Abstract: The main aim of this paper is to show that one can combine Russell's view about the truth conditions of sentences containing definite descriptions with the view that they are referring expressions. Defending this thesis requires a positive account of a referring expression as one which plays a distinctive role in contributing to truth conditions, and releasing the notion from various accretions (that referring expressions must be simple, must have a referent, or require an understander to be en rapport with their referent). The envisaged role in truth conditions is given in terms of subject–predicate sentences: their truth requires there to be something to which the subject-expression refers and of which the predicate is true; falsehood is absence of truth. This is an essentially free logical truth condition, for it treats subject–predicate sentences with subject expressions which lack a referent as false. No distinction is made between subject expressions and referring expressions; and it is argued that although it is easy to state lines of demarcation, it is hard to justify them.

The view that definite descriptions are referring expressions aligns with some commonsensical intuitions, and by hypothesis cannot be accused by Russellians of assigning incorrect truth conditions. At first sight it may seem that the referential view is more easily extended to a unified account of the operation of "the" as it combines with plural predicates and mass terms, but I show that both approaches can give correct truth conditions in a systematic way. The referential view yields homophonic truth conditions, while the Russellian approach yields heterophonic ones. I suggest that this makes the former better adapted to the demands of a theory of meaning, and the latter better adapted to providing an explanation of how there can be a single concept, expressed by "the", which combines with a variety of kinds of expression, including mass terms and plural predicates.

Referring Descriptions

by R. M. Sainsbury

1   Introduction

One of Russell's arguments for his theory of descriptions involves contrasting them with names. According to Russell, names, conceived as a logician should conceive them, are expressions which, if

co-referring, are everywhere substitutable salva veritate, are not involved in semantically significant distinctions of scope, require acquaintance with their referent in order to be understood, and are such that to suppose that they have no bearer would be to suppose that they have no meaning (i.e. cannot be used in a significant sentence). This conception of names serves as a point of convergence for many distinctive features of Russell's logic, metaphysics, epistemology and philosophy of mind. For example, just as acquaintance needs to be a relation immune to the tactics of Descartes' demon, so a name's relation to its referent must be somehow guaranteed; just as, in acquaintance, the mind makes "direct" contact with the world, so a name, in referring to an object, somehow makes "direct" contact with it; just as a name is semantically simple, so its referent, Russell was tempted to think, must be ontologically simple (or "relatively simple").

The purpose of this paper is to see what happens to definite descriptions if we detach ourselves from Russell's conception of reference. I argue that, even if we do not wish to depart from Russellian truth conditions for any sentence, we may still treat definite descriptions as referring expressions in giving a semantic account of our language. By implication, we thereby come to see that Russell's theory of descriptions, as normally understood, involves two separable components: there is the claim about the truth conditions of sentences containing definite descriptions, and this claim is not challenged in this paper; and there is a claim about the semantics of definite descriptions themselves, to the effect that they are not referring expressions but quantifier phrases. It is this second claim which this paper addresses, showing that there is a coherent alternative view consistent with the first claim. The two Russellian claims converge in the framework of distinctively Russellian assumptions, notably those about what is involved in the mind's contact with the world. Remove these assumptions, and it is hard to resist the thought that definite descriptions are complex referring expressions. In the next section, §2, I elaborate this thought. In §3 I consider objections based upon the idea that the relation I call reference does not deserve to be so called. In §4 I take into account the distinction between referential and attributive descriptions, and in §5 I consider whether the referential approach is superior to the Russellian one, especially in the light of the attempt to devise a complete theory of "the", one which includes the cases in which it attaches to a plural or mass predicate.

2   The proposal

Suppose that a good start towards an account of understanding a sentence is to say that it involves knowing the sentence's truth conditions, and that this knowledge can be embodied in a Davidsonian style

semantic theory for a language. In this context, what would be the right thing to say about what is involved in understanding a referring expression? To be sure, understanding it requires that one know its contribution to truth conditions; but what is this? The currently conventional answer is that one must know to which object it refers, where this is glossed so that it entails that there is such an object and that one knows, concerning it, that it is referred to by the expression. This gloss is not compulsory, and leads to a disanalogy between the roles of reference and truth in understanding. It would be mad to think that understanding a sentence involves knowing which truth value it has, for this knowledge is evidently neither necessary nor sufficient for understanding; rather, what must be known are the conditions under which the sentence would be true. The analogous claim for reference would be that understanding a referring expression involves knowing what it would be for the expression to refer, that is, knowing how an object would have to be for the expression to refer to it. Preserving the analogy has the advantage of making empty names unproblematic.

Burge (1974) has developed a truth theory which meets the desiderata of the previous paragraph. An approximate example of the envisaged axioms for simple singular terms is:

1. for all x ("Hesperus" refers to x iff x = Hesperus).

As "Hesperus" occurs in quotation marks, the name belongs to the object language; as it is used alongside the identity sign, it belongs to the metalanguage. Although I have assumed that the metalanguage will contain all the terms the object language does, this is not essential. The general recipe for writing axioms would say something like: a term is to be associated with a reference condition of the form "$x = t$", where "t" is the metalanguage translation of the term in question.

The setting is negative free logic: every atom containing a non-referring term is false, the quantifier rules are appropriately adjusted, and the "law of identity" is only "$\forall x \; x = x$", rather than the truth of every singular sentence of the form "$t = t$". [1] The reference condition and the logic of identity ensure that singular reference is an "at most" relation: if "t" refers to x and to y then $x = y$.[2]

---

[1] There are several different logics called "free", and one reason for which, I suspect, semantic theories in free logical frameworks are less common than they should be is that some free logics are quite unsuitable for semantics. The negative free logic presupposed in this paper is a far cry from what Joseph Almog (this volume) thinks of as free logic: "Another response, imagined by various free logicians, is to extend the range of objectual values so as to make seemingly empty names come out as loaded names after all." The free logic of the present paper does not assign values to empty referring expressions.

[2] Plural reference is also an "at most" relation. Using capital letters for plural variables, if a plural referring expression (like "Ben and Mary" or "the apostles") refers to X and to Y, then $X = Y$. This assumes that a plural referring expression does not refer to some but not all of the things it refers to ("Ben and Mary" does not refer to Ben). This is semantically required, for a

No axiom of this kind could do justice to a semantically complex expression. But the general idea can be implemented, for singular definite descriptions, along the following lines:

2. for all predicates F in the singular, for all x ("the" $\cap$ F refers to x iff x = the satisfier of F).

If the metalanguage does not itself contain complex referring expressions, a Russellian-sounding axiom could be used:

3. for all predicates F in the singular, for all x ("the" $\cap$ F refers to x iff x uniquely satisfies F),

though this would preclude homophonic theorems. To allow for cases in which F contains an anaphoric pronoun (like the "her" in "Sally loves the man who kissed her") the reference condition must be relativized to an assignment, though this will be ignored here. It is at least on the cards that plural reference (e.g. "the men") can be treated in similar ways (see §5.3 below).

These truth conditions set context on one side, and so constitute a far from finished account.[3] The importance of context will not discriminate between Russellian and referential accounts. One can classify context dependence as linguistic (dependent on previous utterances) and non-linguistic (the remaining cases). Non-linguistic dependence is often significant in cases of "underspecification", when more than one thing satisfies the predicate in the description. One form of linguistic dependence is (or is analogous to) anaphora, for example "A man opened a door for a girl. The girl thanked him". The well-known difficulties of doing justice to contextual phenomena appear no greater on the referential account than on the Russellian. The main rival way of dealing with some kinds of contextual variation treats definite descriptions as having a semantics which falls short of truth conditions, these being variously supplied by contextually controlled enrichment, and is thus inconsistent with both Russellian accounts and the present form of referential account.[4]

If the object language contains modal idioms, and the semantics represents these one-dimensionally by quantification over worlds (with domains including trans-world individuals), the axioms become, respectively:

---

sentence whose predicate is "together lifted the boat" should be true iff true of whatever the subject expression refers to. The apparent oddity is reduced by rephrasing in terms of "has X as its referent". I assume a single use of the expression, singular or plural: the shape or sound "Aristotle" can be used to refer to distinct things.

[3] In common with Russellian treatments, they also ignore various arguably distinct uses of descriptions: "generic" uses, exemplified by "The whale is a mammal" (cf. Strawson (1950)), "functional" uses, as in "The President is re-elected annually", and predicative ones, as in "Napoleon was the greatest French general" (cf. Graff (2001)).

1′   for all worlds, w, for all x ("Hesperus" refers$_w$ to x iff at w (x = Hesperus)).

2′   for all worlds, w, for all x ("the" ∩ F refers$_w$ to x iff x = the satisfier$_w$ of F).

3′   for all worlds, w, for all x ("the" ∩ F refers$_w$ to x iff x uniquely satisfies$_w$ F).

The rigidity of names (simple singular referring expressions) is not marked explicitly in the axioms, but will obtain if the metalinguistic translations are rigid. If a metalinguistic term, t, meets the following condition, so does its object language homonym:

for all worlds, w, w', and all things, x:

if t refers$_w$ to x then t refers$_{w'}$ to x.

I assume that a term may refer$_w$ to something which does not exist at w.[5] This is not to assume that there can be names which actually have no bearer, for one could coherently insist, contrary to the claims of the present paper, that any intelligible name must refer to something with respect to the actual world.

As one would expect within a homophonic truth-theoretical approach, (2′) does not tell us explicitly whether descriptions are or are not rigid designators: object language ones are iff metalanguage ones are. Some descriptions are non-rigid if, as seems likely, there is a predicate F, worlds w and w', and distinct things x and y such that the following are true: "at w (the F is x)" and "at w' (the F is y)". Given that "the" can attach to an arbitrary predicate, (3′) ensures that some descriptions are non-rigid if, as seems likely, some predicates have a single object as their extension at one world, and a distinct single object as their extension at some other world.

So, along with other natural assumptions, some descriptions will refer rigidly and some non-rigidly. The rigid reference will be, in Kripke's terms, de facto: the description "happens to use a predicate 'F' that in each possible world is true of one and the same unique object" (1980, p. 21n). It is not possible for a semantically complex expression to be literally de jure rigid, that is one for which "the reference … is *stipulated* to be a single object" (1980, p. 21n), since the semantic properties of a complex must emerge from stipulations concerning their semantically significant simple parts, and cannot be stipulated

---

[4] This alternative, which has been endorsed by, for example Stalnaker (1970), Recanati (1993), Bezuidenhout (1997), certainly merits consideration, but it raises too many additional issues to be considered here.

[5] Evidence for and against the opinion that this was what Kripke originally had in mind is assembled by Kaplan (1989: 569-70).

directly.[6] However, one can if one wishes fashion a "the" whose semantic axiom ensures that the referring expression it helps to form is a rigid designator (see "the$_{DR}$" below, §4); this designator would meet the condition for de jure rigidity, as naturally extended to complex expressions.[7]

The conclusion agrees with intuition: the most natural way to account for the truth of such modal truths as "Someone else might have been the inventor of the zip" is to suppose that "the inventor of the zip" refers to different things with respect to different worlds, depending upon who invented the zip at the world.

An atom is a simple (*n*-ary) predicate concatenated with *n* names. On a traditional view, all complexity other than that found in atoms comes from the essentially sentence-forming logical constants. On a more realistic view of natural language, there are other ways in which semantic complexity is achieved. For predicates, these include predicate modifiers (as in "bright red"). For subject expressions, these include complex subjects like "Ben and Mary". Given the existence of collective predicates, like "together lifted the boat", it is not plausible that this is an example of essentially sentential complexity, for "Ben and Mary together lifted the boat" is not equivalent to "Ben together lifted the boat and Mary together lifted the boat". Even before considering definite descriptions, therefore, the atomic sentence would not represent the basic mode of sentence-formation in natural languages. Rather, we need the notion of a subject–predicate (S–P) sentence. This consists in the concatenation of *n* subject expressions with an *n*-ary predicate. There is no ban on the presence of logical constants within either the subject or the predicate. Thus "Ben and Mary together lifted the boat" is subject–predicate despite arguably having a complex predicate ("together lifted") and two complex subject expressions ("Ben and Mary", "the boat"), one of which even contains an expression generally counted as a logical constant ("and"). A language thought of as built up from atomic sentences differs sharply from one thought of in terms of S–P sentences. If the language has a finite number of simple singular terms and simple predicates, it has a finite number of atomic sentences, but it does not follow that is has a finite number of S–P sentences. For example, if "very" iterates, there is no limit to this list of predicates: "bright red", "very bright red", "very, very bright red" …. Each is fit to form an S–P sentence, by combining with a singular term (and the copula). The unbounded character of S–P sentences is also ensured if "the" can combine with

---

[6] One could define a semantically simple expression as one whose semantic properties can be stipulated. This reduces Kripke's de facto/de jure distinction, taken precisely as he formulated it, to that between the semantically simple and the semantically complex. However, one can reformulate the distinction to avoid this collapse.

arbitrary predicates to form subject expressions; and if "and" can conjoin names without limit to form complex subject expressions.

The notion of subject–predicate sentence I shall work with does not count indefinite descriptions ("a(n) F") and explicit "quantifier phrases" as themselves subject-expressions. This is not based on (at least simple-minded) grammatical congruity tests, but rather on inferential properties; I pass over the details.

Non-modal truth theory for a language with referring descriptions would need a composition axiom which exploits the notion of S–P. The unary singular case will go along these lines:

    4.   An S–P sentence is true iff for some x, S refers to x and x satisfies P.[8]

Achieving homophonic T-sentences will require a logic which ensures the validity of such inferences as:

    5.   From:

          S–P is true iff for some x, $x = t$ and x is G

      infer:

          S–P is true iff t is G.

Extending to the non-unary case is straightforward. Descriptions require an inference in which something like "x = the satisfier of 'F'" is replaced by "x = the F", preparatory to using (5). In the non-modal case, "for all y, Fy iff Gy" is sufficient for

    for all x, $x = $ the F iff $x = $ the G;[9]

and this in turn is enough (via substitution of equivalents) for the inference just mentioned.

3   Reasons for thinking this isn't really reference

On the present view, reference is one essential relation in the composition axiom. It is distinguished from the other relation, satisfaction, by the fact that it introduces an at-most condition: the axioms plus the background logic ensure that for each singular term there is at most one satisfier of its reference condition. In this section, I consider some forms of the objection that reference is more than this.

---

[7]It is not semantic complexity as such which induces non-rigidity: "Ben and Mary" rigidly designates Ben and Mary, if each of "Ben" and "Mary" are rigid designators (cf. Hossack 2000).

[8] Cf. Larson and Segal (1995), p. 119. An earlier source (by more than 600 years) of this truth condition for S–P sentences is William of Ockham, *Summa Logica* (II, 2): "it is sufficient and necessary that the subject and predicate supposit for the same thing" (Freddoso and Schuurman 1998: 86).

[9]It is a nice feature of Burge's system that although, for example "the unicorn = the golden mountain" is false, it is true that, for all x, $x = $ the unicorn iff $x = $ the golden mountain.

(3.1) Referring is something only simple expressions can do.

This is certainly a Russellian view, and very likely it runs deep in his thought. Justifying it from a neutral starting point (if such a thing exists) would be another matter. It confronts the problem of apparent complex referring expressions other than descriptions, like "Ben and Mary", which appear not to be eliminable in every context. And it makes it hard to find a unitary notion of reference common to semantic reference and speaker reference.

One motivation for the position is that the reference of an expression has to be something that can be arbitrarily assigned. The motivation is itself threatened by its unintuitive upshots: the denial both of complex referring expressions other than descriptions and of unity between semantic and speaker reference. (A speaker's choice of words or gestures to refer may be far from arbitrary.)

(3.2) It is necessary and apriori that a referring expression refers.

This thought was central to Russell's conception of a name: by definition, names, name. One can know a truth condition without knowing whether or not the condition is satisfied; and in many cases it will not be satisfied. If the analogy between truth and reference in relation to understanding takes this as its starting point, room is thereby made for the possibility, metaphysical and epistemic, of an expression associated with a reference condition failing to meet that condition, and so being a referring expression which fails to refer. On the view presented here, referring expressions are ones which, in their semantics, purport to refer; whether or not they succeed is another matter, falling in the domain of some science other than semantics. The view makes "names name" analogous to "police police", a specification of what would happen if all went well, rather than of what always in fact happens.

I do not intend to try to provide (and do not know how to provide) a knock-down argument in favour of the free logical conception of reference as opposed to the traditional Russellian one.[10] It seems to me that once people are satisfied that the former is coherent, its advantages for natural language semantics will be obvious. The point of the present paper is to argue that the application of the free logical notion to definite descriptions is indeed coherent.

Russell himself probably wanted it to be apriori, relative to linguistic understanding, whether an expression is a referring expression or not, and so whether or not it refers. This gave him a decisive

---

[10] Some relevant considerations are adduced by Sainsbury (2002: essays IX and XII).

reason for not counting ordinary names as referring expressions. Modern day theorists influenced by Russell in their treatment of names, like Evans and McDowell, do not suppose that this categorization is apriori (else it would be apriori whether or not Homer exists). A semantic theorist nonetheless needs to make the categorization, and so, if he belongs to the classical mould, must be involved in empirical work, including astronomy (to see if "Vulcan" is a referring expression) and the study of ancient societies (to see if "Homer" is a referring expression). The free logical semanticist has an easier task: one who understands a language is thereby equipped to know, without further investigation, which are its referring expressions, though not which of them refer.

A connected objection is that *reference failure engenders lack of truth conditions*. One motivation is that referential success is required for the standard composition axiom to deliver a truth condition, but one needs also to motivate choosing a composition axiom which has this effect, rather than choosing a free logical one, like (4), which does not.

(3.3) Understanding a referring expression involves de re knowledge

Russell thought that understanding a referring expression (that is, a logically proper name) involves being acquainted with its referent, and more recent theorists in this tradition have held something structurally similar, though without commitment to a Russellian conception of acquaintance: understanding a referring expression involves knowledge (de re) of its referent. On a free logical approach, this cannot be so in general, since not all referring expressions have referents. But can it be so for those expressions which do have referents?

On standard views, knowledge of an axiom for a referring expression whose referent is o will be held to be "de re" in one or more of the following respects:

(i) O is such that one who knows the axiom thereby knows something about it.

(ii)   The inference from knowledge of the axiom to the existence of an object, o, about which the knower has knowledge, is formally valid.

(iii)   The axiom can be known only by one who is "en rapport" with o.

Here "en rapport" marks a position for some unspecified but relatively hard to satisfy relation between a knower and an object. Russell's view entails that all three conditions are met, and the "en rapport" relation is acquaintance. Opinions about (ii) and (iii) can reasonably vary even among those otherwise

sympathetic to a Russellian conception of reference,[11] but they should all agreed on (i). The free logical approach can accept (i) for axioms relating to non-empty referring expressions. Hesperus is such that one who knows that for all x ("Hesperus" refers to x iff x is Hesperus) thereby knows something about it; the capital of England is such that one who knows that for all x ("the capital of England" refers to x iff x is the satisfier of "capital of England") thereby knows something about the satisfier of "capital of England", that is, about it (the capital of England). "Thereby" does not mean that there is a general inference to this conclusion, regardless of what referring expression appears in the axiom known, but merely that *this* item of knowledge is or involves knowledge about an object. Russellian and free logical theorists of reference can therefore agree that in some sense (that pointed to by (i)) understanding a semantic axiom for a non-empty referring expression involves de re knowledge of its referent.

(3.4) *Referring expressions are scopeless.*
Since definite descriptions are not scopeless, the objector infers that they are not referring expressions.

Although it is widely held that referring expressions are scopeless, it is harder to find justifications, and such as I can construct depend upon other theses, like (3.2) above or (3.5) below. The proponent of the free logical element of the present proposal will suggest that attributing scope distinctions even to names has some theoretical purpose, whence the existence of scope distinctions for definite descriptions is not enough to show them not to be referring expressions.[12] Here are two examples of phenomena relating to names which can be addressed in terms of scope.

(1) Some people hear "it is necessarily true that Socrates is human" as false, while hearing "Socrates has the property of being necessarily human" as true. This is prima facie evidence that names have scope (in the dialects of these people).

An objection is that this is not a scope distinction but rather one between meta-language modality (masquerading as wide scope) and object-language modality (masquerading as narrow scope). This seems intrinsically dubious, since it is hard to hear the first as something that could be understood quite independently of whether "Socrates is human" is understood. As a scope distinction, it fits into a theory

---

[11] Thus Kaplan (1989: 605) explicitly denies (iii): "acquisition of a name does not in general put us *en rapport* ... with the referent."

which seems quite natural. Let propositions be 0-place properties which exist at every world, let sentences be expressions which refer to propositions, and let the application of "necessarily *F*" to an object be true iff the object has the property expressed by "*F*" at every world at which it exists. "Necessarily, Socrates is human" is then not true, since there is a world at which Socrates does not exist and at which the proposition that he is human exists and is not true; but "Socrates is (necessarily human)" is true, since Socrates possesses the property of being human at every world at which he exists.

(2) One who asserts, as non-fictional, that Pegasus flies, has made some kind of mistake. One simple option is to suppose that, in the service of exposing the myth for what it is, we can truly deny what he says. But we do not want to share an opinion with those who, also in the grip of the myth, assert that Pegasus doesn't fly (but gallops). So, as Ockham implied (Freddoso and Schuurman 1998: 123), "it is not the case that Pegasus flies" is true, and can be used to deny the original assertion, whereas "Pegasus doesn't fly" (understood as "Pegasus is a non-flyer") is false.

In this case it may again be suspected that rather than a scope distinction we have a distinction between metalanguage and object language use; and metalanguage negation is, arguably, quite common (see Horn 1989). But, once again, it is hard to hear "It is not the case that Pegasus flies" as something which can be understood independently of whether or not "Pegasus flies" is understood. Free logical (i.e. Ockhamist) truth conditions naturally embed this as a scope distinction: a subject–predicate sentence is true iff the subject refers to something which has the property expressed by the predicate; negation toggles truth and falsity; "doesn't fly" is the predicate in "Pegasus doesn't fly" but "It is not the case that Pegasus flies" is the negation of the false "Pegasus flies".

The phenomena are complex, and these examples cannot be considered decisive. They show that, although it may be hard conclusively to establish that simple referring expressions can be involved in semantically significant scope distinctions, it will also be hard to establish that this is not so.

It may not be essential to the coherence of the position proposed here that simple referring expressions are scope sensitive. It is not obviously incoherent to hold that simple ones are scopeless but (some or all)

---

[12] The formal language of free logic makes room for semantically significant scope differences of singular terms and operators. The semantic theorist who uses a free logical framework is not obliged to suppose that such scope distinctions are

complex ones scope sensitive.

(3.5) Referring expressions are rigid designators.

Names and descriptions are governed by different substitution conditions in modal contexts (made vivid in slingshot arguments: cf. Neale 1995). Whereas no paradox arises from the supposition that co-referring names are substitutable salva veritate in modal contexts, a paradox does arise from supposing the same about definite descriptions.

A category of expressions all of which are to be treated by the same semantic relation may differ in logical/semantic properties. Even if all predicates are to be treated by means of satisfaction, they differ in their logical properties (e.g. some are "rigid" (have the same extension at every world) and some are not). One might see a range of modal expressions as variants on a single semantic theme (differing, perhaps, in whether they introduce all worlds, or just the logically possible, or just the physically possible, or just the epistemically possible) and so as belonging to a single semantic kind, even if modal equivalence of the different kinds generates different strengths of substitution condition. So we cannot move from the premise (which I do not dispute) that names and definite descriptions are governed by different substitution conditions to the conclusion that they are not of a single semantic kind.

The contrary view about descriptions can be fitted into systematic theory. What substitution of t_1 by t_2 in a context dominated by an operator, ω, requires is the truth of $\omega(t_1 = t_2)$. It so happens that in the case of names, but not descriptions, $t_1 = t_2$ entails its necessitation, so names seem effortlessly to meet the demands for substitution in modal contexts. However, co-referring names are not generally substitutable in hyperintensional contexts, since the relevant ω-dominated identity (e.g. "John believes that Hesperus is Phosphorus") may not be true. If substitution inferences were set out in full, with the appropriately dominated identities among the premises, the contrast between names and descriptions would no longer be at all striking. They would share conditions for successful substitution, but definite descriptions would, more often than names, fail to meet them.

Russell's conception of reference and identity dispensed entirely with ω-qualifications, and the upshot is, as he himself affirmed, that there are no (or almost no) logically proper names in human languages.[13] We

---

exploited by an sentences in natural language.

[13] Not all direct reference theorists would agree with this: cf. Salmon (1986).

get a conception of co-reference that is useful for natural language semantics only by qualifying Russell's pure vision. Once we start to qualify, however, the idea that genuine co-reference can fail to sustain substitution in hyperintensional contexts but must sustain it in modal contexts becomes hard to motivate. By contrast, the view that the substitution inference can be counted upon only if the strength of the identity matches the strength of the context in which substitution is to occur seems highly natural.[14]

4   Referential vs. attributive

Everyone agrees that there is some difference between "referential" and "attributive" uses of definite descriptions, and this raises the question whether the view that definite descriptions are referring expressions can allow for this difference. There are various conceptions of what the referential/attributive difference consists in, and I start by showing how the present proposal can accommodate two versions of the view that there is a semantic ambiguity located in the word "the".

Suppose the single word "the" is ambiguous between an attributive version, "the$_{att}$" and a referential version, "the$_{ref}$"; and suppose these have the following impact upon the truth conditions of whole sentences containing descriptions:

(Att)   "The$_{att}$ F is G" is true iff something is uniquely F and it is G.

(Ref)   if there is a unique x such that u (on a given occasion) intends his utterance to concern x and thinks he can achieve appreciation of this intention by using "the F", "The$_{ref}$ F is G" is true as uttered by u (on that occasion) iff x is G.

The current proposal provides "the" with truth conditions appropriate to "the$_{att}$", but ignores "the$_{ref}$". (Regarding "the" as a referring expression should not be identified with assigning it a contribution to truth conditions which would make it fit to form "referential" descriptions in anything like Donnellan's sense.) Referential descriptions would require a further axiom like:

(2$_{ref}$)   for all x ("the$_{ref}$" $\cap$ F refers to x as uttered by u iff x = the thing which u intends his utterance to concern and such that u thinks he can achieve appreciation of this intention by using "the$_{ref}$ F").[15]

---

[14] Added note: When I first wrote this paper I thought that suitably strengthened identities would validate substitution in all contexts, including hyperintensional ones: in the presence of "John knows that Hesperus = Phosphorus", "John knows that Hesperus is visible" would entail "John knows that Phosphorus is visible". This no longer appears to me to be in general correct. The Postscript suggests a better view.

[15] It is no accident, I believe, that attempts to provide this kind of reference condition fail to do justice to the possible semantic complexity of the predicate (e.g. "man drinking martini").

Nothing in the present proposal prevents the addition of such an axiom, which would recognize "the" as ambiguous in the most straightforward way. It would be somewhat misleading to use the terms "referential" and "attributive" to mark the two senses, for in both "the" is treated as a term which takes a predicate to form a referring expression. It is just that the different "the"s generate different reference conditions. Applying this to one version of Donnellan's original distinction, the sense in which an utterance of "The man drinking martini is drunk" is true, even though no one in the context is drinking martini, can be represented as "The$_{ref}$ man drinking martini is drunk"; the sense in which an utterance of this sentence is false can be represented as "The$_{att}$ man drinking martini is drunk". Both definite descriptions count as referential by the standard adopted here.

The semantic ambiguity just considered is not the only one that might be represented under the heading referential/attributive. Another possible ambiguity is between descriptions which, while in some cases designating rigidly, do so merely "de facto" (as a result of the rigidity of their predicates) and descriptions which de jure rigidly designate, as a result of stipulations concerning the word "the". We have already seen how to handle the former. The following modification of (2) will serve for de jure rigidly designating ones:

> (2$_{DR}$)   for all x ("the$_{DR}$" $\cap$ F refers to x iff x = the actual satisfier of F).

Here "actual" is absolute, taking the evaluation just to our actual world. Nothing in the present proposal prevents the addition of such an axiom, and the upshot is a semantic theory which recognizes at least two types of referring descriptions. One of these might be used to represent two readings Stalnaker (1970: 42) discerns in "The man in the purple turtleneck shirt might have been someone else". A true one could be represented as "Possibly, the$_{DR}$ man in the purple turtleneck shirt is not the$_{att}$ man in the purple turtleneck shirt", and a false one as "Possibly, the$_{DR}$ man in the purple turtleneck shirt is not the$_{DR}$ man in the purple turtleneck shirt".

Not all accounts of the referential/attributive distinction see it as a semantic ambiguity in "the". One cannot dismiss such alternative accounts as thereby irrelevant to present concerns, for these accounts may also have implications for the semantics of "the". This is so both for standard Grice-style accounts of the distinction as pragmatic and non-truth-conditional, based on unambiguously Russellian truth conditions, and for accounts of the distinction as pragmatic and truth conditional, based on a sub–truth-

conditional semantics for the relevant sentences.[16] Theorists of the former kind may see their position as inconsistent with that proposed here, for they offer the pragmatics in defence of Russell's theory of descriptions, which indeed excludes descriptions from the category of referring expressions. However, these theorists generally rely not upon this claim, but upon the Russellian truth conditions for sentences containing definite descriptions. If a central point of this paper is granted, namely that one can combine the view that definite descriptions are referring expressions with the Russellian view of the truth conditions of sentences containing them, there is room for Grice-style pragmatic accounts of referential/attributive which agree that descriptions are referring expressions.

The view that the semantics of definite description sentences are "underdetermined", that is, are sub–truth-conditional, and require pragmatic enrichment to generate either referential or attributive truth conditions, is inconsistent with the view proposed here. Once context is allowed to enter the picture, there is some slack in what constitutes semantic axioms like (2). Perhaps the axiom is just what is on the page, in which case it is a mere template, requiring interpreters to enrich the predicate F in various contextually prompted ways. Or perhaps there are as many axioms as there are such enrichments, so that each axiom delivers something determinate. These possibilities, however, do not match the view under discussion, for they do not allow for variation in truth conditions along a dimension describable as referential/attributive. On the present view, any truth-conditional variation involves distinct words "the". This view accordingly needs to find fault with underdetermination theories, though that will not be attempted here.

5   Logical form and truth conditions.

The referential proposal agrees with Russell about truth conditions (setting aside possible refinements adopted in the light of alleged referential/attributive distinctions).[17] How can the residual disagreement about logical form matter?

(5.1)   Russell and Davidson on logical form

Russell thought of the logical form of a sentence, or rather of a judgement, as made plain by its translation into the language of PM. On this view, the proposal to treat descriptions as complex referring expressions is not a logical form proposal, since the language of PM recognizes no such category. In

---

[16] References are in footnote 4, above.

order to have an intelligible debate, I shall adopt a more Davidsonian conception of logical form (cf. Davidson, 1967, 1970). It was unavailable to Russell but it does justice to a fair amount of his use of the idea of logical form. It is motivated primarily by considerations of semantic theory, taken as a theory which is constrained by the goal of stating something knowledge of which would be enough for understanding the language. On the Davidsonian view, the logical form of a sentence of a natural language is the sentence itself, if it can feature appropriately in a theorem of an interpretive truth theory for the language as a whole; otherwise it is the closest approximation to the sentence which meets this condition of accessibility to semantic theorizing.[18] A positive degree of closeness requires identity of truth conditions, and one candidate is closer to the original than another if it involves fewer modifications. (This is a vague notion which no one has troubled to make precise.)

Within the Davidsonian framework, logical forms (in indexical-free languages) will receive homophonic theorems in the truth theory. This gives another test for divergence of grammatical and logical form, for divergence implies that the official specification of the original sentence's truth conditions will be non-homophonic.

Whereas there is no firm distinction in Russell's approach between philosophical analysis and the finding of logical form, Davidson keeps these two activities distinct. It could be that a word for a concept which features homophonically in semantic theory deserves substantial philosophical analysis, in which our standard way of introducing the concept is replaced by something whose very unfamiliarity is illuminating. I suggest that just this is the case for "the". I will argue that we can gain philosophical illumination by analyzing it in the Russellian fashion, but that from the standpoint of semantic theory it should figure homophonically, and this requires a referential approach.

According to Russell's theory of definite descriptions, their grammatical form is distinct from their logical form. In Davidson's framework, this means that a sentence containing a description is subject to some non-trivial modification before becoming accessible to semantic theory, and that the proper specification of its truth conditions will be non-homophonic. In particular, it may well be that "the phrase [sc. one of the form "the F"] *per se* has no meaning, because in any proposition in which it occurs the proposition, fully expressed, does not contain the phrase, which has been broken up" (Russell 1905: 51).

---

[17] Cf. Burge 1974 (1991, p. 193).

[18] This skates over some exegetical difficulties, relating especially to Davidson (1970).

Russell also speaks of his theory as a reduction (e.g. 1905: 45). One way to express the position is to say that the Russellian logical form of "the F is G" will not contain "the". There are various ways to implement this, but for definiteness let us take the following:

6. The logical form of sentences of the grammatical form "the F is G" is "there is something, x, such that x is F, and anything that is F is identical to x; and x is G".

One contrast between Russell's view and the referential proposal is now stark: the referential proposal sees a sentence of the form "the F is G" as almost its logical form (perhaps bar fiddly modifications to make explicit variables and scope). In starkest formulation:

7. The logical form of sentences of the grammatical form "the F is G" is "the F is G".

We now need to ask how we should choose between (6) and (7).

Please insert table 1 near here

The contrasts in Table 1 make vivid the fact that the pre-theoretical view is that descriptions are referring expressions, for intuitions favour the right-hand claims in every case; we need to be dug out of this view by Russellian considerations. If we think that referring expressions must refer, and that all descriptions should be treated by the same kind of semantic axiom, we progress rapidly to Russell's position. The free logical context removes any logical motivation for the view that singular referring expressions must refer. Reference becomes like Russellian denotation: a referring expression counts as such in virtue of its "form" (and not in virtue of succeeding in referring). Russell's own conception of reference is not one which can be taken seriously by those who believe that genuine names with the same bearer are not everywhere substitutable salva veritate; and it will appear unattractive to those not moved by Russell's hope that thinking about the world involves a demon-proof relation.

The source of the pre-theoretical view that definite descriptions are referring expressions is probably a guess that we apply the same processing strategy to S–P sentences regardless of whether they have names or definite descriptions in subject position. The guess could be tested empirically. If it is correct, then from the point of view of an account of understanding, we do well to treat definite descriptions as referring expressions (provided that by doing so we can reach the right truth conditions). This is consistent with thinking that Russell's account can throw some light on the concept associated with "the". The consistency is like that between the claim that the right semantic axiom for "bachelor" says that this word is satisfied by all and only bachelors, while a possibly instructive analysis says that a bachelor is an unmarried man.

(5.2)   "The" as a determiner

Some of the ideas in Russell's theory of descriptions can be implemented with a logical form closer to surface form (the locus classicus is Neale 1990). "The" can be regarded as a determiner which attaches to a predicate to form a quantifier. Neale says that this is not "an *alternative* to the Theory of Descriptions" but the theory itself stated in a more "congenial" way (Neale 1990: 45). Neale in effect takes the truth conditions for whole sentences as the essential characteristic of Russell's theory of descriptions, independently of the precise logical form which is supposed to generate the sentence-sized truth conditions.

A version of Russell's theory in which "the" is treated as a determiner enables logical form to approach much more closely to grammatical form. On this proposal, the logical form of sentences whose grammatical form is "the F is G" is something like: the x that is F is G. This gives us three views to compare: what I call Russell's, in which the descriptive phrase is "broken up" at the level of logical form, and two views which retain "the" in the metalanguage: the determiner view, and the referential view, both of which offer logical forms close to, or even identical with, grammatical forms.

This paper largely brackets the determiner view. The view agrees with the referential one on sentence-sized truth conditions, but disagrees with it about the mechanisms that generate these: one view fingers quantification, the other reference; one view is set within classical logic, the other within free logic. Whether anything substantive hangs on these differences is not obvious. I will take it for granted that a proper account of "the" will account for its uses as it attaches to plural as well as singular predicates, and also as it attaches to mass terms ("the gold in Zurich"), and that such an account will deal with collective as well as distributive uses of expressions in the remainder of the sentence. A hypothesis, which I here make no attempt to substantiate, is that the determiner view faces difficulties in achieving this level of generality.[19] As I now go on to show, the same does not hold either for the Russellian or the referential views.

(5.3)   Plural and mass definite descriptions

"The" can attach to mass terms (as in "the gold in Zurich") and to plural predicates (as in "the people in Auckland") (cf. Sharvy, 1980). Intuitively, these descriptions are perfectly in order, despite the fact that

---

[19] Neale is sympathetic to the idea that a final account of "the" should be general, and his own account extends naturally to plural distributive cases (1990: 46). He explicitly brackets collective readings (1990: 61).

there are many satisfiers of the predicates "gold in Zurich" and "people in Auckland". Nothing in "the" has any special connection with any of singularity, plurality or masshood, and a full account of the word should explain this neutrality: it should be itself neutral and uniform, allowing the variations in question to emerge from the predicate with which "the" couples.

One who holds that singular definite descriptions are singular referring expressions is likely to hold that plural definite descriptions are plural referring expressions. We need to add plural variables, ones which stand in the kind of position fit to be occupied by plural referring expressions (e.g. "Ben and Mary"). The proposed singular semantic axiom was:

    2   for all x ("the" $\cap$ F refers to x iff x = the satisfier of F).

Its plural correlate, using "X" as a plural variable is:

    2P for all X ("the" $\cap$ F refers to X iff X are the satisfiers of F).

An instance (2P) is:

    "the apostles" refers to the apostles iff the apostles are the satisfiers of "apostles".

Although Mathew and Mark are satisfiers of "apostles" they are not *the* satisfiers of "apostles", so their properties alone are not enough to verify a claim about the apostles.

(2) and (2P) are essentially the same axiom, modulo number. The composition axiom would be extended to include a plural form. The proposal leads to a straightforward analogue of (7):

    8. The logical form of sentences of the grammatical form "the F are G" is "the F are G".

In specifying a minimal departure from Russell's actual theory of descriptions which would allow it to speak to plural descriptions, the first thing is to add plural variables.[20] The extended Russellian can allow that there are plural referring expressions like "Ben and Mary", while denying that plural descriptions are referring expressions. This is analogous to allowing that singular names are referring expressions while denying that the same is true of singular descriptions. In both cases, the conceded referring expressions suffice to introduce the notion of a corresponding kind of variable. A possible Russellian proposal which mirrors (6) is:

    9. The logical form of sentences of the grammatical form "the F are G" is "there are things, X,

---

[20] I assume that there is genuine and irreducible plural reference and plural quantification (persuaded by Hossack 2000). The historical Russell would no doubt have preferred to reduce this to singular reference and, in the first instance, quantification using classes. Hossack suggested that this would not affect the contrast between the two conceptions of logical form which is the theme of §5.

such that *FX*; and *X* are *G*".

This is incorrect, for it would make Mathew and Mark's happiness sufficient for the truth of "the apostles are happy". "*X* are *G*" is true of things which are *G*, so on this proposal "the apostles are happy" is true if Mathew and Mark are happy apostles. Intuitively, we need some notion of maximality: for *the* apostles to be happy is for all of them to be so. But how should this notion be introduced? Not like this:

> 10. The logical form of sentences of the grammatical form "the *F* are *G*" is "there are things, *X*, such that *FX*; and all *X* are *G*".

Suppose the department is composed of Ben, Mary and Joshua. An application of (10) is:

> The logical form of "The members of the department together formed a triangle" is "there are things, *X*, such that *X* are members of the department and all *X* together formed a triangle".

Even if Ben, Mary and Joshua together formed a triangle, Ben and Mary, while members of the department, did not together form a triangle (how could they?). The Russellian needs further resources. He could allow the positions occupied by plural variables to be occupied also by expressions like "all the members of the department taken together". This is the natural way, but in effect it concedes a treatment of plural descriptions as referring expressions. Alternatively, he can introduce something other than quantifier-variable devices. Richard Sharvy (1980) has suggested a single notion which he claims will do the trick. He calls it *part of*, and we could adapt his suggestion to the approach taken here as follows:

> 11. for plural predicates, *F*, the logical form of sentences of the grammatical form "the *F* are *G*" is "there are things, *X*, such that *FX and any Y that are F are part of X*; and *X* are *G*".

Let us call the italicized portion the "maximality condition". "Part of" can also be read as "are among". Sharvy's suggestion avoids both counterexamples: Mathew and Mark's happiness is not enough for the apostles to be happy, because Mathew and Mark by themselves do not meet the maximality condition and so will not constitute a verifying value for *X* in "there are things, *X*"; that the members of the department together formed a triangle does not require Ben and Mary to, for they do not meet the maximality condition. A nice feature of Sharvy's account is that we can recover from (11) the condition for the singular description, since, using singular variables, for *x* to be among *y* is for *x* to be *y*. The condition of unity for singular and plural descriptions is met; the evidently non-homophonic character of the resulting semantic theory cannot (in the present dialectic) be regarded as a defect. Once again, the Russellian approach and the referential approach can both generate the right truth conditions, the former heterophonically, the latter homophonically.

The same holds when we consider mass descriptions. Intuitively, "The gold in Zurich is worth more than

a million dollars" is true, even though there is gold in Zurich worth less. (11) as such does not address these cases, for they are singular, but Sharvy's idea is easily adapted:

> 12. for mass terms in the singular, F, the logical form of sentences of the grammatical form "the F is G" is "there is stuff, m, such that *Fm and any m′ that is F is part of m*; and m is G".

This tells us, correctly, that the gold relevant to the truth of this sentence is just all the gold in Zurich: gold of which any gold in Zurich is part. (The new style of variable, "m", is one fit to be replaced by a mass term.)

On the referential approach, the logical form will be derived from reference axioms. The original (2) could become:

> 13. for mass terms in the singular, F, for all m ("the" $\cap$ F refers to m iff m = the satisfier of F).

There are satisfiers of "gold in Zurich" which are not worth more than a million dollars, for example, the gold in Hans Ernst's wedding ring. However, the gold in this ring is not identical to the satisfier of "gold in Zurich", for the latter includes other gold. And although the ring perhaps satisfies "gold in Zurich" (at any rate, it satisfies "is made of gold and is in Zurich") it is not a permissible value for "m", for such a value needs to be the sort of thing a mass term is true of. So the ring is not wrongly said by (13) to be something to which "the gold in Zurich" refers. This does not show that (13) is adequate, but it removes some grounds for thinking it inadequate.

A proper test for a theory of descriptions is that it should deal in a unified way not only with singular descriptions but with plural and mass descriptions. We have found no reason to think that the Russellian and referential approaches cannot both do this, in their different ways. So an appropriate question would seem to be: which is correct? Or should we say that semantic facts are indeterminate, that there is no fact of the matter which theory is right?

I think we do better to see the two approaches as addressed to different issues. It seems to me likely that the referential approach stays closest to the way in which we actually process language. At least, there is some intuitive support for this, though the issue deserves empirical investigation. Moreover, the referential approach, unlike the Russellian approach, leads to homophonic logical forms, and this seems what is needed by a semantic theory which encodes what we know in understanding language. It would endorse, as impeccable, reports of what is said by uses of sentences containing definite descriptions which re-use those sentences, whereas this would not be so if the Russellian approach were applied to

understanding.

On the other hand, the *part of* (or *among*) relation, which appears essential to a Russellian approach and which stands in the way of homophonic logical forms, provides some kind of explanation or analysis of how there can be a unified concept *the* capable of accepting predicates in the singular and in the plural, and mass and non-mass predicates. It is precisely because the *part of* relation is relatively unfamiliar that it can throw some explanatory light on the unified operation of "the". [21]

The main aim of this paper has been to show that there is room for a genuinely referential theory of descriptions. A subsidiary aim was to show that such a theory is likely to provide logical forms which, from the standpoint of semantic theory, are more accurate than Russellian theories, though the latter have something else to contribute: an explanation of how "the" can exercise its unified function in various different constructions.[22]

Postscript.

I now think that only rigid subject expressions should be counted as referring expressions. My reason is that the speech act of referring involves attempting to single out an entity upon whose states the transworld truth or falsehood of what is said is to turn. When successful, the referent introduced is the object which matters to truth or falsehood, actual or counterfactual, and so should be constant across worlds, and this will be reflected in the way in which reference is handled in the composition axiom.

One impact of the new view is that genuine identities expressed with referring descriptions will be necessary if true, and so the idea of ω-qualified identity can be scrapped, which is all to the good. Another impact is that although Russellian and referential descriptions accounts will (in a sense) coincide on the truth conditions of sentences in non-modal languages, they will diverge with respect to modal languages. An interesting contrast will remain, which is structurally similar to that discussed here: it is between rigid referential descriptions, and Russellian descriptions with predicates rendered rigid, for

---

[21] As Sharvy stresses, it is certainly distinct from the concept introduced by the standard use of "part of" in mereology.

example by the use of a suitable expression for actuality (where this notion is treated as fixed to our actual world). "The actual *F* is *G*" will have as its Russellian logical form something like "Exactly one thing is actually *F* and it is *G*". By contrast, for referential descriptions, there is no need for a difference between surface and logical form. While referential definite descriptions and rigidified Russellian definite descriptions thus diverge in logical form, any sentence which can be seen as containing a referential definite description will be assigned the same overall truth-conditions as if it were seen as containing, rather, a rigidified Russellian definite description; so considerations analogous to those raised in §5 above will be relevant.

References

Almog, Joseph (2003). "The vernacular and the omniscient observer of history." This volume.

Bezuidenhout, Anne (1997). "Pragmatics, semantic underdetermination and the referential/attributive distinction." *Mind* 106: 1–34.

Burge, Tyler (1974). "Truth and singular terms." *Noûs* 8, 309–325. Reprinted in Lambert, Karel (ed.) *Philosophical Applications of Free Logic,* Oxford, New York, 1991, pp. 189–204.

Davidson, Donald (1967). "The logical form of action sentences." In *Essays on Actions and Events*, 105–22, Oxford: Clarendon Press.

Davidson, Donald (1970). "Reply to Cargile." In *Essays on Actions and Events*, 137–148, Oxford: Clarendon Press.

Donnellan, Keith (1966). "Reference and definite descriptions." *Philosophical Review* 77, 203-15.

Graff, Delia (2001). "Descriptions as predicates." *Philosophical Studies* 102.1, 1–42.

Horn, L. R. (1989) *A Natural History of Negation*. Chicago: University of Chicago Press.

Hossack, Keith (2000). "Plurals and complexes." *British Journal for Philosophy of Science* 51.3, 411–443.

Larson, Richard and Gabriel Segal (1995). *Knowledge of Meaning: an Introduction to Semantic Theory*. Cambridge, MA: MIT Press.

Neale, Stephen (1990). *Descriptions*. Cambridge, Mass., London: MIT Press.

Neale, Stephen (1995). "The philosophical significance of Gödel's slingshot." *Mind* 104.416, 761–825.

Ockham, William of (1323). *Summa Logicae*. English translation in Alfred J. Freddoso and Henry Schuurman *Ockham's Theory of Propositions*. South Bend, Indiana: St Augustine's Press, 1998.

Recanati, François (1993). *Direct Reference*. Oxford: Blackwell Publishers.

Russell, Bertrand (1905). "On denoting." *Mind* 14, 479-93. Reprinted in Marsh, R.C. *Logic and Knowledge,* London, Routledge and Kegan Paul, pp. 41–56.

Russell, Bertrand (1919). *Introduction to Mathematical Philosophy*. London: George Allen and Unwin. Ch. 16 reprinted as "Descriptions" in Moore, Adrian (ed.) *Meaning and Reference*, p. 46–55, Oxford: Oxford University Press.

Russell, Bertrand, and Whitehead, Alfred North (1910-13). *Principia Mathematica*. Cambridge: Cambridge University Press.

Sainsbury, R.M. (2002). *Departing From Frege: Essays in Philosophical Logic*. London: Routledge.

Salmon, Nathan (1986). *Frege's Puzzle*. Cambridge, Mass: MIT Press.

Sharvy, Richard (1980). "A more general theory of definite descriptions." *Philosophical Review* 89.4, 607–624.

Stalnaker, Robert (1970). "Pragmatics." *Synthese* 22. Reprinted in his *Context and Content,* Oxford: Oxford University Press, 1999, 31–46.

Strawson, Peter (1950). "On Referring." *Mind* 59, 269-86.