University of Massachusetts Occasional Papers in Linguistics

Volume 26 *University of Massachusetts Occasional Papers in Linguistics -- Vol 23: Issues in Semantics*

Article 11

2000

Embedded Questions: Some Comments on Heim 1994 and on Beck and Rullmann 1999

Yael Sharvit
University of Connecticut

Embedded Questions: Some Comments on Heim 1994 and on Beck and Rullmann 1999

Yael Sharvit

University of Connecticut

This paper argues that we need to distinguish between two variants of 'de dicto' readings of embedded constituent questions, which are referred to here as 'de dicto$_e$' and 'de dicto$_f$'. It also argues that Strong Exhaustivity is a sufficient condition for the availability of three independent readings of sentences with embedded questions – 'de re', 'de dicto$_e$' and 'de dicto$_f$' – in accordance with what is implied in Heim 1994, but contrary to what is implied in Rullmann and Beck 1998, and Beck and Rullman 1999 (henceforth: B&R). This claim is based on intuitions regarding sentences involving various question-embedding verbs.

1. The ANS-operator and Strong Exhaustivity

Heim 1994 shows that the predictions made by Groenendijk and Stokhof 1982, 1984 (henceforth: G&S) regarding constituent questions can be carried over to the H/K system, provided that we enrich this system with a proper semantics of question-embedding verbs. We adopt the following semantics for the verb *know* ('Q' is a question-intension – i.e., a function from worlds to sets of possible answers):

(1) [[*know*]](Q)(a)(w)=1 iff for all worlds w'∈Dox(a)(w), ANS(Q)(w') = ANS(Q)(w).

Dox(a)(w) is the set of a's "doxastic alternatives" in w – the set of worlds w' compatible with what a believes in w. In other words, [[*know*]](Q)(a)(w) is true iff a believes in world w the proposition λw'[ANS(Q)(w') = ANS(Q)(w)].

This semantics of *know* makes crucial reference to 'ANS' – an Answerhood operation, given below:

UMOP23: Issues in Semantics, Kiyomi Kusumoto and Elisabeth Villalta (eds.),209-224.

(2) ANS(Q)(w) is the intersection of the set of propositions $\{p:w\in p \ \& \ p\in Q(w)\}$.[1,2]

Here is how (1) and (2) work. For example, if Bill, Sue and John are the individuals in our domain, **[[*who ran*]]**(w) is {λw'[John ran in w'], λw'[Bill ran in w'], λw'[Sue ran in w']}. If Bill and Sue are the ones who ran in w and John did not run in w, then the following hold:

(3) ANS([[*who ran*]])(w) = the intersection of $\{p:w\in p \ \& \ \exists x[p=\lambda w'[x \text{ ran in } w']]\}$ = λw'[Bill and Sue ran in w']

To know in w who ran is to believe in w the following proposition:

(4) λw'[ANS([[*who ran*]])(w') = λw[Bill and Sue ran in w]]

Heim shows that with this semantics of *know* we can account for Strong Exhaustivity, thus circumventing one of the problems which G&S point out for the H/K semantics. "Strong Exhaustivity" refers to the observation that, for example, *John knows who ran* entails not only that John knows that Bill and Sue ran, if they indeed ran (Weak Exhaustivity), but also that he knows who didn't run. The semantics in (1) predicts this, because one cannot believe the proposition in (4) without believing that John didn't run (assuming that the domain consists of John, Bill and Sue and everyone is acquainted with everyone).

G&S point out that the original H/K approach captures only Weak Exhaustivity. Phrased in terms of ANS, according to the original H/K approach, to know Q in w is to believe in w the proposition ANS(Q)(w) (and if $\{p:w\in p \ \& \ p\in Q(w)\}$ is ∅, it means to believe in w $\lambda w'[\{p:w'\in p \ \& \ p\in Q(w')\}=\emptyset]$). So it wrongly predicts that knowing in w that Bill and Sue ran qualifies as knowing in w who ran. G&S propose instead that the basic denotation of a question is a proposition, rather than a set of propositions. In particular, they propose that **[[*who ran*]]**(w) is the proposition λw'[{x:x ran in w'} = {x:x ran in w}]. World w being the way it is, to claim that John knows in w who ran is to claim that he believes in w the proposition λw'[{x:x ran in w'} = {Sue, Bill}]. But Heim's semantics of *know* allows the H/K theory to capture Strong Exhaustivity just as well. So if we adopt her semantics, this particular criticism of G&S cannot be taken as an argument against a "set-of-propositions" approach to questions.

Notice that in Heim's system the basic denotation of an interrogative sentence is not strongly exhaustive (as it is in the G&S system). Strong Exhaustivity comes from the semantics of the embedding verb (in our case, *know*). As Heim and later B&R show, there are question-embedding predicates, such as the emotive factive *surprise*, which trigger only weakly exhaustive readings, as evidenced by the acceptability of *it surprised*

[1] In view of recent work on questions (e.g., Lahiri 1991, Dayal 1996 and B&R), it may be useful to adopt a definition according to which a ANS(Q)(w) is the most informative true proposition in Q(w) (if there is one).

[2] "$w\in p$" means $w\in\{v:p(v)=1\}$.

me who was at the party but it didn't surprise me who wasn't at the party (see also Berman 1991). They use this observation to argue that G&S's analysis, according to which Strong Exhaustivity is part of the basic semantics of an interrogative sentence, is problematic. In G&S's system there is no way to account for the weakly exhaustive implications of *surprise*. Heim and B&R argue then, that the basic denotation of a question is a set of propositions in the H/K style, and that some question-embedding verbs (e.g., *know*) give rise to strongly exhaustive readings, whereas other question-embedding verbs (e.g., *surprise*) do not.

G&S also point out that the original H/K theory does not capture the 'de re'/'de dicto' distinction. Heim's rendition of *know* takes care of this too, as we will soon see. Moreover, we will see that the strongly exhaustive nature of *know* predicts it to license two kinds of 'de dicto' readings. But Heim does not say anything about the (un)availability of (any kind of) 'de dicto' readings in sentences with question-embedding verbs that are not strongly exhaustive. We will see that all strongly exhaustive question-embedding verbs give rise to three readings – 'de re', 'de dicto$_e$' and 'de dicto$_f$' – but not all verbs that lack Strong Exhaustivity also give rise to these three readings.

2. Three Types of Question-intensions

Consider the sentence in (5), whose main predicate is the strongly exhaustive *know*. (5) has three distinct readings, corresponding to three different interpretations of the *wh*-phrase, as shown in (6)-(8) (I assume, for current purposes only, that STUDENTS and LEFT admit singular as well as plural individuals in their extensions):

(5) John knows which students left.

(6) 'de re'
John knows in w $\lambda w' \lambda p \exists y_e [y \in \text{STUDENTS}(w) \ \& \ p = \lambda w'' [y \in \text{LEFT}(w'')]]$

(7) 'de dicto$_e$'
John knows in w $\lambda w' \lambda p \exists y_e [y \in \text{STUDENTS}(w') \ \& \ p = \lambda w'' [y \in \text{LEFT}(w'')]]$

(8) 'de dicto$_f$'
John knows in w $\lambda w' \lambda p \exists f_{<s,e>} [f \in C \ \& \ \forall w'' \in \text{Dom}(f) [f(w'') \in \text{STUDENTS}(w'')] \ \& \ p = \lambda w'' [f(w'') \in \text{LEFT}(w'')]]$

In (6), the existential quantifier inside the question intension binds a variable over individuals, and the world argument of STUDENTS is the same as the world argument of *know*. A good answer to the embedded question in (6) provides a list of individuals who fit the description "students who left" in the world relative to which *know* is interpreted. If Bill and Sue are the students who actually left, it follows from (6) that John knows that Bill and Sue left, but it does not follow that he knows that they are students.

In (7), the existential quantifier inside the question intension binds a variable over individuals, and the world argument of STUDENTS is bound by a lambda-operator, and is crucially not co-indexed with the world argument of *know*. The difference between the question embedded in (6) and the one embedded in (7) is subtle. The two embedded question-intensions, when applied to w, yield the same set of propositions. A good

answer in w to both embedded questions lists the individuals who fit the description "students who left" in w. The difference between the two readings becomes apparent when the question is a complement of a question-embedding verb such as *know*. In (6), John need not be aware of the student status of the student-leavers, whereas in (7) he is. If Bill and Sue are the students who left, it follows from (7) that John knows that Bill and Sue left and that he knows that they are students. As G&S point out, in the original H/K system to know which students left is to believe that Bill and Sue left (according to this scenario). This says nothing about whether or not the knower believes that Bill and Sue are students. In Heim's system, this problem does not arise. As we will see below, since Heim's semantics of *know* refers to the value of ANS in the subject's doxastic alternatives, we correctly predict (6) and (7) to have different truth conditions.

In (8), the existential quantifier inside the question intension binds a variable over functions from worlds to individuals ('individual concepts'), which are members of the set C of salient functions. A good answer to the embedded question in (8) provides a list of salient descriptions (e.g., *the best student and the tallest student*). If Bill, the best student, and Sue, the tallest student, are the students who left, it follows from (8) that John knows that the best student and the tallest student left, but it does not follow that he knows that Bill and Sue left.[3] We will make the following assumptions. First, when quantifying over individual concepts, we consider only functions that are not constant – that is to say, do not have the same value for every world in their domain (this rules out descriptions such as "the individual John"). Secondly, for any individual x, world w' and property P, if x is in P(w'), then there is at least one salient individual concept f such that f(w') is x and f=λv[P(v) & ...] (this means that if Bill is a student in w', there is a salient description of the form "the... student..." which fits him in w'). Thirdly, for any individual concepts f1 and f2, f1+f2 is the function g such that for every world w, g(w) is f1(w)+f2(w).[4]

The three readings in (6)-(8) are independent of each other. That is to say, none of them entails any of the others, as we will now show. Starting with the 'de re' reading, it is easy to show that this reading does not entail either the 'de dicto$_e$' or the 'de dicto$_f$' reading. Suppose that in world w, Bill (who happens to be the best student) and Sue (who happens to be the tallest student) are all the students who left, and suppose that John knows that out of a certain group of people (who happen to be the students in w), only Bill and Sue left, but he does not know that they are students. In addition, *the best student* (the BEST STUDENT function – the function which maps every world to the individual who is the best student in it, if there is one) and *the tallest student* (the TALLEST STUDENT function) are the salient descriptions that fit Bill and Sue, respectively, in w. There is some sense in which *John knows which students left* is true in w, and some sense

[3] Notice that if 'C' bears a world index, then there is yet another source for a de-re/de-dicto ambiguity, but we will not be concerned with it here. Notice also that Heim 1994 discusses only 'de re' and 'de dicto$_e$' readings, but 'de dicto$_f$' readings have also been discussed in the literature (though not under this name), mainly in connection with "functional" questions (e.g., *which woman does every man love? His mother*; see Engdahl 1986 and G&S), which elicit answers that provide definite descriptions.

[4] These assumptions, of course, are not necessary, but they allow us to simplify the discussion of the 'de dicto$_f$' readings.

in which it is not. This is captured by the fact that the proposition corresponding to the 'de re' reading of *John knows which students left* comes out true in w but the one corresponding to its 'de dicto$_e$' reading and the one corresponding to its 'de dicto$_f$' reading do not. To see this, first note that (9), (10) and (11) give the values of ANS(*which students left ('de re')* according to w)(w), ANS(*which students left ('de dicto$_e$')* according to w)(w), and ANS(*which students left ('de dicto$_f$')* according to w)(w), respectively (from now on we omit the restrictions 'f∈C' and 'w∈Dom(f)' from the 'de dicto$_f$' representations):

(9) $\lambda w''[\text{Bill+Sue} \in \text{LEFT}(w'')] =$
$\text{ANS}(\lambda w'\lambda p \exists y_e[y \in \text{STUDENTS}(w) \ \& \ p = \lambda w''[y \in \text{LEFT}(w'')]])(w)$

(10) $\lambda w''[\text{Bill+Sue} \in \text{LEFT}(w'')] =$
$\text{ANS}(\lambda w'\lambda p \exists y_e[y \in \text{STUDENTS}(w') \ \& \ p = \lambda w''[y \in \text{LEFT}(w'')]])(w)$

(11) $\lambda w''[\text{BEST-STUDENT+TALLEST-STUDENT}(w'') \in \text{LEFT}(w'')] =$
$\text{ANS}(\lambda w'\lambda p \exists f_{<s,e>}[\forall w''[f(w'') \in \text{STUDENTS}(w'')] \ \&$
$p = \lambda w''[f(w'') \in \text{LEFT}(w'')]])(w)$

According to the semantics of *know* we are assuming, for *John knows which students left ('de re')* to be true in w it must be the case that the proposition λw'[Bill+Sue∈LEFT(w')] is the intersection of {p:v∈p & ∃y[y∈STUDENTS(w) & p=λw'[y∈LEFT(w')]]}, where v is any member of Dox(John)(w). According to the scenario described above this is indeed the case. On the other hand, in order for *John knows which students left ('de dicto$_e$')* to be true in w, the same proposition must also be the intersection of {p:v∈p & ∃y[y∈STUDENTS(v) & p=λw'[y∈LEFT(w')]]}, where v is any member of Dox(John)(w). But this is not so, because Bill and Sue are students in w, but not in every world in Dox(John)(w). Likewise, for *John knows which students left ('de dicto$_f$')* to be true in w it must be the case that the proposition λw'[BEST-STUDENT+TALLEST-STUDENT(w') ∈ LEFT(w')] is true in any world v in Dox(John)(w), but this is not so.

It can also be shown that neither the 'de dicto$_e$' reading of *John knows which students left* nor its 'de dicto$_f$' reading entails its 'de re' reading. Suppose that in world w, Bill and Sue – the best student and the tallest student respectively – are all the students who left, and that *the best student* and *the tallest student* are the only relevant descriptions that fit them. Suppose also that John is aware of all of this. Suppose further that Sally is a student who didn't leave, but John mistakenly believes that she is not a student, and that she did leave. There is some sense in which *John knows which students left* is true in w, and some sense in which it is not. This is captured by the fact that both the 'de dicto$_e$' and 'de dicto$_f$' readings of *John knows which students left* come out true in w, but its 'de re' reading does not. Here is why.

Consider a typical world v in Dox(John)(w). According to our scenario, the proposition λw'[Bill+Sue∈LEFT(w')] – the value of ANS in w for *which students left ('de dicto$_e$')* – is the intersection of {p:v∈p & ∃y[y∈STUDENTS(v) & p=λw'[y∈LEFT(w')]]}. Likewise, the proposition λw'[BEST-STUDENT+TALLEST-

STUDENT(w') ∈ LEFT(w')] – the value of ANS in w for *which students left ('de dicto$_f$')* – is the intersection of {p:v∈p & ∃f[∀w'[f(w') ∈ STUDENTS(w')] & p=λw'[f(w')∈LEFT(w')]]}. But λw'[Bill+Sue ∈ LEFT(w')] – the value of ANS in w for *which students left ('de re')* – is not the intersection of {p:v∈p & ∃y[y∈STUDENTS(w) & p=λw'[y∈LEFT(w')]]} (because Sally also left in v, and she is a student in w). Since this holds for every world in Dox(John)(w), both the 'de dicto$_e$' and the 'de dicto$_f$' readings come out true in w, but the 'de re' reading does not.

We have seen that the 'de re' and 'de dicto$_e$' readings of *John knows which students left* are independent of each other. This is so despite the fact that in the world relative to which *know* is interpreted, ANS yields the same output when applied to the 'de re' and the 'de dicto$_e$' question-intensions. The distinctness of the 'de re' and 'de dicto$_e$' readings under *know* results from the assumption that the semantics of *know* refers to the value of the function ANS(relevant question-intension) in the subject's doxastic alternatives. In these worlds ANS may have different values when applied to the 'de re' and the 'de dicto$_e$' question-intensions. In other words, for any two worlds w and w', ANS(question-intension 'de re' according to w)(w) may differ from ANS(question-intension 'de dicto$_e$' according to w)(w').

It turns out that every strongly exhaustive verb V has the property that its semantics refers to the value of ANS(relevant question-intension) in the doxastic alternatives (or, more generally, in the worlds compatible with the attitude associated with V) of the bearer of the attitude. To see why, let us informally define '¬Q' as follow. If Q is a question intension, ¬Q is defined only if Q is of the form λwλp∃x[x∈A & p=λw'[x∈P(w')]] (where x is an individual, A – a set of individuals, and P – a function f whose domain is the set of possible worlds and for every v∈Dom(f), f(v) is the set of individuals that are in P(v)). Wherever defined, ¬Q is λwλp∃x[x∈A & p=λw'[x∉P(w')]]. For any Q, if ¬Q is defined, then for any w' the propositions λw[ANS(Q)(w) = ANS(Q)(w')] and λw[ANS(¬Q)(w) = ANS(¬Q)(w')] are the same proposition.[5] A question-embedding V qualifies as strongly exhaustive iff for any Q such that ¬Q is defined, [[V]](Q)(x_1,…,x_n)(w) entails [[V]](¬Q)(x_1,…,x_n)(w). Therefore, in order for V to qualify as strongly exhaustive its semantics must express an attitude towards a proposition of the form "λw[ANS(Q)(w) = …]". This means that the semantics of every strongly exhaustive question-embedding verb refers to the value of the function ANS(embedded question-intension) in the doxastic alternatives (or in the worlds compatible with the relevant attitude) of the bearer of the attitude. Since in these worlds ANS has different values when applied to the 'de re' and the 'de dicto$_e$' intensions (according to the actual world), it follows that Strong Exhaustivity is a sufficient condition for a verb's ability to distinguish between 'de re' and 'de dicto$_e$' readings.

We also saw that 'de re' readings are distinct from 'de dicto$_f$' readings. This is not unexpected, given that ANS applied to the two question-intensions may have different

[5] If they weren't the same proposition, there would be an individual b in A and a world w" such that b is in P(w") and b is not in P(w").

values in any world (even in the world relative to which the embedding verb is interpreted – compare (9) with (11)). So one might expect the two 'de dicto' readings to be distinct too, because ANS applied to the 'de dicto$_e$' and the 'de dicto$_f$' intensions may also have different values in any world (compare (10) with (11)). However, given that both 'de dicto' readings of *John knows which students left* imply awareness on the part of John of the description contributed by the *wh*-phrase, we need to ask whether there is an entailment relation between these two readings. We will now see that there is no such entailment relation.

The 'de dicto$_f$' reading of *John knows which students left* does not entail its 'de dicto$_e$' reading. Suppose, as before, that in world w Bill and Sue, the best student and the tallest student respectively, are all the students who left (and these descriptions are the only relevant ones). John is aware that the best student and the tallest student left (and that they are the only students who left and that these are the relevant descriptions), he knows who the tallest student is, but he does not know who the best student is (in fact, he believes that Bill is not a student at all). There is some sense in which *John knows which students left* is true in w, and some sense in which it is not. This is explained by the fact that the 'de dicto$_f$' reading comes out true in w, but the 'de dicto$_e$' reading does not. Let us see why.

Consider a typical world v in Dox(John)(w). The proposition λw'[BEST-STUDENT+TALLEST-STUDENT(w') ∈ LEFT(w')] (i.e., the value of ANS(relevant question-intension) in w) is the intersection of {p:v∈p & ∃f[∀w'[f(w') ∈ STUDENTS(w')] & p=λw'[f(w')∈LEFT(w')]]}. Since this holds for all the worlds in Dox(John)(w), the 'de dicto$_f$' reading is true in w. However, for the 'de dicto$_e$' reading to be true in w it must be the case that the relevant output of ANS in w, namely, λw'[Bill+Sue∈LEFT(w')], is the relevant value of ANS in any world in Dox(John)(w). However, this proposition is not the intersection of {p:v∈p & ∃y[y∈STUDENTS(v) & p=λw[y∈LEFT(w)]]} (where v is any world in Dox(John)(w)), because John does not believe that Bill is a student.

On the other hand, the 'de dicto$_e$' reading of *John knows which students left* does not entail its 'de dicto$_f$' reading. Suppose that in world w Bill and Sue, the best student and tallest student respectively, are all the students who left in w, and John, who knows that Bill and Sue are the students who left and that Sue is the tallest student, does not believe that Bill is the best student, so he does not believe that the best student left. There is some sense in which *John knows which students left* is true, and some sense in which it is not. This is captured by the fact that the 'de dicto$_e$' reading of *John knows which students left* comes out true in w, but its 'de dicto$_f$' reading does not. Let us see why.

Consider a typical world v in Dox(John)(w). The proposition λw'[Bill+Sue∈LEFT(w')] – the value of ANS(relevant question-intension) in w – is the intersection of {p:v∈p & ∃y[y∈STUDENTS(v) & p=λw'[y∈LEFT(w')]]}, and therefore the 'de dicto$_e$' reading of *John knows which students left* is true in w. However, the proposition λw'[BEST-STUDENT+TALLEST-STUDENT(w')∈LEFT(w')] – the value

of ANS(relevant question-intension) in w – is not the intersection of {p:v∈p & ∃f[∀w[f(w)∈STUDENTS(w)] & p=λw[f(w)∈LEFT(w)]]}, because John does not believe that Bill is the best student. An example that may illustrate this effect better is the following, where the embedded question is a "functional" question (in the sense of Engdahl 1986 and G&S):

(12) John knows which student every professor invited.

Suppose every professor invited his favorite student. Suppose further, that John is aware of the actual professor-student pairs that are members of {<x,y>:x invited y}, and he is aware that the inviters are professors and that the invitees are students, but he is not aware of the fact that for each inviter, the invitee is his favorite student. I think it is pretty clear that (12) is true in the 'de dicto$_e$' sense and false in the 'de dicto$_f$' sense. John is not aware of one crucial thing – the description that fits each invitee with respect to his/her inviter (or the relation that holds between each inviter and his/her invitee).[6]

The 'de dicto$_f$' reading seems to be what we call a 'de dicto' reading in declarative sentences with propositional attitude verbs (e.g., *Mary expects to meet the finest student*, where Mary need not have an expectation regarding a particular individual), whereas the 'de dicto$_e$' reading seems to arise only in embedded questions. This is perhaps why the distinctness of the 'de dicto$_e$' and 'de dicto$_f$' readings in questions becomes even more apparent when we look at an embedded question which itself contains a propositional attitude verb, as in the following example:

(13) John knows which student Mary expects to meet.

Note that (13) has, among its many readings, the following three:

(14) 'de dicto$_e$'
John knows in w λw'λp∃y[y∈STUDENT(w') &
p=λw"[Mary expects to meet y in w"]]

(15) 'de dicto$_f$1'
John knows in w λw'λp∃f[∀w"[f(w")∈STUDENTS(w")] &
p=λw"[Mary expects to meet f(w') in w"]]

(16) 'de dicto$_f$2'
John knows in w λw'λp∃f[∀w"[f(w")∈STUDENT(w")] &
p=λw"[Mary expects to meet f(w") in w"]]

The 'de dicto$_e$' reading is the one where Mary has an expectation regarding a particular individual, who John knows to be a student, and he knows who that individual is. The 'de dicto$_f$1' reading is the one where Mary has an expectation regarding a particular individual who happens to be, say, the finest student, and John knows this, but he does

[6] Here, the 'de dicto$_f$' sense involves quantification over <s,<e,e>>-functions:
(i) λp∃g[∀<x,w>[g(w)(x)∈STUDENTS(w)] &
p=λw[every professor(w) y invited g(w)(y) in w]]

not necessarily know who the finest student is. The 'de dicto$_f$2' reading is the one where Mary does not necessarily have an expectation regarding a particular individual (but she says to herself something like: "I will probably meet the finest student").

We have seen that while the 'de re' and 'de dicto$_f$' readings arise independently of the embedding predicate, the 'de dicto$_e$' reading arises in embedded contexts. For example, it arises under *know* as a reading distinct from both the 'de re' and the 'de dicto$_f$' readings. Moreover, since the semantics of *know* is strongly exhaustive, it crucially refers to the value of ANS(relevant question-intension)(w') where w' is any doxastic alternative of the subject, and this is the reason why *know* gives rise to these three independent readings, rather than just to the 'de re' and the 'de dicto$_f$' readings. We predict, then, that any question-embedding verb whose semantics refers to the output of ANS in the worlds compatible with the attitude associated with the verb, be it strongly exhaustive or not, will give rise to these three readings.

This prediction is borne out. For example, the verb *wonder* is not strongly exhaustive,[7] as evidenced by the acceptability of *John wondered who left but he did not wonder who didn't leave*. And yet, *wonder* gives rise to the two 'de dicto' readings (in addition to the 'de re' reading). I propose that **[[*wonder*]]**(Q)(x) is true in w iff (a) x believes in w λw'[{p:w'$\in$p & p$\in$Q(w')} $\neq \varnothing$]; and (b) x wants in w λw'[x believes in w' λw"[ANS(Q)(w")=ANS(Q)(w')]. This semantics does not qualify as strongly exhaustive, because believing λw'[{p:w'$\in$p & p$\in$Q(w')} $\neq \varnothing$] does not entail believing λw'[{p:w'$\in$p & p$\in(\neg$Q)(w')} $\neq \varnothing$] (for any Q such that $\neg$Q is defined). So we correctly predict that one can wonder who left without wondering who didn't leave.

But this proposed semantics still predicts *wonder* to give rise to three distinct readings – 'de re', 'de dicto$_e$' and 'de dicto$_f$' – because its truth conditions refer to the value of ANS in the worlds compatible with the relevant attitude (in this case, the worlds compatible with the subject's desires). In these worlds, ANS may yield different outputs when applied to the 'de re', 'de dicto$_e$' and 'de dicto$_f$' intensions. The distinctness of these three readings here is evident. According to the 'de re' reading of *John wonders which students left*, John wonders about actual students, but according to both 'de dicto' readings, he wonders about students in his belief worlds (who need not be actual students). In addition, John may want to know the identities of the students who left, without wanting to be aware of the descriptions that fit them (so he may wonder *which students left ('de dicto$_e$')* without wondering *which students left ('de dicto$_f$')*). The opposite is also true: John may want to be aware of the descriptions that fit the students who left, without wanting to be aware of their identities.

In the next section, we will look at other question-embedding verbs whose semantic properties differ from those of *know* and we will see that they confirm the generalization we have arrived at. Verbs whose semantics refers to ANS in the worlds compatible with the relevant attitude (and this includes all but not only strongly exhaustive verbs) give rise to three independent readings. We will also see that verbs that

[7] Thanks to Klaus Abels, Gidi Avrahami, and Luisa Martí for comments and judgments.

do not have this property give rise to an independent 'de dicto$_f$' reading, but do not distinguish between a 'de re' and a 'de dicto$_e$' reading.

3. *Agree on* and *Surprise*

As mentioned in Section 1, B&R agree with Heim that Strong Exhaustivity is not part of the basic denotation of an interrogative sentence.[8] However, they disagree with Heim regarding the connection between Strong Exhaustivity and the availability of 'de dicto$_e$' readings (which is the only 'de dicto' type she discusses).

As we have seen, in Heim's system the 'de dicto$_e$' reading of a question embedded under *know* comes about, just like Strong Exhaustivity, via the semantics of *know* – a semantics which implies that the subject of *know* believes that the output of ANS is the actual output of ANS. Moreover, due to this semantics, Strong Exhaustivity is predicted to correlate with the availability of 'de dicto$_e$' readings. B&R argue that this is not a good prediction. They claim that 'de dicto$_e$' readings have nothing to do with Strong Exhaustivity. One of their arguments goes like this. Since Heim's semantics for embedded questions crucially relies on the value of ANS in the actual world, it can only work for verbs that look for true answers such as *know*. For verbs such as *agree on*, the true answer in the actual world is irrelevant. For example, if John and Mary falsely believe of Sue that she is a student and that she is the student who left, then we can truthfully utter the sentence *John and Mary agree on which student left* (or *John agrees with Mary on which students left*) but we cannot account for its meaning by assuming that the semantics of *agree on* appeals to the true answer in the actual world (because in the actual world Sue is not a student). Therefore, the grammar must generate 'de dicto' readings independently of Strong Exhaustivity. Furthermore, the 'de dicto' reading is part of the basic question denotation.

I disagree with B&R regarding this point. I think that 'de dicto$_f$' readings are part of the basic question denotation (and as such, are completely independent of the semantics of the embedding verb), but the availability of 'de dicto$_e$' readings depends on the embedding verb. B&R seem to associate Strong Exhaustivity with the requirement that the subject of the question-embedding verb believe that the output of ANS is the output of ANS in the actual world. I understand Strong Exhaustivity differently. The semantics we have used for *know* is strongly exhaustive according to the informal definition given in Section 2, and so is the following proposed semantics for *agree on*:

(17) [[*agree-on*]](Q)(b)(a)(w)=1 iff there is a proposition p such that for all w'∈Dox(a)(w), for all w"∈Dox(b)(w), p=ANS(Q)(w')=ANS(Q)(w").

This semantics predicts that if, for example, John and Mary agree on who left (in the 'de re' sense of *who*), they also agree on who didn't leave.[9] Moreover, being strongly

[8] I thank Sigrid Beck for some important comments regarding this topic.

[9] This is compatible with the semantics B&R propose. I have to add, like B&R, that (17) makes correct predictions as long as we ignore complications that arise due to quantificational variability effects.

exhaustive, the proposed semantics of *agree on* refers to ANS applied to the embedded question intension in the doxastic alternatives of the two individual-arguments of the verb. Therefore, we predict this verb to give rise to three independent readings – 'de re' 'de dicto$_e$' and 'de dicto$_f$'.

As we will soon see, by contrast, a verb like the emotive factive *surprise* does not distinguish between a 'de re' and a 'de dicto$_e$' reading. This is because the semantics of *surprise* does not refer to the output of ANS in the worlds compatible with the relevant attitude (but it does appeal to the actual output of ANS). Let us start with *agree on*. Consider the following sentence:

(18) John agrees with Mary on which students left.

The readers can easily see that the 'de re' reading of (18) does not entail either its 'de dicto$_e$' or its 'de dicto$_f$' readings, and that the 'de dicto$_f$' reading does not entail the 'de re' reading. Let us show here that the 'de dicto$_e$' reading does not entail the 'de re' reading, and that the two 'de dicto' readings to do not entail each other.

Suppose that in world w John and Mary falsely believe that Bill and Sue are students and they also falsely believe of them that they left (and that they are all the students who left). Suppose further that John falsely believes of Sally and Bob that they are not students and he also falsely believes of them that they did leave. Mary, however, knows that Sally and Bob are students who didn't leave. In this case, (18) is true in w in its 'de dicto$_e$' sense, and false in its 'de re' sense. This is because the proposition λw[Bill+Sue$\in$LEFT(w)] is the intersection of {p:v$\in$p & $\exists$y[y$\in$STUDENTS(v) & p=λw[y$\in$LEFT(w)]]}, where v is any member of Dox(John)(w); and it is also the intersection of {p:v'$\in$p & $\exists$y[y$\in$STUDENTS(v') & p=λw[y$\in$LEFT(w)]]}, where v' is any member of Dox(Mary)(w). However, there is no proposition which is simultaneously the intersection of {p:v$\in$p & $\exists$y[y$\in$STUDENTS(w) & p=$\lambda w'$[y$\in$LEFT(w')]]} and of {p:v'$\in$p & $\exists$y[y$\in$STUDENTS(w) & p=$\lambda w'$[y$\in$LEFT(w')]]}, (where v is any member of Dox(John)(w) and v' is any member of Dox(Mary)(w)). This is because Sally and Bob are students who didn't leave in w and in all the members of Dox(Mary)(w), but they did leave in all the members of Dox(John)(w).

Now let us show that the two 'de dicto' readings do not entail each other. Suppose that in world w Mary and John falsely believe two things: that Bill and Sue are students, and that they are all the students who left. Suppose further, that John also believes that Bill is the best student and Sue is the tallest student, but Mary does not. In this case, (18) is true in w in its 'de dicto$_e$' sense and false in its 'de dicto$_f$' sense. This is because the proposition $\lambda w'$[Bill+Sue$\in$LEFT(w')] is the intersection of {p:v$\in$p & $\exists$y[y$\in$STUDENTS(v) & p=$\lambda w'$[y$\in$LEFT(w')]]}, where v is any member of Dox(John)(w); and it is also the intersection of {p:v'$\in$p & $\exists$y[y$\in$STUDENTS(v') & p=$\lambda w'$[y$\in$LEFT(w')]]}, where v' is any member of Dox(Mary)(w). This applies to any member of Dox(John)(w) and any member of Dox(Mary)(w), and therefore, the 'de dicto$_e$' reading is true in w.

However, in order for the 'de dicto$_f$' reading to be true in w, there must be a single proposition which is simultaneously the intersection of {p:v∈p & ∃f[∀w[f(w) ∈ STUDENTS(w)] & p = λw[f(w) ∈ LEFT(w)]]} and of {p:v'∈p & ∃f[∀w[f(w) ∈ STUDENTS(w)] & p = λw[f(w) ∈ LEFT(w)]]} (where v is any world in Dox(John)(w) and v' is any world in Dox(Mary)(w)). But there is no such proposition. In particular, the proposition λw'[BEST-STUDENT+TALLEST-STUDENT(w') ∈ LEFT(w')] is true in any world in Dox(John)(w), but the same does not hold for Dox(Mary)(w). Here too, a "functional" question may bring the distinction out more clearly. For example, *John agrees with Mary on agree which student every professor invited* is false in the 'de dicto$_f$' sense, if John thinks that every professor invited his favorite student, but Mary does not, and yet they have the same professor/student pairs in mind.

Likewise, the 'de dicto$_f$' reading of (18) does not entail its 'de dicto$_e$' reading. Consider a scenario in w where John and Mary falsely believe that the actual professors are students. They both believe that the best student left and the tallest student left (and that no other students left), but they have no beliefs regarding the identities of the best student and the tallest student (we are assuming, as usual, that *the best student* and *the tallest student* are the only relevant descriptions, from John and Mary's points of view). The 'de dicto$_f$' reading comes out true in w but the 'de dicto$_e$' reading does not. The proposition λw[BEST-STUDENT+TALLEST-STUDENT(w) ∈ LEFT(w)] is the intersection of {p:v∈p & ∃f[∀w[f(w) ∈ STUDENTS(w)] & p = λw[f(w) ∈ LEFT(w)]]} and of {p:v'∈p & ∃f[∀w[f(w) ∈ STUDENTS(w)] & p = λw[f(w) ∈ LEFT(w)]]} (where v is any world in Dox(John)(w) and v' is any world in Dox(Mary)(w)). Therefore, the 'de dicto$_f$' reading comes out true in w. However, there is no proposition which is simultaneously the intersection of {p:v∈p & ∃y[y∈STUDENTS(v) & p=λw[y∈LEFT(w)]]} and of {p:v'∈p & ∃y[y∈STUDENTS(v') & p=λw[y∈LEFT(w)]]}, (where v is any member of Dox(John)(w) and v' is any member of Dox(Mary)(w)). This is so because neither John nor Mary has a belief about the identities of the best student and the tallest student. Therefore, the 'de dicto$_e$' reading comes out false in w.

Once again we see that since strongly exhaustive verbs have the property that their semantics refers to the value of ANS(relevant question-intension) in the doxastic alternatives of the bearer(s) of the attitude, they give rise to the three distinct readings under discussion.

Let us now turn to *surprise*. This verb gives rise to a 'de dicto$_f$' reading, but does not distinguish between 'de re' and 'de dicto$_e$' readings. We predict this if we assume that the semantics of *surprise* does not refer to the value of ANS(relevant question-intension) in the doxastic alternatives of the subject. We saw in Section 1 that *surprise* is not strongly exhaustive. So let us check our intuitions regarding (19) against the non-strongly exhaustive truth conditions in (20) (which implies that to be surprised by a question is to expect the negation of its answer; see Berman 1991, Lahiri 1991, Heim 1994 and B&R), and the non-strongly exhaustive truth conditions in (21):[10]

[10] Note that clause (a) – which could alternatively be viewed as a presupposition – makes both (20) and (21) non-strongly exhaustive. (20) would be non-strongly exhaustive even without clause (a).

(19) It surprised John which students were elected for the neighborhood committee.

(20) **[[*surprise*]]**(x)(Q)(w)(t)=1 iff (a) {p:w∈p & p∈Q(w)} is not ∅; and (b) up to time t, x believes in w λw'[¬ANS(Q)(w)(w')]; and (c) after time t, x believes in w ANS(Q)(w).

(21) **[[*surprise*]]**(x)(Q)(w)(t)=1 iff (a) {p:w∈p & p∈Q(w)} is not ∅; and (b) up to time t, x believes in w λw'[ANS(Q)(w') ≠ ANS(Q)(w)]; and (c) after time t, x believes in w λw'[ANS(Q)(w') = ANS(Q)(w)].

In (20), ANS(Q) applies only to w, and crucially not to a doxastic alternative of the subject in w. So according to the proposal in (20), the 'de re' and 'de dicto$_e$' readings of (19) are predicted to come out the same because ANS(*which students were elected ('de re')* according to w) and ANS(*which students were elected ('de dicto$_e$')* according to w) yield the same value when applied to w. Therefore the subject of *surprise* is predicted to believe the same proposition under the 'de re' and the 'de dicto$_e$' readings of the embedded question. In fact, this was precisely G&S's objection to Karttunen's weakly exhaustive semantics of *know* (namely, that it wrongly predicted *know* not to distinguish between 'de re' and 'de dicto$_e$' readings). In (21), on the other hand, ANS(Q) applies to the subject's doxastic alternatives, and therefore the 'de re' and 'de dicto$_e$' readings of (19) are predicted to be distinct from each other. It turns out that (20), which predicts *surprise* to give rise to 'de re' and 'de dicto$_f$' readings only, is what we want.

Let us consider the following scenario in world v. The newly elected neighborhood committee consists of students and non-students. Bill and Sue – the best student and the richest student respectively – are the students who were elected. John discovers to his amazement that Bill and Sue were elected. He is aware that Bill and Sue are students, but up to the time of discovery he thinks they are not popular enough to be elected. At the time of the discovery he also comes to falsely believe that Sally is not a student and to falsely believe that she was elected. In addition, at the time of the discovery, he discovers that Bill and Sue are the best and richest student respectively. He says: "I am surprised that the best student and the richest student were elected, because usually people don't like to have the best student and the richest student, whoever they may be, on the same committee."

Here are the relevant values of ANS. ANS(*which students were elected ('de re'/'de dicto$_e$')* according to v)(v) = λw[Bill+Sue∈BE-ELECTED(w)], and ANS(*which students were elected ('de dicto$_f$')* according to v)(v) = λw[BEST-STUDENT+RICHEST-STUDENT(w) ∈ BE-ELECTED(w)]. Thus, according to (20), (19) is true in v in its 'de re' and 'de dicto$_f$' senses, because John expected both Bill and Sue not to be elected; and he expected the best student and the richest student, whoever they might be, not to be elected. There is no sense in which (19) is false in v according to (20). This prediction matches our intuitions. (21), on the other hand, predicts (19) to be true in v in its 'de dicto$_e$' sense, but false in its 'de re' sense, because after John's discovery, in any world v' in Dox(John)(v) the intersection of {p:v'∈p & ∃y[y∈STUDENTS(v) & p=λw'[y∈BE-ELECTED(w')]]} is not λw'[Bill+Sue∈BE-

ELECTED(w')] (because Sally is also a student in v, and she was elected in v'). But this prediction does not match our intuitions.

Now let us consider the following scenario in world w. The newly elected neighborhood committee consists of students and non-students. Bill and Sue – the best student and the richest student respectively – are the students who were elected. John discovers to his amazement that Bill and Sue were elected. Prior to that discovery, he believes that they are students, but at the time of the discovery, he comes to falsely believe that they are not students. At that time he also comes to believe, correctly, that the best student and the richest student were elected (only he does not know who the best and richest students are). He is surprised by that. He says: "I am surprised that the best student and the richest student were elected, because usually people don't like to have the best student and the richest student, whoever they may be, on the same committee."

According to (20), (19) is true in w in its 'de re' and 'de dicto$_f$' senses, because John expected Bill and Sue not to be elected, and he expected the best and richest students, whoever they may be, not to be elected. There is no sense in which (19) is false in w according to (20). This prediction matches our intuitions, even though after the discovery John is no longer aware of Bill and Sue's student status. On the other hand, according to (21), (19) is true in w in its 'de re' sense and false in its 'de dicto$_e$' sense, because after the time of the discovery, in any world v in Dox(John)(w), the intersection of {p:v∈p & ∃y[y∈STUDENTS(v) & p=λw'[y∈BE-ELECTED(w')]]} is not λw'[Bill+Sue∈BE-ELECTED(w')]. But this prediction does not match our intuitions. We conclude, then, that (20) is the adequate semantics for *surprise*.

Is this really the conclusion we should arrive at? After all, didn't G&S reject Karttunen's semantics of *know* precisely because it didn't account for the fact that the subject may be, but need not be, aware of the student status of the student-leavers in *John knows which students left*? But notice that there is a significant difference between *know* and *surprise*. In the case of *know*, we were able to argue for the existence of three readings, neither of which entailed the other. But *surprise* is different. We crucially cannot come up with a scenario in which (19) is true in its 'de dicto$_e$' sense and false in its 'de re' sense. But one *can* be surprised by which actual students were elected, without being aware of their student status. Therefore, we are forced to say that (19) does not have a 'de dicto$_e$' reading, and that one of the possible situations that make (19) true in its 'de re' sense is one where John is aware of the student status of the actual student-electees. But this is unproblematic, because a 'de re' interpretation is not incompatible with the subject's awareness of the relevant description. For example, the 'de re' reading of *John believes that some student was elected* is compatible with a situation where there is a student x such that John believes that x was elected, and he also believes that x is a student. In short, awareness on the part of the subject, or lack of it, of the description contributed by the *wh*-phrase does not affect the truth of (19) in its 'non-de dicto$_f$' sense.

To test our prediction further, let us briefly look at yet another question-embedding verb, namely, *predict*. In Beck and Rullmann 1999 p. 283 it is proposed that "to predict Q n percent accurately, n percent of the atomic propositions in Q that you

predict have to be true" (where an atomic proposition in Q is one which does not entail any of the other propositions in Q). I think this is a suitable semantics. Notice that this semantics does not refer to the answers to Q in the subject's doxastic alternatives (although it does refer to actual true answers to Q). Therefore, we do not expect this verb to give rise to 'de dicto$_e$' readings. This prediction is borne out. The readers can convince themselves that one cannot predict *which students were elected*, while being aware of the identities of the electees and of their student-status, without predicting which actual students were elected.

4. Conclusion

This paper made the following points. 'De dicto$_e$' readings must be distinguished from 'de dicto$_f$' readings, and our grammar must generate both of them independently of each other. Furthermore, while the availability of 'de re' and 'de dicto$_f$' readings does not depend on the embedding verb, the availability of a 'de dicto$_e$' reading does, because it depends on the embedding verb having a semantics that refers to the value of ANS in the worlds compatible with the relevant attitude. Since a strongly exhaustive semantics has this property, Strong Exhaustivity is a sufficient condition for the availability of 'de dicto$_e$' readings. As a result, all strongly exhaustive verbs (e.g., *know*, *agree on*) are predicted to give rise to all three readings, some non-strongly exhaustive verbs (e.g., *wonder*) are also predicted to give rise to all three readings, and some non-strongly exhaustive verbs (e.g., *surprise*, *predict*) are not.

References:

Beck, Sigrid and Hotze Rullmann (1999). "A Flexible Approach to Exhaustivity in Questions", *Natural Language Semantics* 7/3:249-298.

Berman, Stephen (1991). *On the Semantics and Logical Form of WH-Clauses*, PhD dissertation, University of Massachusetts Amherst.

Dayal, Veneeta (1996). *Locality in WH-Quantification*, Kluwer Academic Publishers.

Engdahl, Elisabet (1986). *Constituent Questions*, Kluwer, Dordrecht.

Groenendijk, Jeroen and Martin Stokhof (1982). "Semantic Analysis of WH-Complements", *Linguistics and Philosophy* 5/2:175-233.

Groenendijk, Jeroen and Martin Stokhof (1984). *Studies in the Semantics of Questions and the Pragmatics of Answers*, PhD dissertation, University of Amsterdam.

Hamblin, C.L. (1973). "Questions in Montague English," *Foundations of Language* 10:41-53.

Heim, Irene (1994). "Interrogative Complements of *Know*", in R. Buchalla and A. Mittwoch (eds.), *Proceedings of the 9th Annual IATL Conference and of the 1993 IATL Workshop on Discourse*, Akademon, Jerusalem, 128-144.

Karttunen, Lauri (1977). "Syntax and Semantics of Questions", *Linguistics and Philosophy*, 1/1:3-44.

Lahiri, Utpal (1991). *Embedded Interrogatives and Predicates that Embed Them*, PhD dissertation, MIT.

Rullmann, Hotze and Sigrid Beck (1998). "Presupposition Projection and the Interpretation of *which*-Questions", in D. Strolovitch and A. Lawson (eds.), *Proceedings of SALT VIII*, Cornell University, 215-232.

Department of Linguistics
University of Connecticut
U-1145
Storrs, CT 06269-1145

sharvit@uconn.edu