the asserted meaning of DOU is shifted from (138a) to (138b). $\mathcal{MB}(\text{DOU}_C)$ takes a $\langle sst, st\rangle$-type sentence as its argument and returns a $\langle sst, st\rangle$-type sentence. The value of the variable C is also shifted from a set of propositions into a set of functions of type $\langle sst, st\rangle$.

(138)   Asserted meaning of DOU

     a. $[\![\text{DOU}_C]\!] = \lambda p_{st}\lambda w_s.p(w) = 1 \wedge \forall q \in \textsc{AntiExcl}(p, C)[O_C^{\text{IE}}(q)(w) = 0]$ (=(64))

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

The above assumptions allow to redefine the RelExcl presupposition as follows. (Here for any modal base $M_{\langle s,st\rangle}$, $C^M := \{\rho(M) \mid \rho_{\langle stt,st\rangle} \in C\}$.) In this definition, to evaluate exclusivity relative to a modal base M', all we need is to apply each of the modalized alternatives to M'.

(139)   *Relativized Exclusivity* (modified from (134))
     For any θ of type $\langle sst, st\rangle$ and M of type $\langle s, st\rangle$, $[\![\mathcal{MB}(\text{DOU}_C)]\!](\theta)(M)(w)$ is defined only if
     $\forall\delta_{\langle sst,st\rangle}[\delta(M) \in \textsc{AntiExcl}(\theta(M), C^M) \wedge \delta(M) \subset \theta(M)$
         $\rightarrow \forall M'[M'_w$ is a minimal subset of M_w s.t. $\delta(M')(w) = 1 \rightarrow O_{C^{M'}}^{\text{IE}}(\delta(M'))(w) = 1]]$.

Next, consider how the variable-free analysis affects Relativized Exhaustivity. As discussed in f.n. 30, the definition in (111), repeated below, requires to extract short answers from the question denotation, because the semantics of sentential answers are dependent on the modal base. Moreover, this definition requires access to value of M, which is difficult from a compositional perspective.

(140)   *Relativized Exhaustivity* (repeated from (111))

     a.   Given modal base M, a question Q is defined in w only if
        $\forall M'_{\langle s,st\rangle}[|M'_w| = 1 \wedge M'_w \subseteq M_w \wedge \exists\alpha[\alpha \in [\![Q]\!]_w^M \wedge \alpha \in [\![Q]\!]_w^{M'}] \rightarrow \textbf{\textit{DEP}}(w, M', [\![Q]\!])]$

     b. $\textbf{\textit{DEP}}(w, M', [\![Q]\!]) := \exists\alpha[\alpha \in [\![Q]\!]_w^{M'} \wedge \forall\beta[\beta \in [\![Q]\!]_w^{M'} \rightarrow [\![Q]\!]^{M'}(\beta) \subseteq [\![Q]\!]^{M'}(\alpha)]]$

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

(141)   *Relativized Exhaustivity* (modified from (140))

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

     b. $\textbf{\textit{DEP}}(w, M', Q) := \exists\theta[w \in \theta(M') \in Q^M \wedge \forall\delta[w \in \delta(M') \in Q^{M'} \rightarrow \theta(M') \subseteq \delta(M')]]$

56

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. 2006. Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry* 37:535–590.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford: Oxford University Press.

Chierchia, Gennaro, and Ivano Caponigro. 2013. Questions on questions and free relatives. Presentation at the 18th Sinn und Bedeutung, University of the Basque Country (UPV/EHU) in Vitoria-Gasteiz, September 2013. https://scholar.harvard.edu/files/chierchia/files/frs_and_qs_iii_2013-9-11.pdf.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In *An international handbook of natural language meaning*, ed., Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 2297–2332. Berlin: Mouton de Gruyter.

Dayal, Veneeta. 1996. *Locality in Wh Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer.

Dayal, Veneeta. 2013. A viability constraint on alternatives for free choice. In *Alternatives in semantics*, ed., Anamaria Fălăuş, 88–122. London: Palgrave Macmillan.

Dayal, Veneeta. 2017. *Questions*. Oxford: Oxford University Press.

Dotlačil, Jakub, and Floris Roelofsen. 2019. Dynamic inquisitive semantics: Anaphora and questions. In *Proceedings of sinn und bedeutung 23*, ed., M.Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, and Cristina Real-Puigdollers, 365–382. https://doi.org/10.18148/sub/2019.v23i1.538.

Dotlačil, Jakub, and Floris Roelofsen. 2021. A dynamic semantics of single-*wh* and multiple-*wh* questions. In *Proceedings of Semantics and Linguistic Theory 30*, ed., Joseph Rhyne, Kaelyn Lamp, Nicole Dreier, and Chloe Kwon, 376–395. Washington, D.C.: LSA. DOI:https://doi.org/10.3765/salt.v30i0.4839.

von Fintel, Kai, Danny Fox, and Sabine Iatridou. 2014. Definiteness as maximal informativeness. In *The art and craft of semantics: A festschrift for Irene Heim, vol. 1*, ed., Luka Crnic and Uli Sauerland, 165–174. MIT Working Papers in Linguistics.

Fox, Danny. 2007. Free choice disjunction and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, ed., Uli Sauerland and Penka Stateva, 71–120. New York: Palgrave Macmillan.

Fox, Danny. 2012. Multiple *wh*-questions: Uniqueness, pair-list and second order questions. MIT

seminar handouts. http://lingphil.mit.edu/papers/fox/uniqueness-pair-list.pdf.

Fox, Danny. 2013. Mention-some interpretations. MIT seminar handouts. http://lingphil.mit.edu/papers/fox/class1-3.pdf.

Fox, Danny. 2018. Partition by exhaustification: Comments on Dayal 1996. In *Proceedings of Sinn und Bedeutung 22*, ed., Uli Sauerland and Stephanie Solt, 403–434. Berlin: Leibniz-Centre General Linguistics. https://semanticsarchive.net/Archive/jQ5MTcyZ/.

Fox, Danny. 2020. Partition by exhaustification: Towards a solution to Gentile and Schwarz's puzzle. Manuscript, MIT. https://semanticsarchive.net/Archive/TljZGNjZ/.

Fox, Danny, and Martin Hackl. 2007. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.

George, B. R. 2011. Question embedding and the semantics of answers. Ph.D. Dissertation, University of California, Los Angeles.

Ginzburg, Jonathan. 1995. Resolving questions, i. *Linguistics and Philosophy* 18:459–527.

Ginzburg, Jonathan, and Ivan Sag. 2000. *Interrogative investigations*. Stanford: CSLI Publications.

Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. *Varieties of formal semantics* 3:143–170.

Hausser, Roland, and Dietmar Zaefferer. 1979. Questions and answers in a context-dependent Montague grammar. In *Formal semantics and pragmatics for natural languages*, ed., Franz Guenthner and Siegfried J. Schmidt, 339–358. Dordrecht: Reidel.

Hausser, Roland R. 1983. The syntax and semantics of English mood. In *Questions and answers*, ed., Ferenc Kiefer, 97–158. Dordrecht: Springer.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.

Heim, Irene. 1994. Interrogative semantics and Karttunen's semantics for *know*. In *Proceedings of the 9th Annual Conference of the Israel Association for Theoretical Linguistics (IATL) and of the Workshop on Discourse*, ed., Rhonna Buchalla and Anita Mitwoch, volume 1, 128–144. Jerusalem: IATL. https://semanticsarchive.net/Archive/jUzYjk1O/Interrogative%2094.pdf.

Hirsch, Aron, and Bernhard Schwarz. 2020. Singular *which*, mention-some, and variable scope uniqueness. In *Proceedings of Semantics and Linguistic Theory 29*, ed., Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 748–767. Washington, D.C.: LSA. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/29.748.

Huang, C.-T. James. 1982. Move WH in a language without wh movement. *The Linguistic Review* 1:369–416.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22:117–185.

Jacobson, Pauline. 2014. *Compositional semantics: An introduction to the syntax/semantics interface*. Oxford: Oxford University Press.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1:181–194.

Klinedinst, Nathan, and Daniel Rothschild. 2011. Exhaustivity in questions with non-factives. *Semantics and Pragmatics* 4:1–23.

Kratzer, Angelika. 1977. What 'must' and 'can' must and can mean. *Linguistics and Philosophy* 1:337–355.

Kratzer, Angelika. 1981. The notional category of modality. In *Words, worlds, and contexts: New approaches in word semantics*, ed., H. Eikmeyer and H. Rieser, 38–74. Berlin: Walter de Gruyter.

Kratzer, Angelika. 1991. Modality. In *Semantics: An international handbook of contemporary research*, ed., A. von Stechow and D. Wunderlich, 639–650. Berlin: de Gruyter.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, ed., Yukio Otsu, 1–25. Tokyo: Hituzi Syobo. Reprinted in 2017, in *Contrastiveness in information structure, alternatives and scalar implicatures*, ed., Lee C., Kiefer F., Krifka M, 123–143. Springer, Cham.

Krifka, Manfred. 2001a. For a structured meaning account of questions and answers. *Audiatur vox sapientiae. A festschrift for Arnim von Stechow*, ed., C. Féry and W. Sternefeld, 52:287–319. Berlin: Akademie Verlag.

Krifka, Manfred. 2001b. Quantifying into question acts. *Natural Language Semantics* 9:1–40.

Kobayashi, Filipe Hisao, and Vincent Rouillard. 2021. High and low uniqueness in singular *wh*-interrogatives. In *Proceedings of Semantics and Linguistic Theory 30*, ed., Joseph Rhyne, Kaelyn Lamp, Nicole Dreier, and Chloe Kwon, 714–733. Washington, D.C.: LSA http://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/30.714.

Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.

Li, Haoze. 2019. Deriving short answers from Hamblin sets: A case study of Mandarin *wh*-conditionals. In *Proceedings of Semantics and Linguistic Theory 29*, ed., Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 138–159. Washington, D.C.: LSA.

Li, Haoze. 2021. Mandarin *wh*-conditionals: A dynamic question approach. *Natural Language Semantics*. Online-first articles.

Li, Jie. 1995. *Dou* and *wh*-questions in Mandarin Chinese. *Journal of East Asian Linguistics* 29:313–323.

Li, Xiaoguang. 1997. Deriving distributivity in Mandarin Chinese. PhD Dissertation, University of California, Irvine.

Lin, Jo-Wang. 1998. Distributivity in Chinese and its implications. *Natural Language Semantics* 6:201–243.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use, and interpretation of language*, ed., Christoph Schwarze Rainer Bäuerle and Arnim von Stechow, 302–323. Berlin: De Gruyter.

Liu, Mingming. 2016a. Varieties of alternatives: Mandarin focus particles. *Linguistics and Philosophy* 40:61–95.

Liu, Mingming. 2016b. Mandarin *wh*-conditionals as interrogative conditionals. In *Proceedings of Semantics and Linguistic Theory 26*, ed., Mary Moroney, Carol-Rose Little, Jacob Collard, and Dan Burgdorf, 814–835. Ithaca, NY: CLC Publications. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/26.814.

Matthewson, Lisa. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9:145–189.

Menéndez-Benito, Paula. 2010. On universal free choice items. *Natural Language Semantics* 18:33–64.

Nicolae, Andreea C. 2013. *Any* questions? Polarity as a window into the structure of questions. Ph.D. Dissertation, Harvard University.

Nicolae, Andreea C. 2015. Questions with NPIs. *Natural Language Semantics* 1:21–76.

Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.

van Rooij, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26:727–763.

van Rooij, Robert. 2004. Utility of mention-some questions. *Research on Language and Computation* 2:401–416.

van Rooij, Robert, and Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13:491–519.

Rooth, Mats. 1996. Focus. In *The handbook of contemporary semantic theory*, ed., Shalom Lappin, 271–297. Oxford: Blackwell Publishers.

Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *which*-questions. In *Semantics and Linguistic Theory 8*, ed., Devon Strolovitch and Aaron Lawson, 215–232. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2811.

Sauerland, Uli, Jan Anderssen, and Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In *Linguistic evidence: Empirical, theoretical, and computational perspectives*, ed., Stephan Kepser and Marga Reis, 413–434. Berlin: Mouton de Gruyter.

Schulz, Katrin, and Robert van Rooij. 2006. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29:205–250.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89:607–624.

Spector, Benjamin. 2007. Modalized questions and exhaustivity. In *Proceedings of Semantics and Linguistic Theory 17*, ed., M. Gibson and T. Friedman, 282–299. Ithaca, NY: CLC Publications. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2962.

Spector, Benjamin. 2008. An unnoticed interpretation for *wh*-questions: Elided answers and weak islands. *Linguistic Inquiry* 39:677–686.

Spector, Benjamin, and Marta Abrusán. 2011. A semantics for degree questions based on intervals: Negative islands and their obviation. *Journal of Semantics* 28:107–147.

Szabolcsi, Anna. 1997b. Quantifiers in pair-list readings. In *Ways of scope taking*, ed. Anna Szabolcsi, 311–342. Dordrecht: Kluwer.

von Stechow, Arnim, and Thomas Ede Zimmermann. 1984. Term answers and contextual change. *Linguistics* 22:3–40.

Theiler, Nadine, Floris Roelofsen, and Maria Aloni. 2018. A uniform semantics for declarative and interrogative complements. *Journal of Semantics* 35:409–466.

Uegaki, Wataru. 2015. Interpreting questions under attitudes. Ph.D. Dissertation, MIT.

Uegaki, Wataru. 2018. On the projection of the presupposition of embedded questions. In *Proceedings of Semantics and Linguistic Theory 28*, ed., Sireemas Maspong, Brynhildur Stefánsdóttir, Katherine Blake, and Forrest Davis, 789–808. Washington, D.C.: LSA. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/28.789.

Uegaki, Wataru. 2021. The existential/uniqueness presupposition of *wh*-complements projects from the answers. *Linguistics and Philosophy* 44:911–951.

Wagner, Michael, Elise McClay, and Lauren Mak. 2013. Incomplete answers and the rise-fall-rise contour. In *Proceedings of the Seventeenth Workshop on the Semantics and Pragmatics of Dialogue (Sem-*

Dial), ed., R. Fernández and A. Isard, 140–149. `http://events.illc.uva.nl/semdial/dialdam/papers/WagnerEtal_dialdam.pdf`.

Xiang, Yimei. 2016. Interpreting questions with non-exhaustive answers. Ph.D. Dissertation, Harvard University.

Xiang, Yimei, and Alexandre Cremers. 2017. Mentions-some readings of plural-marked questions: Experimental evidence. In *Proceedings of North East Linguistics Society 47*, ed., Andrew Lamont and Katerina A. Tetzloff, volume 3, 261–274.

Xiang, Yimei. 2019. Getting quantifying-into questions uniformly: Functionality, domain exhaustivity, and quantificational variability. In *Proceedings of Semantics and Linguistic Theory 29*, ed., Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 160–179. Washington, D.C.: LSA. `https://doi.org/10.3765/salt.v29i0.4610`.

Xiang, Yimei. 2020. Function alternations of the Mandarin particle *dou*: Distributor, free choice licensor, and 'even'. *Journal of Semantics* 37:171–217.

Xiang, Yimei. 2021a. A hybrid categorial approach to question composition. *Linguistics and Philosophy* 44:587–647.

Xiang, Yimei. 2021b. Higher-order readings of *wh*-questions. *Natural Language Semantics* 29:1–45.

Xiang, Yimei. 2021c. Quantifying into *wh*-dependencies: Multiple-*wh* questions and questions with quantifiers. `https://ling.auf.net/lingbuzz/005097`, manuscript, Rutgers University.