A hybrid categorial approach to question composition

Yimei Xiang, Rutgers University

Abstract This paper revisits two fundamental issues in question semantics — what does a question mean, and how is this meaning compositionally derived? Drawing on observations with the distribution of *wh*-words in questions and free relatives as well as quantificational variability effects in question-embeddings, I argue that the nominal meanings of short answers must be derivable from question denotations, which therefore calls for a categorial approach to defining questions, including embedded questions. I provide a novel hybrid categorial approach to compose questions. This approach overcomes the problems with traditional categorial approaches in defining bare *wh*-indefinites, composing multi-*wh* questions, and accounting for coordinations of questions.

Keywords: questions, categorial approaches, short answers, quantificational variability, multi-*wh* questions, question coordinations, free relatives, *wh*-conditionals, mention-some, *wh*-indefinites

1. Introduction

What does a question mean? It is hard to address this issue directly: unlike declaratives, questions cannot be defined by truth values or truth conditions — we don't say that a question is true or a question is false. Instead, studies on question semantics usually start with the relation between questions, answers, and question-containing or question-like constructions (such as question-embeddings, free relatives, and so on). As such, researchers tackle this issue from different entry points and end up with quite different conclusions. Among the classics, starting from the relation between questions and propositional answers, Hamblin-Karttunen Semantics (Hamblin 1973; Karttunen 1977) define a question as a set of propositions, each of which is a possible answer or a true answer to this question. In contrast, starting from the relation between questions and short answers, categorial approaches (Hausser and Zaefferer 1979; Hausser 1983) define questions as lambda (λ-)abstracts, or more precisely, functions over short answers. Motivated by the world dependency relation in interpreting questions under attitudes,[1] Partition Semantics (Groenendijk and Stokhof 1982, 1984) defines questions as partitions of possible worlds.

In the recent studies of question semantics, categorial approaches are less dominant, partially due to their technical deficiencies in composing complex structures, and partially because the relation between questions and short answers can be explained alternatively by syntax. However, independent evidence from the distribution of *wh*-words in free relatives and cases of quantificational variability effects in question-embeddings show that questions, including embedded questions, must be able to supply predicative or nominal meanings, which leaves λ-abstracts the only possible denotations of questions. Hence, this paper proposes a novel hybrid categorial approach to compose questions. This approach carries forward the advantages of traditional categorial approaches and overcomes their technical deficiencies.

The rest of this paper is organized as follows. Section 2 presents the reasons for pursuing categorial approaches. I will start with the debate on short answers in discourse, which motivates categorial approaches in previous literature (§2.1). The core insight drawn from this debate is

[1] Groenendijk and Stokhof (1982, 1984) claim that the interpretation of a question embedded under a factive attitude predicate is world-dependent. Consider the following question-embedding sentence for example: *Jenny knows whether Andy left*. If Andy indeed left, this sentence entails that Jenny knows that Andy left. If Andy actually didn't leave, it entails that Jenny knows that Andy didn't leave.

that only categorial approaches allow nominal or predicative meanings to be derived from question denotations semantically. Further, I will introduce two new pieces of evidence for categorial approaches, drawn on facts about *wh*- free relatives and quantificational variability effects (§2.2). Section 3 takes a step back and reviews the problems with traditional categorial approaches to defining bare *wh*-indefinites, composing multi-*wh* questions, and accounting for question coordinations. Section 4 presents a novel hybrid categorial approach. This approach overcomes the problems with the traditional categorial approaches in defining bare *wh*-indefinites and composing multi-*wh* questions with single-pair readings. Section 5 extends this approach to functional readings and pair-list readings. In responding to the evidence for categorial approaches introduced in section 2.2, section 6 shows that the presented approach easily accounts for *wh*-constructions with predicative or nominal meanings (such as *wh*- free relatives and Mandarin *wh*-conditionals) and quantificational variability inferences. Section 7 addresses the issue of composing question-coordinations, especially question-coordinations in embeddings. Section 8 concludes.

2. Why pursuing a categorial approach?

This section will start by reviewing the original motivation of categorial approaches, which is to capture the relation between matrix questions and short answers (§2.1). Over the past few decades, it remains a topic of debate whether this relation should be ascribed to syntax or semantics. Although this paper takes no position on this debate, knowing it helps to understand the relation between λ-abstracts (i.e., the question denotations assumed by categorial approaches) and question denotations proposed by other canonical approaches, namely, that Hamblin sets and world partitions can be derived from λ-abstracts, but not the other way around.

Next, I will present two new pieces of evidence for pursuing categorial approaches that are independent from the debate on short answers in discourse (§2.2). One is based on a cross-linguistic observation about the distribution gap of *wh*-words in questions and free relatives (§2.2.1). The other draws on facts about two cases of quantificational variability effects in question-embeddings (§2.2.2). These facts show that questions, including especially embedded questions, must be defined in a way that can supply predicative or nominal meanings. This prediction leaves λ-abstracts as the only possible question denotations. For ease of presentation, this section addresses only the "why"-issue — why pursuing a categorial approach. Issues as to "how" a categorial approach captures these facts will be discussed separately in section 6.

2.1. The old debate: short answers in discourse

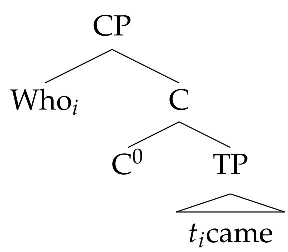
Categorial approaches were originally motivated to capture the relation between matrix questions and their *short answers* (also called *constituent answers*) in discourse semantically. As exemplified in the following, a *full answer* (also called a *clausal answer* or a *propositional answer*) is a full declarative, while a short answer is a constituent that specifies only the new information. To be theory-neutral, this paper avoids terms such as *fragment answers* and *elided answers*, which involve the assumption that short answers are derived syntactically by ellipsis.

(1) Who came?
   a. Mary came. (full answer)
   b. Mary. (short answer)

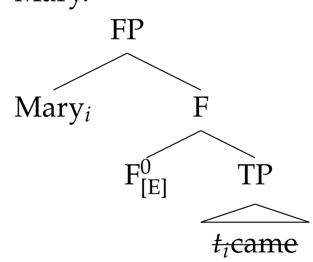
It remains controversial whether a short answer in discourse is a bare nominal (Groenendijk and Stokhof 1982, 1984; Stainton 1998, 2005, 2006; Ginzburg and Sag 2000; Jacobson 2016) or covertly clausal (Merchant 2005). If a short answer is a bare nominal, it should have a nominal interpretation (i.e., be interpreted as an entity in the set denoted by the extension of the *wh*-complement or a generalized quantifier over such entities), which therefore calls for a way to derive such nominal meanings out of a question denotation. Previous works taking this position define the denotations of matrix questions as λ-abstracts, or more precisely, functions that select for the nominal denotation of a short answer as an argument (Groenendijk and Stokhof 1982, 1984; Ginzburg and Sag 2000; Jacobson 2016). In this view, for example, the *wh*-question in (1) denotes the function $\lambda x[\textsf{human}(x).\textsf{came}(x)]$,[2] which can take the human individual *Mary* as an argument. Alternatively, if a short answer is covertly clausal, it should be regarded as an elliptical form of the corresponding full answer and interpreted as a proposition. The ellipsis approach (Merchant 2005) proceeds in two steps. First, the focused constituent *Mary* moves to a left-peripheral position. Next, licensed by a linguistic antecedent provided by the *wh*-question, the rest of the clause gets elided.

(2) Ellipsis approach for short answers (Merchant 2005)

a. Who came?



b. Mary.



Different definitions of questions predict different meanings and derivational procedures of short answers in discourse. Categorial approaches and their variants (Hausser and Zaefferer 1979; Hausser 1983; von Stechow and Ede Zimmermann 1984; von Stechow 1991; Ginzburg 1992; Ginzburg and Sag 2000; Krifka 2001a; Jacobson 2016; among others) define the root denotation of a question as a λ-abstract/ function. Short answers to a question are possible arguments of the λ-abstract denoted by this question, and a full answer is the output of applying this λ-abstract to a short answer.

(3) a. $[\![\textit{who came}]\!] = \lambda x[\textsf{human}(x).\textsf{came}(x)]$

b. $[\![\textit{who came}]\!]([\![\textit{Mary}]\!]) = (\lambda x[\textsf{human}(x).\textsf{came}(x)])(m) = \textsf{came}(m)$

In contrast, Hamblin-Karttunen Semantics (Hamblin 1973; Karttunen 1977) and their descendants (Heim 1995; Cresti 1995; Dayal 1996; Rullmann and Beck 1998; among many others) define the root of a question as a set of propositions, each of which is a possible answer to this question (as assumed by Hamblin) or a true answer to this question (as assumed by Karttunen). Formalizations in (4) exemplify the assumptions of classic Hamblin Semantics. A *wh*-word denotes a set of individuals, as in (4a). A *wh*-question denotes a set of propositions, each of which names an object in the set denoted by the *wh*-word, as in (4b). This set is frequently referred to as a *Hamblin set*. A full answer denotes a singleton set consisting of the proposition denoted by the ordinary value of that answer, as in (4c).

[2] In this paper, λ-terms with presuppositions are represented in the form of $\lambda v\colon \beta.\alpha$ or $\lambda v[\beta.\alpha]$ (where β stands for the definedness condition or the presupposition and α stands for the value description), whichever is easier to read. λ-terms without presuppositions are written in the form of $\lambda v.\alpha$ or $\lambda v[\alpha]$.

(4) a. $[\![who]\!] = \textsf{human}$

b. $[\![who\ came]\!] = \{\hat{}\textsf{came}(x) \mid x \in \textsf{human}\}$

c. $[\![Mary\ came]\!] = \{\hat{}\textsf{came}(m)\}$

Given a proposition set (4b)/(4c), it is technically difficult to recover the constituents that make up the question/answer. More specifically, given a possible worlds analysis of propositional intensions (namely, assuming that a proposition denotes a set of possible worlds where this proposition is true), one won't be able to know that the singleton set (4c) is the interpretation of a declarative made up of a subject *Mary* and a predicate *came*; the exact same set of worlds is equally denoted by the sentence *Mary came alone or came together with someone*, which has a complex predicate *came alone or came together with someone*, or by the sentence *The one among John and Mary that is not John came*, which has a complex subject *the one among John and Mary that is not John*. Likewise, one won't be able to know that the Hamblin set (4b) is an interpretation of a question syntactically made up of *who* and *came*, or to retrieve the λ-abstract and short answers to a question based on this Hamblin set. As such, Hamblin Semantics predicts that short answers can only be derived syntactically by ellipsis. This problem also exists in Karttunen Semantics and the descendants of Hamblin-Karttunen Semantics.

For a concrete example of why λ-abstracts and short answers cannot be derived from Hamblin sets, compare the following three questions. Here the discourse domain is concerned with only two boys, Andy and Billy, and one girl, Cindy. Noun phrases are interpreted *de re* (relative to the actual world @). Abstracts and Hamblin sets for these questions are provided in the (a)-formulas and (b)-formulas, respectively. Short answers are meanings in the domain of an abstract. Observe that these three questions yield the very same Hamblin set but different λ-abstracts and short answers.

(5) Did Andy come or Billy come?

a. $\lambda p\colon p \in \{\hat{}\textsf{came}(a), \hat{}\textsf{came}(b)\}.p$

b. $\{\hat{}\textsf{came}(a), \hat{}\textsf{came}(b)\}$

(6) Which boy came?

a. $\lambda x\colon x \in \{a, b\}.\hat{}\textsf{came}(x)$

b. $\{\hat{}\textsf{came}(a), \hat{}\textsf{came}(b)\}$

(7) The boy among which people came?

a. $\lambda x\colon x \in \{X \mid \textsf{human}_@(X) \wedge \exists! y[y \leq X \wedge \textsf{boy}_@(y)]\}.\hat{}\textsf{came}(\iota y[\textsf{boy}_@(y) \wedge y \leq x])$
$= \lambda x\colon x \in \{a \oplus c, b \oplus c\}.\hat{}\textsf{came}(\iota y[\textsf{boy}_@(y) \wedge y \leq x])$

b. $\{\hat{}\textsf{came}(\iota y[\textsf{boy}_@(y) \wedge y \leq a \oplus c]), \hat{}\textsf{came}(\iota y[\textsf{boy}_@(y) \wedge y \leq b \oplus c])\}$
$= \{\hat{}\textsf{came}(a), \hat{}\textsf{came}(b)\}$

The predictions of Partition Semantics (Groenendijk and Stokhof 1982, 1984, 1990) are a bit more complex, because Partition Semantics treats embedded questions and matrix questions differently. One the one hand, Partition Semantics defines matrix questions as λ-abstracts, as exemplified in (8a). Thus, as in categorial approaches, the relation between matrix questions and their short answers can be modeled semantically as a function-argument relation. On the other hand, in contrast with categorial approaches, Partition Semantics defines embedded questions as partitions of possible worlds, as exemplified in (8b).

(8) a. Matrix question
$[\![who\ came]\!] = \lambda w \lambda x[\textsf{human}(x).\textsf{came}_w(x)]$

b. Embedded question (as in *Jenny knows who came*)
$[\![who\ came]\!] = \lambda w \lambda w'[\lambda x[\textsf{human}(x).\textsf{came}_w(x)] = \lambda x[\textsf{human}(x).\textsf{came}_{w'}(x)]]$

Two world indices belong to the same cell of a partition iff the property denoted by the question nucleus holds for the same set of items in these two worlds. For example, in (8b), w and w' are in the same cell iff the very same set of individuals came in w and w'. With two relevant individuals John and Mary, the partition in (8) is illustrated as in Table 1. Each cell/row stands for a subset of worlds where certain people or nobody came, or equivalently, an exhaustified propositional answer as to *who came*. For instance, the first cell stands for the set of worlds where only John and Mary came, or equivalently, the exhaustified proposition that *only John and Mary came*.

w: only j and m came in w
w: only j came in w
w: only m came in w
w: nobody came in w

Table 1: Partition for *who came*

Since short answers cannot be extracted from partitions, Partition Semantics predicts that nominal short answers cannot be retrieved from the denotation of an embedded question. However, as will be seen in section 2.2.2, contra this prediction, to capture quantificational variability effects in quantified question-embeddings, denotations of embedded questions with non-divisive predicates must be able to supply nominal short answers.

That short answers cannot be extracted from partitions was proved formally in Zimmermann (1985) (details omitted). For intuitive illustration of this proof, consider the following questions. These questions have contrasts with respect to number-marking (i.e., whether the *wh*-item is unmarked or singular-marked or plural-marked, as in (9a, 10a)-vs-(9b, 10b)-vs-(9c, 10c)), polarity (i.e., whether the nucleus is positive or negative, as in (9)-vs-(10)), and exhaustivity (i.e., whether the nucleus involves an exhaustifier or not, as in (9a, 10a)-vs-(11)). The exhaustified questions (11a-b) are in Mandarin because their English counterparts are ungrammatical.

(9) a. Who came?
b. Which person came?
c. Which people came?

(10) a. Who didn't come?
b. Which person didn't come?
c. Which people didn't come?

(11) a. Zhiyou shei lai -le?
only who come -PERF
'Which people x is s.t. only x came?'

b. Zhiyou shei mei lai?
only who not come
'Which people x is s.t. only x didn't come?'

Strikingly, despite the above contrasts, all of these questions yield the very same partition. For a concrete example, compare the two questions (9a) and (9b). The domain of a singular *wh*-item is a proper subset of that of a bare *wh*-word, namely, the former consists of only atomics while the latter also includes sums (Sharvy 1980; Link 1983). With only two relevant individuals John and Mary, the abstracts and Hamblin sets of these two questions are as follows:

(12) a. Who came?
  i. $\lambda x\colon x \in \{j, m, j \oplus m\}.^\wedge\textit{came}(x)$
  ii. $\{^\wedge\textit{came}(j), ^\wedge\textit{came}(m), ^\wedge\textit{came}(j \oplus m)\}$

b. Which person came?
  i. $\lambda x\colon x \in \{j, m\}.^\wedge\textit{came}(x)$
  ii. $\{^\wedge\textit{came}(j), ^\wedge\textit{came}(m)\}$

The abstract and Hamblin set of the number-neutral question (9a) involve pairs and propositions based on sums, while those of the singular-marked question (9b) do not. However, these two questions yield the very same partition, as illustrated in Table 2. In both partitions, the first cell for instance stands for the set of worlds where both John and Mary came. In particular, the one yielded by (9a) is the set of worlds w such that the set of atomic/sum individuals who came in w is $\{j, m, j \oplus m\}$, and the one yielded by (9b) is the set of worlds w such that the set of atomic individuals who came in w is $\{j, m\}$. These two sets of worlds are identical.[3]

Partition yielded by (9a)				Partition yielded by (9b)
w: $\{x \mid w \in c(x)\} = \{j, m, j \oplus m\}$	=	w: only $j \oplus m$ came in w	=	w: $\{x \mid w \in c(x)\} = \{j, m\}$
w: $\{x \mid w \in c(x)\} = \{j\}$		w: only j came in w		w: $\{x \mid w \in c(x)\} = \{j\}$
w: $\{x \mid w \in c(x)\} = \{m\}$		w: only m came in w		w: $\{x \mid w \in c(x)\} = \{m\}$
w: $\{x \mid w \in c(x)\} = \varnothing$		w: nobody came in w		w: $\{x \mid w \in c(x)\} = \varnothing$

Table 2: Partitions for (9a-b)

Likewise, as illustrated in Table 3, the question with negation (10a) and the question with an exhaustifier (11a) also yield the very same partition. In both partitions, the same as the partitions yielded by (9a-b), the second cell stands for the set of worlds where only j came. In particular, in the partition yielded by (10a), this cell is the set of worlds where the set of individuals who didn't come is $\{m\}$. In the one yielded by (11a), this cell is the set of worlds where the individual x such that only x came is j. ($O[c(x)]$ is read as 'only x came'.)

[3] Eagle-eyed readers would notice that the partition yielded by the singular-marked (9b) might not be exactly the one illustrated in Table 2. This question presupposes a uniqueness inference that only one of the relevant individuals came; therefore, it is possible that the partition only considers the the worlds where the uniqueness presupposition is satisfied. This idea is very appealing; for example, in defining the partition of the question *who did John's daughter invite?*, we only consider the worlds where John has a daughter. I thank Christopher Tancredi (pers. comm.) for pointing out this possibility.

An immediate question is how to derive this uniqueness presupposition. Existing work has provided two ways to account for the uniqueness presuppositions of singular-marked questions. The more popular way assumed by Dayal (1996) derives this uniqueness inference by assuming a presupposition that requires the existence of the strongest true proposition in the Hamblin set. This approach does not work here because Hamblin sets cannot be extracted from partitions. The other approach encodes uniqueness or maximality to the question nucleus (*a la* Rullmann and Beck 1998). Existing analyses that assume local maximality does not directly extend to Partition Semantics. The following provides a possible implementation compatible with the assumptions in Partition Semantics. The question *which person came* will first form the property in (ia), where D is a set of relevant atomic individuals. This property carries a conditional maximality presupposition as defined in (ib). Type-shifting this property yields the partition in (ic). This partition is clearly different from the one yielded by (9a) — from (ic), any worlds with two or more people came will be grouped into the cell of worlds where the property is undefined.

(i) a. $\lambda w \lambda x\colon x \in D \wedge \textsc{max}(D, c, w).c_w(x)$
  b. $\textsc{max}(D, c, w) = \exists x \in D[c_w(x)] \rightarrow \exists y \in D[c_w(y) \wedge \forall x \in D[c_w(x) \rightarrow x \leq y]]$
  (If anyone in D came in w, then there is a maximal individual in D who came in w.)
  c. $\lambda w \lambda w'[\lambda x[x \in D \wedge \textsc{max}(D, c, w).c_w(x)] = \lambda x[x \in D \wedge \textsc{max}(D, c, w).c_{w'}(x)]]$

Partition yielded by (10a)				Partition yielded by (11a)
w: $\{x \mid w \notin c(x)\} = \varnothing$		w: only $j \oplus m$ came in w		w: $\{x \mid w \in O[c(x)]\} = \{j \oplus m\}$
w: $\{x \mid w \notin c(x)\} = \{m\}$	$=$	w: only j came in w	$=$	w: $\{x \mid w \in O[c(x)]\} = \{j\}$
w: $\{x \mid w \notin c(x)\} = \{j\}$		w: only m came in w		w: $\{x \mid w \in O[c(x)]\} = \{m\}$
w: $\{x \mid w \notin c(x)\} = \{j, m, j \oplus m\}$		w: nobody came in w		w: $\{x \mid w \in O[c(x)]\} = \varnothing$

Table 3: Partitions for (10a) and (11a)

Hence, abstracts and Hamblin sets cannot be recovered from partitions. In other words, given the partition of a question, we cannot extract out the possible short/propositional answers or recover the question nucleus.

2.2. Independent arguments for categorial approaches

This paper does not take a position on the syntax or semantics of short answers in discourse. But, the above discussion teaches us that the nominal meanings of short answers can only be derived from λ-abstracts, not from Hamblin sets or partitions. In the rest of this paper, the term "short answer", unless specified, refers to the nominal meaning of a bare short answer. In what follows, I will present two independent arguments for categorial approaches, drawn on observations with free relatives and quantificational variability effects. These observations show that meanings equivalent to the bare nominal denotations of short answers must be derivable from the root denotation of a question.

2.2.1. Caponigro's generalization and free relatives

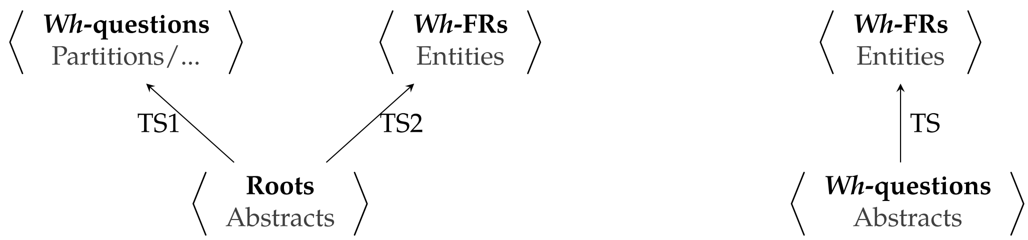
The meaning of a *wh*- free relative (*wh*-FR henceforth) is systematically equivalent to the nominal meaning of the/a complete true short answer of the corresponding *wh*-question. On the one hand, a *wh*-FR is interpreted exhaustively if the corresponding *wh*-question admits only exhaustive answers. For example in (13a), the FR *what Jenny bought* refers to the exclusive set of items that Jenny bought.[4] On the other hand, a *wh*-FR takes an existential reading if the corresponding *wh*-question admits a "mention-some" reading and can be properly addressed by specifying one of its true answers. For example, in (13b), the FR *where he could get help* refers to one of the places where John could get help. (See more discussions and references on "mention-some" in section 4.2.)

(13) a. Mary ate [$_{\text{FR}}$ what Jenny bought].
b. John went to [$_{\text{FR}}$ where he could get help].

The similarities in meaning and form between *wh*-FRs and *wh*-questions make it quite appealing to treat these two constructions uniformly. Previous studies have provided two options to unify these two constructions, illustrated in Figure 1. The nodes each represents a ⟨**construction**, meaning⟩ pair. **Option I** derives a *wh*-question and its corresponding *wh*-FR from the same root via two independent operations. For example, Partition Semantics (Groenendijk and Stokhof 1982, 1984, 1990) assumes that an embedded *wh*-question and its FR counterpart have the same *wh*-root which

[4]Due to the common exhaustive readings, FRs as such are also called "definite FRs", as opposed to "indefinite FRs", which take indefinites-like readings. Unlike the existential FR in (13b), whose non-exhaustive/existential reading is licensed by the presence of a possibility modal, indefinite FRs must be licensed by an external existential operator and usually occur as complements of existential verbs (mainly the equivalents of existential *be* and existential *have*).

denotes an abstract. This abstract can be the input to different type-shifting (TS) rules, yielding different construction-meanings: in Figure 1a, TS1 turns an abstract into a partition of possible worlds and forms a question, while TS2 turns an abstract into an entity and forms a FR. **Option II** treats *wh*-FRs as derivatives of *wh*-questions. As illustrated in Figure 1b illustrates a categorial approach that adopts this option, defining questions as abstracts, and deriving FRs out of questions via one single TS-rule.



(a) Option I (b) Option II

Figure 1: Options to unify *wh*-questions and *wh*-FRs

Which of the options is more plausible? A cross-linguistic generalization due to Caponigro (2003, 2004) supports Option II:

(14) **Caponigro's Generalization** (Caponigro 2003, 2004)
If a language uses the *wh*-strategy to form both questions and FRs, the *wh*-words found in FRs are always a subset of those found in questions. Never the other way around. Never some other arbitrary relation between the two sets of *wh*-words.

This generalization was firstly made based on 28 languages from Indo-European, Finno-Ugric, and Semitic families.[5] It also extends to Tlingit and Haida (Cable 2005), Nieves Mixtec and Melchor Ocampo Mixtec (Caponigro et al. 2013), and Chuj (Kotek and Erlewine 2018). For example, cross-linguistically, *how*-words can be used in questions but not in FRs. Language-specifically, Spanish *qué* 'what' and Estonian *millal* 'when' can be used in questions but not in FRs, as exemplified in (15) and (16). Sometimes, the use of a *wh*-word is blocked only in particular types of FRs. For example in (17), English *who* can be perfectly used in object-FRs and questions, but is deviant in subject-FRs.

(15) Spanish *qué* 'what' (Caponigro 2003: ex. 14)
a. Pregunté **qué**/ lo-que cocinaste.
asked.1SG what/ the.N.S-COMP cooked.2SG
'I asked what you cooked.'

[5] The 28 languages attested by Caponigro (2003, 2004) are:

(i) a. INDO-EUROPEAN:
i. Germanic: Bavarian, Dutch, English, Standard German, Swiss German, West Flemish, Yiddish;
ii. Romance: Catalan, French, Italian, Brazilian Portuguese, European Portuguese, Romanian, Sardinian, Spanish;
iii. Slavic: Bulgarian, Macedonian, Polish, Russian, Serbo-Croatian, Slovenian;
iv. Albanian, Modern Greek
b. FINNO-UGRIC: Estonian, Finnish, Hungarian.
c. SEMITIC: Modern Hebrew, Modern Moroccan Arabic.

b. Comí *qué/ lo-que cocinaste.
ate.1SG *what/ the.N.S-COMP cooked.2SG
'I ate what you cooked.'

(16) Estonian *millal* 'when' (Caponigro 2003: ex. 51)

a. Ma küsisin sult, [$_{Q}$ **millal** Maria saabus].
I asked you when Maria arrived
'I asked you when Maria arrived.'

b. * Ma lahkusin (siis), [$_{FR}$ **millal** Maria saabus].
I left then when Maria arrived.
'I left when Maria arrived.'

(17) English *who* (Caponigro 2004: ex. 40)

a. I will marry [$_{FR}$ who you choose].

b. I don't know [$_{Q}$ who couldn't sleep enough].

c. ?? [$_{FR}$ Who couldn't sleep enough] felt tired the following morning.

To account for the cross-linguistic distributional gap of *wh*-words in FRs relative to questions, analyses that follow Option I must stipulate that the TS rule for deriving FRs is available only for a subset of the inputs in language after language, whereas that the TS rule for deriving questions never is. This stipulation is clearly undesirable — it would be very mysterious why these TS rules must be used in this way cross-linguistically.

Alternatively, treating *wh*-FRs as derivatives of *wh*-questions naturally predicts this gap (Chierchia and Caponigro 2013).[6] With Option II, Caponigro's Generalization is predicted as long as the TS-operation that turns questions into FRs is partial and is licensed only under particular linguistic or non-linguistic conditions. For example, Cecchetto and Donati (2015: chap. 3) attributes part of the gap to a syntactic operation called "relabeling" (details omitted). What determines the distributional gap of *wh*-words can vary from language to language. For the purpose of this paper, all we need is to make sure that the derivation of a *wh*-FR is strictly more complex than the derivation of its question counterpart. The basic idea is that a *wh*-FR is derived by combining an answerhood-operator-like determiner to a question/abstract-denoting root. I will return to the derivations in section 6.1.1.

To sum up, two facts suggest that *wh*-FRs are derivative forms of *wh*-questions, and hence that the denotations of *wh*-questions should be able to supply nominal or predicative meanings that the corresponding *wh*-FRs have. First, in meaning, a FR is semantically equivalent to a/the complete true short answer of the corresponding question. Second, in form, there is a cross-linguistic distributional gap of *wh*-words in FRs relative to questions.

2.2.2. Quantificational variability effects

Existing arguments for function-like denotations of questions are mostly made for matrix questions (Groenendijk and Stokhof 1990; Jacobson 2016). However, drawing on evidence from quantificational

[6]Chierchia and Caponigro (2013) adopted a Hamblin-Karttunen Semantics of questions and proposed an analysis of *wh*-FRs in line with Option II. However, as pointed out by Ede Zimmermann (pers. comm. to Gennaro Chierchia and Ivano Caponigro), this analysis is technically infeasible since it requires extracting abstracts and short answers out of Hamblin sets.

variability (QV) effects in question-embeddings, I show that embedded questions also should have function-like denotations.

As first observed by Berman (1991), question-embeddings modified by a quantity adverbial (e.g., *mostly, partly, for the most part, in part*) are subject to QV effects. For example, the sentences in (18a-b)/(19a-b) intuitively imply the quantificational inference (henceforth called the "QV inference") in (18c)/(19c).[7]

(18) a. Jenny mostly knows who came.
b. For the most part, Jenny knows who came.
c. $\rightsquigarrow$ *For most of the individuals who came, Jenny knows that they came.*

(19) a. Jenny partly knows who came.
b. In part, Jenny knows who came.
c. $\rightsquigarrow$ *For part of the individuals who came, Jenny knows that they came.*

Typically, in paraphrasing a QV inference, the quantificational domain of the matrix quantity adverbial can be formed by the true atomic propositional answers of the embedded question (Lahiri 1991, 2002; Cremers 2016; compare Beck and Sharvit 2002). As schematized in (20), with respect to a given set of propositions, a contained proposition is atomic iff it does not entail any other propositions in this set. For the question *who came*, the atomic propositional answers are of the form $\hat{}\textsf{came}(x)$ where x is an atomic human individual. In case exactly five individuals *abcde* came, the quantification domain of the matrix quantity adverbial is thus the set of atomic propositions $\{\hat{}\textsf{came}(x) \mid x \in \{a,b,c,d,e\}\}$.

(20) Atomic propositions (Cremers 2016: chap. 5)
$\textsc{At}(Q_{\langle st,t\rangle}) = \{p \mid p \in Q \wedge \forall q \in Q[p \subseteq q \rightarrow q = p]\}$

As such, analyses that follow Hamblin-Karttunen Semantics predict schematizations of QV inferences as follows, where $[\![\mathrm{Q}]\!]$ stands for the Hamblin set of the embedded question Q:

(21) The propositional answer-based account (modified from Lahiri 2002)
The QV inference of *Jenny mostly knows Q* is as follows:
$\lambda w.\, \textsc{Most}\ p\ [w \in p \in \textsc{At}([\![\mathrm{Q}]\!])][\textsf{know}_w(j,q)]$
(For most p such that p is a true atomic proposition in $[\![\mathrm{Q}]\!]$, Jenny knows p.)

In the LF derivation, the embedded question moves to the matrix clause, so that the matrix quantity adverbial *mostly* can access the Hamblin set of the embedded question directly (Lahiri 1991, 2002).

(22) [[who came]$_i$ mostly [Jenny knows t_i]]

[7] These quantified question-embedding sentences are ambiguous. For example, (18a-b) can also be read as follows:

(i) a. Most of what Jenny knows is the answer to the question *who came*.
b. What Jenny knows about the question *who came* is almost complete.

Lahiri (2002) calls (ia) a focus-affected reading, where the matrix quantity adverbial quantifies over the set of propositions that are compatible with Jenny's knowledge (details omitted). I'm not aware of any discussions on reading (ib). This reading is true in the following scenario:

> In reality, only Andy came. Jenny knows that either Andy or Billy came. Moreover, based on some reliable clue, she is more inclined to believe that the person who came is Andy.

This paper will not discuss these two readings and will only be interested in the QV reading.

The propositional answer-based account is only concerned with question-embeddings with responsive predicates (i.e., predicates that can take both interrogative and declarative complements, such as *know* and *remember*). Beck and Sharvit (2002) provide an extended account using sub-questions. Observing QV effects in question-embeddings with limited cases of rogative predicates (i.e., predicates that admit interrogative complements and reject declarative complements, such as *depends on*, cf. *wonder* and *ask*), Beck and Sharvit argue that the quantificational domain of the matrix quantity adverbial should be a set of sub-questions of the embedded question, paraphrased as in (23a). It's easy to see that this inference is equivalent to (23b).

(23) For the most part, who will be admitted depends on this committee.

a. $\rightsquigarrow$ *For most Q, Q is a relevant sub-question of the form "whether x will be admitted", Q depends on the committee.*

b. $\rightsquigarrow$ *For most individuals x, whether x will be admitted depends on the committee.*

The sub-question-based account is also compatible with Hamblin-Karttunen Semantics — all one needs to do is to shift each atomic propositional answer of the embedded question into a *whether*-question. Beck and Sharvit define sub-questions of a *wh*-question as *whether*-questions containing a possible atomic answer to this *wh*-question, and assume that whether a sub-question is relevant is determined by the meaning of the question-embedding predicate. In particular, if the question-embedding predicate is veridical, a sub-question is relevant only if it contains a true answer to the embedded question. Using the notations in (21), I schematize Beck and Sharvit's account as follows:[8]

(24) Sub-questions-based account (modified from Beck and Sharvit 2002)
The QV inference of *Jenny mostly knows Q* is as follows:
$\lambda w.\, \textsc{Most}\ Q'\ [\exists p[w \in p \in \textsc{At}([\![\mathrm{Q}]\!]) \wedge Q = \textsf{whether-}p]][\textsf{know}_w(j, Q')]$
(For most Q' of the form *whether-p* such that p is a true atomic proposition in $[\![\mathrm{Q}]\!]$, Jenny knows Q'.)

To sum up, the propositional answer-based account and the sub-question-based account are both compatible with Hamblin-Karttunen Semantics. Both accounts link the formation of the quantificational domain of the matrix quantity adverbial with the extraction of atomic propositional answers from the embedded question. In particular, in the propositional answer-based account, this domain is directly made up of atomic propositional answers; in the sub-question-based account, this domain is made up of sub-questions defined based on atomic propositional answers. As such, both accounts crucially rely on a proper treatment of atomic propositional answers.

However, there are two challenging cases of getting desired atomic propositional answers. One case involves embedding questions with non-divisive predicates, and the other involves embedding questions with pair-list readings. The following introduces these two challenging cases and explains why they cannot be solved by proposition-based accounts of QV effects. See section 6.2 for solutions of these two cases using the proposed hybrid categorial approach.

Challenge 1. Questions with a non-divisive predicate In a question-embedding sentence modified by a quantity adverbial, if the predicate of the embedded question is non-divisive, the

[8] The original formalization by Beck and Sharvit (2002) is a bit different, since they define an embedded question as a Karttunen-intension (of type $\langle s, stt\rangle$), namely, a relation that maps a world to the set of true propositional answers to this question in that world. But this difference is not crucial for this paper.

domain restriction of the matrix quantity adverbial cannot be formed based on the propositional answers of the embedded question.

(25) A predicate P is **divisive** iff $\forall x[P(x) \rightarrow \forall y \leq x[y \in \mathrm{Dom}(P) \rightarrow P(y)]]$
(Whenever P holds of something x, it also holds of every subpart of x for which P is defined.)

Intuitively, for the sentences in (26), the domain of the matrix quantity adverbial is a set of individuals that are atomic subparts of the unique true short answer of the embedded question.[9]

(26) a. Jenny knows for the most part which students formed the bassoon quintet.
b. For the most part, Al knows which soldiers surrounded the fort.
c. Jenny mostly knows which professors formed the committee.

For example, in (26a), if the bassoon quintet was made of five students *abcde*, the unique true short answer of the embedded question is the plural individual $a \oplus b \oplus c \oplus d \oplus e$, and the sentence itself is true if Jenny knows the identities of at least three of the atomic parts of this plural individual. The QV inference can thus be paraphrased as "for most x such that x is an atomic individual among *abcde*, Jenny knows that x is in the group of the students who formed the bassoon quintet." Alternatively, if we follow the propositional answer-based definition in (21), the quantificational domain of the matrix quantity adverbial would contain at most one member because the propositional answers of the embedded question are mutually exclusive. Even though the propositions in the Hamblin set $\{\hat{}\textsf{f.t.b.q}(x) \mid {}^{*}\textsf{student}_{@}(x)\}$ are semantically independent from each other and thus are all atomic, only $\hat{}\textsf{f.t.b.q}(a \oplus b \oplus c \oplus d \oplus e)$ is true. The QV inference predicted by a propositional answer-based account would be as follows: "for most p in the set $\{\hat{}\textsf{f.t.b.q}(a \oplus b \oplus c \oplus d \oplus e)\}$, Jenny knows p.' This inference is problematic in two respects. First, empirically, the quantificational domain of *mostly* must be non-singleton. As seen in (27), *mostly* cannot be grammatically associated with a definite singular noun or with a single name.

(27) {The students, *the student, *John} mostly went to the party.
≈ *Most of {the students, *the student, *John} went to the party.*

Second, the proprosition-based QV inference remains problematic even if the quantificational domain is allowed to be singleton. For (26a) to be true under this analysis, Jenny would have to know most propositions in this singleton set. This can only be achieved by knowing the unique proposition in the set: 1 out of 1 counts as most, but 0 out of 1 does not. This implies that (26a) can only be true in the situation envisioned if Jenny knows the entire membership of the bassoon quintet, which goes against our intuition.[10] This problem with yielding a singleton quantificational

[9] Examples (26a-b) are taken from Schwarz (1994) and Williams (2000) and have been discussed by Lahiri (2002: 215). I thank Alexandre Cremers (pers. comm.) for bringing these data to my attention.

[10] Christopher Tancredi (pers. comm.) points out a related case that does not involve a collective predicate but casts a similar challenge to the proposition-based account. The sentence in (i) implies that Jenny correctly identifies at least three of the five students who came.

(i) Jenny mostly knows which five students came.

The propositional answer-based account cannot define the QV inference properly. Propositional answers to the embedded question are of the form $\hat{}\textsf{came}(x)$, where x is the plurality of some five students, not an atomic student. If the five students who came are *abcde*, the quantificational domain of *mostly* predicted by the propositional answer-based account would be a singleton set $\{\hat{}\textsf{came}(a \oplus b \oplus c \oplus d \oplus e)\}$. Alternatively, defining the quantificational domain of *mostly* as a set domain also applies to the sub-question-based account, because this account requires to define sub-questions based atomic true propositional answers.

Williams (2000) salvages the propositional answer-based account by interpreting the embedded question with a "sub-divisive" reading. He stipulates that the embedded question in (26a) provides propositional answers of a sub-distributive form "x is part of a group that formed the bassoon quintet" where x is an atomic student. Williams derives this reading by assigning the determiner *which* a collective semantics, as schematized in the following:

(28) Jenny knows which students formed the bassoon quintet.
$\approx$ 'Jenny knows which atomic students x are s.t. x is part of the group of students who formed the bassoon quintet.'

a. $[\![\textit{which}]\!] = \lambda A.\lambda P.\{\lambda w.\exists y \in A[y \geq x \wedge P_w(y)] \mid x \in \textsc{At}(A)\}$

b. $[\![\textit{which students}_{@}\ \textit{f.t.b.q.}]\!] = \{\lambda w.\exists y \in {}^{*}\textsf{stdt}_{@}[y \geq x \wedge \textsf{f.t.b.q.}_w(y)] \mid x \in \textsc{At}({}^{*}\textsf{stdt}_{@})\}$
({x is part of the group of students y s.t. y formed the bassoon quintet | x is an atomic student})

Since this approach attributes sub-divisiveness to the lexicon of *which*, it predicts that sub-distributive readings should also be available in matrix *which*-questions. However, contra the prediction, the expected sub-divisive reading is not observed in the corresponding matrix questions. Compare the following sentences:

(29) a. Who is part of the students that formed the bassoon quintet, for example?
b. Which students formed the bassoon quintet, # for example?

(30) a. Who is part of the professors who formed the committee, for example?
b. Which professors formed the committee, # for example?

The use of a partiality marker *for example* (or alternatively, *give me an example* or *show me an example*) indicates that the speaker is tolerant of incomplete true answers, and thus it presupposes the existence of such an answer. The presupposition is schematized in the following, where $\textsc{Ans}(w)([\![\mathrm{Q}]\!])$ stands for the set of complete true answers to Q in w (see section 4.2 for discussions on answerhood):[11]

(31) **Definiteness conditions of partiality markers**
$[\![\text{Q, for example}]\!]^w$ is defined only if there is a proposition p such that

of individuals, we can paraphrase the QV inference as follows: "for most x such that x is an atomic individual among *abcde*, Jenny knows that x came."

[11]Condition (a) excludes answers that are always partial. For example, disjunctive answers like (ia) and negative answers like (ib) are asymmetrically entailed by the complete true answer but are not accepted by the given question.

(i) Who formed a team, for example?
a. # The two girls or the two boys.
b. # Not the two girls.

One might argue that propositions in the Hamblin set of a question are all potentially complete, and therefore that condition (a) can be written more generally as $p \in [\![\mathrm{Q}]\!]$. However, recent studies on *wh*-questions find that some *wh*-questions can have propositional answers formed out of Boolean disjunctions (Spector 2007, 2008; Fox 2013) and conjunctions (Xiang 2016: section 1.6). In particular, Xiang (2016: chap. 2) shows that the answer space of the question in (i), where the *wh*-phrase combines with a non-divisive collective predicate, should be closed under Boolean conjunction and disjunction.

a. $\exists w'[p \in \textsc{Ans}(⟦\text{Q}⟧)(w')]$
(*p* is potentially a complete answer to Q.)

b. $\exists q \in \textsc{Ans}(w)(⟦\text{Q}⟧)[p \subset q]$
(*p* is asymmetrically entailed by a complete true answer to Q.)

As seen in (32), the partiality marker *for example* cannot combine with a question that can have only one true answer. The only way to interpret these sentences is to let *for example* quantify over a set of speech acts, read as "for example, tell me which boy came/ whether it is raining/ ...".

(32) a. Which boy came, # for example?
b. Is it raining, # for example?
c. Did you vote for Andy or Billy, # for example?
d. Zhiyou shei lai -le, # ju-ge lizi? (Mandarin)
only who come -PERF give-CL example
Intended: 'Which people *x* are such that **only** *x* came, # for example?'

Thus, the infelicity of using *for example* in (29b) and (30b) suggests that the embedded questions in (26) admit only collective readings and have only one true answer. From this I conclude that the seeming sub-divisive meaning in their QV inferences comes from the bigger question-embedding constructions, not from the embedded questions.

Challenge 2. Multi-*wh* questions with pair-list readings This challenge is theory-internal to Dayal (1996, 2017), who adopts a Hamblin-Karttunen Semantics and defines single-*wh* questions and multi-*wh* questions uniformly as sets of propositions. As seen in (33), multi-*wh* questions are ambiguous between single-pair readings and pair-list readings. Under a single-pair reading, the question in (33) presupposes that there is exactly one boy-invite-girl pair and asks the addressee to specify this pair. Under a pair-list reading, it expects that there are multiple boy-invite-girl pairs and requests the addressee to list all these pairs.

(33) Which boy invited which girl?
a. Andy invited Mary. (Single-pair)
b. Andy invited Mary, Billy invited Julia, Charlie invited Daphne. (Pair-list)

Dayal (1996, 2017) analyzes the single-pair reading of a multi-*wh* question as denoting a set of atomic propositions, as in (34a), and the pair-list reading as denoting a set of conjunctive propositions, as in (34b). In particular, for the pair-list reading, she proposes a function-based approach and derives these conjunctive propositions based on functions from the domain of the subject-*wh* (viz., the set of atomic boys) to the domain of the object-*wh* (viz., the set of atomic girls). For example, the conjunctive answer $\mathit{invt}(b1, g1) \wedge \mathit{invt}(b2, g1) \wedge \mathit{invt}(b3, g1)$ is based on the function $f = [b1 \to g1, b2 \to g1, b3 \to g1]$. (See Xiang 2016: chap. 5 for a review of this approach.)

(34) a. $⟦\textit{which boy invited which girl}⟧_{\text{single-pair}}$
$$= \left\{ \begin{array}{lll} \mathit{invt}(b1, g1), & \mathit{invt}(b2, g1), & \mathit{invt}(b3, g1), \\ \mathit{invt}(b1, g2), & \mathit{invt}(b2, g2), & \mathit{invt}(b3, g2), \\ ... & & \end{array} \right\}$$

b. $[\![\textit{which boy invited which girl}]\!]_{\text{pair-list}}$

$$= \left\{ \begin{array}{l} \textsf{invt}(b1, g1) \wedge \textsf{invt}(b2, g1) \wedge \textsf{invt}(b3, g1) \\ \textsf{invt}(b1, g1) \wedge \textsf{invt}(b2, g2) \wedge \textsf{invt}(b3, g1) \\ \ldots \end{array} \right\}$$

Compared with competing approaches, especially family-of-question approaches (Hagstrom 1998, Fox 2012, Nicolae 2013, Kotek 2014) which analyze the pair-list reading of a multi-*wh* question as denoting a family of sub-questions (e.g., $[\![34\text{b}]\!]_{\langle stt,t\rangle} = \{[\![\text{which girl did } x \text{ invite?}]\!] \mid \textsf{boy}_{@}(x)\}$), Dayal's approach manages to keep the semantic type of multi-*wh* questions low (uniformly of type $\langle st, t\rangle$). Thus, Dayal advantageously leaves space to tackle more complex structures (such as *wh*-triangles, which have more than two *wh*-items).

However, Lahiri (2002) points out that the use of conjunction in deriving pair-list readings has unwelcome consequences in predicting QV effects. In (35), if paraphrased based on propositions, the quantity adverbial *mostly* would have to quantify over a set of non-conjoined atomic propositions of the form 'boy x invited girl y'. To derive this QV inference, the atomic propositions must be kept alive and should not be mashed under conjunction — once two propositions are conjoined, there is no way to retrieve them back.[12]

(35) Jenny mostly knows which boy invited which girl.
$\rightsquigarrow$ *For most p such that p is a true proposition of the form 'boy x invited girl y', Jenny knows p.*

In contrast, categorial approaches get these QV inferences easily. Treating the multi-*wh* question as a function from boy-girl pairs to propositions, we can define the quantification domain of *mostly* as a set of boy-girl pairs and paraphrase the QV inference as follows: "for most atomic boy-girl pairs $\langle x, y\rangle$ such that x invited y, Jenny knows that x invited y."

To sum up, in a quantified question-embedding sentence, if the embedded question has a non-divisive collective predicate or takes a pair-list reading, the domain of the matrix quantity adverbial cannot be formed out of propositional answers of the embedded question. Rather, this domain should be a set of atomic individuals or pairs of atomic individuals. In categorial approaches, these individuals or individual pairs can be extracted as atomic subparts of the complete true short answer of the embedded question. I will return to a formal solution in section 6.2.

3. Traditional categorial approaches and their problems

3.1. Assumptions of traditional categorial approaches

Categorial approaches (also called functional approaches) define questions as functions from "short answers" (i.e., the missing information of open sentences) to sentences, as exemplified in (36). This idea originates from Cohen (1929) and is later popularly adapted to Lambda calculus (Hausser and Zaefferer 1979; Hausser 1983). The assumed function-like denotations are thus commonly referred to as "λ-abstracts". In traditional categorial approaches, the *wh*-determiner is commonly treated as the λ-operator, as in (37a), and *wh*-phrases as identity functions over predicates, as in (37b-d).

[12] In a colloquium talk at MIT, Dayal (2016) removes the conjunctive closure and analyzes the root denotation of a multi-*wh* question as denoting a family of sets of propositions, in the spirit of family-of-question approaches. This revision manages to keep the atomic propositions alive but sacrifices the advantage of keeping the semantic type of questions low.

(36) *Wh*-questions

a. $[\![who\ came]\!] = \lambda x_e\colon \textsf{human}(x).\textsf{came}(x)$

b. $[\![who\ bought\ what]\!] = \lambda x_e \lambda y_e\colon \textsf{human}(x) \wedge \textsf{thing}(y).\textsf{bought}(x, y)$

(37) *Wh*-determiner and *wh*-phrases

a. $[\![wh\text{-}]\!] = \lambda A_{\langle e,t\rangle} \lambda P_{\langle e,t\rangle} \lambda x_e\colon A(x).P(x)$

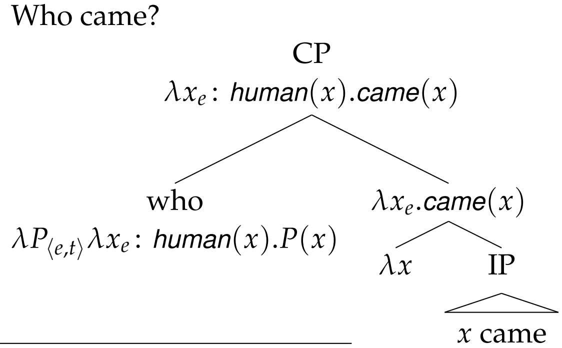
b. $[\![who]\!] = \lambda P_{\langle e,t\rangle} \lambda x_e\colon \textsf{human}(x).P(x)$

c. $[\![what]\!] = \lambda P_{\langle e,t\rangle} \lambda x_e\colon \textsf{thing}(x).P(x)$

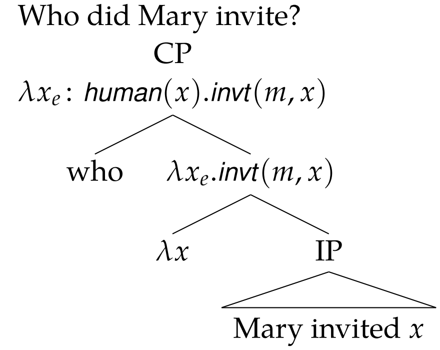
d. $[\![which\ student]\!] = \lambda P_{\langle e,t\rangle} \lambda x_e\colon \textsf{stdt}(x).P(x)$

Traditional categorial approaches do not attempt to derive the meaning of a question compositionally. However, based on the assumed question denotations and *wh*-lexical entries, single-*wh* questions can be composed by standard functional application. As illustrated in (38), *who* moves to the specifier of CP. Applying *who* to a total predicate denoted by its sister node returns a partial predicate defined only for human individuals, which are simply possible short answers to this question.[13] Since traditional categorial approaches are not interested in the extension/intension distinction, here the sister node of the *wh*-word is defined extensionally as a predicate.

(38) a. Who came?



b. Who did Mary invite?



[13] Alternatively, in the context of defending Direct Compositionality and Variable-free Semantics, Jacobson (1995, 1999) shows that a single *wh*-question can be composed directly without assuming LF or abstractions over variables. Her idea is roughly as follows. First, the *wh*-trace denotes an identity function over individuals, as in (ia). Second, the verb *invite* and the Montague-lifted subject *Mary* are shifted by the "Geach rule", indicated by boldface superscript **g**, as in (ib) and (ie). (The Geach Rule: for any function F of type $\langle a, b\rangle$, $F^{\mathbf{g}}$ is a function of type $\langle ca, cb\rangle$ where $F^{\mathbf{g}} = \lambda V_{\langle c,a\rangle} \lambda C_c.F(V(C))$.) Finally, composing the strings directly via left-to-right Function Application yields the wanted λ-abstract.

(i) Structure: [$_{\text{CP}}$who [$_{\text{IP}}$Mary [invited t]]]

a. $[\![t]\!] = \lambda y_e.y$

b. $[\![invite]\!]^{\mathbf{g}} = \lambda f_{\langle e,e\rangle} \lambda y_e \lambda x_e.\textsf{invt}(x, f(y))$

c. $[\![invite]\!]^{\mathbf{g}}([\![t]\!]) = \lambda y_e \lambda x_e.\textsf{invt}(x, y)$

d. $\textsc{Lift}([\![Mary]\!]) = \lambda P_{\langle e,t\rangle}.P(m)$

e. $(\textsc{Lift}([\![Mary]\!]))^{\mathbf{g}} = \lambda R_{\langle e,et\rangle} \lambda x_e.(\lambda P_{\langle e,t\rangle}.P(m))(R(x)) = \lambda R_{\langle e,et\rangle} \lambda x_e.R(x)(m)$

f. $[\![\text{IP}]\!] = (\textsc{Lift}([\![Mary]\!]))^{\mathbf{g}}([\![invite]\!]^{\mathbf{g}}([\![t]\!])) = \lambda x_e.\textsf{invt}(m, x)$

g. $[\![who]\!] = \lambda P_{\langle e,t\rangle} \lambda x_e\colon \textsf{human}(x).P(x)$

h. $[\![\text{CP}]\!] = [\![who]\!]((\textsc{Lift}([\![Mary]\!]))^{\mathbf{g}}([\![invite]\!]^{\mathbf{g}}([\![t]\!]))) = \lambda x_e\colon \textsf{human}(x).\textsf{invt}(m, x)$

The composition of English multi-*wh* questions (or any sentences with in-situ *wh*-phrases) involves more complications. Consider the question *who bought what*. Since Variable-free Semantics assumes Direct Compositionality, the object *wh*-phrase need to be interpreted in-situ and compose directly with the verb *buy*. However, if *wh*-phrases are uniformly defined as identity functions over predicates, $[\![buy]\!]^{\mathbf{g}}([\![what]\!])$ does not return the right meaning. One possible way around this problem is to treat *in situ wh*-items as identity functions over individuals (just like traces). I will explore possible implementations in future research.

Relatedly, the structured meaning approach (von Stechow and Ede Zimmermann 1984; von Stechow 1991; Ginzburg 1992; Ginzburg and Sag 2000; Krifka 2001a) is also a variant of categorial approaches. This approach defines questions as function-domain pairs $\langle Q, D\rangle$, where Q is the function-like denotation of a question as assumed by categorial approaches but with an unrestricted domain, and D is the domain restriction.

(39) Which student came?
  a. Denotation assumed by canonical categorial approaches: $\lambda x : \mathit{stdt}(x).\mathit{came}(x)$
  b. Denotation assumed by structured meaning approaches: $\langle \lambda x.\mathit{came}(x), \mathit{stdt}\rangle$

Observe that the above two denotations are "notationally equivalent" — the function-domain pair (39b) provides exactly the same information as the corresponding domain-restricted function (39a). Hence, in this paper, arguments made for categorial approaches to extracting short answers also apply to structured meaning approaches.

3.2. Problems with traditional categorial approaches

In recent studies of question semantics, categorial approaches are less commonly used than the alternative approaches (especially Hamblin-Karttunen Semantics) due to the their deficiencies in defining bare *wh*-indefinites and composing complex structures.

3.2.1. Problem 1: Bare *wh*-indefinites

First of all, defining the *wh*-determiner as a λ-operator, categorial approaches cannot account for the indefinite use of *wh*-expressions, especially for languages that use bare *wh*-words as indefinites.

Cross-linguistically, *wh*-words are often used to form indefinites, henceforth called *wh*-indefinites. As reported by Haspelmath (1997: pp. 174), 64 languages out of his 100-language sample have *wh*-indefinites. In principle, as in Table 4, languages with *wh*-indefinites can be classified into four types on the basis of the morphological forms of their *wh*-words in questions and existential statements. In reality, however, natural languages with *wh*-indefinites mostly fall into Types I and II. *Wh*-indefinites in Type I languages are formed out of *wh*-words together with additional morphology, henceforth called "complex *wh*-indefinites", while those in Type II languages have a bare form, henceforth called "bare *wh*-indefinites". Type III and Type IV languages, where *wh*-words are morphologically complex when used in interrogatives, are unattested.[14]

Language type	Are *wh*-words morphologically marked in ...		Representatives
	... interrogatives?	... existential statements?	
I	No	Yes	Hebrew, Japanese
II	No	No	Dutch, Chuj
III	Yes	No	(unattested)
IV	Yes	Yes	(unattested)

Table 4: Language-type by morphological forms of *wh*-words in questions and existential statements

[14] The only known example of a Type III language is Esperanto, a constructed international auxiliary language created in the late 19th century. Esperanto uses *kiu* and *kio* for 'who' and 'what' while using *iu* and *io* for 'someone' and 'something'. Since Esperanto is not a natural language, it is not interesting to consider it in discussing the typology of *wh*-indefinites.

Previous studies on the semantic relation between *wh*-words and *wh*-indefinites mostly draw on observations from Type I languages which have complex *wh*-indefinites. For example, Hebrew uses *mi* as 'who' and *ma* as 'what' in questions, but *mi-Sehu* as 'someone' and *ma-Sehu* as 'something' in existential statements. This affixation pattern is found in many languages such as Romanian, Bulgarian, and Basque.

(40) Hebrew *mi* 'who' and *ma* 'what' (Itai Bassi pers. comm.)

a. **Mi** ba?
who come
'Who is coming?'

b. **Mi-Sehu** ba.
who-SEHU come
'someone is coming.'

c. **Ma** hu mevi?
what he bring
'What is he bringing?'

d. Hu mevi **ma-Sehu**.
he bring what-SEHU
'he is bringing something.'

In other languages with complex *wh*-indefinites, the extra morphology found in *wh*-indefinites can also be attached to the entire sentence to form a question. For example, in Japanese, the *-ka* morpheme in *wh*-indefinites (such as in *dare-ka*) also occurs in questions, as seen in (41a-b). Here, however, the *-ka* morpheme is not attached to the *wh*-word but rather to the entire interrogative CP. It is also used in polar questions, as in (41c). Hence, Japanese *wh*-words per se are morphologically unmarked in questions. See the uniform analysis of *-ka* in Uegaki (2018) and references therein.

(41) Japanese *dare* 'who' (Uegaki 2018: ex. 2 and ex. 34)

a. [$_{DP}$ **dare-ka**] -ga hashitta.
who-KA -NOM ran
'Someone ran.'

b. [$_{CP}$ **dare**-ga hashitta-**ka**] oshiete.
who-NOM ran-KA tell
'Tell me who ran.'

c. Hanako-ga hashitta-**ka**?
Hanako-NOM ran-KA
'Did Hanako come?'

For these complex *wh*-indefinites, it is plausible to attribute their existential meaning to operations external to the lexical interpretations of the corresponding *wh*-words. For example, Bittner (1994) assumes that an existential operator shifts a predicate-denoting *wh*-word into an existential indefinite. Similarly, adopting the definitions of *wh*-words assumed by traditional categorial approaches, one can derive complex *wh*-indefinites as in (42): a *wh*-word denotes a function from partial predicates to partial predicates, and it is shifted into an existential quantifier via the application of an $\mathcal{E}$-operator.

(42) Using traditional categorial approaches

a. $[\![who]\!] = \lambda P_{\langle e,t \rangle} \lambda x_e : \textsf{human}(x).P(x)$

b. $\mathcal{E} = \lambda \pi_{\langle et,et \rangle} \lambda P_{\langle e,t \rangle}.\exists x \in \text{Dom}(\pi(P))[\pi(P)(x)]$

c. $\mathcal{E}([\![who]\!]) = \lambda P_{\langle e,t \rangle}.\exists x \in \text{Dom}(\lambda x[\textsf{human}(x).P(x)])[\textsf{human}(x).P(x)]$
$= \lambda P_{\langle e,t \rangle}.\exists x \in \textsf{human}[P(x)]$

Other than applying an existential operator to *wh*-words, a similar strategy to get existential statements is to place the existential closure at a clausal level. Representatives include Beck (2006), Shimoyama (2006), Kotek (2014), and Uegaki (2018). These works adopt Hamblin Semantics in defining and composing questions and sentences.

While the analysis just outlined is plausible for the complex *wh*-indefinites in Type I languages, the bare indefinite use of *wh*-words in Type II languages needs to be analyzed differently. In these languages, *wh*-words are morphologically unmarked in both interrogatives and existential statements. This language type is widespread in Haspelmath's 100-language sample. As Haspelmath (1997: chap. 7) reports, out of the studied 64 languages whose indefinites are morphologically similar to their *wh*-words, there are 31 languages whose *wh*-words occur in their bare form when functioning as indefinites in non-interrogative sentences. Moreover, based on an aggregate survey of approximately 150 languages, Gärtner (2009) compiled a list of 62 languages showing such an interrogative-vs-indefinite ambiguity. Since one of the goals in this paper is to unify the meanings and derivations of FRs and questions, it is particularly interesting to look at languages with both bare *wh*-indefinites and (definite) *wh*-FRs. Representatives include Dutch, German, Russian, Slovene, and Chuj.

(43) Dutch *wat* 'what'

a. **Wat** heb je gegeten?
what have you eaten
'What have you eaten?'

b. Je hebt **wat** gegeten.
You have what eaten
'You have eaten something.'

c. Ik heb gegeten [$_{\text{FR}}$ **wat** jij gekookt had].
I have eaten what you cooked had
'I have eaten what you cooked.'

(44) Chuj *mach* 'who' (Kotek and Erlewine 2018, 2019)

a. **Mach** ix-/ø-ulek'-i?
who prfv-B3-come-itv
'Who came?'

b. Ol-/ø-w-il **mach**.
prosp-B3-A1s-see who
'I will see someone.'

c. Tato tz-/ø-/ø-il **mach**, /ø-/ø-al t'a hin.
if impf-B3-A2s-see who B3-A2-say prfv B1s
'If you see someone, let me know.'

d. Ix-/ø-in-mak [$_{\text{FR}}$ **mach** ix-/ø-ulek'-i].
prfv-B3-A1s-hit who prfv-B3-come-itv
Intended: 'I hit the person who came.'

Bare *wh*-indefinites in Type II languages pose a great challenge to traditional categorial approaches — "It is extremely unlikely that zero-grammaticalization should happen so often, and so systematically." (Haspelmath 1997: pp. 174) Hence, for these bare *wh*-indefinites, the existential flavor needs to be part of their lexical meaning, not obtained by the application of an external existential operator. In other words, in languages with bare *wh*-indefinites, the interrogative meaning of a *wh*-word should be essentially identical to the existential indefinite meaning. Definitions of *wh*-words proposed by traditional categorial approaches cannot capture this relation.

So far, we have discussed three *wh*-constructions: *wh*-questions, *wh*-FRs, and existential statements with *wh*-indefinites. The following summarizes the relationship of these three constructions and the involved *wh*-expressions:

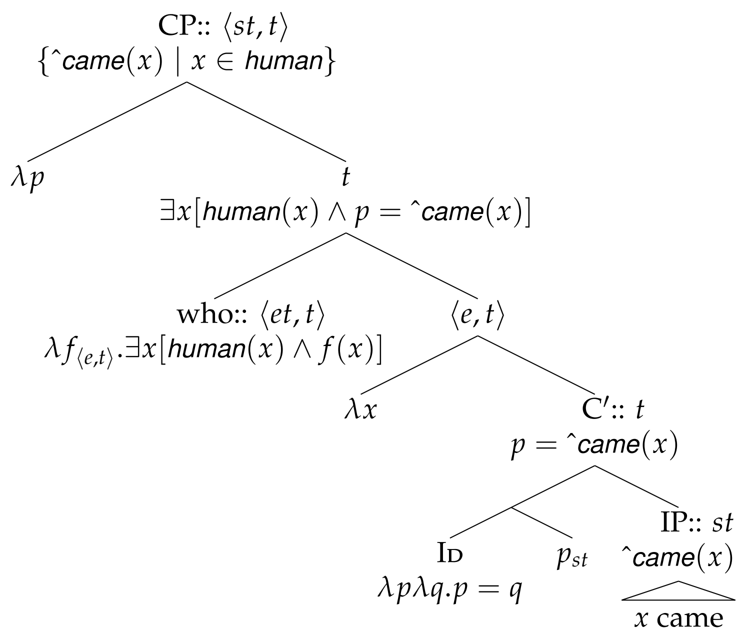
***Wh*-questions versus *wh*-FRs:** The similarities in meaning between *wh*-questions and *wh*-FRs (especially on the distribution of mention-some/existential readings) and Caponigro's Generalization on the distributional gap of *wh*-words suggest that *wh*-FRs are derivatives of *wh*-questions. This is a CP-level relation. To capture this relation, we should define *wh*-questions as abstracts and hence pursue a categorial approach.

***Wh*-words versus bare *wh*-indefinites:** The morphological equivalence between *wh*-words in questions and bare *wh*-indefinites in existential statements suggests that *wh*-words need to be semantically equivalent to or close to *wh*-indefinites. This is a word-level relation. To capture this relation, for languages with bare *wh*-indefinites, we should treat the existential meaning as part of the lexicon of a *wh*-word. Definitions of *wh*-words assumed by traditional categorial approaches are compatible with complex *wh*-indefinites, but not with bare *wh*-indefinites.

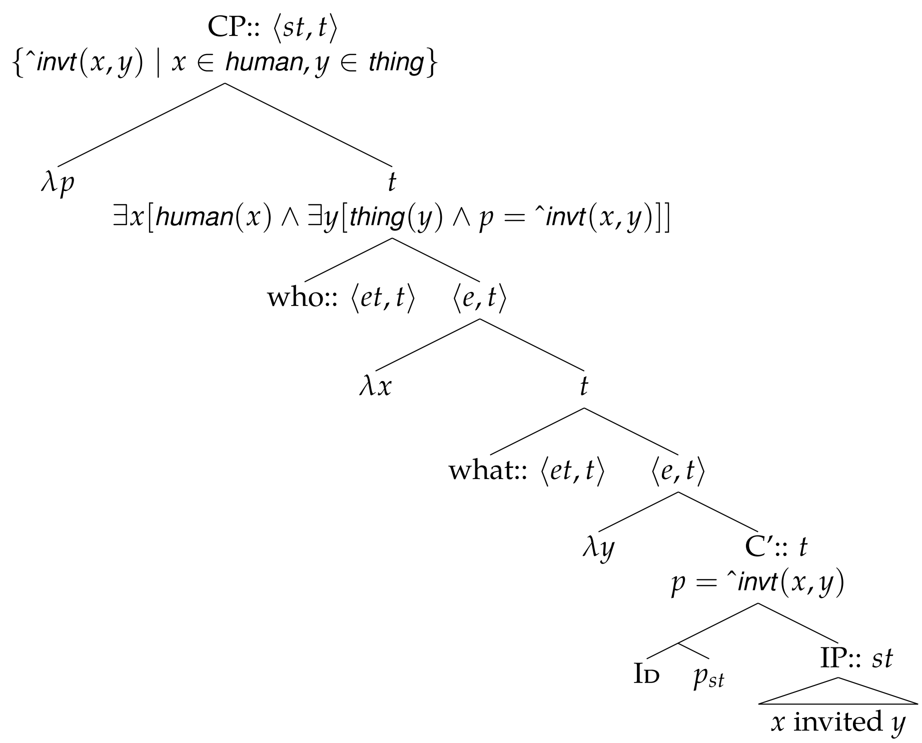
3.2.2. Problem 2: Composing multi-*wh* questions

As seen in (33) in section 2.2.2, multi-*wh* questions are ambiguous between single-pair readings and pair-list readings. While the derivation of a pair-list reading is known to be challenging (see Xiang 2016: chap. 5 and Dayal 2017: chap. 4 for overviews), the composition of a single-pair reading is straightforward in most canonical approaches to question semantics. For example, for Heim (1995), who assigns Karttunen Semantic type interpretations to Government and Binding (GB-)style LFs, the compositions of the single-*wh* question (45) and the multi-*wh* question (46) proceed uniformly. In this derivation, the Proto-Question Rule assumed in Karttunen's original semantics is ascribed to an identity (Id)-function, which creates an identity relation between a propositional variable and the question nucleus (viz., the IP part). *Wh*-words are defined as existential generalized quantifiers. (For example, *who* is analyzed as semantically identical to *someone* and is of type $\langle et, t\rangle$). They undertake quantifier raising to [Spec, CP] and "quantify-into" the identity relation created by the Id-function. Finally, abstracting the first argument p of the Id-function across the existential quantifier(s) denoted by the *wh*-word(s) returns a set of propositions (i.e., the Hamblin set). It can be nicely observed that, in both single-*wh* and multi-*wh* questions, *wh*-words combine with their sister nodes (of type $\langle e, t\rangle$) uniformly via standard Function Application.

(45) Who came?

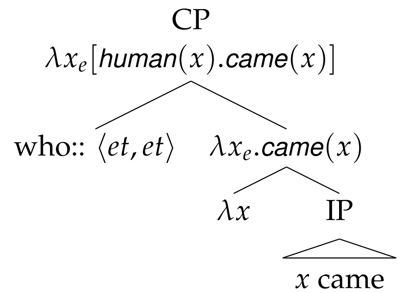


(46) Who bought what?

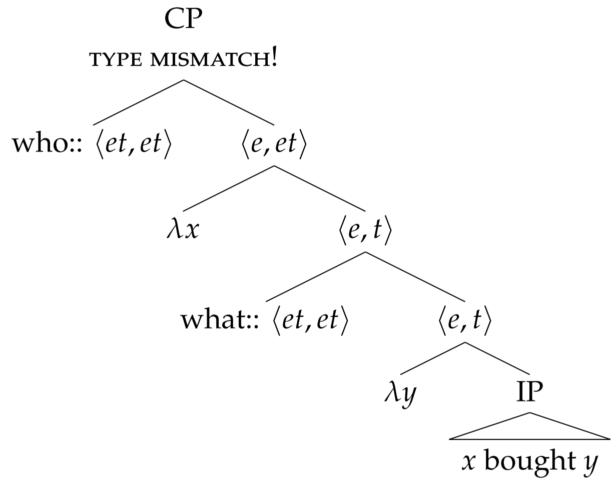


In categorial approaches, the composition of a single-*wh* question also involves *wh*-movement and Function Application. In (47) (repeated from (38)), *who* moves to [Spec, CP], applies to a predicate (of type $\langle e,t\rangle$) denoted by its sister, and returns a predicate defined only for human individuals. However, it is difficult to extend this simple derivation to the composition of a multi-*wh* question. For example, the composition of the LF in (48) suffers a type mismatch: the higher-*wh* *who* selects for an argument of type $\langle e,t\rangle$, while its sister node is of type $\langle e,et\rangle$. These two nodes cannot be combined via Function Application or Predicate Modification. (See George 2011: §2.4.2 for a solution using *tuple types*.)

(47) Who came?



(48) Who bought what?



3.2.3. Problem 3: Coordinations of questions

Conjunction and disjunction are standardly defined as meet '$\sqcap$' and join '$\sqcup$', which coordinate meanings of the same conjoinable type (Partee and Rooth 1983, Groenendijk and Stokhof 1989). (In the following formalizations, the symbols '$\wedge$' and '$\vee$' are reserved for coordinating truth values, and A' stands for the interpretation of a syntactic expression A.)

(49) **Conjoinable types**

a. t is a conjoinable type.

b. If τ is a conjoinable type, then $\langle \sigma, \tau\rangle$ is a conjoinable type for any type σ.

(50) a. **Meet**

$$A \sqcap B = \begin{cases} A \wedge B & \text{if } A \text{ and } B \text{ are of type } t \\ \lambda x. A(x) \sqcap B(x) & \text{if } A \text{ and } B \text{ are of some other conjoinable type } \tau \\ \text{undefined} & \text{otherwise} \end{cases}$$

b. **Join**

$$A \sqcup B = \begin{cases} A \vee B & \text{if } A \text{ and } B \text{ are of type } t \\ \lambda x. A(x) \sqcup B(x) & \text{if } A \text{ and } B \text{ are of some other conjoinable type } \tau \\ \text{undefined} & \text{otherwise} \end{cases}$$

See the following expressions for a simple illustration. An intransitive verb (of type $\langle e, t\rangle$) cannot be coordinated with a transitive verb (of type $\langle e, et\rangle$), as seen in (51a-b). A proper name *Jenny* (of type e) can be shifted to type $\langle et, t\rangle$ via Montague-lift and hence can be coordinated with the generalized quantifier *every student* (of type $\langle et, t\rangle$), as in (51c). But, it cannot be coordinated with a common noun such as *student* (of type $\langle e, t\rangle$) because proper names and common nouns are of different types regardless of the application of Montague-lift, as in (51d).

(51) a. jump and run — $\textit{jump}_{\langle e,t\rangle} \sqcap \textit{run}_{\langle e,t\rangle}$

b. * jump and look for — $\#\textit{jump}_{\langle e,t\rangle} \sqcap \textit{look-for}_{\langle e,et\rangle}$

c. Jenny and every student — $\textsc{Lift}(\textit{Jenny})_{\langle et,t\rangle} \sqcap \textit{every student}_{\langle et,t\rangle}$

d. * Jenny and student — $\#\textsc{Lift}(\textit{Jenny})_{\langle et,t\rangle} \sqcap \textit{student}_{\langle e,t\rangle}$
$\#\textit{Jenny}_e \sqcap \textit{student}_{\langle e,t\rangle}$

Assigning different semantic types to different questions, categorial approaches have difficulties in analyzing coordinations of question. As shown in (52), the single-*wh* question *who came* and the multi-*wh* question *who bought what* can be naturally conjoined or disjoined and embedded under an interrogative-embedding predicate. But, categorial approaches treat these two questions as of type $\langle e, t\rangle$ and $\langle e, et\rangle$, respectively. This treatment conflicts with the standard view that only items of the same conjoinable type can be coordinated. Due to this problem, many studies that acknowledge the merits of categorial approaches in analyzing matrix questions do not assume abstract/function-like denotations for embedded questions. For example, Groenendijk and Stokhof (1990) and Jacobson (2016) assume that matrix questions denote λ-abstracts, while taking embedded questions to denote partitions of possible worlds or Hamblin sets.

(52) a. John asked Mary [[who came] and/or [who bought what]].
b. John knows [[who came] and/or [who bought what]].

Moreover, even if the coordinated questions are of the same conjoinable type, traditional categorial approaches do not predict the correct reading. Consider the sentence in (53) for example. Categorial approaches define Q_1 as the set of individuals who voted for Andy, and Q_2 as the set of individuals who voted for Billy, as in (53a-b). By the standard definition of conjunction, the meet of these two sets is their intersection, namely, the set of individuals who voted for both Andy and Billy, as in (53c). In consequence, categorial approaches incorrectly predict (53) to mean that Jenny knows who voted for both Andy and Billy.[15]

(53) Jenny knows [$_{Q_1}$ who voted for Andy] and [$_{Q_2}$ who voted for Billy].
a. $[\![Q_1]\!] = \lambda x.\textit{vote-for}(x, a)$
b. $[\![Q_2]\!] = \lambda x.\textit{vote-for}(x, b)$
c. $[\![Q_1]\!] \sqcap [\![Q_2]\!] = (\lambda x.\textit{vote-for}(x, a)) \sqcap (\lambda x.\textit{vote-for}(x, b))$
$= \lambda x[\textit{vote-for}(x, a) \wedge \textit{vote-for}(x, b)]$
$= [\![\textit{who voted for Andy and Billy}]\!]$

4. A hybrid categorial approach

Previous sections have presented motivations for pursuing a categorial approach and problems with traditional categorial approaches. This section proposes a novel hybrid categorial approach to compose *wh*-questions. This approach has three basic components.

A. **Question denotation**
Root denotations of questions are topical properties, namely, functions from nominal

[15]Hamblin-Karttunen Semantics also has this problem: if questions are defined as sets of propositions, and conjunction is treated standardly as meet, the conjunction of two questions would be analyzed as the intersection of two sets of propositions. This prediction is clearly incorrect. For example, in (53), if Andy and Billy are two candidates vying for the same position and voters are only allowed to vote for at most one of them, the two coordinated questions have no propositional answer in common. A common solution of this problem is to allow the conjunction to be applied point-wise, as defined in (i): a point-wise conjunction of two sets of propositions returns a set of conjunctive propositions.

(i) **Point-wise conjunction $\cap_{\text{PW}}$**
Given two sets of propositions α and β, then $\alpha \cap_{\text{PW}} \beta = \{a \cap b \mid a \in \alpha, b \in \beta\}$

Inquisitive Semantics restores the standard treatment of conjunction. See details in Ciardelli et al. (2013), Ciardelli and Roelofsen (2015) and Ciardelli et al. (2017).

short answers to propositional answers. (§4.1)

B. **Answerhood**
A topical property can input different answerhood operations and yield either nominal or propositional answers. (§4.2)

C. **Composition**
Wh-phrases are existential quantifiers in their lexical interpretations, but are type-shifted into polymorphic domain restrictors in question composition. (§4.3)

(A) and (B) achieve the advantages of traditional categorial approaches in retrieving short answers, and (C) overcomes their problems in defining bare *wh*-indefinites and composing multi-*wh* questions with single-pair readings. An extension to multi-*wh* questions with pair-list readings will be presented separately in section 5. Discussions on coordinations of questions are postponed to section 7, because the solution is independent from the assumptions introduced in this section.

4.1. Questions as topical properties

I define the root denotation of a question as a *topical property*, namely, a λ-abstract/function ranging over propositions.[16] For example, the topical property of the question in (54) is a function that maps an atomic student *Jenny* to the proposition *that Jenny came*.

(54) Which student came? Jenny.

a. $\boldsymbol{P} = \lambda x \colon \mathit{stdt}_{@}(x) = 1.\hat{}\,\mathit{came}(x)$

b. $\boldsymbol{P}(j) = \hat{}\,\mathit{came}(j)$

Short answers, propositional answers, and partitions can be easily derived out of topical properties. Let $\boldsymbol{P}$ stand for a topical property. Short answers are elements in the domain of $\boldsymbol{P}$. Propositional answers are propositions obtained by applying $\boldsymbol{P}$ to its possible arguments. A partition is a relation between worlds such that the true short answers to this question are identical in these worlds.[17]

(55) Given a question Q with a topical property $\boldsymbol{P}$, we have:

a. Short answers to Q
$\mathrm{Dom}(\boldsymbol{P})$

b. Propositional answers to Q
$\{\boldsymbol{P}(\alpha) \mid \alpha \in \mathrm{Dom}(\boldsymbol{P})\}$

[16]The term "topical property" was firstly used by Chierchia and Caponigro (2013) to describe the meaning of a *wh*-FR derived out of a question root. This way of calling something a 'property' is a bit unconventional. Standardly, a property is a meaning of type $\langle s, et \rangle$, a function from possible worlds to predicates (viz., sets of entities).

[17]The relevant partition can also be defined based on propositional answers or complete true answers. See (63) for definitions of answerhood-operators Ans and Ans^S.

(i) a. Defining **partition** based on true propositional answers
$\lambda w \lambda w'[Q_w = Q_{w'}]$, where $Q_w = \{\boldsymbol{P}(\alpha) \mid \alpha \in \mathrm{Dom}(\boldsymbol{P}) \wedge w \in \boldsymbol{P}(\alpha)\}$

b. Defining **partition** based on complete true short answers
$\lambda w \lambda w'[\text{Ans}^S(w)(\boldsymbol{P}) = \text{Ans}^S(w')(\boldsymbol{P})]$

c. Defining **partition** based on complete true propositional answers
$\lambda w \lambda w'[\text{Ans}(w)(\boldsymbol{P}) = \text{Ans}(w')(\boldsymbol{P})]$

c. Partition of possible worlds of Q
$\lambda w \lambda w'[\boldsymbol{P}_w = \boldsymbol{P}_{w'}]$, where $\boldsymbol{P}_w = \{\alpha \mid w \in \boldsymbol{P}(\alpha)\}$

For example, for the singular-marked question in (54), two worlds w and w' are in the same cell iff the very same set of atomic students came in w and w'. With two relevant students Jenny and Mary, the partition of (54) is illustrated in Table 5. For instance, the first cell stands for the set of worlds where only Jenny and Mary came, or equivalently, where the true short answers of the question *which student came* include only the two atomic individuals *Jenny* and *Mary*.

w: only j and m came in w		w: $\{\alpha \mid w \in \boldsymbol{P}(\alpha)\} = \{j, m\}$
w: only j came in w	=	w: $\{\alpha \mid w \in \boldsymbol{P}(\alpha)\} = \{j\}$
w: only m came in w		w: $\{\alpha \mid w \in \boldsymbol{P}(\alpha)\} = \{m\}$
w: nobody came in w		w: $\{\alpha \mid w \in \boldsymbol{P}(\alpha)\} = \varnothing$

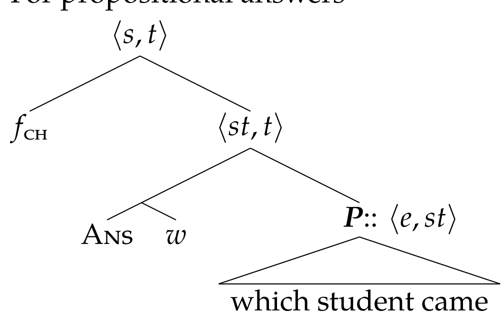
Table 5: Partition for *which student came*

As such, any information that is derivable from the Hamblin set or the partition of a question is also derivable from the topical property of this question. Hence, defining questions as topical properties, we can still capture the objective of Hamblin-Karttunen Semantics in modeling the relationship between questions and propositional answers, as well as the objective of Partition Semantics in modeling world-dependency in question-embeddings.

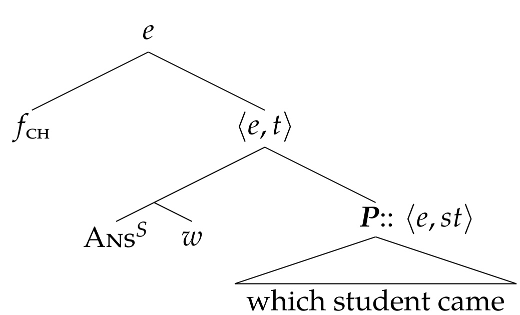
4.2. Answerhood

The topical property of a question directly combines with an answerhood-operator, returning a set of complete true answers. These answers can be propositional or nominal/short, depending on whether the employed answerhood-operator is Ans or Ans^S. (The superscript 'S' in Ans^S stands for 'short'). Finally, a choice function f_{CH} is applied to pick out one of these complete true answers. If a question has only one complete true answer, the output set of employing Ans is a singleton set, and employing f_{CH} returns the unique member of this set. The derivational procedures of the answers are illustrated as follows:

(56) a. For propositional answers



b. For short answers



Note that answerhood is concerned about what *meaning* (as opposed to *expression*) constitutes a good answer to a question. Moreover, the Ans/Ans^S -operator and f_{CH}-function are not necessarily syntactically present in the structure. They can be incorporated into an interpretation rule (for matrix questions) or the lexicon of a question-embedding predicate (for embedded questions).

An important feature of the proposed framework is that the procedure of question formation itself has no stage that creates a Hamblin set or a Karttunen set. Instead, answerhood-operators directly operate on the topical property and hence they can access any information that is retrievable from the topical property, especially the property domain (or equivalently, the possible short answers). This consequence makes the proposed analysis significantly different from George (2011) and a recent analysis by Champollion et al. (2015) using Inquisitive Semantics. The latter two analyses also start with a λ-abstract, but then use a question-formation operator to convert this abstract into a partition or a Hamblin set. In these two proposals, an answerhood-operator or any operation outside a question cannot interact with the domain of the λ-abstract.

In defining complete true answers, I adopt Fox's (2013) view that complete answers can be non-exhaustive, and that a question can have multiple complete true answers. Compared with the more commonly adopted answerhood-operators which require complete answers to be exhaustive (Heim 1994; Dayal 1996), Fox's answerhood-operator leaves space for analyzing mention-some readings of questions. To see what it is meant by "mention-some", compare the following questions:

(57) Who went to the party?
(*w*: *Only John and Mary went to the party.*)
a. John and Mary.
b. John did .../ $\quad\rightsquigarrow$ *I don't know who else did.*
L H* L-H%
c. # John did.\ $\quad\rightsquigarrow$ *Only John did.*
H* L-L%

(58) Who can chair the committee?
(*w*: *Only John and Mary can chair; single-chair only.*)
a. John can.\ $\quad\not\rightsquigarrow$ *Only John can chair.*
b. John and Mary.\
c. John or Mary.\

The simple question in (57) requires the addressee to specify all the actual attendants to the party, as in (57a). If the addressee can only provide a non-exhaustive answer, she would have to indicate the incompleteness of her answer in some way, such as marking the answer with a prosodic rise-fall-rise contour (indicated by '.../'), as in (57b). If an incomplete answer is not properly marked, as in (57c) which takes a falling tone (indicated by '\'), it gives rise to an undesired exhaustive inference. In contrast, as in (58), questions with a possibility modal (henceforth called $\Diamond$-questions) admit non-exhaustive answers (Groenendijk and Stokhof 1984) — the question in (58) can be naturally answered by specifying one of the chair candidates, as in (58a). Crucially, while being non-exhaustive, the answer (58a) does not need to carry an incompleteness-marking intonation: it does not yield an exhaustivity inference even if taking a falling tone. To distinguish answers like (58a) from other non-exhaustive answers, we call (58a) a "mention-some answer", questions that admit mention-some answers "mention-some questions", and readings under which a question admits mention-some answers "mention-some readings".

A proper treatment of mention-some is crucial to the issues concerned in this paper, because the mention-some/mention-all contrast is systematically observed in *wh*-constructions with predicative or nominal interpretations. For example, in (59), a *wh*-FR takes an existential reading iff its interrogative counterpart is mention-some (Chierchia and Caponigro 2013). Likewise, as exemplified in

(60), a Mandarin *wh*-conditional — a construction consisting of two *wh*-clauses and expressing an inclusion relation between the short answers of the interrogative counterparts of these two clauses — takes an existential reading iff the form of the antecedent *wh*-clause resembles a mention-some question (Liu 2016). Using a categorial approach, once we understand the constraints on the distribution of mention-some readings in matrix questions, we can easily explain the distribution of existential readings in these predicative/nominal constructions. I will return to the derivations of their interpretations in section 6.1.[18]

(59) **Free relatives**

a. John ate what Mary cooked for him.
⇝ *John ate **everything** that Mary cooked for him.*

b. John went to where he could get help.
⇝ *John went to **some** place where he could get help.*

(60) **Mandarin *wh*-conditionals**

a. Ni qu-guo nar, wo jiu qu nar.
you go-EXP where, I JIU go where
'Where you have been to, I will 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
'Where I can buy liquor, I will go where.'
Intended: 'I will go to **some** place where I can buy liquor.'

Fox (2013) proposes that a true answer to a question is complete as long as it is not asymmetrically entailed by any true answers to this question. Following Hamblin-Karttunen Semantics and defining a question root as a Hamblin set (written as 'Q'), Fox defines the answerhood-operator as in (61):

(61) $\textsc{Ans}_{Fox}(w)(Q) = \{p \mid w \in p \in Q \wedge \forall q[w \in q \in Q \rightarrow q \not\subset p]\}$ (Fox 2013)
($\{p \mid p$ is a true proposition in Q, and p is not asymmetrically entailed by any true propositions in $Q\}$)

[18] A similar distributional pattern of mention-some is observed with question-answer (QA-)clauses in American Sign Language (Davidson et al. 2008; Caponigro and Davidson 2011). A QA-clause is uttered by a single signer. It consists of two parts, namely, a question constituent which looks like an interrogative clause conveying a question, and an answer constituent which resembles a propositional answer or a short answer to that question. As shown below, just like their corresponding discourse-level question-answer pairs in (a), the answer constituent of each QA-clause in (b) resembles a mention-some answer iff the question constituent resembles a $\Diamond$-question.

(i) (Context: *John bought a book, a CD, and a DVD.*)

a. Signer A: JOHN BUY WHAT? Signer B: #BOOK.
'John bought what?' 'Book.'

b. JOHN BUY WHAT, # BOOK.
'What John bought is a book.'

(ii) (Context: *There are two coffee places nearby, Starbucks and Peet's.*)

a. Signer A: CAN FIND COFFEE WHERE? Signer B: STARBUCKS.
'Where can you find coffee?' 'Starbucks.'

b. CAN FIND COFFEE WHERE, STARBUCKS.
'You can find coffee at Starbucks.'

When the concerned question takes a mention-some reading, $\textsc{Ans}_{Fox}(w)(Q)$ consists of multiple true mention-some answers to this question; otherwise it is a singleton set consisting of only the strongest true answer to this question. Due to the scope of this paper and the complexities of deriving mention-some readings, I will not dive into the details. See Xiang (2016: chap. 2) for a review of Fox (2013) and a new explanation to the mention-some/mention-all ambiguity compatible with Fox's definition of answerhood. Here the following are the predicted complete true answers for the questions in (57) and (58) following Fox (2013) and Xiang (2016: chap. 2):

(62) a. For the mention-all question (57) :
$\textsc{Ans}_{Fox}(w)(Q) = \{\textsf{John and Mary went to the party}\}$

b. For the $\Diamond$-question (58) with a mention-some reading:
$\textsc{Ans}_{Fox}(w)(Q) = \{\textsf{John alone can chair the committee, Mary alone can chair the committee}\}$

Adapting Fox's answerhood-operator to the proposed hybrid categorial approach, I define the answerhood-operators as follows. The major revision is replacing a Hamblin set Q with a topical property $\boldsymbol{P}$, which can supply both propositional answers and short answers.[19]

(63) **Answerhood-operators**

a. For short answers

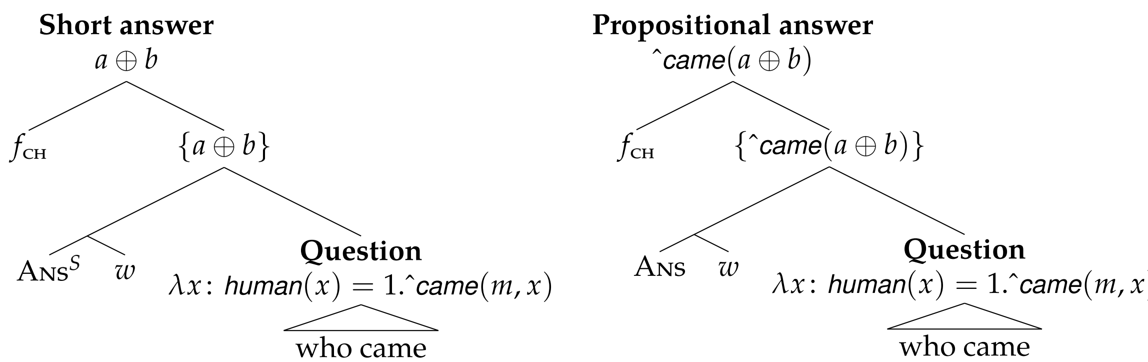
$$\textsc{Ans}^S(w)(\boldsymbol{P}) = \left\{\alpha \;\middle|\; \begin{array}{l} \alpha \in \mathrm{Dom}(\boldsymbol{P}) \wedge w \in \boldsymbol{P}(\alpha) \wedge \\ \qquad \forall\beta \in \mathrm{Dom}(\boldsymbol{P})[w \in \boldsymbol{P}(\beta) \rightarrow \boldsymbol{P}(\beta) \not\subset \boldsymbol{P}(\alpha)] \end{array}\right\}$$

b. For propositional answers

$$\textsc{Ans}(w)(\boldsymbol{P}) = \{\boldsymbol{P}(\alpha) \mid \alpha \in \textsc{Ans}^S(w)(\boldsymbol{P})\}$$

Trees in (64) exemplify the applications of $\textsc{Ans}$ and $\textsc{Ans}^S$ to single-*wh* questions. Again, the trees illustrate the derivations of answers, not the LFs of matrix questions. The answerhood-operators and choice functions are semantically active but not necessarily syntactically present.

(64) (w: *Only Andy and Billy came.*)



Next, turn to multi-*wh* questions with single-pair readings. Strictly speaking, the derived root denotation $\boldsymbol{P}$ in (65a) is not a topical property by the definition in section 4.1; it is a function from

[19] In (63), propositional answers are defined in terms of short answers. This definitions are made in line with the tradition of categorial approaches, which regards short answers as default answers. However, technically, this analysis would be no different if it starts by defining propositional answers and then defines complete true short answers as meanings in the domain of the topical property such that combining the topical property with these meanings derive the complete true propositional answers.

atomic boys to a topical property of atomic girls. Moreover, its domain is not a set of short answers to a multi-*wh* question, and its range is not a set of propositions. Then, how can we derive answers from this $\boldsymbol{P}$? The simplest solution that I have seen so far is to use *tuple types*, an idea developed by George (2011: Appendix A). George writes an *n*-ary sequence as $(x_1; x_2; ...; x_n)$ which has a tuple type $(\tau_1; \tau_2; ...; \tau_n)$, and then equates a function from a denotation of tuple type $(\tau_1; \tau_2; ...; \tau_n)$ to type σ, namely, of type $\langle(\tau_1; \tau_2; ...; \tau_n), \sigma\rangle$, with a denotation of type $\langle\tau_1\langle\tau_2\langle...\langle\tau_n, \sigma\rangle...\rangle\rangle\rangle$. For instance, $\langle e, \langle e, st\rangle\rangle$ is identified with $\langle(e; e), st\rangle$. Following this idea, we can consider the abstract $\boldsymbol{P}$ in (65a) as a property of duple-sequences of an atomic boy and an atomic girl and write its domain as (65b-i). Then answerhood-operations proceed regularly.

(65) Which boy invited which girl?
(*w*: *John invited only Mary; no other boy invited any girl.*)

a. Topical property
$\boldsymbol{P} = \lambda x \lambda y\colon \mathsf{boy}_@(x) = 1 \wedge \mathsf{girl}_@(y) = 1.\,\hat{}\,\mathsf{invt}(x, y)$

b. Possible short answers and complete true short answers

i. $\mathrm{Dom}(\boldsymbol{P}) = \{(x; y) \mid x \in \mathsf{boy}_@, y \in \mathsf{girl}_@\}$

ii. $\textsc{Ans}^S(w)(\boldsymbol{P}) = \{(j; m)\}$

c. Possible propositional answers and complete true propositional answers

i. $\{\boldsymbol{P}(\alpha) \mid \alpha \in \mathrm{Dom}(\boldsymbol{P})\} = \{\hat{}\,\mathsf{invt}(x, y) \mid x \in \mathsf{boy}_@, y \in \mathsf{girl}_@\}$

ii. $\textsc{Ans}(w)(\boldsymbol{P}) = \{\hat{}\,\mathsf{invt}(j, m)\}$

4.3. Composition of questions

This section focuses on single-*wh* questions and multi-*wh* questions with single-pair readings. For functional readings and pair-list readings, see section 5.

4.3.1. Assumptions on the formal theory

I will follow Heim and Kratzer (1998) in assuming that semantic composition takes place at the syntactic level of Logical Form (LF). I also adopt the view of Huang (1982) that *wh*-movement can take place either overtly at surface structure or covertly at LF. While not entirely necessarily, these assumptions are handy for dealing with cases where a *wh*-item is not overtly fronted.[20]

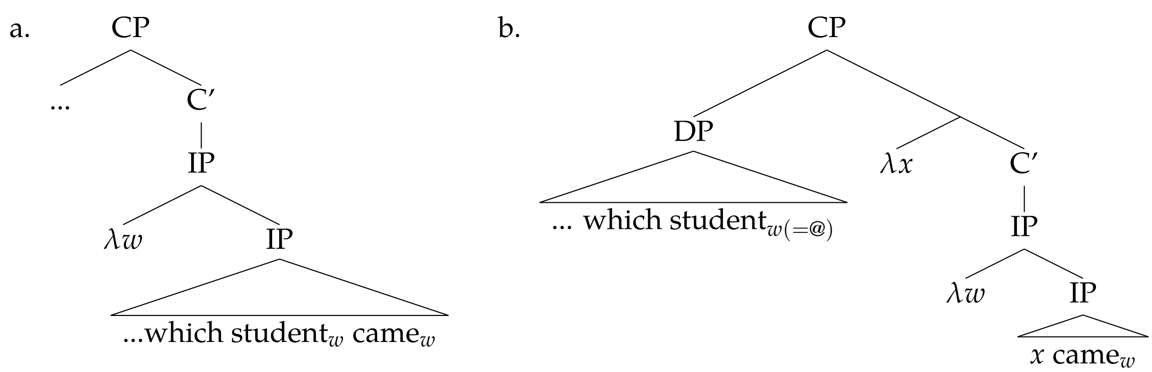
An LF representation of the object language (e.g., English) is **translated** into a metalinguistic representation with an unambiguous model-theoretic interpretation. The metalanguage I use for translation is the Two-sorted Type Theory (Ty2) of Gallin (1975). (See Gamut 1991: section 5.8 for a nice brief introduction.) Due to its convenience in handling intensional meanings, Ty2 is commonly used in modeling question semantics (Groenendijk and Stokhof 1984: chapter 2, a.o.). Compared with Montague's Intensional Logic (IL), Ty2 is different in that it introduces *s* (the type of possible worlds) as a basic type (just like *e* and *t*), and in that it uses variables and constants of type *s* which

[20]LF is assumed to resolve the (seeming) syntax-semantics mismatch on scoping and binding — using covert syntactic operations, the linear order of strings at LF can be different from that at the surface structure and be made more in parallel with semantics. Nevertheless, the seeming syntax-semantics mismatch can be alternatively resolved without assuming an abstract level like LF. For example, scoping can be done by type-shifting operations (Hendriks 1993; Shan and Barker 2006; a.o.), and binding can be done via the combinatory rules in Variable-free Semantics (Jacobson 1999). Hence, for the issues concerned in this paper, it is not necessary to commit to an LF theory.

can be thought of as denoting possible worlds. For example, the English common noun *student* is translated into $\textsf{student}_w$ in Ty2, where $\textsf{student}$ is a constant of type $\langle s, et\rangle$ and w a variable of type s. As such, Ty2 can make direct reference to worlds and allows quantification and abstraction over world variables. Basically, for a translation α shared by Ty2 and IL, $\lambda w.\alpha$ and $\alpha(w)$ in Ty2 correspond to the intension $\hat{}\alpha$ and the extension $\check{}\alpha$ in IL, respectively. For ease of presentation, I will use λw and $\hat{}$ interchangeably for intensionalization. (See Gamut 1991: p. 136 for the correspondence of the model-theoretic interpretations of these two notions.) Next, a Ty2 formula is **interpreted** with respect to a model $M = \langle D, W, I\rangle$ (D: a set of individuals, W: a set of worlds, I: an interpretation function) and an assignment function g. For example, for the Ty2-translation $\textsf{student}_w$, the predicate $\textsf{student}$ is interpreted as $I(\textsf{student})$ and the world variable w as $g(w)$.

I assume that every predicative word (noun, verb, or predicative adjective) in the object language carries a world argument, filled by general rules in Ty2. This assumption makes it easy to think of interpreting object language expressions as interpreting Ty2 translations. The world variable binder λw is introduced at the topmost edge of IP. The tree diagrams in (66) illustrate general LF schema for *wh*-questions in English. (Here '...' stands for silent syntactic material to be introduced later.) In (66a), before *wh*-movement, the noun *student* and the verb *came* carry the same world variable w, and both occurrences of w are bound by the λw-operator introduced right above IP. In (66b), as the *wh*-phrase is moved across the λw-binder, the variable w carried by *student* becomes unbound, yielding a *de re* reading. To highlight that the world variable of the *wh*-complement is free (or at least can be free), I will replace it with the constant @ which stands for the actual world. (This paper considers only *de re* readings and interprets *wh*-complements uniformly relative to the actual world. See Beck and Rullmann (1999) and Sharvit (2002) for ways of deriving *de dicto* readings.)

(66) Which student came?



4.3.2. Composing basic *wh*-questions

The next goal is to derive the root denotations of *wh*-questions compositionally. I will map the tree diagram (66b) to the following topical property (repeated from (54a)):

(67) Which student came?
$\boldsymbol{P} = \lambda x\colon \textsf{stdt}_{@}(x) = 1.\hat{}\textsf{came}(x)$

It is easy to see that the *wh*-phrase serves as a function domain restrictor — *which student* combines with a function defined for any entities and returns a more restrictive function only defined for atomic students. Moreover, the composition overcomes the problems with traditional categorial approaches in getting bare *wh*-indefinites (§3.2.1) and composing multi-*wh* questions (§3.2.2). More precisely, *wh*-expressions, especially those having a bare indefinite use, need to be interpreted existentially, but in questions, *wh*-expressions serve as domain restrictors of properties and need to be able to combine with functions of various types. As such, all we need is an operation that shifts an existential quantifier into a type-flexible function domain restrictor.

Observe in (67) that the domain of $\boldsymbol{P}$ is equivalent to the set denoted by the *wh*-complement *student*. Defining *which student* as a simple existential indefinite (Karttunen 1977; Heim 1995), we can extract out the set $[\![student_@]\!]$ by employing the type-shifter Be. As schematized in (68), Be shifts an existential quantifier to the set it ranges over (Partee 1986).

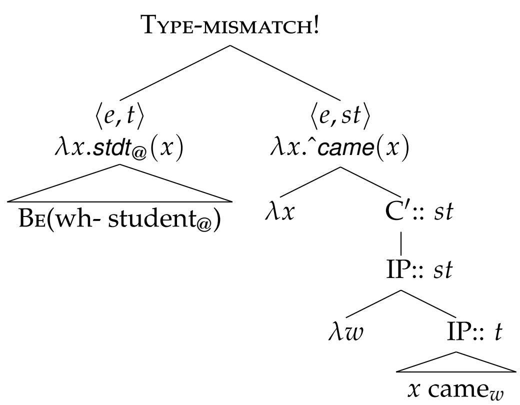
(68) a. $[\![\textit{which student}_@]\!] = \lambda f_{\langle e,t\rangle}.\exists x \in \textsf{stdt}_@[f(x)]$ (To be revised in (79))

b. $\textsc{Be} = \lambda\pi\lambda z.\pi(\lambda y.y = z)$

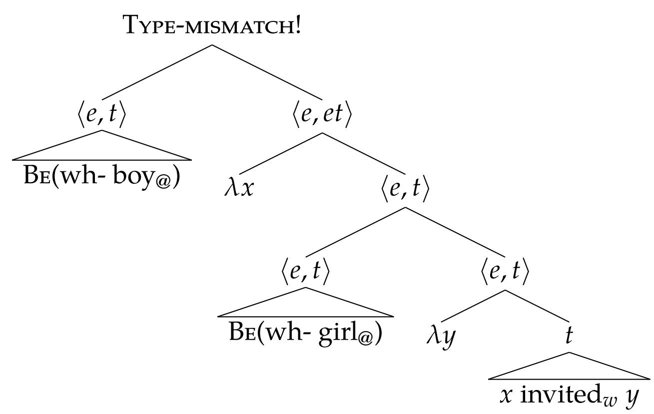
c. $$\begin{aligned}\textsc{Be}([\![\textit{which student}_@]\!]) &= \lambda z[(\lambda f_{\langle e,t\rangle}.\exists x \in \textsf{stdt}_@[f(x)])(\lambda y.y = z)] \\ &= \lambda z.\exists x \in \textsf{stdt}_@[(\lambda y.y = z)(x)] \\ &= \lambda z.\exists x \in \textsf{stdt}_@[x = z] \\ &= \{z \mid z \in \textsf{stdt}_@\} \\ &= \textsf{stdt}_@\end{aligned}$$

The next step is to combine the property domain $\textsc{Be}([\![\textit{wh- student}_@]\!])$ with the property nucleus derived from the sister of 'Be(wh- student)' at LF (i.e., $\lambda x.{}^\wedge\textsf{came}(x)$). It would be nice to be able to compose these two pieces via Predicate Modification (Heim and Kratzer 1998). However, Predicate Modification is not applicable here because of a type mismatch. First, as shown in (69), 'Be(wh-student$_@$)' is extensional (of type $\langle e,t\rangle$) while its sister node is intensional (of type $\langle e,st\rangle$). Second, a more severe problem arises in deriving the single-pair reading of a multi-*wh* question. For example, even if we neglect the extension-vs.-intension mismatch and do not assume abstraction of worlds at the edge of IP, the composition in (70) still suffers a type mismatch: 'Be(wh- boy$_@$)' is of type $\langle e,t\rangle$, but its sister node is of type $\langle e,et\rangle$. The two nodes cannot be combined via Function Application or Predicate Modification. In short, Predicate Modification is too restrictive in that it only be used to combine nodes of the same type.

(69) Which student came?



(70) Which boy invited which girl?



Rather than appeal to Predicate Modification, I introduce a new type-shifter BeDom, which shifts a bare *wh*-indefinite (or any existential quantifier) π into a **polymorphic domain restrictor**. As schematized in (71), $\textsc{BeDom}(\pi)$ applies to a function θ and returns a similar function P with a more restrictive domain — compared with that of θ, the domain of P is further restricted by the set $\textsc{Be}(\pi)$.[21] Type-wise, to ensure that the output function has a non-empty domain, the only needed restriction is that the domain of the input function θ have the same type as $\textsc{Be}(\pi)$ — if π is of type $\langle at, t\rangle$, then $\textsc{BeDom}(\pi)$ is defined for any function of an arbitrary type that starts with a.[22]

(71) **Type-shifter $\textsc{BeDom}$** (for shifting bare *wh*-indefinites)
$\textsc{BeDom}(\pi) = \lambda\theta_\tau.\iota P_\tau[[\mathrm{Dom}(P) = \mathrm{Dom}(\theta) \cap \textsc{Be}(\pi)] \wedge \forall\alpha \in \mathrm{Dom}(P)[P(\alpha) = \theta(\alpha)]]$

See (72) for a simple illustration. Assume that among the three relevant individuals abc, only a and b are students. A *came*-function θ defined for any atomic or sum individual in the domain of discourse maps each such individual x to the proposition $\hat{}\,\mathit{came}(x)$. Applying the domain restrictor $\textsc{BeDom}([\![\textit{which student}_@]\!])$ to θ excludes the individuals that are not in the set $\textsc{Be}(\text{which student}_@)$ (i.e., it excludes the individuals that are not atomic students) from the domain, and retains the x-to-$\hat{}\,\mathit{came}(x)$ mapping for the remaining individuals. In consequence, this application returns a partial *came*-function defined only for atomic students.

(72) (*Among the three relevant individuals abc, only ab are students in the actual world.*)

a. $$\theta = \begin{bmatrix} a \to \hat{}\,\mathit{came}(a), & a\oplus b \to \hat{}\,\mathit{came}(a\oplus b), & a\oplus b\oplus c \to \hat{}\,\mathit{came}(a\oplus b\oplus c) \\ b \to \hat{}\,\mathit{came}(b), & b\oplus c \to \hat{}\,\mathit{came}(b\oplus c), & \\ c \to \hat{}\,\mathit{came}(c), & a\oplus c \to \hat{}\,\mathit{came}(a\oplus c), & \end{bmatrix}$$

b. $$\textsc{BeDom}([\![\textit{which student}_@]\!])(\theta) = \begin{bmatrix} a & \to & \hat{}\,\mathit{came}(a) \\ b & \to & \hat{}\,\mathit{came}(b) \end{bmatrix}$$

To see how the composition of single-*wh* and multiple-*wh* questions works out in practice, consider the tree diagrams in (73) and (74). In both cases, $\textsc{BeDom}$ is applied to each *wh*-phrase, and the type-shifted *wh*-phrases move to [Spec, CP] to check off their [+*wh*] feature. The $\textsc{BeDom}$-shifter can be viewed either as the interpretation of a syntactic DP-adjunct or as a purely semantic type-shifting rule. The $\textsc{BeDom}$-shifter turns a *wh*-phrase into a polymorphic domain restrictor (of type $\langle\tau,\tau\rangle$, where τ stands for an arbitrary type starting with e): the output partial property P has the identical semantic type as the input property θ. In (73), '$\textsc{BeDom}$(which student$_@$)' applies to a *came*-property defined for any entities, and returns a more restrictive *came*-property only defined for atomic students. Likewise, in (74), '$\textsc{BeDom}$(which boy$_@$)' applies to a total function of type $\langle e, \langle e, st\rangle\rangle$, and returns a partial function of type $\langle e, \langle e, st\rangle\rangle$ defined only for atomic boys. Superior to traditional categorial approaches, since $\textsc{BeDom}$(*wh*P) can combine with expressions of any type starting with e, this way of composition does not suffer type mismatch.

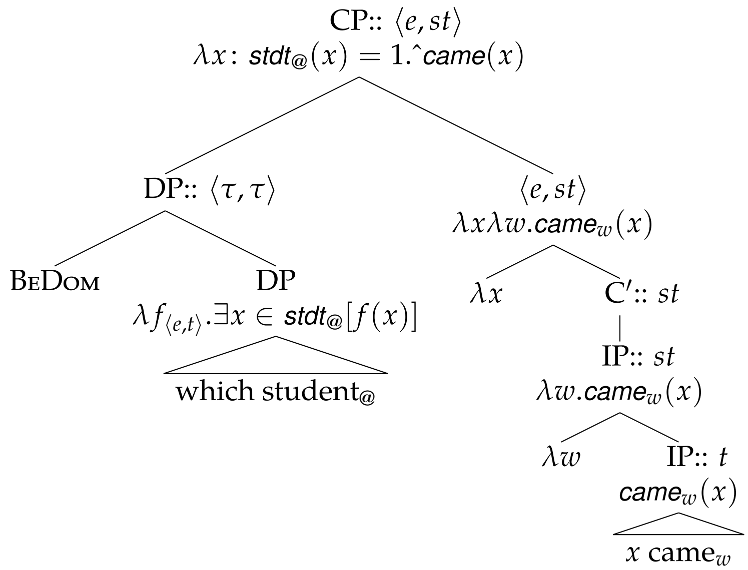
[21] The definition of $\textsc{BeDom}$ in (71) handles the case of bare *wh*-indefinites, which are existential in the lexicon. For complex *wh*-indefinites, however, their existential force might come from external existential operators (section 3.2.1). In such a case, we can alternatively adopt the definitions of *wh*-expressions assumed by traditional categorial approaches (e.g., $[\![\textit{which student}_@]\!] = \lambda P_{\langle e,t\rangle}\lambda x_e\colon \mathit{stdt}_@(x).P(x)$) and define the $\textsc{BeDom}$-operator as in (i): for any function f, $\textsc{BeDom}(f)$ restricts any input function with the domain of the function f.

(i) **Type-shifter $\textsc{BeDom}$** (for shifting complex *wh*-indefinites)
$\textsc{BeDom}(f) = \lambda\theta_\tau.\iota P_\tau[[\mathrm{Dom}(P) = \mathrm{Dom}(\theta) \cap \mathrm{Dom}(f)] \wedge \forall\alpha \in \mathrm{Dom}(P)[P(\alpha) = \theta(\alpha)]]$

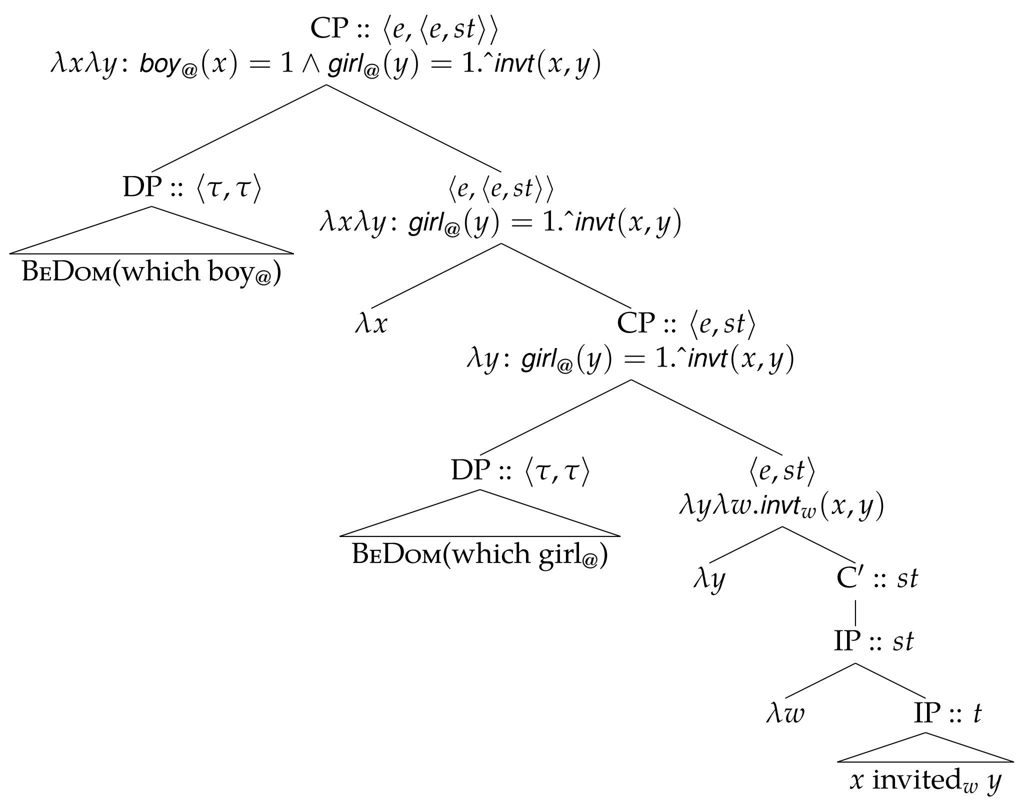
For the rest of the paper, I consider only bare *wh*-indefinites and use the $\textsc{BeDom}$-operator in (71).

[22] Here I consider the case where π is simply typed, namely, all the elements in $\textsc{Be}(\pi)$ have the same type. However, as will be seen in (79), the restriction of a *wh*-determiner can be polymorphic. If π is polymorphic, the only needed type restriction with $\textsc{BeDom}(\pi)$ would be that the elements in the domain of the input function θ have the same type as some elements in $\textsc{Be}(\pi)$.

(73) Which student came?



(74) Which boy invited which girl? (Single-pair reading)



4.4. Interim summary

To sum up, the proposed hybrid categorial approach has three basic ingredients.

First, the root denotation of a question (matrix or embedded) is a topical property, which can supply both nominal short answers and propositional answers.

Second, an answerhood-operator, which might be encoded within an interpretation rule or a question-embedding predicate, directly operates on the topical property and returns a set of complete true answers. The returned answers can be nominal or propositional, depending on the employed answerhood-operator. In particular, answerhood is defined following Fox (2013), which allows mention-some answers to be complete answers.

Third, in languages with bare *wh*-indefinites, this root denotation is derived as follows: (i) *wh*-phrases are existential quantifiers in the lexicon, but in questions they are shifted into polymorphic domain restrictors via the application of a BeDom-operator; (ii) moving BeDom(*wh*P) to [Spec, CP] yields a partial property that is defined for only elements in the quantification domain of the *wh*-phrase.

5. Functional readings and pair-list readings

Traditional categorial approaches didn't make an attempt to derive pair-list readings of multi-*wh* questions. This section will first discuss functional readings and then present a function-based approach to pair-list readings.

Function-based approaches to question semantics were firstly proposed to deal with pair-list readings of questions with a universal quantifier ($\forall$-questions henceforth) and then were extended to multi-*wh* questions (Engdahl 1980, 1986; Groenendijk and Stokhof 1984; Chierchia 1993; Dayal 1996, 2017). The following examples illustrate individual readings and basic functional readings of single-*wh* questions and $\forall$-questions. Individual readings expect answers that specify a particular individual, while functional readings expect answers that specify a Skolem function of type $\langle e, e\rangle$ defined for the subject-name (e.g., *Andy*) or individuals in the domain of the subject-quantifier (e.g., individuals in the domain of *every boy*).

(75) Which girl did Andy/ every boy invite?
  a. Mary. (Individual answer)
  b. His girlfriend. (Functional answer)

Dayal (1996, 2017) argues that pair-list answers to a multi-*wh* question involve a functional dependency between the quantification domains of the *wh*-items. For instance, the pair-list reading of (76) asks about a function $\mathbf{f}$ from atomic boys to atomic girls such that each boy x invited the girl $\mathbf{f}(x)$, as illustrated in (77).[23] One way to identify this function is to list all the boy-invite-girl pairs, which is therefore a pair-list answer.

(76) Which boy invited which girl?
  $\approx$ 'For which function $\mathbf{f}$ from $\textsf{boy}_@$ to $\textsf{girl}_@$ is such that x invited $\mathbf{f}(x)$?'

[23] Dayal (1996, 2002, 2017) claims that multi-*wh* questions with pair-list readings are subject to domain exhaustivity, such as that the multi-*wh* question (76) presupposes that every considered boy invited some girl. By contrast, I argue that such domain exhaustivity is involved in questions with universal quantifiers, but not in multi-*wh* questions. Compare the following two sentences. In the given context, the quantification domain of the subject-*wh*/quantifier is greatly larger than that of the object-*wh*. The contrast between (ia-b) shows that the multi-*wh* question in (ia) is not subject to domain exhaustivity, or at least that its domain exhaustivity effect, if any, is much less robust than that of the $\forall$-question in (ib).

(i) (*w*: *100 candidates are competing for three job openings.*)
  a. $\surd$ Guess which candidate will get which job.
  b. # Guess which job every/each candidate will get.

(77) (w: *Andy invited Mary, Billy invited Jenny, Clark invited Sue; no other boy invited any girl.*)

$$\mathbf{f} = \begin{bmatrix} a & \rightarrow & m \\ b & \rightarrow & j \\ c & \rightarrow & s \end{bmatrix}$$

Adopting this idea, I treat pair-list readings of multi-*wh* questions as special functional readings. This section will firstly discuss the derivation of a basic functional reading by enriching the *wh*-lexicon, and then show the composition of pair-list readings for multi-*wh* questions.[24]

5.1. Functional readings

In section 4.3.2, I treated a *wh*-phrase as an existential quantifier ranging over the set denoted by the extension of the *wh*-complement, repeated in (78). To get functional readings, I propose that the set that '*which*-A' ranges over is polymorphic — this set consists of not only individuals in A (of type e) but also Skolem functions to A (of type $\langle e, e\rangle$).[25] The meaning of *which* is then modified as in (79).

(78) **Lexical entries of *wh*-items** (Old definition)

a. $[\![which]\!] = \lambda A_{\langle e,t\rangle}\lambda P_{\langle e,t\rangle}.\exists x \in \textsf{stdt}_@[P(x)]$

b. $\textsc{Be}([\![which\ student_@]\!]) = \textsf{stdt}_@$

(79) **Lexical entries of *wh*-items** (Revised definition)

a. $[\![which]\!] = \lambda A \lambda P.\exists x \in (A \cup \{\mathbf{f} \mid \text{Range}(\mathbf{f}) \subseteq A\})[P(x)]$

where $\text{Range}(\mathbf{f}) \subseteq A$ iff $\forall x \in \text{Dom}(\mathbf{f})[\mathbf{f}(x) \in A]$

b. $\textsc{Be}([\![which\ student_@]\!]) = \textsf{stdt}_@ \cup \{\mathbf{f} \mid \text{Range}(\mathbf{f}) \subseteq \textsf{stdt}_@\}$

In a *wh*-question, the semantic type of the topical property is determined by the type of the highest *wh*-trace, as shown in (80). If the *wh*-item moves directly from the *in situ* position and leaves only an individual trace, the obtained topical property is a function over individuals, yielding an individual reading. If the movement of a *wh*-item leaves a functional trace, the obtained topical property is a function over Skolem functions.[26]

[24] I will not discuss pair-list readings in questions with quantifiers, since they involve further complications. See Xiang (2016: chap. 6), Dayal (2017: chap. 4), and Ciardelli and Roelofsen (2018) for recent overviews on composing questions with quantifiers.

[25] Some *wh*-questions admit also higher-order readings, under which they expect answers naming a generalized quantifier (Spector 2007, 2008; Fox 2013; Xiang 2016). As such, we expect that in those cases the domain of the *wh*-item contains also Skolem functions to generalized quantifiers (of type $\langle e, \langle et, t\rangle\rangle$), or generalized quantifiers ranging over Skolem functions (of type $\langle\langle ee, t\rangle, t\rangle$. For example in (i), the disjunction in the answer is interpreted under the necessity modal *have to*. This reading requires the disjunctive answer to be interpreted as a Boolean disjunction over two Skolem functions and the (highest) *wh*-trace to be of type $\langle e, \langle et, t\rangle\rangle$ (or $\langle\langle ee, t\rangle, t\rangle$, so that the disjunction can be semantically reconstructed to a position under the necessity modal. For a compositional derivation see Xiang (2016: section 5.4.1).

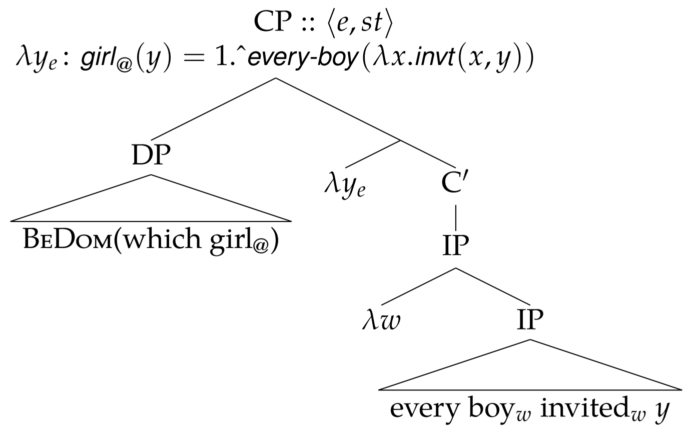
(i) Speaker A: 'Who does Andy have to invite?'
Speaker B: 'His mother or sister. [The choice is up to him.]'

[26] More precisely, the meaning of a functional answer is intensional (of type $\langle e, se\rangle$) — it specifies a function from individuals to individual concepts. A more precise schematization for the topical property of the functional reading (80b) is as follows:

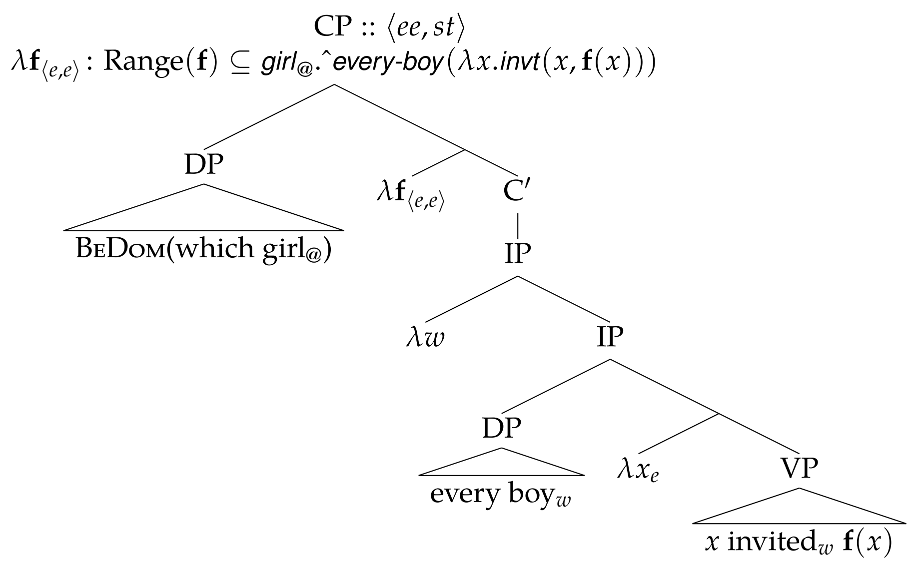
(i) $\lambda \mathbf{f}_{\langle e,se\rangle} : \forall y \forall w[\mathbf{f}(y)(w) \in \textsf{girl}_@].\lambda w[\forall x[\textsf{boy}_w(x) \rightarrow \textsf{invt}_w(x, \mathbf{f}_w(x))]]$

(80) Which girl did every boy invite?

a. Individual reading: 'Which girl x is such that every boy invited x?'



b. Functional reading: 'Which function $\mathbf{f}$ to $\textit{girl}_{@}$ is such that every boy x invited $\mathbf{f}(x)$?'



The functional reading in (80b) involves double-binding — the *wh*-phrase binds a functional trace, while the subject-quantifier *every boy* binds the argument variable of this functional trace. Existing literature has offered two approaches to get this binding relation. One assumes a complex functional trace with multiple indices (Groenendijk and Stokhof 1984; Chierchia 1993; Dayal 1996, 2017), and the other uses a combinatory rule that closes off the anaphoric dependency between the arguments of a transitive verb *inivte* (Jacobson 1994, 1995, 1999).[27] For compatibility with the

[27] The Variable-free Semantics approach developed by Jacobson does not assume complex functional traces or use indices at all. As seen in footnote 13, even in a basic *wh*-question, the *wh*-trace is functional — it is interpreted as an identity function over individuals. As such, there is no need to stipulate a contrast between simple (individual) traces and complex (functional) traces. Instead, as seen in the following, the contrast between individual and functional readings comes from the selection of shifting rules locally applied to the verb predicate. In (ib), applying the **z**-rule to the verb predicate *invite* closes off the anaphoric dependency between the object-trace and the subject-quantifier, yielding a functional reading.

(i) Structure: $[_{\text{CP}}$[which girl] $[_{\text{IP}}$[every boy] [invited t]]]

a. Individual reading: $[\![\textit{which girl}]\!]([\![\textit{every boy}]\!]^{\mathbf{g}}([\![\textit{invite}]\!]^{\mathbf{g}}([\![t]\!])))$

b. Functional reading: $[\![\textit{which girl}]\!]([\![\textit{every boy}]\!]^{\mathbf{g}}(([\![\textit{invite}]\!]^{\mathbf{z}})^{\mathbf{g}}([\![t]\!])))$

assumed formal theory, this paper follows the former approach and assumes the structure in (81). Here the *wh*-trace t_i^j carries two indices — the functional index i is bound by *which girl* and is interpreted as a functional variable of type $\langle e, e\rangle$, and the basic argument index j is bound by *every boy* and is interpreted as an individual variable of type e.

(81) $[_{\text{CP}}$ [which girl]$_i$ $[_{\text{IP}}$ [every boy]$_j$ [t_j invited t_i^j]]

5.2. Pair-list readings

Let's return to multi-*wh* questions with pair-list readings. I argue that pair-list readings of multi-*wh* questions are special functional readings — the question denotation is a property for functions from a subset of the subject-*wh* domain to the object-*wh* domain. The following tree illustrates the derivation of a root denotation. In this derivation, the most important assumption is that the movement of the object-*wh* leaves a functional trace, whose argument is bound by the subject-*wh*.

I will not go through the computation of (i) step by step. Type-shifting rules and lexical interpretations needed for the computation are gathered in the following:

(ii) Type-shifting rules

a. The **g**-rule
For any function F of type $\langle a, b\rangle$, $F^{\mathbf{g}}$ is a function of type $\langle ca, cb\rangle$ where $F^{\mathbf{g}} = \lambda V_{\langle c,a\rangle}\lambda C_c.F(V(C))$.

b. The **z**-rule
For any function F of type $\langle a, eb\rangle$, $F^{\mathbf{z}}$ is a function of type $\langle ea, eb\rangle$ where $F^{\mathbf{z}} = \lambda f_{\langle e,a\rangle}\lambda x_e.F(f(x))$.

(iii) Lexical interpretations for the individual reading

a. $[\![t]\!] = \lambda y_e.y$

b. $[\![invite]\!]^{\mathbf{g}} = \lambda f_{\langle e,e\rangle}\lambda y_e\lambda x_e.\textsf{invt}(x, f(y))$

c. $[\![every\ boy]\!]^{\mathbf{g}} = \lambda R_{\langle e,et\rangle}\lambda z_e.\forall x[\textsf{boy}(x) \to R(z)(x)]$

d. $[\![which\ girl]\!] = \lambda P_{\langle e,t\rangle}\lambda x_e : \textsf{girl}(x).P(x)$

(iv) Lexical interpretations for the functional reading

a. $[\![t]\!] = \lambda f_{\langle e,e\rangle}.f$

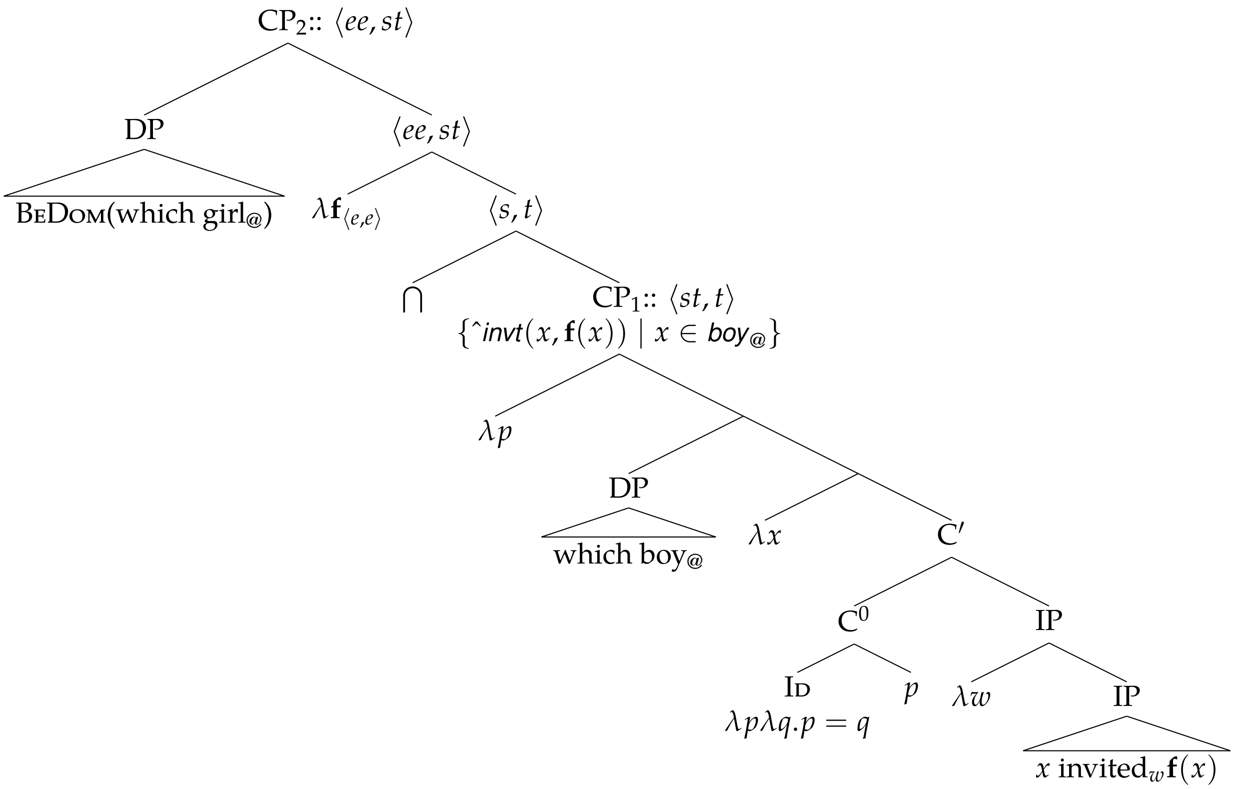
b. $[\![invite]\!]^{\mathbf{z}} = \lambda f_{\langle e,e\rangle}\lambda x_e.\textsf{invt}(x, f(x))$

c. $([\![invite]\!]^{\mathbf{z}})^{\mathbf{g}} = \lambda g_{\langle e,e\rangle}\lambda x_e.\textsf{invt}(x, g(x))$

d. $[\![every\ boy]\!]^{\mathbf{g}} = \lambda\theta_{\langle ee,et\rangle}\lambda f_{\langle e,e\rangle}.\forall x[\textsf{boy}(x) \to \theta(f)(x)]$

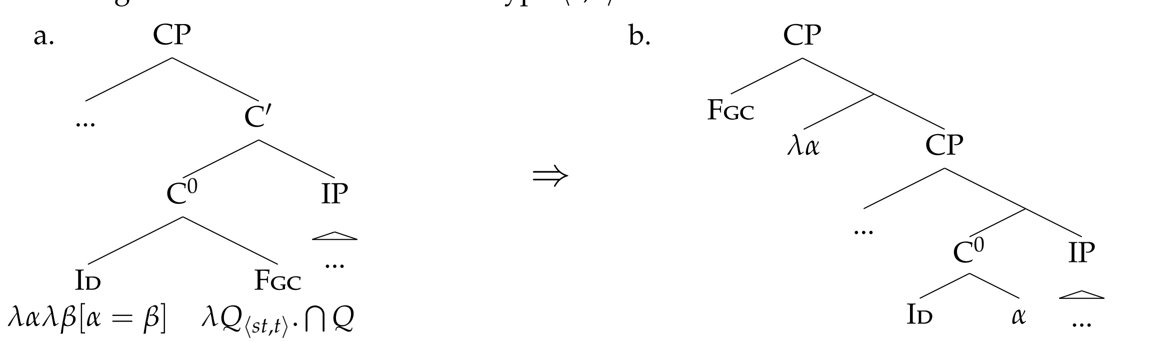
e. $[\![which\ girl]\!] = \lambda P_{\langle ee,t\rangle}\lambda f_{\langle e,e\rangle} : \text{Range}(f) \subseteq \textsf{girl}.P(f(x))$

(82) Which boy invited which girl? (Pair-list reading)



This LF involves two layers of interrogative CPs. The embedded CP_1 denotes a set of propositions (roughly, the Hamblin set of the single-*wh* question "which boy x invited $\mathbf{f}(x)$?"), compositionally derived based on the regular GB-transformed LF for Karttunen Semantics (see (45) in §3.2.2). This set of propositions is immediately closed by a $\bigcap$-closure, returning a conjunctive proposition. This $\bigcap$-closure can be considered as a *function graph creator* (Fgc) in the sense of Dayal (2017). Accordingly, the abstraction of the first argument of Id is due to a type-driven movement of the Fgc-operator:

(83) **Movement of the Fgc-operator**
The Id function requires its two arguments to be of the same semantic type. Since IP denotes a proposition, interpreting Fgc *in situ* yields a type-mismatch. Hence, Fgc moves to the left edge of CP and leaves a trace of type $\langle s, t\rangle$.



Next, moving the object-*wh* 'BeDom(which girl)' to [Spec, CP_2] leaves a functional trace within IP and forms a property of functions, just like what we saw with the basic functional reading in (80a). This topical property, as schematized in (84b), is defined for functions ranging over atomic girls, and it maps each such function to a conjunctive proposition that spells out the graph of this function. Note that here $\boldsymbol{P}$ restricts the range of $\mathbf{f}$, but not the domain of $\mathbf{f}$; in other words, any function mapping to $\textit{girl}_@$ counts as a possible "short" answer to this question. Applying $\boldsymbol{P}$ to each of those functions point-wise returns the set of possible propositional answers, as listed in (84b). Crucially, as a function, $\mathbf{f}$ maps every item in its domain to one and only one girl; therefore, there is no possible answer of the form $^\wedge\textit{invt}(a,m) \cap {}^\wedge\textit{invt}(a,j)$ where one boy is paired with multiple girls. Finally, as in (84c), applying the Ans-operator and the Ans^S-operator to the topical property return the singleton set containing the pair-list answer and the singleton set containing the corresponding boy-to-girl function, respectively.

(84) Which boy invited which girl? (Pair-list reading)

a. Root denotation of the multi-*wh* question

$$\begin{aligned}\boldsymbol{P} &= \iota P[\mathrm{Dom}(P) = \{\mathbf{f} \mid \mathrm{Range}(\mathbf{f}) \subseteq \textit{girl}_@\} \wedge \forall\alpha \in \mathrm{Dom}(P)[P(\alpha) = \bigcap[\![\mathrm{CP}_1]\!]]] \\ &= \iota P[\mathrm{Dom}(P) = \{\mathbf{f} \mid \mathrm{Range}(\mathbf{f}) \subseteq \textit{girl}_@\} \wedge \\ &\qquad \forall\alpha \in \mathrm{Dom}(P)[P(\alpha) = \bigcap\{^\wedge\textit{invt}(x,\alpha(x)) \mid x \in \textit{boy}_@\}]] \\ &= \lambda\mathbf{f}_{\langle e,e\rangle} : \mathrm{Range}(\mathbf{f}) \subseteq \textit{girl}_@ . \bigcap\{^\wedge\textit{invt}(x,\mathbf{f}(x)) \mid x \in \textit{boy}_@\}\end{aligned}$$

b. Possible propositional answers

(*Consider only two boys, Andy and Billy, and two girls, Mary and Jenny.*)

$$\begin{aligned}\{\boldsymbol{P}(\mathbf{f}) \mid \mathbf{f} \in \mathrm{Dom}(\boldsymbol{P})\} &= \left\{ \bigcap\{^\wedge\textit{invt}(x,\mathbf{f}(x)) \mid x \in \textit{boy}_@\} \mid \mathrm{Range}(\mathbf{f}) \subseteq \textit{girl}_@ \right\} \\ &= \left\{ \begin{array}{ll} ^\wedge\textit{invt}(a,m) & ^\wedge\textit{invt}(a,m) \cap {}^\wedge\textit{invt}(b,m) \\ ^\wedge\textit{invt}(a,j) & ^\wedge\textit{invt}(a,m) \cap {}^\wedge\textit{invt}(b,j) \\ ^\wedge\textit{invt}(b,m) & ^\wedge\textit{invt}(a,j) \cap {}^\wedge\textit{invt}(b,m) \\ ^\wedge\textit{invt}(b,j) & ^\wedge\textit{invt}(a,j) \cap {}^\wedge\textit{invt}(b,j) \end{array} \right\}\end{aligned}$$

c. Complete true answers

(w: *Andy invited Mary, Billy invited Jenny; no other boy invited any girl.*)

i. $\text{Ans}(w)(\boldsymbol{P}) = \{^\wedge\textit{invt}(a,m) \cap {}^\wedge\textit{invt}(b,j)\}$

ii. $\text{Ans}^S(w)(\boldsymbol{P}) = \left[\begin{array}{c} a \to m \\ b \to j \end{array} \right]$

6. Applications

Defining a question as a topical property, the proposed hybrid categorial approach can easily retrieve the short answers to an embedded question. This section responds to the two new arguments for pursuing categorial approaches introduced in section 2.2. I will show a simple derivation of *wh*-FRs (§6.1.1) and extend the discussion to Mandarin *wh*-conditionals (§6.1.2). Next, I will account for QV effects using the presented hybrid categorial approach (§6.2).

6.1. *Wh*-constructions with predicative or nominal meanings

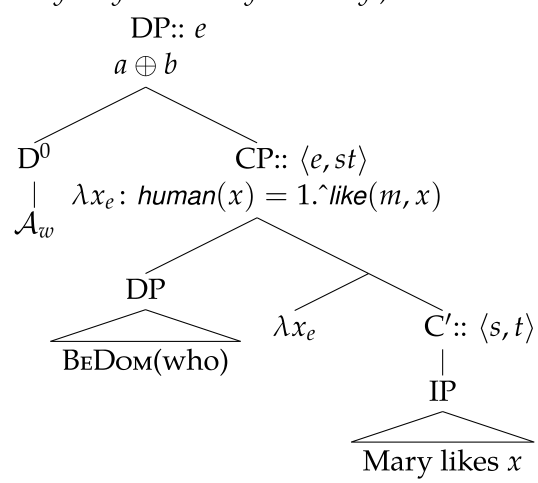
6.1.1. *Wh*- free relatives

The meaning of a *wh*-FR is systematically equivalent to the nominal meaning of the/a complete true short answer to the corresponding *wh*-question. In particular, a *wh*-FR is interpreted existentially if the corresponding *wh*-question admits a "mention-some" reading, and is interpreted exhaustively otherwise. The following examples are repeated from (59):

(85) a. John ate what Mary cooked for him.
$\rightsquigarrow$ *John ate* ***everything*** *that Mary cooked for him.*
b. John went to where he could get help.
$\rightsquigarrow$ *John went to* ***some*** *place where he could get help.*

Caponigro's Generalization on the cross-linguistic distribution gap of *wh*-words teaches us that *wh*-FRs are derivatives of *wh*-questions (section 2.2.1). In the proposed hybrid categorial approach, defining questions as topical properties makes it simple to derive the nominal meaning of a *wh*-FR from a *wh*-question. I assume that a *wh*-FR is a DP with the following LF:

(86) (w: *Among the relevant individuals, Mary only likes Andy and Billy.*)
John invited [$_{\text{FR}}$ who Mary likes].



In this LF, an $\mathcal{A}$-determiner selects for an interrogative CP-complement and returns an entity-denoting DP. The CP-complement denotes a topical property, derived in the same way as it is in the corresponding matrix question. The $\mathcal{A}$-determiner is a mixture of a choice function and an Ans^S-operator. It picks out a complete true short answer to the question denoted by the CP-complement, which is therefore the denotation of the *wh*-FR.

(87) $[\![\mathcal{A}]\!] = \lambda w \lambda \boldsymbol{P}. f_{\text{CH}}[\text{Ans}^S(w)(\boldsymbol{P})]$

With the assumed derivation of *wh*-FRs, Caponigro's Generalization is predicted as long as the use of the $\mathcal{A}$-determiner is partial and can be blocked for whatever reason. Compared with earlier accounts which define questions as λ-abstracts but *wh*-FRs as definite descriptions (e.g., Jacobson 1995; Caponigro 2003, 2004),[28] the proposed analysis easily captures the existential readings: if the CP-complement takes a mention-some reading, the application of the $\mathcal{A}$-operator will return one of the short mention-some answer, yielding an existential reading.

[28] Jacobson (1995) and Caponigro (2003, 2004) follow traditional categorial approaches and interpret *wh*-words as functions from predicates to predicates (e.g., $[\![\mathit{who}]\!] = \lambda P \lambda x[\mathit{human}(x) \wedge P(x)]$), analyzing the CP-complement extensionally as a predicate (e.g., $[\![\mathit{who\ Mary\ likes}]\!]^w = \lambda x[\mathit{human}(x) \wedge \mathit{like}_w(m, x)]$). Moreover, the D head is interpreted as a σ-closure operator (Link 1983). It selects the unique maximal element in the extension of the CP-complement.

(i) σ**-closure** (Link 1983)
$\sigma = \lambda A : \exists x[x \in A \wedge \forall y[y \in A \rightarrow y \leq x]].\iota x[x \in A \wedge \forall y[y \in A \rightarrow y \leq x]]$
(For any set A, return the maximal element in A, defined only if this maximal element exists.)

6.1.2. Mandarin *wh*-conditionals

In Mandarin, a *wh*-conditional is a conditional made up of two *wh*-clauses. In syntax, the two *wh*-clauses use the very same *wh*-words or *wh*-phrases, as shown in (88).

(88) a. Shei xian dao, shei xian chi.
who first arrive, who first eat.
'Whoever arrives first, he eats first.'

b. Ni qu nar, wo qu nar.
you go where, I go where.
'Wherever you go, I will go there.'

a′. * Shei xian dao, shei xian chi sha.
who first arrive, who first eat what

b′. * Shei qu nar, wo qu nar.
who go where, I go where

The *wh*-items in these two clauses can serve distinct syntactic roles. For example, in (89a), the *wh*-word *shei* 'who' serves as the object in the antecedent but the subject in the consequent.

(89) a. Ni xuan shei, shei daomei.
you pick who, who unlucky
'Whomever you pick is unlucky.'

b. Shei xuan wo, wo yaoqing shei.
who pick I, I invite who
'I will invite whomever picks me.'

It is also possible to have multiple *wh*-items in each clause, for example:

(90) a. Shei mai sha, shei chi sha.
who buy what, who eat what.
Intended: 'Whoever buys whatever eats whatever.'

b. Shei rang shei mai sha, shei gei shei mai sha -de qian.
who let who buy what, who give who buy what -DE money.
Intended: 'Whoever$_i$ asks whoever$_j$ to buy whatever gives whoever$_j$ the money for buying whatever.'

In semantics, as seen in the aforementioned examples, a *wh*-conditional usually expresses a universal or exhaustive condition: every true short answer of the antecedent *wh*-clause is also a true short answer of the consequent *wh*-clause. Note that the exhaustivity requirement does not apply in the other direction. As shown in (91), it is possible that some true short answer of the consequent *wh*-clause is not a true short answer of the antecedent clause, or equivalently, that the complete true short answer of the antecedent *wh*-clause is only a partial true short answer of the consequent clause.[29]

[29]Surprisingly, *wh*-conditionals made up of degree questions seem to be bi-directionally exhaustive. The following sentence conveys a command that one should take exactly the amount of the food that he will eat.

(i) Chi duoshao, na duoshao.
eat how.much, take how.much
'How much [food] you will eat, how much [food] you take.' (Modified from Liu 2016)

One way to account for this seeming bi-directional exhaustivity is to assume that answers of degree questions are all exclusive. This assumption is supported by the infelicity of using a partiality-marker like *for example* in a degree questions, as exemplified in (ii). As seen in section 2.2.2, *for example* presupposes the existence of a partial true answer. The infelicity in (ii) is predicted if the answers are all exclusive — for example, answers to (iib) are of the form "John ran exactly *n*-fast".

(91) Ni xiang jian shui, wo jiu yaoqing shui. Dan wo ye hui yaoqing qita-ren.
You want meet who, I JIU invite who. But I also will invite other-person
'Whomever you want to see, I will invite him. But I will also invite some other people.'

Interestingly, analogous to an existential *wh*-FR, a *wh*-conditional takes an existential reading if the antecedent *wh*-clause resembles a mention-some question (Liu 2016), as seen in (92).

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

Using the proposed hybrid categorial approach, I propose to treat the two *wh*-clauses as questions and define the semantics of a *wh*-conditional as a condition on the short answers of the two questions. The syntactic requirement that the two *wh*-clauses must use the same *wh*-items is ascribed to a presupposition that the topical properties denoted by the two *wh*-clauses have the same domain. The truth conditions of *wh*-conditionals are thus schematized as follows uniformly, where $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are topical properties denoted by Q_1 and Q_2, respectively:

(93) **Mandarin conditionals**
$[\![Q_1, Q_2]\!] = \lambda w\colon \text{Dom}(\boldsymbol{P}_1) = \text{Dom}(\boldsymbol{P}_2).\forall w'[\text{Acc}(w', w) \rightarrow w' \in \boldsymbol{P}_2(f_{\text{CH}}[\text{A}\textsc{ns}^S(w')(\boldsymbol{P}_1)])]$
(Some complete true short answer to Q_1 is a true short answer to Q_2 in every accessible world; defined only if the topical properties denoted by Q_1 and Q_2 have the same domain.)

This definition predicts the following: a *wh*-conditional takes an existential reading iff its antecedent *wh*-clause Q_1 takes a mention-some reading. For instance, the universal *wh*-conditional (88a) has a mention-all antecedent: $\text{A}\textsc{ns}^S(w)(\boldsymbol{P}_1)$ is a singleton consisting of only the sum of the individuals arriving first in w, and thus (88a) means that whoever arrives first will eat first. In contrast, the existential *wh*-conditional (92) has a mention-some antecedent: $\text{A}\textsc{ns}^S(w)(\boldsymbol{P}_1)$ denotes a set of places where I can buy liquor, at least one of which is a true short answer to the consequent.

6.2. Getting quantificational variability effects

Recall the two accounts of QV inferences that are compatible with Hamblin-Karttunen Semantics ($[\![Q]\!]$ stands for the Hamblin set of Q):

(94) The QV inference of *Jenny mostly knows Q*.
a. By the propositional answer-based account
$\lambda w.\ \text{M}\textsc{ost}\ p\ [w \in p \in \text{A}\textsc{t}([\![Q]\!])][\textit{know}_w(j, q)]$
(For most p such that p is a true atomic proposition in $[\![Q]\!]$, Jenny knows p.)
b. By the sub-question-based account
$\lambda w.\ \text{M}\textsc{ost}\ Q'\ [\exists p[w \in p \in \text{A}\textsc{t}([\![Q]\!]) \wedge Q' = \textit{whether-p}]][\textit{know}_w(j, Q')]$
(For most Q' of the form *whether-p* such that p is a true atomic proposition in $[\![Q]\!]$, Jenny knows Q'.)

(ii) a. How much food will you eat, # for example?
b. How fast did John run, # for example?

However, to assess this assumption, we should also consider its predictions regarding phenomena such as negative islands and modal obviation (Fox and Hackl 2007; Abrusán 2007; Abrusán and Spector 2011). I leave this issue open.

A more precise schematization that extends to mention-some questions is given in (95a-b). In both formulas, the choice function variable is bound globally. $\textsc{Ans}(w)(\llbracket Q \rrbracket)$ denotes the set of complete true answers to Q in w, defined following Fox (2013). When Q takes a mention-some reading, $\textsc{Ans}(w)(\llbracket Q \rrbracket)$ consists of multiple true mention-some answers to Q, and otherwise it is a singleton set consisting of the unique strongest true answer to Q.

(95) The QV inference of *Jenny mostly knows Q*.

a. By the propositional answer-based account
$\lambda w. \textsc{Most}\ p\ [p \in \textsc{At}(\llbracket Q \rrbracket) \wedge f_{\textsc{ch}}[\textsc{Ans}(w)(\llbracket Q \rrbracket)] \subseteq p][\textsf{\textit{know}}_w(j,q)]]$
(For most propositions p such that p is an atomic propositional answer to Q entailed by some particular complete true answer to Q, Jenny knows p.)

b. By the sub-question-based account
$\lambda w. \textsc{Most}\ Q'\ [\exists p[p \in \textsc{At}(\llbracket Q \rrbracket) \wedge f_{\textsc{ch}}[\textsc{Ans}(w)(\llbracket Q \rrbracket)] \subseteq p \wedge Q' = \textsf{\textit{whether-}}p]][\textsf{\textit{know}}_w(j,Q')]$
(For most Q' of the form *whether-p* such that p is an atomic propositional answer to Q entailed by some particular complete true answer to Q, Jenny knows Q'.)

The revised definitions in (95) can nicely explain the infelicity of using a quantity adverbial in (96), where the embedded question takes a mention-some reading:

(96) Speaker A: "I'm so sleepy. I need to get some coffee."
Speaker B: "Jenny (#mostly/#partly) knows where you can get coffee."

The matrix quantity adverbial *mostly/partly* requires a non-singleton quantification domain, while a mention-some answer of the embedded question in (96) names only one atomic place and supplies only a singleton quantification domain.

As seen in section 2.2.2, the propositional answer-based account and the sub-question-based account both require forming a proper set of atomic propositional answers, which however is challenging in the following two cases. First, in (97) and (98) (repeated from (26a) and (26c)) where the predicates of the embedded questions are non-divisive, the quantification domains of *for the most part* and *mostly* cannot be formed by atomic propositional answers (in the sense of Lahiri (2002) and Cremers (2016)) or by sub-questions (in the sense of Beck and Sharvit (2002)).

(97) Jenny knows for the most part which students formed the bassoon quintet.
$\rightsquigarrow$ *For most x such that x is part of the group of students who formed the bassoon quintet, Jenny knows that x is part of the group of the students who formed the bassoon quintet.*

(98) Jenny mostly knows which professors formed the committee.
$\rightsquigarrow$ *For most x such that x is part of the group of professors who formed the committee, Jenny knows that x is part of the group of professors who formed the committee.*

Second, in (99) (repeated from (35)) where the embedded multi-*wh* question takes a pair-list reading, atomic propositions of the form "boy x invited girl y" cannot be recovered out of the conjunctions of these propositions. This technical difficulty challenges the analysis of Dayal (1996, 2017), which follows Hamblin-Karttunen Semantics and analyzes the embedded multi-*wh* question as denoting a set of conjunctive propositions.

(99) Jenny mostly knows which boy invited which girl.

a. $\rightsquigarrow$ *For most p such that p is a true proposition of the form 'boy x invited girl y', Jenny knows p.*

b. $\rightsquigarrow$ *For most boy-girl pairs* $\langle x, y\rangle$ *such that* x *invited* y*, Jenny knows that* x *invited* y*.*

Defining a question as a topical property whose domain can supply short answers, the proposed hybrid categorial approach can easily retrieve the short answers of the embedded question and define the quantification domain of *mostly* based on short answers:

(100) Deriving the quantification domain of *mostly* in the QV inference of *Jenny mostly know Q*: Given a complete short answer α of the embedded question (viz., $\alpha = f_{\text{CH}}[\text{Ans}^S(w)(\boldsymbol{P})]$ where $\boldsymbol{P}$ is the topical property denoted by embedded question Q and w is the evaluation world), the domain of *mostly* is simply $\text{At}(\alpha)$.

With this quantification domain, there are two ways to define the QV inference, as schematized in the following. Again, choice function variables are all globally bound.[30]

(101) Let $\boldsymbol{P} = [\![\text{Q}]\!]$. The **QV inference** of *Jenny mostly knows Q* is:

a. Definition I
$\lambda w.\text{Most } x[x \in \text{At}(f_{\text{CH}}[\text{Ans}^S(w)(\boldsymbol{P})])][\textit{know}_w(j, \boldsymbol{P}(x))]$
(For most x s.t. x is an atomic subpart of some particular complete true answer to Q, Jenny knows $\boldsymbol{P}(x)$.)

b. Definition II
$\lambda w.\text{Most } x[x \in \text{At}(f_{\text{CH}}[\text{Ans}^S(w)(\boldsymbol{P})])][\textit{know}_w(j, \lambda w'.x \leq f'_{\text{CH}}[\text{Ans}^S(w')(\boldsymbol{P})])]$
(For most x s.t. x is an atomic part of some particular complete true short answer to Q, Jenny knows that x is a part of some particular complete true short answer to Q.)

In Definition I, the scope of Most is simply 'Jenny knows the propositional answer based on the atomic short answer x'. This way of deriving QV inferences only works in cases where the topical property $\boldsymbol{P}$ of the embedded question is divisive, as exemplified in (102).

(102) Jenny mostly knows [$_{\text{Q}}$ which students came].
(w: *Among the relevant students, only abc came.*)

a.	$\boldsymbol{P} = \lambda x_e: \textit{*stdt}_{@}(x) = 1.\char`\^\textit{came}(x)$	Topical property of Q
b.	$\text{Ans}^S(w)(\boldsymbol{P}) = \{a \oplus b \oplus c\}$	Complete true short answer to Q
c.	$\text{At}(a \oplus b \oplus c) = \{a, b, c\}$	Quantification domain of *mostly*
d.	$\lambda w.\text{Most } x[x \in \{a, b, c\}][\textit{know}_w(j, \char`\^\textit{came}(x))]$	The QV inference

In Definition II, what Jenny knows is a sub-divisive inference, namely, the proposition that x is a part of certain complete true short answer to Q, as exemplified in (103).[31] This definition extends to cases where the property of the embedded question is non-divisive.

(103) Jenny mostly knows [$_{\text{Q}}$ which professors formed the committee].
(w: *The committee was formed by three professors abc.*)

[30] In principle, Definition I can be eliminated because Definition II can be used not only in questions with a non-divisive predicate but also any question where Definition I applies. Cremers (2018) adopts this approach and provides a compositional derivation for the QV inference in Definition II.

[31] The sub-divisive QV inference derived in (103) is pretty much the same as the the sub-divisive QV inference derived by Williams (2000) (see (28) in section 2.2.2). However, while Williams (2000) ascribes the sub-divisiveness to the lexicon of the *wh*-determiner and inevitably over-generates sub-divisive readings in matrix questions, my analysis has room to attribute the sub-divisiveness to operators or shifting rules external to the root denotation of the embedded questions.

a. $\boldsymbol{P} = \lambda x_e\colon$ *$\textit{*prof}_{@}(x) = 1.\hat{}\textit{f.t.c}(x)$* Topical property of Q

b. $\textsc{Ans}^S(w)(\boldsymbol{P}) = \{a \oplus b \oplus c\}$ Complete true short answer to Q

c. $\textsc{At}(a \oplus b \oplus c) = \{a, b, c\}$ Quantification domain of *mostly*

d. $\lambda w.\textsc{Most}\ x[x \in \{a,b,c\}][\textit{know}_w(j, \lambda w'.x \leq f'_{\textsc{ch}}[\textsc{Ans}^S(w')(\boldsymbol{P})])]$ The QV inference

A similar sub-divisive reading is observed with the indirect mention-some question (104). Contrary to the case of (96), *mostly* can be felicitously used in (104). The reason is that each mention-some answer of the embedded question in (104) names a group of individuals, which therefore can supply a non-singleton quantification domain for *mostly*.

(104) John mostly knows [who can serve on the committee]$_{\text{mention-some}}$.

The intuition behind the proposal is simple. First, knowing a question basically means correctly identifying the meaning (an individual or a proposition) that constitutes a complete true answer to this question. And naturally, as Definition II says, mostly knowing a question means correctly identifying the majority of the individuals that constitute a complete answer to this question. Second, Definitions I and II are analogous to the following (a) and (b) sentences which involve quantification over individuals. Crucially, as seen in (105a), the quantifier *most professors* cannot directly combine with the non-divisive collective predicate *formed the committee* (in contrast to the case of the distributive predicate *served on the committee* which describes a similar event); instead, the scope of *most* has be divisive. This restriction is parallel to the requirement that the QV inference of the question-embedding sentence (103) has to be paraphrased with a sub-divisive inference.

(105) a. * Most professors formed the committee.

b. Most professors are among the professors who formed the committee.

(106) a. Most professors served on the committee.

b. Most professors are among the professors who served on the committee.

Now turn to the case of pair-list readings. While Dayal follows Hamblin-Karttunen Semantics and treats the root denotation of a multi-*wh* question as a set of propositions, the proposed hybrid categorial approach treats the root denotation as a property of Skolem functions, which can supply short answers. Hence, in the proposed account, the quantification domain of *mostly* can be formed based on the short answers: given a function $\mathbf{f}$ such that $\mathbf{f}$ is a complete true short answer of the embedded question, the quantificational domain of the matrix quantificational adverb *mostly* is the set of functions that are atomic subsets of $\mathbf{f}$. I define atomic functions as follows: a function is atomic iff its domain is a singleton set containing only an atomic item, or equivalently, the supremum of its domain is an atomic element. Notice that $\mathbf{f}_2$ is atomic, even though a is paired with a non-atomic element.

(107) **Atomic functions**

a. A function $\mathbf{f}$ is atomic iff $\bigoplus \mathrm{Dom}(\mathrm{f}')$ is atomic.

i. Examples of atomic functions:
$\mathbf{f}_1 = [\ a \rightarrow m\ ]$
$\mathbf{f}_2 = [\ a \rightarrow m \oplus j\ ]$

ii. Examples of non-atomic functions:
$\mathbf{f}_3 = [\ a \rightarrow m, \quad b \rightarrow j\ ]$
$\mathbf{f}_4 = [\ a \oplus b \rightarrow j\ ]$

b. $\textsc{At}(\mathbf{f}) = \{\mathrm{f}' \mid \mathrm{f}' \subseteq \mathbf{f} \text{ and } \bigoplus \mathrm{Dom}(\mathrm{f}') \text{ is atomic}\}$

The QV inference of (99) is derived as follows following Definition I. The complete true short answer of the embedded question is a function, and its atomic subparts are atomic functions.

(108) Jenny mostly knows [$_{\text{Q}}$ which boy invited which girl].
(*w*: *Andy, Billy, and Clark invited Jenny, Mary, and Sue, respectively*.)

a. Topical property of Q
$\boldsymbol{P} = \lambda \mathbf{f}\colon \text{Range}(\mathbf{f}) \subseteq \textsf{\textit{girl}}_@. \bigcap\{{}^\wedge \textsf{\textit{invt}}(x, \mathbf{f}(x)) \mid x \in \textsf{\textit{boy}}_@\}$

b. Complete true short answers to Q
$$\textsc{Ans}^S(w)(\boldsymbol{P}) = \left\{ \begin{bmatrix} a \to m \\ b \to j \\ c \to s \end{bmatrix} \right\}$$

c. Quantification domain of *mostly*
$$\textsc{At}\left(\begin{bmatrix} a \to m \\ b \to j \\ c \to s \end{bmatrix} \right) = \left\{ \begin{array}{l} [a \to m] \\ [b \to j] \\ [c \to s] \end{array} \right\}$$

d. The QV inference
$$\lambda w.\textsc{Most}\ \mathrm{f}' \left[\mathrm{f}' \in \left\{ \begin{array}{l} [a \to m] \\ [b \to j] \\ [c \to s] \end{array} \right\} \right] \left[\textsf{\textit{know}}_w(j, \boldsymbol{P}(\mathrm{f})) \right]$$
$$= \lambda w.\textsc{Most}\ \mathrm{f}' \left[\mathrm{f}' \in \left\{ \begin{array}{l} [a \to m] \\ [b \to j] \\ [c \to s] \end{array} \right\} \right] \left[\textsf{\textit{know}}_w(j, \bigcap\{{}^\wedge \textsf{\textit{invt}}(x, \mathrm{f}'(x)) \mid x \in \textsf{\textit{boy}}_@\}) \right]$$
$$= \lambda w.\textsc{Most}\ \mathrm{f}' \left[\mathrm{f}' \in \left\{ \begin{array}{l} [a \to m] \\ [b \to j] \\ [c \to s] \end{array} \right\} \right] \left[\textsf{\textit{know}}_w(j, {}^\wedge \textsf{\textit{invt}}(x, \mathrm{f}'(x))) \right]$$
(Jenny knows most of the following boy-invite-girl pairs: a invited m, b invited j, and c invited s.)

Defining the QV inference following Definition II yields the same consequence. As in (109), the scope of Most can be defined as involving a sub-divisive inference:[32]

(109) $$\lambda w.\textsc{Most}\ \mathrm{f}' \left[\mathrm{f}' \in \left\{ \begin{array}{l} [a \to m] \\ [b \to j] \\ [c \to s] \end{array} \right\} \right] \left[\textsf{\textit{know}}_w(j, \lambda w'.\mathrm{f}' \leq f_{\text{CH}}[\textsc{Ans}^s(w')(\boldsymbol{P})]) \right]$$
(For most functions f' in $\{[a \to m], [b \to j], [c \to m]\}$, Jenny knows that f' is a subpart of some particular complete true short answer to Q.)

With three relevant girls mjs, this sub-divisive inference is true iff in every world w' such that w' is compatible with Jenny's belief, the complete true short answer to the embedded multi-*wh* question Q at w' is one of the seven functions listed in Table 6.

[32]This way of formalizing QV inferences for questions with pair-list readings is more advantageous if the embedded question is a $\forall$-question. Intuitively, the QV inference in (i) is similar to that of the multi-*wh* question in (108).

(i) Jenny mostly knows [$_{\text{Q}}$ which girl every boy invited]$_{\text{pair-list}}$.

a. $\rightsquigarrow$ *For most p such that p is a true proposition of the form 'boy x invited girl y', Jenny knows p.*

b. $\rightsquigarrow$ *For most boy-girl pairs $\langle x, y\rangle$ such that x invited y, Jenny knows that x invited y.*

However, different from pair-list readings in questions with multiple *wh*-phrases, pair-list readings in questions with a universal quantifier are subject to domain exhaustivity (see footnote 23). Therefore, the topical property of the embedded $\forall$-question in (i) is only defined for functions that are defined for every atomic boy, and is undefined for the atomic functions in (108c). As such, the QV inference in (i) cannot be formulated following Definition I.

	$\begin{bmatrix} a \rightarrow j \\ b \rightarrow j \\ c \rightarrow s \end{bmatrix}$	$\begin{bmatrix} a \rightarrow s \\ b \rightarrow j \\ c \rightarrow s \end{bmatrix}$
$\begin{bmatrix} a \rightarrow m \\ b \rightarrow m \\ c \rightarrow s \end{bmatrix}$	$\begin{bmatrix} a \rightarrow m \\ b \rightarrow j \\ c \rightarrow s \end{bmatrix}$	$\begin{bmatrix} a \rightarrow m \\ b \rightarrow s \\ c \rightarrow s \end{bmatrix}$
$\begin{bmatrix} a \rightarrow m \\ b \rightarrow j \\ c \rightarrow m \end{bmatrix}$	$\begin{bmatrix} a \rightarrow m \\ b \rightarrow j \\ c \rightarrow j \end{bmatrix}$	

Table 6: Illustration of (109)

Table 6 illustrates a partition of possible worlds of the embedded multi-*wh* question *which boy invited which girl*. Each cell in this table stands for the set of worlds where the named boy-to-girl function is the complete true short answer to this question. The union of these cells represents the sub-divisive inference $\lambda w'.\mathrm{f}' \leq f_{\text{CH}}[\textsc{Ans}^s(w')(\boldsymbol{P})]$, or equivalently, the set of worlds w' such that most of the functions in $\{[a \rightarrow m], [b \rightarrow j], [c \rightarrow m]\}$ (viz., the complete true short answer to Q in the actual world) are parts of the complete true short answer to Q at w'. From this table, it can be easily observed that the sub-divisive QV inference (109) is equivalent to the QV inference (108d).

7. Coordinations of questions

Recall one of the major criticisms to categorial approaches: questions of different kinds are assigned different semantic types, which makes it difficult to account for question coordinations, especially coordinations in embeddings. Given this problem, existing works follow categorial approaches toly for defining matrix questions. For example, Groenendijk and Stokhof (1990) and Jacobson (2016) define matrix questions as λ-abstracts while defining embedded questions as partitions of possible worlds or Hamblin sets. As such, coordinations of embedded questions can be treated standardly as (point-wise) meet/join, while the seeming coordinations of matrix questions can be analyzed as coordinations of speech acts (Krifka 2001b). To account for QV effects in question-embeddings, however, the proposed hybrid categorial approach defines embedded questions also as λ-abstracts/ topical properties. For example, the two coordinated questions in the embedding sentence (110) are treated as of type $\langle e, st\rangle$ and $\langle e, \langle e, st\rangle\rangle$, respectively.

(110) Jenny knows [who came] and [who bought what]

One appealing way of thinking would be to type-shift the coordinated questions into expressions of the same conjoinable type, such as shifting them into sets of propositions or partitions of possible worlds. However, as seen in section 2.2.2 and 6.2, to get QV inferences in question-embeddings, a matrix quantification adverbial needs to have access to the complete true short answers of the embedded question. Therefore, the interrogative complement of the question-embedding predicate has to denote something out of which we can retrieve the short answers. Options include the root question denotation $\boldsymbol{P}$, the set of complete true short answers $\textsc{Ans}^S(w)(\boldsymbol{P})$, and the intension of this answer set $\lambda w.\textsc{Ans}^S(w)(\boldsymbol{P})$. Regardless of the choice we take, the coordinated questions cannot be of the same conjoinable type.

I argue that question coordinations are generated not with meet or join, but rather with generalized quantifiers, in line with an idea briefly mentioned by Krifka (2011) in a review of categorial approaches. To be more specific, when used to coordinate two questions, *and*/*or* does not directly coordinate the denotations of the two questions, but rather coordinates two predication operations (of type t).

Recall that conjunction and disjunction are traditionally treated as meet and join, respectively, which are only defined for conjoinable expressions, namely, expressions of a semantic type of the form $\langle ...t \rangle$ (Partee and Rooth 1983, Groenendijk and Stokhof 1989). If A' and B' are of a non-conjoinable type or of different types, meet and join cannot proceed. For such cases, I propose that *A and/or B* can be interpreted as a generalized quantifier, defined as in (111b)/(112b).[33] (In the formalizations, '$\wedge$' and '$\vee$' are reserved for coordinating truth values, while '$\sqcap$' and '$\sqcup$' are reserved for meet and join. See definitions of meet and join in (50).)

(111) A conjunction "A and B" is ambiguous between (a) and (b):

a. **Meet**
$[\![A \text{ and } B]\!] = A' \sqcap B'$; defined only if A' and B' are of the same conjoinable type.

b. **Generalized Boolean conjunction**
$[\![A \text{ and } B]\!] = A' \barwedge B' = \lambda\alpha[\alpha(A') \wedge \alpha(B')]$

(112) A disjunction "A or B" is ambiguous between (a) and (b):

a. **Join**
$[\![A \text{ or } B]\!] = A' \sqcup B'$; defined only if A' and B' are of the same conjoinable type.

b. **Generalized Boolean disjunction**
$[\![A \text{ or } B]\!] = A' \,\bar{\vee}\, B' = \lambda\alpha[\alpha(A') \vee \alpha(B')]$

The generalized Boolean conjunction $A' \barwedge B'$ universally quantifies over a possibly polymorphic set $\{A', B'\}$ and selects for an item with an ambiguous type as its scope. Equivalently, in set-theoretic notations, $A' \barwedge B'$ denotes the family of sets such that each set contains both A' and B' (formally: $\{\alpha \mid A' \in \alpha, B' \in \alpha\}$). The generalized Boolean disjunction $A' \,\bar{\vee}\, B'$ is analogous.

To see how this approach works in practice, consider the composition of (113). The question coordination denotes a generalized Boolean conjunction, labeled as QP. It undergoes QR and moves to the left edge of the matrix clause, yielding a wide scope reading of *and* relative to the embedding-predicate *know*. (Domain conditions of the topical properties are neglected.)[34]

[33] Note that the following expressions are different:

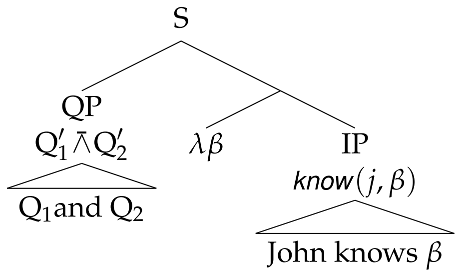
(i) a. $a \,\bar{\vee}\, b \,\bar{\vee}\, a \oplus b = \lambda P[P(a) \vee P(b) \vee P(a \oplus b)]$
b. $(a \,\bar{\vee}\, b) \,\bar{\vee}\, a \oplus b = \lambda\theta[\theta(a \,\bar{\vee}\, b) \vee \theta(a \oplus b)]$

[34] There are two non-trivial technical issues worth noting. First, the proposed definition of generalized Boolean conjunction/disjunction differs from the one given by Partee and Rooth (1983) in accounting for NP-coordination and extending to question coordination as in Groenendijk and Stokhof (1984) and Szabolcsi (1997, 2016). This analysis treats conjunction/disjunction as meet/join over Montague-lifted conjuncts/disjuncts. The case of conjunction is as schematized in (i): Montague lift shifts an expression of type τ to a generalized quantifier of type $\langle\langle\tau, t\rangle, t\rangle$, and then meet conjoins two generalized quantifiers.

(i) NP/question-coordinations as coordinations of two quantifiers
$[\![A \text{ and } B]\!] = \textsc{Lift}(A') \sqcap \textsc{Lift}(B') = (\lambda P.P(A')) \sqcap (\lambda P.P(B'))$

In this analysis, the meet operation requires $\textsc{Lift}(A')$ and $\textsc{Lift}(B')$ to be of the same conjoinable type and hence A' and B' to be of the same conjoinable type. Hence, it is not helpful for solving the type-mismatch problem in question coordinations.

(113) John knows [$_{Q_1}$ who came] and [$_{Q_2}$ who bought what].



$$\begin{aligned}
[\![\mathrm{S}]\!] &= (\mathrm{Q}'_1 \barwedge \mathrm{Q}'_2)(\lambda\beta.\mathit{know}(j,\beta)) \\
&= (\lambda\alpha[\alpha(\mathrm{Q}'_1) \wedge \alpha(\mathrm{Q}'_2)])(\lambda\beta.\mathit{know}(j,\beta)) \\
&= (\lambda\beta.\mathit{know}(j,\beta))(\mathrm{Q}'_1) \wedge (\lambda\beta.\mathit{know}(j,\beta))(\mathrm{Q}'_2) \\
&= \mathit{know}(j,\mathrm{Q}'_1) \wedge \mathit{know}(j,\mathrm{Q}'_2) \\
&= \mathit{know}(j,\lambda x.\hat{}\mathit{came}(x)) \wedge \mathit{know}(j,\lambda x\lambda y.\hat{}\mathit{bought}(x,y))
\end{aligned}$$

(J knows who came, and J knows who bought what.)

The proposed analysis of question coordinations yields the following prediction: in an indirect question where the question-embedding predicate embeds a coordination of questions, this coordination can only take scope above the question-embedding predicate. This prediction cannot be validated or falsified based on sentences like (113). The predicate *know* is divisive — knowing the conjunction of two questions is semantically equivalent to knowing the questions individually (formally: $\mathit{know}(j, Q_1 \textit{ and } Q_2) \Leftrightarrow \mathit{know}(j, Q_1) \wedge \mathit{know}(j, Q_2)$), and therefore the wide scope reading and the narrow scope reading of the disjunction yield the same truth conditions. To evaluate this prediction, we can replace the conjunction with a disjunction, or *know* with a non-divisive predicate. The observations in what follows support the prediction.

First, in (114) and (115), the disjunctions clearly take scope above the question-embedding verb *know*. For the these sentences to be true, John needs to know the complete true answer of at least one of the involved questions, as described in (114a) and (115a). Conversely, if what John believes is just a disjunctive inference as in (114b) and (115b), we cannot conclude (114) and (115) to be true.

(114) John knows who invited Andy or who invited Billy.
(w: *Mary invited both Andy and Billy, and no one else invited Andy or Billy.*)
  a. $\surd$ John knows that Mary invited Andy, or John knows that Mary invited Billy.
  b. $\times$ John knows that Mary invited Andy or Billy (or both).

(115) John knows whether Mary invited Andy or whether Mary invited Bill.
(w: *Mary invited both Andy and Billy.*)
  a. $\surd$ John knows that Mary invited Andy, or John knows that Mary invited Billy.
  b. $\times$ John knows that Mary invited Andy or Billy.

In comparison, observe that the question-embedding sentence (116) is ambiguous. The embedded disjunction of two declaratives admits both a narrow scope reading and a wide scope reading. The narrow scope reading is derived when the disjunction is interpreted as a join/union of two

Second, I also must admit that the proposed composition in (113) is not flawless — the sister node of the question coordination does not have a fixed semantic type, which is not permitted by the *simple type theory* (Church 1940) employed in Montagovian Compositional Semantics. Specifically, the function denoted by this node (viz., $\lambda\beta.\mathit{know}(j,\beta)$) should be able to take both Q'_1 (of type $\langle e, st\rangle$) and Q'_2 (of type $\langle e, est\rangle$) as arguments, yielding conflicting requirements on the type of the abstracted variable β. One solution to this problem is to assume that the abstracted variable β has a *sum type*, a type that can be one of multiple possible options. For example, if D_a is conceived as the set of items of type a and D_b as the set of items of type b, then $D_{a|b}$ is the set of items of type $a|b$, or equivalently, $D_a \cup D_b$. Under these assumptions, in (113), the β variable is of the sum type $\langle e, st\rangle|\langle e, est\rangle$, and then the embedding-predicate *know* is a polymorphic function of the type $\langle\langle e, st\rangle|\langle e, est\rangle, et\rangle$. I thank Danny Fox, Floris Roelofsen, Simon Charlow, and Manuel Križ for helpful discussions on this issue. All errors are mine.

propositions, as in (116a). The wide scope reading arises if the disjunction is read as a generalized conjunction that quantifies over a set of two propositions, as in (116b).[35]

(116) John knows [$_{S_1}$ Mary invited Andy] or [$_{S_2}$ Mary invited Billy].

a. *know* $\gg$ *or*

i. $[\![S_1 \textit{ or } S_2]\!] = \mathrm{S}'_1 \sqcup \mathrm{S}'_2$

ii. $[\![\textit{John knows } S_1 \textit{ or } S_2]\!] = \textsf{know}(j, \mathrm{S}'_1 \sqcup \mathrm{S}'_2)$

b. *or* $\gg$ *know*

i. $[\![S_1 \textit{ or } S_2]\!] = \mathrm{S}'_1 \bar{\vee} \mathrm{S}'_2 = \lambda\alpha[\alpha(\mathrm{S}'_1) \vee \alpha(\mathrm{S}'_2)]$

ii. $[\![\textit{John knows } S_1 \textit{ or } S_2]\!] = \textsf{know}(j, \mathrm{S}'_1) \vee \textsf{know}(j, \mathrm{S}'_2)$

Second, conjunctions of questions embedded under non-divisive predicates admit only wide scope readings. The predicate *be surprised (at)* is non-divisive. In (117), the agent being surprised at the conjunction of two propositions does not necessarily imply that the agent is surprised at each atomic proposition: $\textsf{surprise}(j, p \wedge q) \not\Rightarrow \textsf{surprise}(j, p) \wedge \textsf{surprise}(j, q)$.

(117) John is surprised that [Mary went to Boston] and [Sue went to Chicago]. (He expected that them would go to the same city.)
$\not\rightsquigarrow$ *John is surprised that Mary went to Boston.*

Nevertheless, when embedding a conjunction of questions, *be surprised (at)* seemingly takes only a divisive reading. For example, (118) expresses that John is surprised at the complete true answer of each involved question.

(118) (w: *Only Mary went to Boston, and only Sue went to Chicago.*)
John is surprised at [$_{Q_1}$ who went to Boston] and [$_{Q_2}$ who went to Chicago].

a. $\rightsquigarrow$ *John is surprised at who went to Boston.*

b. $\rightsquigarrow$ *John is surprised that Mary went to Boston.*

The seeming divisive reading in (118) is predicted by the proposed analysis: the conjunction of questions is a generalized Boolean conjunction, which can only scope above *be surprised (at)*. A schematized derivation is as follows:

(119) a. $[\![Q_1 \textit{ and } Q_2]\!] = \mathrm{Q}'_1 \bar{\wedge} \mathrm{Q}'_2 = \lambda\alpha[\alpha(\mathrm{Q}'_1) \wedge \alpha(\mathrm{Q}'_2)]$

b. $[\![\textit{John is surprised at } Q_1 \textit{ and } Q_2]\!] = \textsf{surprise}(j, \mathrm{Q}'_1) \wedge \textsf{surprise}(j, \mathrm{Q}'_2)$

There are some seeming counterexamples. As Groenendijk and Stokhof (1989) observe, in (120), the disjunction of questions can freely take scope above or below the embedding predicate *wonder*. The wide scope reading implies that the speaker knows that Peter wants to know the answer to one of the two questions, but she is unsure which one this is. The narrow scope reading says that Peter will be satisfied as long as he gets an answer to one of the questions involved, no matter which one.

(120) Peter wonders [whom John loves] or [whom Mary loves].

[35]To observe this ambiguity, we have to drop the complementizer *that*. Adding only one occurrence of *that* after *knows* yields a very strong preference for narrow scope disjunction. Adding two occurrences of *that*, one at the beginning of each disjoined clause, yields only the reading with wide scope disjunction. The scope ambiguity described in (116) is clearer in languages that do not have overt complementizers (e.g., Chinese).

So, how can we derive such narrow scope disjunction readings? Based on a long-standing intuition, we can decompose the intensional predicate *wonder* into 'want to know' at LF (Karttunen 1977, Guerzoni and Sharvit 2007, Uegaki 2015: chap. 2). Under this assumption, the seeming narrow scope reading is actually an "intermediate" scope reading: Peter wants it to be the case that he knows whom John loves or that he knows whom Mary loves. Such a reading arises when the disjunction undergoes QR and gets interpreted between *want* and *know*, as illustrated in (121b).

(121) Peter **wants to know** $[_{Q_1}$whom John loves] or $[_{Q_2}$whom Mary loves].

a. $[[Q_1 \text{ or } Q_2]\ \lambda\beta$ [Peter wants to know β]] (*or* $\gg$ *want to know*)

b. [Peter wants $[[Q_1 \text{ or } Q_2]\ \lambda\beta$ [to know β]]] (*want* $\gg$ *or* $\gg$ *know*)

To sum up, coordinations of questions are generalized quantifiers quantifying over possibly polymorphic sets. In a question coordination, the conjunction/ disjunction coordinates two predications (of type *t*), not directly the root denotations of the conjoined/ disjoined questions. This proposal is supported by the mandatory wide scope or intermediate scope readings of question-coordinations in embeddings.[36]

8. Conclusions

The primary goal of this paper has been to revive categorial approaches to question semantics. Two new pieces of evidence, namely, Caponigro's generalization on the distribution of *wh*-words in questions and FRs, and cases of QV inferences in question-embeddings, suggest that question denotations must be able to supply nominal meanings of short answers. This requirement leaves abstracts (or more precisely, functions from short answers) as the only possible denotations of questions.

I argued that questions should be defined as topical properties and proposed a hybrid categorial approach to compose those topical properties. This approach overcomes the problems and insufficiencies with traditional categorial approaches in defining *wh*-words (especially *wh*-words in languages that use bare *wh*-words as indefinites) and with composing multi-*wh* questions (with single-pair readings or pair-list readings). I also extended this approach to deriving functional readings and pair-list readings of questions.

I proposed that question coordinations are generalized quantifiers quantifying over possibly polymorphic sets. This assumption is supported by the lack of narrow scope readings of question-coordinations.

Acknowledgement [To be added ...]

[36]Despite of the facts, it is unclear, in theory, why the standard meet/join readings are not available in question-coordinations. I thank an anonymous reviewer for pointing out this puzzle.

References

Abrusán, Márta. 2007. Contradiction and Grammar: The Case of Weak Islands. Doctoral Dissertation, MIT.

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

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.

Beck, Sigrid, and Hotze Rullmann. 1999. A flexible approach to exhaustivity in questions. *Natural Language Semantics* 7:249–298.

Beck, Sigrid, and Yael Sharvit. 2002. Pluralities of questions. *Journal of Semantics* 19:105–157.

Berman, Stephen Robert. 1991. On the semantics and logical form of *wh*-clauses. Doctoral Dissertation, University of Massachusetts, Amherst.

Bittner, Maria. 1994. Cross-linguistic semantics. *Linguistics and Philosophy* 17:53–108.

Cable, Seth. 2005. Free relatives in Tlingit and Haida: Evidence that the mover projects. Manuscript, MIT.

Caponigro, Ivano. 2003. Free not to ask: On the semantics of free relatives and *wh*-words crosslinguistically. Doctoral Dissertation, University of California Los Angeles.

Caponigro, Ivano. 2004. The semantic contribution of *wh*-words and type shifts: evidence from free relatives crosslinguistically. In *Proceedings of SALT 14*, ed. Robert B. Young, 38–55.

Caponigro, Ivano, and Kathryn Davidson. 2011. Ask, and tell as well: Question–answer clauses in American Sign Language. *Natural language semantics* 19:323–371.

Caponigro, Ivano, Harold Torrence, and Carlos Cisneros. 2013. Free relative clauses in two mixtec languages. *International Journal of American Linguistics* 79:61–96.

Cecchetto, Carlo, and Caterina Donati. 2015. *(Re) labeling*, volume 70. MIT Press.

Champollion, Lucas, Ivano Ciardelli, and Floris Roelofsen. 2015. Some questions in typed inquisitive semantics. Handout at the Workshop of Questions in Logic and Semantics.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1:181–234.

Chierchia, Gennaro, and Ivano Caponigro. 2013. Questions on questions and free relatives. Handout at Sinn und Bedeutung 18.

Church, Alonzo. 1940. A formulation of the simple theory of types. *The journal of symbolic logic* 5:56–68.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen. 2013. Inquisitive semantics: a new notion of meaning. *Language and Linguistics Compass* 7:459–476.

Ciardelli, Ivano, and Floris Roelofsen. 2015. Alternatives in Montague grammar. In *Proceedings of Sinn und Bedeutung*, volume 19, 161–178.

Ciardelli, Ivano, and Floris Roelofsen. 2018. An inquisitive perspective on modals and quantifiers. *Annual Review of Linguistics* 4.

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler. 2017. Composing alternatives. *Linguistics and Philosophy* 40:1–36.

Cohen, Felix S. 1929. What is a question? *The monist* 39:350–364.

Cremers, Alexandre. 2016. On the semantics of embedded questions. Doctoral Dissertation, École normale supérieure, Paris.

Cremers, Alexandre. 2018. Plurality effects in an exhaustification-based theory of embedded questions. *Natural Language Semantics* 26:193–251.

Cresti, Diana. 1995. Extraction and reconstruction. *Natural Language Semantics* 3:79–122.

Davidson, Kathryn, Ivano Caponigro, and Rachel Mayberry. 2008. Clausal question-answer pairs: Evidence from asl. In *Proceedings of 27th West Coast Conference on Formal Linguistics*, ed. Natasha Abner and Jason Bishop, volume 27, 108–115.

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

Dayal, Veneeta. 2002. Single-pair versus multiple-pair answers: Wh-in-situ and scope. *Linguistic Inquiry* 33:512–520.

Dayal, Veneeta. 2016. List answers through higher order questions. Colloquium talk at MIT, February 2016.

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

Engdahl, Elisabet Britt. 1980. The syntax and semantics of questions in Swedish. Doctoral Dissertation, University of Massachusetts.

Engdahl, Elizabet. 1986. *Constituent questions*. Dordrecht: Reidel.

Fox, Danny. 2012. Multiple wh-questions: uniqueness, pair-list and second order questions. Class notes for MIT seminars.

Fox, Danny. 2013. Mention-some readings of questions. *MIT seminar notes* .

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

Gallin, Daniel. 1975. Intensional and higher-order modal logic, volume 19 of north-holland mathematics studies.

Gamut, LTF. 1991. *Logic, language, and meaning. vol 2: Intensional logic and logical grammar*. Chicago University Press.

Gärtner, Hans-Martin. 2009. More on the indefinite-interrogative affinity: The view from embedded non-finite interrogatives. *Linguistic Typology* 13:1–37.

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

Ginzburg, Jonathan. 1992. Questions, queries and facts: A semantics and pragmatics for interrogatives. Doctoral Dissertation, Stanford university.

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

Groenendijk, Jeroen, and Martin Stokhof. 1982. Semantic analysis of *wh*-complements. *Linguistics and Philosophy* .

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

Groenendijk, Jeroen, and Martin Stokhof. 1989. Type-shifting rules and the semantics of interrogatives. In *Properties, types and meaning*, 21–68. Springer.

Groenendijk, Jeroen, and Martin Stokhof. 1990. Partitioning logical space. Annotated handout.

Guerzoni, Elena, and Yael Sharvit. 2007. A question of strength: on NPIs in interrogative clauses. *Linguistics and Philosophy* 30:361–391.

Hagstrom, Paul Alan. 1998. Decomposing questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of language* 10:41–53.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Clarendon.

Hausser, Roland, and Dietmar Zaefferer. 1979. Questions and answers in a context-dependent Montague grammar. In *Formal semantics and pragmatics for natural languages*, 339–358. Springer.

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

Heim, Irene. 1994. Interrogative semantics and Karttunen's semantics for *know*. In *Proceedings of IATOML1*, volume 1, 128–144.

Heim, Irene. 1995. Notes on questions. MIT class notes for Semantics Proseminar.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Blackwell Textbooks in Linguistics.

Hendriks, Herman. 1993. Studied flexibility: Categories and types in syntax and semantics. Doctoral Dissertation, ILLC, University van Amsterdam.

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

Jacobson, Pauline. 1994. Binding connectivity in copular sentences. In *Semantics and Linguistic Theory*, volume 4, 161–178.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In *Quantification in natural languages*, ed. Angelika Emmon Bach, Kratzer and Barbara Partee, 451–486. Springer.

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

Jacobson, Pauline. 2016. The short answer: implications for direct compositionality (and vice versa). *Language* 92:331–375.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and philosophy* 1:3–44.

Kotek, Hadas. 2014. Composing questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Kotek, Hadas, and Michael Yoshitaka Erlewine. 2018. Non-interrogative wh-constructions in chuj (mayan). In *Proceedings of the Workshop on the Structure and Constituency of the Languages of the Americas (WSCLA)*, volume 21, 101–115.

Kotek, Hadas, and Michael Yoshitaka Erlewine. 2019. *Wh*-indeterminates in Chuj (Mayan). *Canadian Journal of Linguistics* 64:62–101.

Krifka, Manfred. 2001a. For a structured meaning account of questions and answers. *Audiatur vox sapientia. a festschrift for Arnim von Stechow* 52:287–319.

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

Krifka, Manfred. 2011. Questions. In *Semantics: An international handbook of natural language meaning*, ed. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, volume 1, 1742–1785. Walter de Gruyter.

Lahiri, Utpal. 1991. Embedded interrogatives and predicates that embed them. Doctoral Dissertation, Massachusetts Institute of Technology.

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

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. De Gruyter.

Liu, Mingming. 2016. Mandarin *wh*-conditionals as interrogative conditionals. In *Semantics and Linguistic Theory*, volume 26, 814–835.

Merchant, Jason. 2005. Fragments and ellipsis. *Linguistics and philosophy* 27:661–738.

Nicolae, Andreea Cristina. 2013. Any questions? Polarity as a window into the structure of questions. Doctoral Dissertation, Harvard University.

Partee, Barbara, and Mats Rooth. 1983. Generalized conjunction and type ambiguity. In *Meaning, use, and interpretation of language*, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 334–356. Blackwell Publishers Ltd.

Partee, Barbara H. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. J. Groenendijk, D. de Jongh, and M. Stokhof, 357–381. Foris, Dordrecht.

Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *which*-questions. In *Semantics and Linguistic Theory*, volume 8, 215–232.

Schwarz, Berhard. 1994. Rattling off questions. University of Massachusetts at Amherst.

Shan, Chung-chieh, and Chris Barker. 2006. Explaining crossover and superiority as left-to-right evaluation. *Linguistics and Philosophy* 29:91–134.

Sharvit, Yael. 2002. Embedded questions and 'de dicto' readings. *Natural Language Semantics* 10:97–123.

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

Shimoyama, Junko. 2006. Indeterminate phrase quantification in japanese. *Natural Language*

Semantics 14:139–173.

Spector, Benjamin. 2007. Modalized questions and exhaustivity. In *Proceedings of SALT 17*.

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

Stainton, Robert J. 1998. Quantifier phrases, meaningfulness "in isolation", and ellipsis. *Linguistics and Philosophy* 21:311–340.

Stainton, Robert J. 2005. In defense of non-sentential assertion. *Semantics versus Pragmatics* 383–457.

Stainton, Robert J. 2006. Neither fragments nor ellipsis. In *The syntax of nonsententials: Multidisciplinary perspectives*, ed. Eugenia Casielles Ljiljana Progovac, Kate Paesania and Ellen Barton, 93–116. John Benjamins Publishing.

Szabolcsi, Anna. 1997. Quantifiers in pair-list readings. In *Ways of scope taking*, ed. Anna Szabolcsi, 311–347. Springer.

Szabolcsi, Anna. 2016. Direct vs. indirect disjunction of wh-complements, as diagnosed by subordinating complementizers.

Uegaki, Wataru. 2015. Interpreting questions under attitudes. Doctoral Dissertation, Massachusetts Institute of Technology.

Uegaki, Wataru. 2018. A unified semantics for the japanese q-particle *ka* in indefinites, questions and disjunctions. *Glossa: a journal of general linguistics* 3.

von Stechow, Arnim. 1991. Focusing and backgrounding operators. *Discourse particles* 6:37–84.

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

Williams, Alexander. 2000. Adverbial quantification over (interrogative) complements. In *The Proceedings of the 19th West Coast Conference on Formal Linguistics (WCCFL 19)*, 574–587.

Xiang, Yimei. 2016. Interpreting questions with non-exhaustive answers. Doctoral Dissertation, Harvard University Cambridge, Massachusetts.

Zimmermann, Thomas Ede. 1985. Remarks on groenendijk and stokhof's theory of indirect questions. *Linguistics and Philosophy* 8:431–448.