# Definiteness as Maximal Informativeness* 

Kai von Fintel, Danny Fox, and Sabine Iatridou<br>Massachusetts Institute of Technology


#### Abstract

We argue that definites are interpreted as denoting the maximally informative object that falls under the relevant predicate.


Given Irene Heim's well-known contributions to the domain of (in)definiteness as well as that of presuppositions, we hope to show our respect for her, as well as our gratefulness for everything she has taught us, with a few thoughts on these topics.

## 1 An influential analysis

This very short paper is concerned with the way the definite article the works. Consider the two definite phrases in (1). The singular one in (1a) presupposes that there is one girl in the relevant domain (existential presupposition) and that there is no more than one girl (uniqueness presupposition). The phrase refers to the unique girl. The plural definite in (1b), by contrast, presupposes existence but not uniqueness. The phrase refers to the entire collection of girls.
(1) a. the girl
b. the girls

There are two influential proposals on the meaning of the definite article that try to unify the singular and plural occurrences: Sharvy (1980) and Link (1983). Regardless of their differences, they share the insight that the definite article the presupposes that the extension of its complement has a unique maximal member based on an ordering of the elements:

[^0](2) $\llbracket$ the $\phi \rrbracket$ (where $\phi$ is of type $\langle\alpha, t\rangle$ ) is defined only if there is a unique maximal object $x$ st. $\phi(x)$ is true (based on an ordering on elements of type $\alpha$ ). The reference of the $\phi$ (when defined) is this unique maximal element.

Let's unpack this a bit. A "maximal object" is an object that, with respect to a certain ordering relation, has no other object ordered higher than it. Note that unique existence of a maximal object is not in general entailed by the existence of an ordering. It could be that there are two maximal objects, if there are two objects that have nothing ordered above them. But the proposal in (2), encodes as part of the meaning of the definite article that there is a unique maximal object.

The idea in applying this to the cases in (1) is that the ordering involved is based on the part/whole relationship between individuals. Both the singular and the plural NPs in (1) have denotations in the domain of individuals. Singular NPs are only true of atomic individuals, that is individuals which either do not have parts, or if they do have parts, those parts are not in the denotation of the noun. For example, while the denotation of the NP girl might have parts, none of the parts can be characterized as girl. ${ }^{1}$

Now, if there were two atomic individuals that are girls, neither would be part of the other, so they would both be maximal. So, the only time there will be a uniquely maximal atomic girl is when there is one girl (the existential presupposition) and only one girl (the uniqueness presupposition).

The denotation of a plural NP results from the closure of the singular NP under sum formation. Again, the domain has to have individuals in it (the existential presupposition). Let us say that it consists of three girls, $a, b, c$. The denotation of the plural NP girls consists of the atomic girls $a, b, c$ and any sum of two or more of them: $a+b, b+c, a+c, a+b+c$. The maximal object with respect to the part/whole ordering is $a+b+c$ and there is no other maximal object. Applying the definite article to a plural NP then, yields the uniquely maximal object, $a+b+c$.

By the above rationale, we have a uniqueness presupposition in the case of the singular NP, but not in the case of the plural.

## 2 Our Alternative

We argue for a different meaning for the definite article:
(3) a. $\llbracket$ the $\phi \rrbracket$ is defined in $w$ only if there is a uniquely maximal object $x$, based on the ordering $\geq_{\phi}$, such that $\phi(w)(x)$ is true. The reference of the $\phi$ (when defined) is this maximal element.
b. For all $x, y$ of type $\alpha$ and property $\phi$ of type $\langle s,\langle\alpha, t\rangle\rangle, x \geq_{\phi} y$ iff $\lambda w \cdot \phi(w)(x)$ entails $\lambda w . \boldsymbol{\phi}(w)(y)$.

Basically (3a,b) only differs from (2) in the nature of the criterion that imposes the ordering. The ordering in (2) looks at the individuals the predicate $\phi$ is true of and at their part/whole ordering. The ordering in $(3 a, b)$ looks at the same individuals but then looks at the information contained in claiming that they are $\phi$-individuals. The criterion for the ordering is informativeness.

[^1]The uniquely maximal object is the one that creates the most informative true proposition. The most informative proposition is defined as the proposition that entails all other relevant propositions (aka as the "strongest proposition").

Since the maximal object by $\geq_{\phi}$ is the one that yields the most informative true proposition of the form $\lambda w \cdot \phi(w)(x)$, we will call this object "the most informative object given $\phi$ " or sometimes simply "the most informative object".

## 3 Comparing the analyses

To compare our proposal and the Sharvy/Link proposal in (2), we will go through certain sample cases.

The first desired result is accounting for the fact that there is a uniqueness presupposition in the case where the definite article combines with a singular NP but not when it combines with a plural NP. We saw above how the Sharvy/Link account derives this fact.

We predict uniqueness for singular definite descriptions in a very similar way to Sharvy and Link: with a singular predicate we would be comparing the informativeness of claiming that atomic individual $a$ falls under the predicate to parallel claims about other atomic individuals. None of these claims logically entail the others, so there can only be a maximally informative object if there is a single one.

We predict uniqueness to disappear for plural definite descriptions such as the one in (1b), for the same reasons as Sharvy and Link: a plural property of individuals $\phi$, like girls, which is true of a plurality of individuals will necessarily be true of its atomic parts, since it is a distributive property. Hence, the largest element by the ordering of informativity will be the largest element by the part-whole ordering.

Assume that Ivy, Olivia, and Emma are the only girls. Then the girls will denote the maximal plurality made up of Ivy+Olivia+Emma. This is predicted by (2) because this is the maximal girl-plurality. That is, it is the object that is ordered highest by the part/whole criterion. It contains all other objects that girls is true of.

We derive the same result on our account: the plurality made up of those three individuals is the most informative object: it is this plurality from which we can deduce the "girlness" of all other girls. Any smaller plurality would be less informative. For example, from the plurality Olivia+Emma being girls we cannot deduce that Ivy is a girl.

So both the Sharvy/Link proposal and our proposal pick out the same individual in the case of (1b). In fact, the Sharvy/Link proposal and ours make the same prediction for a number of cases: all the properties that are upward monotone with respect to informativity. By "upward monotone" we mean here that a property yields a more informative claim the "bigger" the object is that it is predicated of. Conversely, "downward monotone" will mean that the "smaller" the object is, the more informative the claim is that the property holds of it.

Properties of degrees such as $\lambda d$. Miranda is $d$ tall or $\lambda n$. Miranda has $n$ many children are upward monotone: saying that Miranda is 1.65 m is more informative than saying that she is $1.60 \mathrm{~m} .{ }^{2}$ Saying Miranda has 4 children is more informative than saying that she has 3 . On our account, definite descriptions such as Miranda's height or the number of children that Miranda

[^2]has will refer (in any world $w$ ) to the maximal informative object in the extension of the property in $w$. Therefore, they will pick out 1.65 and 4 , respectively.

The Sharvy/Link account picks out the same objects for these definite descriptions: It will pick out 4 and not 3 because it is more highly ordered on the relevant ordering on the set of degrees. And for the same reason it will pick out 1.65 over 1.60 .

We propose that the Sharvy/Link theory got the right results only because the focus was limited to upward monotone properties (in the sense defined above). The different predictions between the Sharvy/Link account and ours start showing up once we look at properties that are not upward monotone with respect to informativity.

Consider then properties that are downward monotone with respect to informativity. These are cases where the smallest amount/object is most informative. Here is such a case:
(4) the amount of walnuts sufficient to make a pan of baklava

Propositions of the form d-much walnuts is sufficient to make a pan of baklava become more informative the smaller $d$ is. We thus correctly predict that the definite description in (4) should refer to the minimum amount of walnuts that would yield a true proposition, i.e. to the minimum amount that would suffice for baklava baking. ${ }^{3}$

On the other hand, according to the Sharvy/Link account, the definite description in this sentence should be undefined (and the sentence should thereby suffer form presupposition failure). The reason is that (given the monotonicity) there can be no maximal amount of walnuts that is sufficient to make a pan of baklava: if an amount of walnuts, $f$, suffices to make a pan of baklava, so does any amount larger than $f$.
(These kinds of examples are familiar from Beck and Rullmann (1999), where they play a crucial rule in thinking about maximality in questions. Beck and Rullmann argue that the relevant notion of maximality of questions is maximal informativeness. We are obviously indebted to their work.)

So we see that a definite description of the form the $\phi$ alternates between referring to the minimal or the maximal individual in the extension depending on the monotonicity with respect to informativity of property $\phi$. We get a maximality effect when $\phi$ is upward monotone and a minimality effect when $\phi$ is downward monotone as in (4).

Once the principle is clear, it is easy to construct further cases showing a minimality effect: consider, for example, the number of Greek soldiers who together can destroy the Trojan army. For Sharvy/Link, this would again result in a presupposition failure because there is no maximal number of Greek soldiers that together can destroy the Trojan army: if a certain number of soldiers can destroy the Trojan army, any larger number of soldiers can too. For us, on the other hand, the description will pick out the minimal number of soldiers that together can destroy the Trojan army,

[^3]because that is the most informative such number: once we know that number we can deduce that all larger numbers would also do. ${ }^{4}$

A particularly nice way to see the effects of variable direction of informativity comes from a pair of predicates in Aloni 2007: "When told John can spend 150 euro, people normally conclude that John cannot spend more. But when told John can live on 150 euro, they conclude that John cannot live on less." Using these predicates in definite descriptions thus gives rise to two different directions of strength of informativity:
(5) a. the amount of money John can spend (on vacation)
b. the amount of money John can live on

The definite description in (5a) refers to the maximum amount of money John can spend because that is the most informative object. The definite description in ( $5 b$ ) refers to the minimum amount of money John can live on, because that is the most informative object in this case. ${ }^{5}$

Finally, let us go to properties that are non-monotone with respect to informativity. What would it mean to be non-monotone with respect to informativity? It would mean that there are no inferences to be drawn either way: from $\phi(x)$ we can not infer anything about any other $\phi(y)$. If so, in order to have a unique maximally informative object, there would have to be just a single object falling under $\phi$. As soon as there is more than one such object, we predict presupposition failure. We will see that the Sharvy/Link account makes again a different prediction.

Let us set up a particular context. Imagine that you are trying to fit books ( $x, y, z, w, v \ldots$ ) on shelves of various size $(a, b, c \ldots)$. Suppose that book $x$ together with book $y$ fit perfectly on shelf $a$ (that is, $x$ and $y$ together fill exactly the space provided by $a$, nothing more, nothing less), and books $x, y$, and $z$ together fit perfectly on shelf $b$. Suppose also that no other combination of books fits perfectly on any other shelf. In this context, consider sentence (6).

[^4](6) \#Pass me the books that together fit perfectly on a shelf!

On our account, the definite description the books that together fit perfectly on a shelf suffers from presupposition failure. We believe that this is the actual judgment for this sentence. The reason that our account predicts presupposition failure is that the predicate books that together fit perfectly on a shelf is not monotone. From $x+y+z$ fitting perfectly together on a shelf, we cannot conclude about any other combination of books that it will fit perfectly together on a shelf. ${ }^{6}$

On the other hand, for Sharvy/Link this NP should be acceptable: it should refer to the largest collection of books that fit perfectly on a shelf, namely, $x+y+z$.

To summarize then so far: in upward monotone environments, our account and the Sharvy/Link account make the same predictions. In downward monotone environments, we correctly predict that the definite description will pick out the minimal object/amount but Sharvy/Link predict a presupposition failure. In non-monotone environments where the NP predicate is true of more than one (plural) object, we correctly predict a presupposition failure, whereas Sharvy/Link predict that the definite description should meet no problem whatsoever as long as the objects have a part/whole structure.

## 4 Temporal definites

An alternation similar to the one between minimality and maximality shows up in the domain of times as well. ${ }^{7}$ Consider the following two definite descriptions:
a. January 5th 1999 is the date until which Bill will be living in Boston.
b. January 5th 1999 is the date since which Bill has been living in Boston

In (7a) the definite description the date until which Bill will be living in Boston picks out the latest date until which Bill will be living in London.

On the other hand, in (7b) the definite description the date since which Bill has been living in Paris picks out the first day since which Bill has been living in Paris.

For sure, this is partly the result of the meaning of until and since. That is, 'forward looking in time" and "backward looking in time" is part of the lexical semantics of until and since. But what is not part of their lexical semantics is the "latest day" in the case of until and the "earliest day" in the case of since. We can see this in the following:

[^5](8) a. For sure he will be in Boston until the wedding. Maybe he will be here after that as well.
b. For sure he has been in Boston since the wedding but maybe he was in town long before that as well.
If the "the latest day" and the "earliest day" were part of the meaning of these items, (8a) and (8b) would be contradictions or at any rate, infelicitous. So we conclude that "earliest" and "latest" are not part of the meaning of the the temporal prepositions. ${ }^{8}$ Instead, the interval that is set up by these items seems to have the same behavior as numerals, which specify a lower bound only:
(9) For sure Miranda has 3 children, but maybe she has more.

We saw earlier how we derived that the number of children that Miranda has picks out the maximal number, as that is the most informative. So unlike (9), (10) is very odd: ${ }^{9}$
(10) \#(For sure) the number of children that Miranda has is three but maybe it is more. ${ }^{10}$

The contrast beween (9) and (10) is due to the fact that Miranda having at most 3 children is part of a scalar implicature in the first clause in (9) and this implicature can be and is cancelled by the second clause. But in (10) it is part of the semantics, given that it is contributed by the meaning of the definite description.

In that light, compare $(8 \mathrm{a}, \mathrm{b})$ to the following, which are much worse:
(11) a. \#January 5, 1999 is the date until which Bill will be living in Boston but maybe he will stay beyond that.
b. \#January 5, 1999 is the date since which Bill has been living in Boston but maybe he was here before that.

In brief, we do not need to and cannot (on the basis of (8a) and (8b)) stipulate 'latest' and 'earliest' in the semantics of the temporal operators until or since. In both cases, the definite descriptions in (7a) and (7b) refer to the most informative time that satisfies the relevant property ( $\lambda t$. Bill had been living in Boston until $t$, in (7a) and $\lambda t$. Bill has been living in Paris since $t$, in (7b). The difference, once again, has to do with the monotonicity of the property.

Finally, we should note that while our account delivers correct interpretations for these temporal definite descriptions, the Sharvy/Link account would fail quite miserably, since it would require, for example, that there be a unique day since which Bill has been living in New York for (7b) to be well-formed. That is simply not the case: if Bill has been living in New York since January 5, 1999, then he has also been living in New York since February 1, 2000, etc. What matters is that the day when he started living in New York is the most informative, not the unique day falling under the description.

[^6]
## 5 A potential problem

Imagine a final exam that consists of 20 questions, five of which $\left(q_{1}-q_{5}\right)$ are designated as necessary and sufficieng for a passing grade. That is, imagine that if one fails to answer even one of $q_{1}-q_{5}$ correctly one fails the test and if one answers all of $q_{1}-q_{5}$ correctly one passes the test, with the grade determined by the answers given to the remaining questions $\left(q_{6}-q_{20}\right)$. If such is the test, the definite description in (12), it seems to us, would refer to the designated five questions, technically the sum of $q_{1}-q_{5}$ (henceforth, just $q_{1}-q_{5}$ ).
(12) The questions such that if one answers all of them correctly one passes the test ${ }^{11}$

At first glance, this seems favorable to our analysis. One would be forgiven for thinking that surely the plurality of $q_{1}-q_{5}$ is the most informative object here and that this is so even though there are bigger pluralities of questions that are also such that if one answers all of them correctly one passes the test, all the way up to the biggest plurality of relevance, $q_{1}-q_{20}$. Surely, if answering $q_{1}-q_{5}$ correctly lets one pass the test, it follows from that that answering supersets of those five questions correctly should also let one pass the test. ${ }^{12}$ So, it would seem, our analysis predicts that the minimal set of questions that lets one pass the test will be the value of the definite description in (12).

Unfortunately, things are not that simple. The property that the is applied to in (12) is a conjunction of the predicate $\lambda w . \lambda x . x$ are questions in $w$ and the predicate $\lambda w . \lambda x$. in $w$, if one answers all of $x$ correctly, one passes the test. Now, the second conjunct is indeed downward monotone in our sense and thus would give us the right handle on why the minimal set of questions is the most informative object. But the problem is that the first conjunct is a standard upward monotone predicate. And thus, the conjunction of the two predicates is, of course, non-monotonic, which leaves us without an explanation for the observed interpretation of (12).

What we would like to do next is to show how this problem could be resolved with machinery that has been suggested to account for scope conflicts that arise in intensional contexts (specifically for a de re interpretation of a quantificational phrase that can be demonstrated to have narrow scope relative to an intensional operator, so called third readings, see von Fintel and Heim 2011, Chapter 9 ). Specifically assume that the property in (12) could receive the following rendition, where $w_{0}$ is (the world denoted by) a free variable inside the NP questions.
(13) $\lambda x . \lambda w . x$ are questions in $w_{0}$ and in $w$, if one answers all of $x$ correctly, one passes the test

To see that this would resolve our problem, take two (plural) individuals $X$ and $Y$ such that both are questions in $w_{0}$. For such individuals the entailment in (13) hold iff the entailment in (14) does:

[^7](14) $\lambda w . X$ are questions in $w_{0}$ and in $w$, if one answers all of $X$ correctly, one passes the test entails
$\lambda w . Y$ are questions in $w_{0}$ and in $w$, if one answers all of $Y$ correctly, one passes the test
(15) if one answers all of $X$ correctly, one passes the test
entails
if one answer all of $Y$ correctly, one passes the test
More generally if $X$ and $Y$ are both members of a set $A$, and $B$ is a function from worlds to sets, we get the following:

```
\(\{w: X \in A \cap B(w)\} \subseteq\{w: Y \in A \cap B(w)\}\) iff
\(\{w: X \in B(w)\} \subseteq\{w: Y \in B(w)\}\)
iff \(X \geq Y\) (when \(B\) involves entailment from sets to subsets)
iff \(Y \geq X\) (when \(B\) involves entailment from sets to supersets)
```

We would thus like to suggest that non-local binding of world variables (inside the NP questions) is what resolves our problem and explains the meaning of (12). Whether or not this suggestion is correct is something that we will have to leave for further research. ${ }^{13}$

## References

Aloni, Maria. 2007. Free choice and exhaustification: An account of subtrigging effects. Proceedings of Sinn und Bedeutung 11:16-30.
Beck, Sigrid, and Hotze Rullmann. 1999. A flexible approach to exhaustivity in questions. Natural Language Semantics 7:249-298.
Erlewine, Michael Yoshitaka, and Isaac Gould. 2014. Interpreting Japanese head-internal relative clauses. West Coast Conference on Formal Linguistics (WCCFL) 31:160-169.
von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. Journal of Semantics 16:97-148.
von Fintel, Kai, and Irene Heim. 2011. Intensional semantics. URL http://mit.edu/fintel/ fintel-heim-intensional.pdf, unpublished lecture notes.
von Fintel, Kai, and Sabine Iatridou. 2005. Since since. URL http://mit.edu/fintel/ fintel-iatridou-2005-since.pdf, ms, MIT.
von Fintel, Kai, and Sabine Iatridou. 2007. Anatomy of a modal construction. Linguistic Inquiry 38:445-483.

[^8]Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. Linguistics and Philosophy 29:537-586.
Iatridou, Sabine. 2014. About determiners on event descriptions, about time being like space (when we talk), and about one particularly strange construction. Natural Language Semantics 22:219-263.
Keshet, Ezra. 2010. Situation economy. Natural Language Semantics 18:385-434.
Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Meaning, use, and interpretation of language, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, Grundlagen der Kommunikation und Kognition, 302-323. de Gruyter.
Schlenker, Philippe. 2012. Informativity-based maximality conditions. Snippets 26:18-19.
Schwarzschild, Roger. 1996. Pluralities. Number 61 in Studies in Linguistics and Philosophy (SLAP). Kluwer.
Sharvy, Richard. 1980. A more general theory of definite descriptions. The Philosophical Review 89:607-624.


[^0]:    *This short paper is a collaborative work long in the making. Ideas towards it are found in von Fintel and Iatridou (2005) and there is a short report on this work in Fox and Hackl (2006, pp.548f). There have also been reports by Schlenker (2012) and Erlewine and Gould (2014). We are grateful to Luka Crnič and an anonymous reviewer for very helpful comments, not all of which we were able to do full justice to in the current state of the project.

[^1]:    ${ }^{1}$ The alternative we will propose in section 2 is committed (as far as we can see) to the idea that atomicity is defined independently of the predicate girl. Specifically we will need it to hold for all $X, Y$ such that $X \geq Y$ that the proposition $\lambda w$. $X$ is a girl or girls in $w$ entails the proposition $\lambda w$. Y is a girl or girls in $w$. This follows as long as plurality involves closure under sum-formation of elements that are inherently atomic.

[^2]:    ${ }^{2}$ In the literature on degree semantics, the predicates that correspond to downward entailing properties are often referred to as "downward scalar".

[^3]:    ${ }^{3}$ There are two interesting points that we won't focus on here. (i) Note that sufficient to make a pan of baklava does not mean that as soon as you have that amount of walnuts you have thereby made baklava, as one might have thought if educated in the language of logic (where "sufficient condition for $p$ " means a condition that guarantees the truth of $p$ ). This is a general feature of the use of sufficient in natural language. (ii) Note that one could use necessary in (4) and convey the very same meaning. This otherwise puzzling fact (necessary and sufficient surely don't mean the same thing) follows from our proposal. We can refer here to relevant discussion in von Fintel and Iatridou (2007).

[^4]:    ${ }^{4}$ This is not entirely innocuous. It is not obvious that larger numbers can, in fact, be deduced. After all, there could be circumstances where adding soldiers to an otherwise victorious army will (by necessity) reverse the outcome of the war. (Assume for example that the war is lengthy and that all the soldiers need to be fed, else there would be a revolt. Under such circumstances an efficient army with 100 soldiers might win the war, but it is conceivable that an army with 1000 soldiers will necessarily loose.) A full account of the minimality effect, would have to demonstrate that the property with which the definite article combines is downward monotone despite this counter-argument.

    One path towards an account would be to develop a semantics under which together $P$ holds of a plural individual $X$ the moment it holds of a sub-part of $X$. If there is a plurality $Y$ of 100 Greek soldiers such that $Y$ together beat the Trojan army, it would follow, under such a semantics, that any plurality $X$ larger than $Y$ beat the Trojan army. The intuition that this is not the case would then have to be attributed to something outside the basic semantics, probably to a Scalar Implicature. Reasons to think that a Scalar Implicature might be involved can be found in the observation that sometimes the intuition is different. (E.g. The books in Widener library together surely contain the solution to our problems can be true even if many books do not contribute to the solution in any way. Complicated issues related to "team credit" arise here which we haven't thought through seriously.) However, pursuing such an account would require quite a bit of work and in particular in the context of this paper would require that something special be said to derive the non-monotonicity assumed in our analysis of (6) below (for example that a necessary exhaustivity operator is present). Another path would be to modify $\geq_{\phi}$ so that it does not make reference to logical entailment but to some other notion of informativity. We will, unfortunately, have to leave the matter to another occasion.
    ${ }^{5}$ Here too, a complete account would require a semantics for the properties that combine with the definite article, showing that it is upward monotone in (5a) and downward monotone in (5b).

[^5]:    ${ }^{6} \mathrm{~A}$ reviewer points out that judgments change if the two sets of books (those that fit perfectly on shelf $a$ and those that perfectly on shelf $b$ ) are mutually exclusive, that is if there's no overlap between the two sets. We tentatively agree and think that this follows from our proposal, once we observe that the predicate together fit perfectly on a shelf can be interpreted distributively relative to a cover (in the sense of Schwarzschild (1996)) each member of which is non-atomic.
    (i) The books in this pile (all) fit perfectly on a shelf.

    We further point out that predicates that resist distributivity altogether, e.g., as a team can win the competition, yield a clear uniqueness presupposition:
    (ii) Introduce me to the athletes who as a team can win the competition.

    To our ears, (ii) yields presupposition failure in a context where there are two disjoint sets of athletes each of which can win the competition as a team.
    ${ }^{7}$ This section is related to discussion in von Fintel and Iatridou (2005).

[^6]:    ${ }^{8}$ There are additional arguments in Iatridou (2014).
    ${ }^{9}$ Similarly:
    (i) \#Miranda's height is 1.60 but maybe it is 1.65 m
    ${ }^{10}$ With for sure and on the assumption that for sure and maybe have the same modal flavor, (10) is a contradiction. Without for sure, (10) is presumably an instance of Moore's Paradox and has whatever defects such instances have.

[^7]:    ${ }^{11}$ We suspect that definite descriptions of this sort would be more easily formed in languages that are more comfortable with resumptive pronouns, as suggested by the following Hebrew example:
    (i) Ha-Se?elot Se mi se ?one naxon ?al kulan ?over et ha-bxina hen $q_{1}-q_{5}$ The-questions that who that answers correctly on all-them passes acc-the-test are $q_{1}-q_{5}$.
    ${ }^{12}$ In effect, we are taking for granted that the conditional in (12) is Strawson downward entailing with respect to its first argument, as argued for in von Fintel 1999. Strictly speaking, we would then suggest that (3b) be modified with 'entail' replaced with 'Strawson entail'.

[^8]:    ${ }^{13}$ One initial reason to think that our suggestion is wrong is Keshet's (2010) observation that NPs and their modifiers should be interpreted in the same world and time. However, it seems to us that Keshet's observation, though correct for the modification structures he looked at, does not extend to relative clauses:
    (i) Speaker A: Mary only cares about her physics grades (and frankly she doesn't care about them that much). So why is she even looking at the questions in Math?
    Speaker B: Oh, that's because she mistakenly believes that the questions in math are questions in physics. And that's why she wants to answer correctly the questions in math that are crucial for passing the test.
    On the face of it, it seems that the NP questions in math can receive a de re interpretation, while the relative clause receives a de dicto interpretation. So while we do not have any independent evidence for the suggestion we are making here, we don't think that there are reasons to reject it based on Keshet's observations.

