Complete and true: A uniform analysis for mention some and mention all[1]

Yimei Xiang — Harvard University

Abstract. This paper provides a uniform analysis for the interpretations of indirect mention-some questions and indirect mention-all questions. The main goal of this paper is to characterize the readings that are sensitive to false answers, which are usually called "intermediately exhaustive" readings in the case of mention-all. To capture mention-some grammatically, I adopt Fox's (2013) view that "completeness" amounts to Max-informativity, not exhaustiveness. Next, I argue that the "sensitivity to false answers" in direct questions is a matter of quality, not a result of exhaustification (compare Klinedinst & Rothschild 2011). Finally, I present a principled explanation as to why some false answers are more tolerated than the others.

Keywords: Questions, exhaustivity, mention-some, false answers

1. Introduction

Most *wh*-questions admit only exhaustive answers. For example, to properly answer (1), the addressee needs to specify all the attendants to the party, as in (1a), which we call a "mention-all (MA) answer". If the addressee can only provide a non-exhaustive answer like (1b), he would have to indicate an ignorance inference in some way, such as marking the answer with a prosodic rise-fall-rise contour (indicated by '.../'); if (1b) is not properly marked, such as taking a falling tone (indicated by '\'), it would yield an undesired exhaustivity inference.

(1) Who came the party? *(w: only John and Mary came to the party.)*

a. John and Mary did.

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%

In contrast, $\Diamond$-questions, namely *wh*-questions containing a possibility modal, admit both exhaustive and non-exhaustive answers. For instance, (2) can be naturally answered by specifying one or all of the chair candidates. Crucially, the non-exhaustive answer (2b) does not need an ignorance mark: it does not yield an exhaustivity inference even if it takes a falling tone. Due to this difference, we call (2b) a "mention-some (MS) answer" while (1b) a "partial answer". Questions admitting and rejecting MS answers are called "MS questions" and "MA questions", respectively.

[1] I thank Emmanuel Chemla, Gennaro Chierchia, Alexandre Cremers, Danny Fox, Martin Hackl, Andreas Haida, Floris Roelofsen, Jesse Snedeker, Wataru Uegaki, and the audiences at LangCog at Harvard Psychology, LFRG at MIT Linguistics, ImPres1 at ZAS, XPRAG 2015, and SuB 20 for comments and discussions.

(2) Who can chair the committee? *(w: only John and Mary can chair; one chair only.)*

a. John and Mary can.\

b. John can.\ $\not\rightsquigarrow$ *Only John can chair.*

Earlier works notice two forms of exhaustivity involved in interpreting indirect MA questions, namely *weak exhaustivity* (Karttunen 1977) and *strong exhaustivity* (Groenendijk & Stokhof 1984). Consider (3) for instance. The weakly exhaustive (WE) reading only requires John to know the MA answer as to *who came*, while the strongly exhaustive (SE) reading also requires John to know the MA answer as to *who didn't come*. Recent works (Klinedinst & Rothschild 2011, Spector & Egré 2015, Uegaki 2015, Cremers & Chemla 2016) start to consider an intermediate form of exhaustivity: stronger than WE but weaker than SE, the intermediately exhaustive (IE) reading requires John to know the MA answer as to *who came* and <u>have no false belief as to *who came*</u>. I call the underlined condition "be sensitive to false answers".

(3) John knows who came. *(w: among the three considered individuals abc, only ab came.)*

a. John knows that a and b came. **WE**

b. John knows that a and b came; and John knows that c did **not** come. **SE**

c. John knows that a and b came; and **not** [John believes that c came]. **IE**

WE and SE have relatively limited distributions (Heim 1994, Guerzoni & Sharvit 2007, Nicolae 2013, Uegaki 2015, a.o.). In general, indirect questions with a non-factive verb (e.g., *tell*, *predict*) cannot take SE, while those with a factive verb (e.g., *know*, *remember*) cannot take WE. In contrast, as experimentally validated by Cremers & Chemla (2016), IE readings are available to most indirect questions, including those with a non-factive verb as well as those with a cognitive factive.

George (2013) observes that indirect MS questions also have readings sensitive to false answers, which are similar to the IE readings of indirect MA questions. Consider the scenario described in (4): Italian newspaper is available at Newstopia but not PaperWorld; both John and Mary know a true MS answer as to *where one can buy an Italian newspaper* (viz., *at Newstopia*), but Mary also believes a false answer, namely *that one can buy an Italian newspaper at PaperWorld*. Intuitively, there is a prominent reading under which (4a) is true while (4b) is false.

(4)

Italian newspaper available at ...	*Newstopia?*	*PaperWorld?*
Facts	*Yes*	*No*
John's belief	Yes	?
Mary's belief	Yes	Yes

a. John knows where one can buy an Italian newspaper. True

b. Mary knows where one can buy an Italian newspaper. False

It is debatable whether the reading described above for (4a-b) is exhaustive (see section 3.1.2). To be theory neutral, for both MA questions and MS questions, I call the readings that are sensitive to false answers "FA-sensitive readings". I divide the truth conditions of an FA-sensitive reading into two parts, namely *Completeness* and *FA-sensitivity*, roughly described in (5).

(5) *John told us Q.*
  a. John told us a complete true answer of Q. Completeness
  b. John does not tell us any false answer of Q. FA-sensitivity

The goal of this paper is to characterize the truth conditions of FA-sensitive readings. The crucial claims of the following sections are summarized as follows.

§2. Completeness amounts to Max-informativity, rather than exhaustiveness (Fox 2013).

§3. (i) FA-sensitivity is concerned with all types of false answers, not only those that are possible complete answers. (ii) FA-sensitivity is a matter of "quality", rather than a consequence of exhaustification (compare Klinedinst & Rothschild 2011). (iii) For indirect questions with an emotive factive, FA-sensitivity collapses under strong factivity.

§4. Experiments show (i) that FA-sensitivity is also concerned with false denials, and (ii) that FA-sensitivity exhibits asymmetries that vary by question-type.

§5. The asymmetry of FA-sensitivity is determined by the Principle of Tolerance.

2. Completeness

2.1. Completeness as exhaustiveness

Earlier works on questions consider only exhaustive answers as complete answers (Groenendijk & Stokhof 1984, Dayal 1996, a.o.). Since MS answers are not exhaustive, works following this line attribute the acceptability of MS to pragmatic factors: MS answers are partial answers that are sufficient for the conversational goal behind the question (Groenendijk & Stokhof 1984, van Rooij 2004, Schulz & van Rooij 2006). Consider the typical MS question *where can I get gas*. If the goal is just to find a local place to get gas, the addressee only needs to name one local gas station; if the goal is to investigate the local gas market, the addressee needs to list out all the local gas stations.

I agree that pragmatics plays a role in distributing MS in several respects; for instance, if a question is semantically ambiguous between MS and MA, a goal that calls for an exhaustive answer blocks MS. But, I doubt that pragmatics is restrictive enough to predict the limited distribution of MS. In the following, I provide two empirical arguments against the pragmatic account of MS. Both of the arguments are related to *mention-intermediate (MI) answers*. Those answers are, as the name implies, non-exhaustive answers that are stronger than MS answers. I show that the pragmatic view cannot capture the differences between MS and MI: contrary to the case of MS, MI is unacceptable in root questions and embedded questions. **First**, MI answers must be ignorance-marked, even though they are informative enough to satisfy the question goal. For instance, assume that the goal of (6) is to find one qualified person to chair the committee. The MS answer (6a) does not have to be ignorance-marked. In contrast, while being sufficient for the pragmatic goal, the MI answer (6b), which names more than one but not all of the chair candidates, must to be ignorance-marked, otherwise it would yield an undesired exhaustivity inference. More generally, the obligatory ignorance-mark on (6b) suggests that whether an answer of a $\Diamond$-question can be read non-exhaustively is primarily determined by the grammatical structure of this answer, not the question goal: if not ignorance-marked, an individual answer like (6a) can be non-exhaustive, while a conjunctive answer like (6b) admits only an exhaustive reading.

(6) Who can chair the committee? *(w: only John, Mary, and Sue can chair; one chair only.)*
  a. John.\ $\not\rightsquigarrow$ *Only John can chair.*
  b. John and Mary.../
  b′. # John and Mary.\ $\rightsquigarrow$ *Only John and Mary can chair.*

Second, interpretations of indirect questions suggest that good answers are always "mention one (group)" or "mention all (groups)", as exemplified in (7a) and (7b), respectively. The conversational goal of a question, however, can be any "mention N (groups)" where N is a number in the available range. For instance, assume that the dean wants to meet with three chair candidates so as to make plans for the committee, then the goal of the embedded question in (7) would be "mention three". A pragmatic account predicts (7) to take the mention-three reading (7c), which however is infeasible. A semantic account does not have this prediction: complete answers derived from the possible logical forms of an MS-question are either mention one or mention all, not intermediate.

(7) John knows who can chair the committee.
  a. For some x such that x can chair, John knows that x can chair. $\surd$
  b. For every x, if x can chair, John knows that x can chair. $\surd$
  c. For some xyz such that xyz each can chair, John knows that xyz each can chair. $\times$

2.2. Completeness as Max-informativity

To capture the availability of MS grammatically, Fox (2013) weakens the definition of completeness and proposes that any *maximally informative* (MaxI) true answer counts as a complete true answer. Given a set of propositions α, the strongest member of α is the unique member that entails all the members of α, while the MaxI members of α are the ones that are not asymmetrically entailed by any members of α. Consider (8a) and (8b) for illustrations. Q_w stands for the set of true answers in w. Underlining highlights their MaxI true answers. The basic *wh*-question (8a) has and can only have one MaxI true answer, namely the MA answer. While the $\Diamond$-question (8b) has two MaxI true answers, both of which are MS answers.

(8) a. Who made the swimming team? *(w: only a and d made the swimming team.)*
$Q_w = \{a \text{ made the team}, d \text{ made the team}, \underline{a \oplus d \text{ made the team}}\}$
b. Where can Sue get a bottle of wine? *(w: wine is only available at store a and d.)*
$Q_w = \{\underline{\Diamond(\text{Sue get a bottle from } a)}, \underline{\Diamond(\text{Sue get a bottle from } d)}\}$

I schematize Fox's basic idea as in (9), using Hamblin-Karttunen semantics of questions (Hamblin 1973, Karttunen 1977): the ANS-operator applies to the Hamblin set Q and the evaluation world w, returning the set of MaxI members of the Karttunen set Q_w.

(9) $\text{ANS}(Q)(w) = \text{MaxI}(Q_w)$, where $\text{MaxI} = \lambda\alpha.\{p : p \in \alpha \wedge \forall q \in \alpha[q \not\subset p]\}$

Compared with the earlier accounts on completeness, Fox's account leaves space for MS: it allows a non-exhaustive answer to be a good answer and a question to have multiple good answers. Nevertheless, Fox's account still misses some good MS answers. For instance in (10), both (10b-c) are intuitively good MS answers; but with a monotonic predicate *serve on the committee*, (10b) is asymmetrically entailed by (10c). Thus, Fox incorrectly predicts (10b) to be a partial answer.

(10) Who can serve on the committee? *(w: the committee can be made up of G+D or G+D+J)*
a. $\times$ Gennaro. b. $\surd$ Gennaro and Danny. c. $\surd$ Gennaro, Danny, and Jim.

Consider what (10b) precisely means. Intuitively, it means that *to form the committee, it is possible to have only Gennaro and Danny serve on the committee*. This reading involves exhaustivity scoping beneath the possibility modal *can*. To capture this intuition, I propose that the weak modal *can* embeds a covert exhaustivity O-operator associated with the *wh*-trace. This O-operator has a meaning approximating to the exclusive focus particle *only*: it affirms the prejacent and negates the alternatives that are not entailed by the prejacent. Moreover, the modal base of the teleological modal verb *can* is restricted to the set of worlds where the question goal is reached.

(11) $O(p) = \lambda w.p(w) \wedge \forall q \in \mathcal{A}lt(p)[p \not\subseteq q \rightarrow \neg q(w)]$ (Chierchia et al. 2013)
(p is true, any alternatives of p not entailed by p are false.)

The O-operator creates a non-monotonic environment with respect to the *wh*-trace, which therefore breaks up the entailment relation from (10c) to (10b) and preserves both (10b-c) as good answers. Moreover, the embedded O evokes local exhaustivity and rules out (10a): it is false that *to form the* *committee, it is possible to have only Gennaro serve on the committee.* Now, the answer space of an MA question and that of an MS question can be illustrated as in (12) and (13), respectively. In (12), an entailment relation holds consistently from the top to the bottom, as indicated by arrows; while in (13), all the answers are logically independent.[2]

(12) Who served on the committee?

$$\begin{tikzcd}
 & f(a \oplus b \oplus c) \arrow[dl] \arrow[d] \arrow[dr] & \\
f(a \oplus b) \arrow[d] \arrow[dr] & f(a \oplus c) \arrow[dl] \arrow[dr] & f(b \oplus c) \arrow[dl] \arrow[d] \\
f(a) & f(b) & f(c)
\end{tikzcd}$$

(13) Who can serve on the committee?

$$\begin{array}{ccc}
 & \Diamond O f(a \oplus b \oplus c) & \\
\Diamond O f(a \oplus b) & \Diamond O f(a \oplus c) & \Diamond O f(b \oplus c) \\
\Diamond O f(a) & \Diamond O f(b) & \Diamond O f(c)
\end{array}$$

The Completeness Condition of *John told us Q*, regardless of whether Q is MS or MA, can be uniformly stated as *John told us a MaxI true answer of* Q, as schematized below. It does not matter whether the existential semantics is attributed by an existential closure or a choice function.

(14) $\lambda w.\exists \phi \in \textsc{Ans}(Q)(w)[\textbf{told}_w(j, \phi)] = \lambda w.\exists \phi \in \text{MaxI}(Q_w)[\textbf{told}_w(j, \phi)]$

3. FA-sensitivity

3.1. The exhaustification-based approach

3.1.1. FA-sensitivity in MA questions

Klinedinst & Rothschild (2011) (K&R) account for IE readings using exhaustifications: exhaustifying (15a) yields an inference entailing (15b). Formally, K&R assume that the ordinary value of (15) is its WE reading, and that IE is derived by exhaustifying the WE inference. Exhaustification affirms the WE inference and negates all the propositions of the form "John told us ϕ" where ϕ is a possible MA answer of *who came* and is not entailed by the true MA answer of *who came*.

(15) John told us who came.

a. If x came, John told us that x came.

[2] This paper considers only individual answers and questions with distributive predicates. See Xiang (2016) for discussions on higher-order answers and questions with collective predicates. The basic idea is as follows: the live-on set of *who* consists of not only individuals of type e but also generalized disjunctions and conjunctions (e.g., $a \oplus b \wedge c \oplus d = \lambda P_{est} \lambda w_s.P_w(a \oplus b) \wedge P_w(c \oplus d)$); therefore, the answer space of (1) is closed under conjunction.

(1) Who formed a team? *(w: ab formed a team, cd formed a team)*
$Q_w = \{f(a \oplus b), f(c \oplus d), f(a \oplus b) \wedge f(c \oplus d)\}$

b. If x didn't come, John didn't say to us that x came.

(16) a. $[\![\text{who came}]\!] = \lambda w \lambda w'.\forall x[\textbf{came}_w(x) \to \textbf{came}_{w'}(x)]$

b. $[\![p]\!] = \lambda w.\textbf{told}_w(j, \lambda w'.\forall x[\textbf{came}_w(x) \to \textbf{came}_{w'}(x)])$ **WE**

(John told Mary the MA answer as to *who came*$_w$)

c. $\mathcal{A}lt(p) = \{q \mid \exists w''[q = \lambda w.\textbf{told}_w(j, \lambda w'.\forall x[\textbf{came}_{w''}(x) \to \textbf{came}_{w'}(x)])]\}$

($\{q \mid \exists w''[q =$ John told Mary the MA answer of *who came*$_{w''}$$]\}$)

d. $[\![O(p)]\!] = \lambda w.p(w) \wedge \forall q \in \mathcal{A}lt(p)[p \not\subseteq q \to \neg q(w)]$ **IE**

(λw. John only told$_w$ us the true MA answer as to *who came*$_w$)

The WE inference of an indirect MA question amounts to the Completeness condition. Thus, using Hamblin-Karttunen semantics, we can re-schematize K&R's idea as follows.

(17) John told us Q.

a. $[\![p]\!] = \lambda w.\exists \phi \in \textsc{Ans}(Q)(w)[\textbf{told}_w(j, \phi)]$ **WE**

b. $\mathcal{A}lt(p) = \{\lambda w.\exists \phi \in \alpha[\textbf{told}_w(j, \phi)] \mid \exists w'[\alpha = \textsc{Ans}(Q)(w')]\}$

$= \{\lambda w.\exists \phi \in \textsc{Ans}(Q)(w')[\textbf{told}_w(j, \phi)] \mid w' \in W\}$

c. $[\![O(p)]\!] = \lambda w.p(w) \wedge \forall q \in \mathcal{A}lt(p)[p \not\subseteq q \to \neg q(w)]$ **IE**

3.1.2. FA-sensitivity in MS questions

In an indirect MS question like (18), there are two possible positions to place the O-operator: one position is immediately above the scope part of the existential closure, called "local exhaustification"; the other is above the existential closure, called "global exhaustification". In the following, I show that neither of the options derives the desired the FA-sensitivity inference.

(18) John told us [$_Q$ where we could get gas].

a. $\exists\phi$ [ϕ is a true MS answer of Q] [O [John told us ϕ]] **Local exhaustification**

b. O [$\exists\phi$ [ϕ is a true MS answer of Q] [John told us ϕ]] **Global exhaustification**

Local exhaustification is apparently infeasible. This operation yields the following truth conditions: first, John told us an MS answer as to *where we could get gas*; second, John didn't give us any answer that is not entailed by this MS answer. The second condition is too strong. For instance, if what John said was *we could get gas at place a and somewhere else*, which is strictly stronger than any MS answer, the sentence (18) would be predicted to be false, contra the fact.[3]

[3] One might suggest to stipulate that the local exhaustifier negates only false inferences. This option is however technically difficult and conceptually circular.

The option of global exhaustification seems to have a better chance of yielding the desired FA-sensitivity inference. As Danny Fox and Alexandre Cremers p.c. to me independently, *innocently exclusive exhaustification* (Fox 2007) yields an inference that is very close to the FA-sensitivity condition. While the regular exhaustifier O negates all the excludable alternatives (i.e., the alternatives that are not entailed by the prejacent of the exhaustifier), the innocently exclusive exhaustifier O_{IE} negates only *innocently (I)-excludable alternatives*. For a proposition p, an alternative q is I-excludable iff $p \wedge \neg q$ is consistent with negating any excludable alternative(s) of p.

(19) a. $\text{Excl}(p) = \{q : q \in \mathcal{A}lt(p) \wedge p \not\subseteq q\}$
b. $\text{IExcl}(p) = \{q : q \in \mathcal{A}lt(p) \wedge \neg\exists q' \in \text{Excl}(p)[[p \wedge \neg q] \rightarrow q']\}$
c. $O_{\text{IE}}(p) = p \wedge \forall q \in \text{IExcl}(p)[\neg q]$

Using innocent exclusion avoids negating propositions of the form "John told us ϕ" where ϕ is a true MS answer or a disjunction involving at least one true MS answer as a disjunct. Consider (20) for instance. Using innocent exclusion, global exhaustification proceeds as follows. The prejacent of O_{IE} is a disjunction that coordinates all the true MS answers, as schematized in (20b). ϕ_a is short for the proposition *we could get gas at place* a. Alternatives are propositions of the form "John told us a member of α" where α is a possible set of complete answers, as list in (20c). Among these alternatives, only $\textbf{told}(j, \phi_c)$ is I-excludable. Hence, empolying O_{IE} yields a very appealing inference (20d): *John told us a true MS answer of Q, and didn't give us any false MS answer of Q*.

(20) John told us [$_Q$ where we could get gas].
(w: among the considered places abc, only ab sold gas)
a. O_{IE} [$_S$ $\exists\phi$ [ϕ is a true MS answer of Q] [John told us ϕ]]
b. $[\![S]\!] = \lambda w.\exists\phi \in \textsc{Ans}(Q)(w)[\textbf{told}_w(j,\phi)] = \textbf{told}(j,\phi_a) \vee \textbf{told}(j,\phi_b)$
c. $\mathcal{A}lt(S) = \{\lambda w.\exists\phi \in \alpha[\textbf{told}_w(j,\phi)] \mid \exists w'[\alpha = \textsc{Ans}(Q)(w')]\}$
$$= \left\{ \begin{array}{lll} \textbf{told}(j,\phi_a), & \textbf{told}(j,\phi_a) \vee \textbf{told}(j,\phi_b), & \textbf{told}(j,\phi_a) \vee \textbf{told}(j,\phi_b) \vee \textbf{told}(j,\phi_c) \\ \textbf{told}(j,\phi_b), & \textbf{told}(j,\phi_a) \vee \textbf{told}(j,\phi_c), & \\ \textbf{told}(j,\phi_c), & \textbf{told}(j,\phi_b) \vee \textbf{told}(j,\phi_c), & \end{array} \right\}$$
d. $[\![O_{\text{IE}}(S)]\!] = [\textbf{told}(j,\phi_a) \vee \textbf{told}(j,\phi_b)] \wedge \neg\textbf{told}(j,\phi_c)$

3.2. Problems with the exhaustification-based account

3.2.1. Problem 1: FA-sensitivity is not a scalar implicature

Treating FA-sensitivity as a logical consequence of exhaustifying Completeness amounts to saying that FA-sensitivity is a scalar implicature (SI) of Completeness. Nevertheless, FA-sensitivity inferences do not behave like SIs. **First**, FA-sensitivity inferences are easily generated even in downward-entailing contexts. In (21a), appearing within the antecedent of a conditional, the scalar item *some* (unless focus-marked) does not evoke an SI. This is so because strengthening the antecedent weakens the entire conditional and violates the *Strongest Meaning Hypothesis* (Chierchia et al. 2013; Fox & Spector to appear) for exhaustifications: the use of an exhaustifier is marked if it gives rise to a reading that is equivalent to or weaker than what would have resulted in its absence. In (21b), however, while uttered as the antecedent of a conditional, the indirect question *Mary knows which speakers went to the dinner* still evokes an FA-sensitivity inference.

(21) a. If [Mary invited some of the speakers to the dinner], I will buy her a coffee.
$\not\rightsquigarrow$ If Mary invited some but **not all** speakers to the dinner, I will buy her a coffee.
b. (w: *Barbara and Irene went to the dinner, but Uli didn't.*)
If Mary knows which speakers went to the dinner, I will buy her a coffee.
$\rightsquigarrow$ If [Mary knows that Barbara and Irene went to the dinner] $\wedge$
not [Mary believes that Uli went to the dinner], I will buy her a coffee.

Second, FA-sensitivity inferences are not cancelable. In (22a), the SI *that Mary did not invite all of the speakers to the dinner* can be easily cancelled, while in (22b) the FA-sensitivity inference *it is not the case that Mary believes that Uli went to the dinner* cannot be cancelled.

(22) a. A: "Did Mary invite some of the speakers to the dinner?"
B: "Yes. Actually she invited all of them."
b. (w: *Barbara and Irene went to the dinner, but Uli didn't.*)
A: "Does Mary know which speakers went to the dinner?"
B: "Yes. #Actually also she believes that Uli went to the dinner."

One might suggest that FA-sensitivity inferences are special species of SIs which are mandatorily evoked and exceptionally robust. To assess this assumption, let us compare FA-sensitivity inferences with SIs that are mandatorily evoked in presence of the overt exhaustifier *only*. In (23) for instance, since the scalar item *some* is associated with *only*, its SI patterns like FA-sensitivity inferences: this SI can be generated within the antecedent of a conditional and cannot be cancelled.

(23) a. If [Mary invited only SOME$_F$ of the speakers to the dinner], I will buy her a coffee.
$\rightsquigarrow$ If Mary invited some but **not all** speakers to the dinner, I will buy her a coffee.
b. A: "Did Mary invite only SOME$_F$ of the speakers to the dinner?"
B: "Yes. # Actually she invited all of them."

Nevertheless, a difference arises in negative sentences. In (24b), associating *only* with the focused item over negation evokes a positive implicature, namely an indirect SI: *only* negates the negative alternative $\neg\phi_{\text{male}}$, yielding an indirect SI ϕ_{male}, as schematized in (24c).

(24) a. Mary **only** invited some [female]$_F$ speakers to the dinner.
$\rightsquigarrow$ Mary did not invite any male speakers to the dinner. $\neg\phi_{\text{male}}$

b. Mary **only** did <u>not</u> invite any [female]$_F$ speakers to the dinner.
$\rightsquigarrow$ Mary did invite some male speaker(s) to the dinner. ϕ_{male}

c. $O\neg\phi_{\text{female}} = \neg\phi_{\text{female}} \wedge \neg\neg\phi_{\text{male}} = \neg\phi_{\text{female}} \wedge \phi_{\text{male}}$

If the FA-sensitivity inference were a mandatory SI, we would analogously predict that a negated indirect question like (25b) takes the LF (25c) and evokes an indirect SI $\textbf{told}(m, \phi_{\text{uli}})$, namely the negation of the FA-sensitivity inference, contra the fact. Note that here the exhaustifier cannot be placed below negation, due to the Strongest Meaning Hypothesis.

(25) (w: *Barbara and Irene went to the dinner, but Uli didn't.*)

a. Mary told us which speakers went to the dinner.
$\rightsquigarrow$ Mary did not tell us that Uli went to the dinner. $\neg\textbf{told}(m, \phi_{\text{uli}})$

b. Mary did <u>not</u> tell us which speakers went to the dinner.
$\not\rightsquigarrow$ Mary told us that Uli went to the dinner. $\textbf{told}(m, \phi_{\text{uli}})$

c. O **not** [Mary told us [$_Q$ which speakers went to the dinner]]

3.2.2. Problem 2: FA-sensitivity is concerned with partial answers

So far, the alternative set used by the exhaustification-based account includes only propositions that are possible complete answers. Hence, exhaustifying the Completeness condition only yields the requirement of avoiding false answers that are possible complete answers. The FA-sensitivity condition, however, requires to avoid all types of false answers, including those that can never be complete. For instance, (26) and (27) are intuitively false in the given scenarios, which suggests that the FA-sensitivity condition is also concerned with disjunctive partial answers like $\phi_c \vee \phi_d$.

(26) John told us where we could get gas. [Judgement: FALSE]

a. Fact: a and b sold gas; c and d didn't.

b. John said to us: "a, b, and somewhere else sell gas, which might be either c or d."

(27) John told us who came. [Judgement: FALSE]

a. Fact: a and b came; c and d didn't come.

b. John said to us: "a, b, and someone else came, who might be either c or d."

Moreover, interpretations of indirect MS questions show that FA-sensitivity is also concerned with false denials, which also are always partial. As seen in section 1, George (2013) has discussed false answers that are *over-affirming* (OA), namely overly affirming a possible answer that is false in the evaluation world: Mary incorrectly believes that Italian newspapers are available at store B. Correspondingly, we should also check false answers that are *over-denying* (OD), namely denying a possible answer that is true in the evaluation world: Sue incorrectly believes that Italian newspapers are unavailable at store C. The truth value of (28c) reflects whether FA-sensitivity is concerned with OD: if OD is involved in FA-sensitivity, then there should be a reading under which (28a) is true while (28c) is false. It is a bit hard to judge whether (28c) is true or false (see explanation in section 5), but my experiments in section 4 do show that OD is involved in FA-sensitivity: (28c) received significantly less acceptances than (28a).

(28)

Italian newspaper available at ...	*A?*	*B?*	*C?*	FA-type
Facts	*Yes*	*No*	*Yes*	
John's belief	Yes	?	?	
Mary's belief	Yes	**Yes**	?	OA
Sue's belief	Yes	?	**No**	OD

a. John knows where one can buy an Italian newspaper. True
b. Mary knows where one can buy an Italian newspaper. False
c. Sue knows where one can buy an Italian newspaper. True or False?

Notice that, from indirect MA questions, we cannot tell whether FA-sensitivity is concerned with OD. In (29) for instance, the requirement of avoiding OD can be understood in two different ways. One way is to treat this requirement simply as a logical consequence of Completeness, given that (29a) entails (29c). The other way is to treat this requirement as part of FA-sensitivity and group it together with the condition (29b), given that both (29b-c) are concerned with false answers. Previous and other ongoing studies on FA-sensitivity (K&R 2011, Uegaki 2015, Roelofsen et al. 2014) take the former option; they predict that FA-sensitivity is only concerned with false answers that are possibly complete answers. But given that FA-sensitivity is concerned with OD in indirect MS questions, we should accordingly take the second option for indirect MA questions.

(29) *John knows who came.*
a. if x came, John believes that x came.
b. if x didn't come, not [John believes that x came] Avoiding OA
c. if x came, not [John believes that x didn't come]. Avoiding OD

One might suggest to enlarge the alternative set based on the condition of *Relevance*: a proposition p is relevant to a question Q iff p is equivalent to the union of some cells of the partition yielded by Q (Heim 2011). This move, however, does not work for the exhaustification-based approach; it yields bad consequence in interpreting indirect MS questions. For instance in (30), it rules in not only inferences as to telling a false answer, like those in (30a-c), but also inferences as to telling a true answer that is strictly stronger than an MS answer, such as (30d). Once (30d) is added into the alternative set, an exhaustification-based account would incorrectly predict (30) to be false in a discourse where John told us multiple accessible gas stations.

(30) John told us where we could get gas. *(w : a and b sell gas; c and d do not.)*
  a. OA: $\textbf{told}(j, \phi_c)$, $\textbf{told}(j, \phi_d)$
  b. OD: $\textbf{told}(j, \neg\phi_a)$, $\textbf{told}(j, \neg\phi_b)$
  c. Partial: $\textbf{told}(j, \phi_c \vee \phi_d)$
  d. MA or MI: $\textbf{told}(j, \phi_a \wedge \phi_b)$

3.3. My analysis: A quality-based approach

I propose that FA-sensitivity is simply a matter of "Quality": only make true contributions.[4] Take (31) for instance, where Q can be either MA or MS. The FA-sensitivity condition of this indirect question is concerned with all types of false answers relevant to Q, not just those that can be complete. $\textsc{Rel}(Q)$ stands for the set generated from closing the Hamblin set Q under propositional connectives (negation, disjunction, and conjunction). For instance, if $Q = \{p, q\}$, then $\textsc{Rel}(Q) = \{p, q, \neg p, \neg q, p \wedge q, p \vee q, p \wedge \neg q, ...\}$. This FA-sensitivity condition does not negate any propositions about telling a true answer of Q, and hence it is free from the problem that we saw in (30).

(31) John told us Q.
  a. $\lambda w.\exists\phi \in \textsc{Ans}(Q)(w)[\textbf{told}_w(j, \phi)]$ **Completeness**
  (λw. John told$_w$ us a complete true answer of Q in w.)
  b. $\lambda w.\forall\phi \in \textsc{Rel}(Q)[\textbf{told}_w(j, \phi) \rightarrow \phi(w)]$ **FA-sensitivity**
  (λw. Every Q-relevant proposition that John told$_w$ us is true in w.)

In case that the question-embedding verb is factive, I predict that FA-sensitivity will collapse under factivity. For instance in (32), the emotive factive *be surprised* triggers a factive presupposition c *came*. Locally accommodating this presupposition does not change Completeness, but turns FA-sensitivity into a tautology. More concretely, (33b) is true as long as the factive presupposition is accommodated under negation, and (33c) is not implied because global accommodation causes presupposition failure.

(32) John is surprised at Q.
  a. $\lambda w.\exists\phi \in \textsc{Ans}(Q)(w)[\textbf{surprised}_w(j, \phi) \wedge \underline{\phi(w)}]$ **Completeness**
  (λw. John is surprised$_w$ at a complete true answer of Q in w)
  b. $\lambda w.\forall\phi \in \textsc{Rel}(Q)[\textbf{surprised}_w(j, \phi) \wedge \underline{\phi(w)} \rightarrow \phi(w)]$ **FA-sensitivity**
  (λw. every Q-relevant proposition that surprises$_w$ John and is true in w is true in w)

[4] I leave it open whether this condition is a grammatical constraint or a Gricean maxim.

(33) John is surprised at who came. *(w: among the considered individuals abc, only ab came.)*

a. $\rightsquigarrow$ John is surprised that ab came. $\textbf{surprise}(j, \phi_a \wedge \phi_b)$

b. $\rightsquigarrow$ it is not the case that John is surprised that c came. $\neg[\textbf{surprise}(j, \phi_c) \wedge \phi_c]$

c. $\not\rightsquigarrow$ John isn't surprised that c came. $\neg\textbf{surprise}(j, \phi_c)_{\phi_c}$

Puzzles arise in cases of cognitive factives. Spector & Egré (2015) speculate that the FA-sensitive (viz. IE) reading of (34) should be paraphrased as (34c) rather than (34a-b): to be more specific, in paraphrasing the FA-sensitivity inference, the factive verb *know* should be replaced with its non-factive counterpart *believe*, and the factive presupposition should be ignored.

(34) John knows who came. *(w: consider three individuals abc; only a and b came.)*

a. $\times$ $\textbf{know}(j, \phi_a \wedge \phi_b) \wedge \neg\textbf{know}(j, \phi_c)_{\phi_c}$

b. $\times$ $\textbf{know}(j, \phi_a \wedge \phi_b) \wedge \neg[\textbf{know}(j, \phi_c) \wedge \phi_c]$

c. $\surd$ $\textbf{know}(j, \phi_a \wedge \phi_b) \wedge \neg\textbf{believe}(j, \phi_c)$

We need to explain two puzzles. **First**, why is that (34c) is more preferable than (34a-b)? The answer is simple: (34a) suffers presupposition failure, and (34b) is a tautology; therefore, whenever allowed, it is better to "deactivate" the factive presupposition of *know* in paraphrasing the FA-sensitivity inference. **Second**, why is that the FA-sensitivity inference of (33) keeps the factive presupposition of *be surprised* and accommodates it locally, contrary to the case in (34)? This contrast correlates with the general distinction between emotive factives and cognitive factives as presupposition triggers, as exemplified in (35): the factive presupposition triggered by the cognitive factive *discover* is defeasible, while that triggered by the emotive factive *regret* is not.

(35) a. If someone *regrets* that I was mistaken, I will admit that I was wrong.
$\rightsquigarrow$ The speaker was mistaken.

b. If someone *discovers* that I was mistaken, I will admit that I was wrong.
$\not\rightsquigarrow$ The speaker was mistaken.

Earlier works have argued that emotive factives are strong triggers, while cognitive factives are weak triggers (Karttunen 1971, Stalnaker 1974). Recent theoretical and experimental works (Romoli 2014, Romoli & Schwarz to appear) argue that the presuppositions of soft triggers are actually scalar implicatures. The contrast between hard and soft triggers is far beyond the scope of this article, but whatever accounting for this contrast can also explain the contrast between (33) and (34) with respect to the FA-sensitivity inferences.

4. Experiments

The primary goal of the following experiments is to investigate whether false answers with OD are involved in the condition of FA-sensitivity. The experiment results show that OD is indeed involved in FA-sensitivity, and that FA-sensitivity exhibits asymmetries that vary by question-type.

4.1. Design

Did ... make the swimming team? *Could Susan buy a bottle of red wine at ...?*	a	b	c	d	Ans-type
Fact	*Yes*	*No*	*No*	*Yes*	
A1	**No**	?	No	Yes	OD
A2	?	No	No	Yes	MS
A3	Yes	?	No	Yes	MA
A4	Yes	**Yes**	?	Yes	OA

Table 1: Design of Exp-MA and Exp-MS

Exp-MA K&R (2011) conducted a survey to establish the existence of IE. They stipulated that four individuals $abcd$ tried out for the swimming team, and that only ad made the team. Four sets of predictions (A1-A4 in Table 1) were made as to whether each individual made the team. For instance, A1 means that the agent predicted that d but not a nor c made the swimming team and that the agent was uncertain whether b made it. Next, they asked the participants to judge whether or not each prediction *correctly predicted who made the swimming team*. Each combination of responses corresponds to a reading of the indirect MA question x *predicted who made the swimming team*. For instance, the participants who chose IE would ideally accept A3 and reject the rest responses.

K&R were not particularly interested in OD. They removed the participants who accepted A1/A2 (viz., the participants who were tolerant of incompleteness) from their analysis. But this survey is helpful for studying the sensitivity to false answers in indirect questions: A1 and A4 represent answers with OD and answers with OA, respectively; A1 incorrectly predicted that a did not make the team, and A4 incorrectly predicted that b made the team. A2-A3 have no false predictions, but A2 violates Completeness. I renamed A1-A4 as OD/MS/MA/OA and re-analyzed the raw data.[5]

[5]See here (`http://users.ox.ac.uk/~sfop0300/questionsurvey/`) for the raw data. This survey has no fillers. Thus I excluded only participants who were (i) non-native speakers, (ii) rejected by Amazon Mechanical Turk (MTurk), or (iii) with missing responses. 107 participants (out of 193) were kept in my analysis.

Exp-MS I conducted a similar experiment for MS-questions on MTurk:[6] among the four liquor stores $abcd$ at Central Square, only ad sold red wine; Susan asked her local friends *where she could buy a bottle of red wine at Central Square* and received four responses (A1-A4 in Table 1). Participants were asked to identify whether each response correctly answered Susan's question. Note here that A2 satisfies Completeness, contrary to the case in Exp-MA.

4.2. Results and discussions

Figure 1 and Figure 2 summarizes the proportions of acceptances by ANSWER in Exp-MA and Exp-MS, respectively. N stands for the sample size.

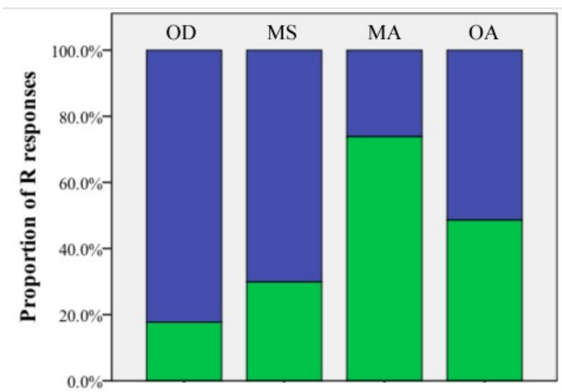

Figure 1: Proportion of acceptances by ANSWER in Exp-MA ($N = 107$)

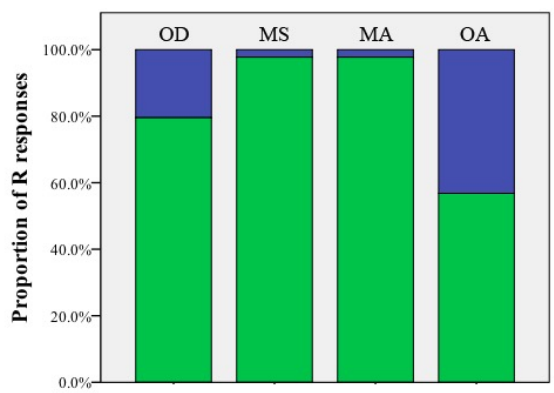

Figure 2: Proportion of acceptances by ANSWER in Exp-MS ($N = 88$)

FA-sensitivity For every two answers in each experiment, I fitted a logistic mixed effects model predicting responses by ANSWER.[7] All the models, except the one for MS versus MA in Exp-MS, reported a significant effect. These significant effects, especially the ones for OD versus MS/MA in Exp-MS, show that FA-sensitivity is concerned with both OA and OD.

Asymmetries of FA-sensitivity Compared with OD, OA received significantly more acceptances in Exp-MA ($\hat{\beta}$ = 1.0952, p<.001) but significantly less acceptances in Exp-MS ($\hat{\beta}$ = -0.7324, p<.005). These results suggest asymmetries with respect to the sensitivity to OA and OD: OA is more tolerated than OD in MA questions, but less tolerated than OD in MS questions.

[6] In Exp-MS, the four target items (A1-A4) and two fillers were randomized into 10 lists. I recruited 100 participants on MTurk. All the participants were required to have completed 90 HITs with the number of HITs approved no less than 50. All IP address were tied to the U.S. Based on the filler accuracy (100%), native language (English), and the completion rate (fully completed exactly one HIT), I kept 88 participants out of 100.

[7] A1 and A4 were coded as -1 and 1, respectively. Fomula: glmer(Choice $\sim$ Item + (1|WorkerId), data = mydata, family = binomial (link="logit"), verbose = TRUE)

What causes these asymmetries? One might argue that OD is less tolerated than OA in MA questions because OD even does not satisfy Completeness. But, the participants in Exp-MA who were tolerant of incompleteness (viz., the participants who accepted both MS and MA, N=28) rejected OD significantly more than OA (binomial test: 89%, p<.05). In other words, OD is consistently less tolerated than OA in MA questions, regardless of whether Completeness is concerned. Therefore, the asymmetries of FA-sensitivity vary by question-type, not result from Completeness.

5. Explaining the asymmetries of FA-sensitivity: Principle of Tolerance

I propose that a false answer is tolerated if it is not misleading: each response brings an update to the answer space, such as removing the incompatible answers or adding the entailed answers. If the questioner accepts this response, he would take any MaxI answer of the new answer space as a resolution and make decisions accordingly. If none of these MaxI answers leads to an improper decision (such as making the questioner go somewhere for gas where however has no gas), this response could be tolerated, even if it contains false information. For a MaxI answer not leading to an improper decision, it has to provide enough information that a complete true answer would do.

Formally, I propose that an answer is tolerated iff it satisfies the *Principle of Tolerance*, as defined in (36). In the following, I elaborate how this principle captures the asymmetries of FA-sensitivity.

(36) **Principle of Tolerance** An answer p is tolerated iff accepting p yields an answer space s.t. every MaxI member of this answer space entails a complete/MaxI true answer.

Figure 3 illustrates **the asymmetry of FA-sensitivity in MA questions**. The letter f stands for the predicate *made the swimming team* and $a/b/c$ for relevant individuals (e.g., $f(a) = \lambda w.\ a$ *made*$_w$ *the swimming team*). Arrows indicate entailments. The shaded answers are the ones that entail the bottom-left answer $f(a)$. Underlining marks the MaxI answers of each answer space.

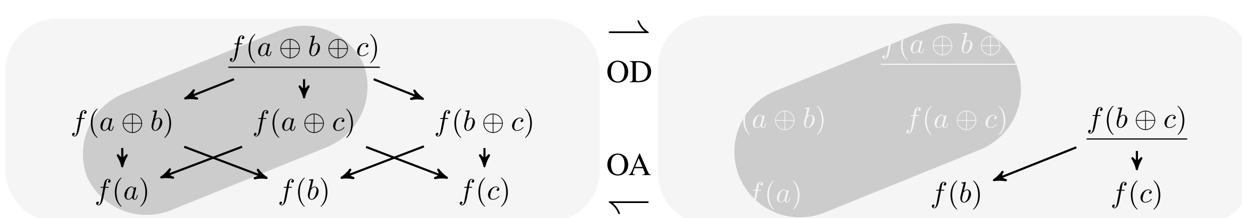

Figure 3: OA and OD in "who made the swimming team?"

OA is tolerated. Assume that only the unshaded answers are true, then the question has a unique MaxI true answer $f(b\oplus c)$. Due to the entailment relation among the answers, overly affirming $f(a)$ brings in all the shaded answers. The unique MaxI member of the updated answer space, namely

$f(a \oplus b \oplus c)$, entails the unique MaxI true answer $f(b \oplus c)$. In contrast, **OD is not tolerated**. Assume that all the present answers are true, then the question has a unique MaxI true answer $f(a \oplus b \oplus c)$. Due to the entailment relation among the answers, overly denying $f(a)$ subsequently excludes all the shaded answers. The MaxI member of the updated answer space, namely $f(b \oplus c)$, does not entail the unique MaxI true answer $f(a \oplus b \oplus c)$.

Figure 4 illustrates **the asymmetry of FA-sensitivity in MS-questions**. The letter f stands for the predicate *serve on the committee* and $a/b/c$ for relevant individuals. Due to the non-monotonicity of the local O-operator (see section 2.2), all the present answers are semantically independent; hence, the bottom-left answer is only entailed by itself (shaded).

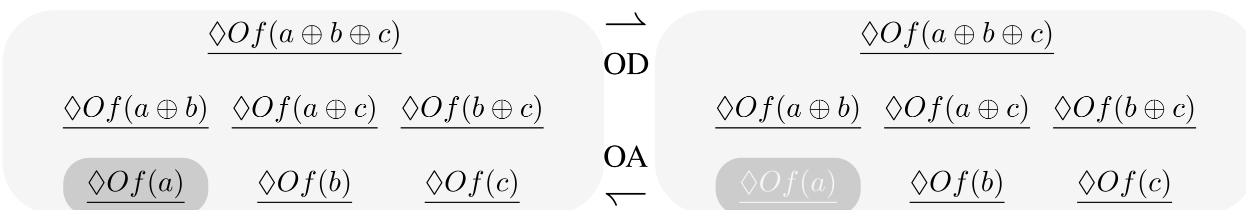

Figure 4: OA and OD in "who can serve on the committee?"

OA is not tolerated. Assume that only the unshaded answers are true, then all of the unshaded answers are MaxI true answers. Overly affirming $\Diamond Of(a)$ only adds $\Diamond Of(a)$ itself to the answer space. $\Diamond Of(a)$ is a MaxI member in the updated answer space, but it does not entail any MaxI true answers. In contrast, **OD is tolerated**. Assume that all the present answers are true, then all of them are MaxI true answers. Overly denying $\Diamond Of(a)$ only removes $\Diamond Of(a)$ itself from the answer space. All the remaining answers are MaxI members of the updated answer space, and each of them entails a MaxI true answer, namely itself.

References

Chierchia, Gennaro, Danny Fox and Benjamin Spector (2013). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. Maienborn, K. von Heusinger, and P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 3, pp. 2297–2331. Mouton de Gruyter.

Cremers, Alexandre and Emmanuel Chemla (2016). A psycholinguistic study of the exhaustive readings of embedded questions. *Journal of Semantics 33*(1), 49–85.

Dayal, Veneeta (1996). *Locality in WH quantification: Questions and relative clauses in Hindi*, Volume 62. Boom Koninklijke Uitgevers.

Fox, Danny (2013). Mention some readings of questions. MIT class notes.

Fox, Danny and Benjamin Spector (to appear). Economy and embedded exhaustification. *Natural Language Semantics*.

George, Benjamin R. (2011). *Question embedding and the semantics of answers*. Ph. D. thesis, University of California Los Angeles.

George, Benjamin R. (2013). Knowing-"wh", mention-some readings, and non-reducibility. *Thought: A Journal of Philosophy 2*(2), 166–177.

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

Guerzoni, Elena and Yael Sharvit (2007). A question of strength: on NPIs in interrogative clauses. *Linguistics and Philosophy 30*(3), 361–391.

Hamblin, Charles L. (1973). Questions in Montague English. *Foundations of language 10*(1), 41–53.

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

Karttunen, Lauri (1971). Some observations on factivity. *Research on Language & Social Interaction 4*(1), 55–69.

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

Klinedinst, Nathan and Daniel Rothschild (2011). Exhaustivity in questions with non-factives. *Semantics and Pragmatics 4*(2), 1–23.

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

Roelofsen, Floris, Nadine Theiler and Maria Aloni (2014). Embedded interrogatives: The role of false answers. In *Presentation at Questions in Discourse workshop*.

Romoli, Jacopo (2015). The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics 32*(2), 173–219.

Romoli, Jacopo and Florian Schwarz (2015). An experimental comparison between presuppositions and indirect scalar implicatures. In *Experimental perspectives on presuppositions*, pp. 215–240. Springer.

Schulz, Katrin and Robert van Rooij (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy 29*(2), 205–250.

Spector, Benjamin (2005). Exhaustive interpretations: What to say and what not to say. In *Unpublished paper presented at the LSA workshop on Context and Content*.

Spector, Benjamin and Paul Egré (2015). A uniform semantics for embedded interrogatives: An answer, not necessarily the answer. *Synthese 192*(6), 1729–1784.

Stalnaker, Robert (1974). Pragmatic presuppositions. In Stalnaker, Robert (Ed.), *Context and Content*, pp. 47–62. Oxford University Press.

Uegaki, Wataru (2015). *Interpreting questions under attitudes*. Ph. D. thesis, Massachusetts Institute of Technology.

Xiang, Yimei (2016). Solving the dilemma between uniqueness and mention-some. In *Proceedings of North East Linguistic Society 46*.