The Semantics of Definite and Indefinite Noun Phrases

Irene Roswitha Heim
2011 EDITION: A NOTE FROM THE TYPESETTERS

Irene Heim's seminal dissertation from 1982 has been freely available in its original format for some time, but we have often longed for a retypeset and fully searchable version of this important work. One day, we therefore embarked on the project of typesetting it ourselves and this is the finished product.

With permission from Irene Heim, we now make this new version available to interested readers. We have in general attempted to stay true to Heim's original formatting, but we have made some stylistic changes when we judged that this would improve readability. Should you find any typographical errors in the text or in the formalisms, please do let us know so that we can correct them. For citations, please refer to the 2011 edition.

Anders J. Schoubye
Ephraim Glick
Irene R. Heim  
*The Semantics of Definite and Indefinite Noun Phrases*  
M.A., Ludwig-Maximilians-Universität München,  
Ph.D., University of Massachusetts  
Directed by Professor Barbara Partee

Submitted to the Graduate School of the University of Massachusetts in  
partial fulfillment of the requirements for the degree of  

**DOCTOR OF PHILOSOPHY**

September, 1982  
Linguistics
PREFACE

In November 1978, a workshop was held at the University of Massachusetts whose title was “Indefinite Reference” and whose topic Barbara Partee described in a circular that started as follows:

One standard view among logicians is that indefinite noun phrases like 'a tall man' are not referring expressions, but quantifier phrases, like 'every man,' 'no man,' and 'most men.' Yet in many respects, indefinite noun phrases seem to function in ordinary language much like definite noun phrases or proper names, particularly with respect to the use of pronouns in discourse. This may be simply a matter of sorting out semantics from pragmatics, but there is not to our knowledge any currently available theory that simultaneously characterizes the logical or truth-functional properties of indefinite noun phrases and accounts for their 'discourse-reference' properties ...

A second installment of the same workshop took place in March 1979, and I was then so preoccupied by the topic that I immediately wrote a paper called “Toward a unified semantics of definite and indefinite noun phrases,” in which I treated all definites and indefinites as variables. In the following months I grew disaffected with that paper because it did not tell a convincing story about the reasons why definites and indefinites, both being variables, were so unlike each other. I wished that all the differences could somehow be made to follow from a difference in the presuppositions of definites and indefinites, But it was only in the summer of 1980 that I began to make sense of this idea. Then the writing of this dissertation took almost two more years. It still feels like a rough draft, and I hope the reader will view it as such and excuse the redundancies, inconsistencies, and omissions, not to speak of mistakes.

The workshop at which my dissertation topic was laid out for me was just one of the many stimulating and instructive events that contributed to my education. I was lucky to be first a student and then a regular guest at one of the liveliest linguistics departments in Germany, at the University of Konstanz, and I was even luckier to become moreover a student at one of the liveliest linguistics departments in the world, at the University
of Massachusetts in Amherst. I am greatly indebted to the people who have made these places what they are, especially to Arnim von Stechow in Konstanz, and to Emmon Bach and Barbara Partee in Amherst. My thinking about linguistics and semantics in general, and about the topics of this dissertation in particular, has been influenced most fundamentally by the writings of Noam Chomsky, Robert Stalnaker, David Lewis, and Angelika Kratzer, and by the teaching of Edwin Williams and Alan Prince. I also learned a lot from conversations with Arnim von Stechow, Emmon Bach, and Barbara Partee.

While I was working on this dissertation, Barbara Partee often drew my attention to relevant examples and the pertinent literature. Arnim von Stechow thought seriously about every stage of the theory I was developing, always with the attitude of someone who was struggling with the same questions and wanted to see answers that he could believe. Hans Kamp also became a source of moral support: When I first learned of his paper “A theory of truth and semantic representation” in September 1980 and found it to contain ideas so strikingly similar to my own, I did not welcome the competition. But the encouraging aspect of not being alone with my ideas has prevailed in the long run, especially since Hans, when he came to see my work, overlooked its deficiencies generously and emphasized its positive contributions to our common goals. Angelika Kratzer has helped me tremendously. Her comments on earlier drafts caused me to reconsider some basic decisions both about the claims I wanted to make and about the form in which I wanted to present them, and I think this has led to considerable improvements. There were various other people who offered good suggestions and criticisms at one occasion or another. I am grateful to them all, but especially to the four I have mentioned.

I would also like to express my gratitude to the friends that have accompanied me through the years during which I labored over this work. Lisa Selkirk bears most of the responsibility for the fact that these years were much less agonizing than I had feared. Others have contributed to this as well, in particular Elisabet Engdahl, Gertrud Niggl, and Angelika Kratzer. Murvet ENC and Nomi Erteschik-Shir helped me through the final stages.

A long and sad story could be told about the typing of this manuscript. I want to thank Kris Dean, who had the bad luck of wasting innumerable hours of beautiful typing and who remained her most helpful through the bitter end. I also appreciate Daniel Flickinger’s help with the proof-reading.
ABSTRACT

Logical semanticists have standardly analyzed definite and indefinite descriptions as quantifiers, and definite pronouns as variables. This dissertation explores an alternative analysis, according to which all definites and indefinites are quantifier-free, i.e., consist of an essential free variable and the descriptive predicate (if any). For instance, “A cat arrived” is analyzed as “cat(x) ∧ arrived(x).” The existential force that such sentences carry in their unembedded (and some of their embedded) uses is attributed not to the indefinite article, but to principles that govern the interpretation of variables in general.

The primary motivation for this variable-analysis is that it leads to a straightforward account of indefinites serving as antecedents for pronouns outside their scope, as in the so-called “donkey sentences.” These are handled by combining the variable-analysis of indefinites with a treatment of quantifiers as basically unselective, drawing on work by David Lewis.

Since a variable-analysis of both definites and indefinites prima facie obliterates the distinctions between the two, it must be accompanied by a new theory of the definite-indefinite contrast. This theory must account for the different conditions under which definites and indefinites can get bound, and for the exclusive capacity of definites for deixis and anaphora. All these differences can be predicted if the uniform semantic analysis of definites and indefinites is supplemented by suitable assumptions about their contrasting felicity conditions (presuppositions): Felicitous definites must be “familiar” variables, felicitous indefinites must be “novel” variables.

Familiarity may be a matter of having an antecedent in the text, or else of having a contextually salient referent. The underlying concept of familiarity that unites these two cases is best captured in an enriched semantic theory, which includes a further level of analysis, here called the “file” level. A major part of the dissertation is devoted to developing and motivating such a theory of “file change semantics,” and to making precise its relation to conventional truth-conditional semantics.
CONTENTS

I Problems Concerning Indefinites and Anaphora 9
1 Do Indefinites Refer? .................................................. 11
   1.1 Russell’s view ....................................................... 11
   1.2 Anaphoric pronouns as bound variables .................. 14
   1.3 Anaphoric pronouns as picking up a speaker’s reference. 17
   1.4 Anaphoric pronouns as disguised definite descriptions . 24
   1.5 Anaphoric pronouns and the ambiguity hypothesis .... 29
   1.6 Summary .......................................................... 34
2 Problems with Donkey Sentences ......................... 35
   2.1 Donkey anaphora as variable binding .................... 37
      2.1.1 Egli’s proposal .............................................. 39
      2.1.2 Evaluation of Egli’s proposal ......................... 42
      2.1.3 Smaby’s proposal ......................................... 48
   2.2 Donkey sentences and pragmatic accounts of anaphora . 51
   2.3 Donkey anaphora and disguised definite descriptions ... 52
      2.3.1 Pro and contra the uniqueness implication .......... 58
      2.3.2 Conditionals .................................................. 63
   2.4 Donkey sentences and the ambiguity hypothesis ....... 68
   2.5 Summary .......................................................... 69
3 Donkey Anaphora in Game-Theoretical Semantics ....... 70

II Indefinites as Variables 81
1 Adaptability of Indefinites and Adverbs of Quantification ... 81
2 Logical Forms for English Texts with Indefinites and Quantifiers .................................................. 86
3 Semantic Interpretation of Logical Form .................... 101
   3.1 Semantic categorization of elements of logical form .... 101
   3.2 An extensional semantics, based on satisfaction ....... 105
   3.3 From satisfaction to truth ...................................... 108
4 Invisible Necessity Operators ..................................... 112
   4.1 Indefinites bound by overt modal operators .......... 113
   4.2 Indefinites bound by invisible modal operators ....... 121
   4.3 A note on generic and other restrictive indefinites .... 126
5 Indefinites and Constraints on Scope and Anaphora ...... 129
   5.1 The notions of binding and anaphoric relatedness .... 130
   5.2 Constraints on scope and coindexing ...................... 132
   5.3 Coindexing and scope constraints in view of the variable analysis of indefinites. ......................... 140
   5.4 The exceptional behavior of “specific” indefinites. .... 147
Chapter I

PROBLEMS CONCERNING INDEFINITES AND ANAPHORA IN LOGICAL SEMANTICS

Logical semantics, i.e., the analysis of natural language in terms of concepts of modern logic, is an enterprise as old as modern logic itself. Logical semanticists have been using such notions as quantification, variable binding, and scope in trying to characterize the semantics of words and constructions in natural languages. In the realm of noun phrase semantics, for instance, they have studied the quantificational force of various determiners and have analyzed certain semantic relations between noun phrases and pronouns as variable binding.

While some of the earlier work in this tradition was directed at purposes of no concern to the linguist (as when philosophers hoped to expose their colleagues’ fallacious arguments more convincingly by applying logical analysis to them), there has been increasing interest in logical semantics as a contribution to linguistic theory, more specifically, a contribution to a theory of one subsystem of grammar, the (or one of the) semantic component(s). Seen in this way, logical semantics is not just about algorithms for translating one language into another language, the former natural, the latter artificial. Rather, it implies a commitment to the hypothesis that the grammar of a natural language does in fact include a component in which expressions are analyzed as quantifiers, as variables, as having a certain scope, as standing in binding relations to each other, and the like. In other words, the logical semanticist hypothesizes that the ability to assign to English sentences the logical analyses he or she is proposing is part of knowing English.

By and large, logical semantics has been a very successful enterprise, especially with regard to such topics as the semantics of determiners and the semantics of pronouns, including such issues as are central to this dissertation. Analyses in terms of quantifiers and variables not only have captured the intuitive truth conditions of sentences involving those expressions. They
have also made it possible to formulate certain principles that relate the structural properties of sentences to the range of readings of which they admit. The existence of these principles is an entirely contingent fact about natural language, and the finding that they make reference to aspects of the logical analysis of a sentence gives justification (of a sort that could not be anticipated in the early days of logical semantics) to the hypothesis that grammars of natural languages do in fact assign logical analyses to sentences.

Notwithstanding the successes alluded to, there remains a residue of data that are widely believed to have not received an entirely satisfactory treatment in logical semantics. This chapter is about those data and the attempts at accounting for them that have been made. Actually, the scope of the chapter is more limited than that in at least two respects. Firstly, I will focus only on examples that exhibit pronominal anaphora to an indefinite antecedent which is singular and contains the indefinite article "a/an." I will not consider the various types of plural indefinites which show to some extent an analogous behavior. (The limitation to singular NPs will in fact be maintained throughout the dissertation.) I will also avoid talking about certain related phenomena involving a definite description as the antecedent (but see Chapter II, section 6). These omissions will help to keep the discussion perspicuous, but they may also have permitted me to jump to conclusions that will not survive future scrutiny.

Secondly, I will not include in this chapter all existing approaches to the problems under discussion, not even all that are known to me and that fall under the general orientation of logical semantics. Most notably, I will exclude David Lewis’ proposal in “Adverbs of Quantification” (1975) from the discussion here, because I see it as a direct precursor of my own proposal and want to present it as such in Chapter II. I will also exclude Hans Kamp’s paper, “A Theory of Truth and Semantic Representation” (1981), whose leading ideas are very similar to my own and were developed independently at approximately the same time. There will be a few brief remarks on Kamp’s work in comparison to mine in later chapters, but no attempt will be made to present arguments that would bear on a choice between the two. I hope this will be possible at some future time. An approach I will discuss, though without coming to a definitive judgment about it, is the game-theoretical approach (also a variety of “logical semantics,” as I would apply that label) that Jaakko Hintikka and Lauri Carlson advocate in their article, “Conditionals, Generic Quantifiers, and other Applications of Subgames” (1979). This approach differs rather fundamentally from the ones to be discussed in this chapter, and one also does not perceive it as related to my own approach. I will append what little I have to say about it to the very end of the chapter.
Given all these limitations, it should be clear that the purpose of this chapter cannot possibly be to convince the reader that the problems addressed in it have remained unsolved to this day. No such conclusion could be based on a survey of proposals that neglects precisely the most promising ones. Rather, one of the purposes is to review some facts that will later enter into the motivation of my proposal, and another purpose is to point out weaknesses in some, though not all, of the proposals that are currently being advocated. As for the latter aim, let me add another caveat: The weaknesses that I will find in some of the proposals discussed are by no means dramatic, and they do not justify rejecting them out-of-hand. It remains to be seen what alternatives there are, and at what cost they overcome the weaknesses in question. And, needless to say, this evaluation by comparison must be based on more criteria than mere coverage of data. So this chapter cannot possibly convey all of the reasons I see for choosing the position I will in fact adopt over the available alternatives. More reasons will have to emerge in the course of the following chapters.

1 Do Indefinites Refer?

There is a familiar set of arguments to the effect that indefinites cannot be referring expressions, which I will review very briefly. Then there is one commonly used argument which appears to favor the opposite conclusion. This argument relies on certain assumptions concerning the nature of anaphoric relations between an indefinite antecedent and an anaphoric element. I will devote the bulk of this section to clarifying those assumptions and the argument based on them, and to various attempts at invalidating it that have been made.

1.1 Russell’s view

What is the meaning of a noun phrase with the indefinite article, e.g., the noun phrase “a dog” as it occurs in the following sentences?

(1) A dog came in.
(2) John is friends with a dog.

According to a most widely held view, originally advocated by Russell1, such indefinite NPs (for short: “indefinites”) have the meanings of existential

---

1 In various writings, especially Russell (1919, Ch. 16).
quantifiers. (1) is thus analyzed as meaning that the set of individuals that are both dogs and came in is not empty. And (2) is taken to mean that the set of individuals that are both dogs and friends with John is non-empty. On this view, indefinites do not refer: “a dog” in (1) does not denote any particular dog any more than “every dog” in (3) does:

(3) Every dog came in.

Arguments in favor of Russell's view include the following:² Consider a sentence like:

(4) John is friends with a dog, and Mary is friends with a dog.

If “a dog” referred to any particular dog, then (4) would have to be understood as saying about that dog, the referent of “a dog,” that John is friends with it and so is Mary. But (4) is clearly consistent with a situation where no dog is a mutual friend of John's. Russell's analysis predicts that: According to it, (4) is true whenever neither the set of dog-friends of John nor the set of dog-friends of Mary is empty.

This argument, though repeated often, tacitly relies on an extremely dubious premise, which was first exposed by Strawson: It presupposes that “a dog” either refers to one and the same thing on each occasion where it is uttered, or else it never refers to anything at all. A third possibility, viz., that each occurrence of “a dog” refers to something, but not the same thing in each case, is plainly ignored. The argument is thus on a par with the following attempt to show that personal pronouns do not refer: Suppose that “he” referred to some individual x. Then “He likes him” would have to mean that x likes himself, which it does not mean. Ergo, “he” does not refer to anything. What this argument, as well as the analogous one about indefinites, does show is that the meaning of the expression “he” (or the expression “a dog”) is not to be identified with an individual that is supposed to be that expression's referent. But it is a huge step from this insight to Russell's view that indefinites are not referring expressions. Other familiar arguments pro Russell are harder to resist. Consider (5), the negation of (1):

(5) It is not the case that a dog came in.

If “a dog” referred to anything, we would expect (5) to be denying that that thing came in, just like (6) denies that the referent of the name “Fido” came in:

² These arguments are found in Quine (1960) and Kaplan (1970), among other places.
(6) It is not the case that Fido came in.

But (5) expresses a much stronger claim than that: It says that no dog whatsoever came in. And this happens to be precisely what Russell’s analysis predicts: Given that (1) means that the set of dogs that came in is not empty, (5), the negation of (1), means that it is not the case that the set of dogs that came in is not empty, or equivalently: (5) means that the set of dogs that came in is empty.

Next, consider:

(7) Every child owns a dog.

If "a dog" referred to anything, then (7) should amount to claiming about that thing that every child owns it, just as (8) claims that every child owns Fido.

(8) Every child owns Fido.

Again, Russell’s analysis of (7) is preferable. It predicts that (7) means that for every child \( x \), the set of dogs owned by \( x \) is non-empty. Considerations like these seem to have convinced not only several generations of philosophers, but also a great many linguistic semanticists that are currently working in either the “Montague Grammar” or the REST framework.\(^3\) Even those who are not convinced that all uses of indefinites are properly analyzed by Russell’s theory have mostly conceded that it captures at least one of the readings that an indefinite can take on.

It has often been pointed out (e.g., in Strawson (1952)) as a difficulty for Russell’s analysis of indefinites that indefinites can serve as antecedents for anaphoric pronouns. For instance, (1) can be continued as in (9):

(9) A dog came in. It lay down under the table.

The pronoun “it” in this text presumably refers to something. But what does it refer to? The first answer that comes to mind is that it refers to the same dog as “a dog” in the preceding sentence. But for that answer to make any sense, “a dog” itself must, contra Russell, refer to some dog. This consideration of course does not prove that "a dog" in (9) refers; it merely shows that, if we assume that “a dog” in (9) refers, we can give a simple and sensible account of what “it” means in a text like this, whereas the role of “it” would seem unaccounted for and mysterious under Russell’s assumption that “a dog” does not refer. So we need not yet conclude that the Russellian

---

\(^3\) Montague (1974), Barwise and Cooper (1981), May (1977), to mention only a few.
view is wrong; but we have to acknowledge that its advocates owe us an alternative account of anaphoric pronouns with indefinite antecedents if we are to remain convinced.

There are various ways in which Russelians have risen to this challenge, and I will sketch the most important lines of approach in the next few sections.

1.2 Anaphoric pronouns as bound variables

Example (9) remains puzzling as long as we accept Russell's analysis of the indefinite while presuming that the pronoun must refer. But why should it? It is well recognized that reference is not the only function that natural language pronouns can serve, and that there are cases where they clearly must be understood as bound variables. (9) might be just another case like that. Indeed, we predict intuitively correct truth conditions for this text if we analyze it as saying that the set of dogs that came in and lay down under the table is non-empty, i.e., as equivalent to:

\[(9') \exists x (x \text{ is a dog } \land x \text{ came in } \land x \text{ lay down under the table})\]

This analysis was proposed by Geach.\(^4\) It implies as a general moral that the proper unit for the semantic interpretation of natural language is not the individual sentence, but the text. (9') provides the truth condition for the bisentential text as a whole, but it fails to specify, and apparently even precludes specifying, a truth condition for the individual sentence "It lay down under the table," which appears as part of (9).

Geach's analysis is quite unpopular these days and is sometimes held to have been amply refuted by the following considerations: First of all, how is this analysis supposed to deal with dialogues like the following (example (10) is due to Strawson).\(^5\)

\[(10)\]
\[\begin{align*}
 a. & \quad \text{A man fell over the edge.} \\
 b. & \quad \text{He didn't fall; he jumped.}
\end{align*}\]

\[(11)\]
\[\begin{align*}
 a. & \quad \text{A dog came in.} \\
 b. & \quad \text{What did it do next?}
\end{align*}\]

With (10), where the two speakers contradict each other, an analysis of the whole dialogue as a unit could only amount to assigning it the necessarily false truth condition:

\(^4\) Geach (1962, 126ff).

\(^5\) Strawson (1952, 187).
(10') \[ \exists x (x \text{ is a man} \land x \text{ fell over the edge} \land \neg x \text{ fell over the edge} \land x \text{ jumped over the edge}) \]

One can hardly object to the claim that this is indeed the truth condition of the dialogue as a whole. But that just goes to show that there is rather limited interest in an analysis that is confined to assigning truth conditions to whole discourses only. It is intuitively clear about (10) that each speaker has expressed a contingent proposition and that we can imagine circumstances under which we say, e.g., that A said something false and B said something true. There has to be something wrong with a theory which implies that neither utterance has any truth conditions of its own. (11) emphasizes basically the same point. In this case, B’s utterance of course does not have truth conditions, since it is a question. Rather, it has a question-meaning, whatever that may be (several plausible proposals are in existence, among them that question meanings are sets of propositions). The relevant point is that Geach’s approach to anaphoric pronouns would force us to look for a meaning of the dialogue as a whole, but one cannot even imagine what type of meaning that would be.

Let us be clear about the force of this objection before we move on to the next. Examples like the ones just considered do not show that Geach’s analysis makes false predictions; they merely show that it does not make enough predictions. In particular, it falls short of predicting truth conditions for the full range of units of discourse that are capable of truth or falsity in the pretheoretical sense. So the objection is a challenge to, not a refutation of, the Geachian position, and it stands as an objection only as long as no one comes up with a truth criterion for sentence-sized parts of utterances that is consistent with Geach’s treatment of intersentential anaphora as variable binding. I emphasize this point, because I will later have to face the same objection with respect to my own approach, and will then try to argue that the challenge can indeed be met (see Chapter II, section 3, and Chapter III, section 3, below).

An additional, independent objection to Geach’s analysis is due to Evans\(^6\) who claims that it even predicts inadequate truth conditions in some of the cases where a multisentential discourse basically does amount to the assertion of one complex proposition. His most striking examples are those involving a plural indefinite that serves as an antecedent to a plural pronoun. Plural indefinites presumably mean something like “there are at least two.”

\[ (12) \text{ John owns some sheep.} \]

---

\(^6\) Evans (1980, 339).
16 | Problems Concerning Indefinites and Anaphora

Apparently (12) is true if and only if the set of sheep that John owns has at least two members. Given this, what truth conditions would Geach’s analysis lead us to expect for (13)?

(13) John owns some sheep. Harry vaccinated them.

By analogy with the singular examples, (13) should amount to the assertion that the set of sheep that John owns and that Harry vaccinated contains at least two members. Consequently, (13) ought to count as true in a situation where John owns six sheep and Harry vaccinates three of them, but does not vaccinate the other three. But intuitively, we are inclined to consider (13) false in this situation. We tend to read (13) as meaning that John owns some sheep and Harry vaccinated all the sheep that John owns.

Evans even goes so far as to claim that a Geachian analysis predicts inadequate readings in cases with singular indefinites. He says that if such an analysis were correct, then we should be able “to report the non-emptiness of the class of Welsh doctors in London by saying:

(14) There is a doctor in London and he is Welsh.”

But (14) is undoubtedly inappropriate for conveying the message. Contrast this with (15), which would be appropriate:

(15) There is a doctor in London who is Welsh.

Geach’s problem is that he analyzes (14) in exactly the same way as (15) and thus fails to explain the difference. (I will take a closer look at the contrast between (14) and (15) below and will suggest that Evans himself does not have a satisfactory explanation for it either. But that is an independent point, and it does not change the fact that the observed contrast casts doubt on the correctness of Geach’s analysis.)

There is a third objection that is often brought forward against analyses which, like Geach’s, rely on attributing to indefinites a scope that may encompass as much as an entire text. The objection is that an analysis of this type necessitates ad hoc qualifications to what would otherwise be a uniform theory of structural constraints on quantifier scope. (This is of course not an absolute inadequacy of the analysis under consideration, but merely a relative disadvantage in comparison with certain competing analyses, which we have yet to discuss.) The objection draws on the observation that many of the examples where an existential antecedent supposedly binds an anaphoric

---

7 Evans (1980, 343)
pronoun do not have grammatical analogues with other quantifiers in the antecedent position. The contrast between (9) on the one hand and (16) or (17) on the other is typical:

(9) A dog came in. It lay down under the table.
(16) Every dog came in. It lay down under the table.
(17) No dog came in. It lay down under the table.

Neither (16) nor (17) permits an interpretation where “it” is bound by the quantified NP in the preceding sentence. (In fact, these texts cannot be felicitously used unless a referent for “it” is fixed in some way that is completely independent of the utterance of the preceding sentence.) The generalization behind this fact is that an unembedded sentence is always a “scope-island,” i.e. a unit such that no quantifier inside it can take scope beyond it. This generalization (which is just a special case of the structural restrictions on quantifier-scope and pronoun-binding that have been studied in the linguistic literature\(^8\)) is only true as long as the putative cases of pronouns bound by existential quantifiers under Geach’s analysis are left out of consideration. Therefore, if Geach’s analysis of (9) as an instance of variable binding is adopted, the conditions have to be modified in appropriate ways to predict the seemingly exceptional behavior of existential quantifiers. To be sure, this point can only be considered an objection to the Geachian analysis if it is suspected that such a modification would result in a considerable loss of explanatory power of the overall theory, and that this can be avoided by adopting a different analysis. I postpone the issue until section 2.1.2., where I’ll suggest that the objection does not hold up in light of a more fully developed version of Geach’s approach.

1.3 Anaphoric pronouns as picking up a speaker’s reference

A radically different approach to the relation between an indefinite antecedent and a pronoun anaphoric to it was originally suggested by Grice.\(^9\) The suggestion was taken up by Kripke (1977), on whose paper I will base my discussion. Let us recall once more how the puzzle originally presented itself (e.g., in Strawson’s or Geach’s exposition of it): On the one hand, anaphoric

---

\(^8\) Various theories of scope-island restrictions and restrictions on bound-variable readings of pronouns are found in Rodman (1976), Reinhart (1976), May (1977), and Higginbotham (1980).

\(^9\) Probably in a talk that Grice gave at some conference in 1971, and perhaps also in the 1967 William James Lectures on “Logic and Conversation.”
pronouns seem to pick up the reference of their antecedents; on the other hand, Russell seems to have shown that indefinites do not have a reference. Therefore, they should not be able to serve as antecedents for anaphoric pronouns, but in fact, they are.

Unlike Geach, who attempted, to get out of the dilemma by denying that anaphora need have anything to do with “picking up a reference” previously made, Kripke thinks that there is not really a dilemma at all; there only appears to be one because two distinct notions of reference have been confounded in the statement of the puzzle: “speaker’s reference” and “semantic reference,” as he calls them. Whether an utterance of an expression has semantic reference, and what its semantic referent is, is purely a matter of the rules that define the language in which the utterance is made. It is in this sense of “reference” that Russell’s claim about indefinites is to be understood: No utterance of an indefinite has semantic reference. But this does not mean that no such utterance has speaker’s reference. Whether there is speaker’s reference and what the speaker’s referent is depends on the intentions that the speaker of the utterance happens to have. The speaker’s referent is that individual which the speaker “wishes to talk about” (or “has in mind”) on the occasion of the utterance. The notion applies to utterances of indefinites as follows:

When a speaker asserts an existential quantification, \((\exists x)(\phi x \land \psi x)\), it may be clear which thing he has in mind as satisfying ‘\(\phi x\)’ and he may wish to convey to his hearers that that thing satisfies ‘\(\phi x\)’. In this case, the thing in question (which may or may not actually satisfy ‘\(\phi x\)’) is called the ‘speaker’s referent’ when he makes the existential assertion. (Kripke, 1977, 17)

So it is perfectly conceivable for an utterance that lacks semantic reference to have speaker’s reference nevertheless, and this might even be the typical situation with indefinites.

Only one additional assumption is needed now to reconcile Russell’s analysis of indefinites with the naive view that anaphoric pronouns pick up the reference of their antecedents: If we assume that they pick up the speaker’s reference of their antecedents, i.e., that an anaphoric pronoun’s semantic referent is (or at least, may be) the speaker’s referent of its antecedent, then examples like (9) are naturally accounted for, with Russell’s claims unaffected and the naive identification of anaphora with coreference basically vindicated. In Kripke’s words (1977, 26, n32):

Often one hears it argued against Russell’s existential analysis of indefinite descriptions that an indefinite description may be anaphorically referred to by a pronoun that seems to preserve the reference of the indefinite description. I am not sure that these phenomena do conflict
with the existential analysis. (I am not completely sure there are some that don't, either. [sic]) In any event, many cases can be accounted for (given a Russellian theory) by the facts that: (i) existential statements can carry a speaker’s reference; (ii) pronouns can refer to the speaker’s referent.

As Kripke furthermore notes, it is not at all an unnatural assumption about pronouns that they should be able to get their reference in this way, which depends on pragmatic rather than purely semantic factors. After all, pronouns are also capable of getting their reference by virtue of such purely pragmatic factors as an object's (perceptual or merely associative) salience.

Kripke, if I understand his remark correctly, deliberately refrained from claiming that all instances of a pronoun anaphoric to an indefinite can be accounted for by adopting his suggestion. In fact, the scope of his proposal may be much more limited than is often thought. Consider the following sentence:

(18) A dog has been rummaging in the garbage can.

I may utter this sentence in a situation where I have not witnessed the event directly, but am merely inferring what happened on the basis of the mess I am seeing. I have no idea which dog is responsible, so there presumably is no speaker’s referent in Kripke’s sense. Nevertheless, I can naturally go on to utter (19):

(19) It has torn open all the plastic bags.

The pronoun “it” in this utterance, certainly an anaphoric pronoun in the traditional sense of “anaphoric,” cannot be analyzed as picking up a previous speaker’s reference, because there was none. (Neither can it be analyzed as a bound variable in the scope of its antecedent, unless we want to burden ourselves with an analysis more or less like Geach’s after all.)

Lewis (1979) sketches another essentially Gricean treatment of anaphoric expressions (pronominal and other) with indefinite antecedents that is similar to Kripke’s, but might have more potential of extending even to cases like our example involving the utterance of (18) and (19). According to Lewis, a pronoun may refer to whatever object is maximally salient in the situation of its utterance. Anaphoric pronouns are a special instance of this, as one method of raising the salience of an object is by producing a suitable utterance. An utterance may be suited to raising an object's salience by containing an expression which (semantically) refers to that object; but it may also be so suited for reasons other than its involving any reference. Thus an existential statement may raise an object's salience in the following way:
It is worth mentioning another way to shift comparative salience by conversational means. I may say 'A cat is on the lawn' under circumstances in which it is apparent to all parties to the conversation that there is some one particular cat that is responsible for the truth of what I say, and for my saying it. Perhaps I am looking out of the window, and you rightly presume that I said what I did because I saw a cat; and further (since I spoke in the singular) that I saw only one. What I said was an existential quantification; hence, strictly speaking, it involves no reference to any particular cat. Nevertheless it raises the salience of the cat that made me say it. Hence this newly-most-salient cat may be denoted, by brief indefinite descriptions, or by pronouns, in subsequent dialogue: 'No, it's on the sidewalk.' 'Has Bruce noticed the cat?' ... Thus although indefinite descriptions – that is, idioms of existential quantification – are not themselves referring expressions, they may raise the salience of particular individuals in such a way as to pave the way for referring expressions that follow. (Lewis, 1979, 180)

Now the particular example that Lewis discusses in this passage is probably one where Kripke's notion of speaker's reference applies: The speaker who is looking out the window and seeing a cat certainly had that cat "in mind as satisfying 'x is a cat' " – although I find it somewhat less obvious that he "wished to convey to his hearers that that cat was on the lawn," or at least I can think of a way of reading the latter as ascribing him a wish that he might well have lacked. Anyway, let us assume this is a case of Kripkean speaker's reference. Then it would not seem to matter much whether we, account for the anaphoric pronoun that follows in Kripke's or in Lewis' terms. Either one would be satisfactory.

However, as we reconsider the example about (18), (19) in light of Lewis' remark, it might not present a problem at all. It could be argued that my utterance of (18) does raise the salience of a particular dog, namely the dog that "is responsible for the truth of what I say, and for my saying it," in other words, the dog that did rummage in the garbage. Just because I do not know anything else about this dog (what it looks like, what its name is, who it belongs to, ...) does not mean it cannot become the most salient dog in the situation. So the anaphoric pronoun in (19) could after all be accounted for in much the same way as in the unproblematic cases of indefinites uttered with a speaker's reference.

What if my utterance of (18) is false though, i.e., if the mess I am contemplating is in fact the result of a strong blast of wind? Which object has then been raised in salience? The answer is presumably: none. Thus, if I uttered (18) falsely, then I did not refer to anything when uttering "it" in (19), and my utterance of (19) did not express any proposition at all (although it would have expressed one, had the world in which my utterance took place been as I thought it was). I have no intuitive disagreement with this answer.
What, on the other hand, if my utterance of (18) were true, but true because a couple of dogs were in fact the offenders? Again, no dog can then have been promoted to maximal salience by my utterance, and the subsequent utterance of (19) must be considered a case of failure of reference, with no proposition expressed. Intuitions are quite unclear in cases like this. Sometimes I find it plausible that the utterance of (19) should be judged true just in case one of the dogs that rumbled in the garbage can also tore open all the plastic bags, and it should be judged false if neither one of the dogs that rumbled in the garbage can also tore open all the plastic bags. But that is certainly not the way the utterance was meant by the speaker; if she had only thought of the possibility that more than one dog was involved (which she obviously did not) she would not have committed herself to the completely unwarranted claim that only one of them tore open the bags while the other one stood by. At other times, I tend to think that the whole discourse, including (19), should be judged true, regardless of how many dogs rumbled in the garbage can and tore open the bags, as long as any did. But I certainly would not uphold this judgment if the discourse had run as follows:

(18) (as above) [A dog has been rumbled in the garbage can.]
(20) It was a very patient dog, as you can see from the fact that it tore open every single bag.

I think there is little point in scrutinizing intuitions in cases of this sort any further. They are insecure enough for us to accept a theory that predicts automatic truthvaluelessness whenever the existential antecedent is verified by multiple instances, as long as there is not anything else that is wrong with that theory. Against the Grice-Kripke-Lewis account of intersentential anaphora involving indeinite antecedents, I would like to bring to bear a consideration concerning minimal pairs like the following:10

(21) a. I dropped ten marbles and found all of them, except for one. It is probably under the sofa.
   b. I dropped ten marbles and found only nine of them. It is probably under the sofa.

We might imagine the first sentence of (21b) to be uttered in a situation which is as suited as any to make the tenth, missing marble salient (say, everybody is engaged in the search). Still, the utterance does not seem to be a successful means of raising the salience of the tenth marble to a

---

10 Examples are due to Barbara Partee.
degree of salience that would suffice for the subsequent pronoun to refer to it. By contrast, an utterance of the first sentence of (21a) does have that capacity. If we adopt Lewis’ position on the logical analysis of indefinites and on the pragmatic preconditions for successful pronominal anaphora, then we are compelled to conclude that the salience-shifting potential of an utterance is not predictable from its truth-conditions and the surrounding circumstances alone; it moreover depends on how the utterance is worded. For example, notice that the first sentences of (21a) and (21b) coincide in truth conditions, and that the situations in which they are uttered are by hypothesis identical. (If we took Kripke’s, rather than Lewis’ view, a slightly different, but essentially analogous, conclusion would be inevitable. I will stick to Lewis’ terms, but take it that my point carries over.)

The conclusion that was just drawn raises a question: How exactly does wording influence the capacity of an utterance to raise an individual’s salience? Which properties of the sentence uttered are crucial, and what are the relevant generalizations? The phenomenon that (21) illustrates strikes me as a systematic one, and it calls for a systematic account. A first attempt to formulate the generalization of which our example is an instance might be this:

A necessary condition for an utterance of a sentence S to promote an object x to maximal salience is that S contain either an NP that refers to x or a singular indefinite NP whose predicate is true of x.

This would discriminate between (21a) and (21b) in that the first sentence of (21a) contains the singular indefinite “one,” here to be read as “one of the marbles,” whose predicate amounts to “is one of the marbles” and is true of the tenth marble, the intended referent of the subsequent pronoun. The first sentence of (21b), on the other hand, contains neither a suitable singular indefinite, nor an expression that refers to the missing marble, so it is ruled out by the above generalization as a means of making that marble maximally salient.

If this generalization, or a more or less refined version of it, holds in fact true, we should wonder why. Does it need to be stipulated, or can it be explained? It seems to me that under the assumptions that Lewis is committed to, there is no way of relating the ability of singular indefinites to contribute to shifts in salience to any other characteristic property they have. Now it is not inconceivable that the salience-shifting potential of an expression is indeed a property that natural languages specify independently of the logical and other properties of that expression. But we would ceteris paribus prefer a theory which predicts it to correlate with other properties.
We have to note in this connection that the need for a separate stipulation disappears if we abandon Lewis’ or Kripke’s account of cross-sentential anaphora and take, say, Geach’s approach. For Geach, the relevant reading of (21a) involves binding of the pronoun “it” by the text-scope existential quantifier “one (of the marbles).” (21b) cannot share that reading because there simply is no existentially quantifying NP to bind the pronoun. The contrast falls out from Geach’s assumptions about the logical analysis of indefinites and their relation to anaphoric pronouns, with no additional stipulations needed. Also, if we assume, contra Lewis and Kripke, that some occurrences of indefinites refer and that intersentential anaphora relations involving indefinite antecedents arise when the pronoun picks up the reference of its antecedent (a view that I have attributed to Strawson), the contrast in (21) can be accounted for much more naturally: (21b) simply does not contain an expression that refers to the missing marble and from which the pronoun could pick up that reference.

Much the same considerations apply to contrasting pairs like the following:

(22) a. John owns a bicycle. He rides it daily.  
b. ? John is a bicycle-owner. He rides it daily.

(23) a. John has a spouse. She is nice.  
b. ? John is married. She is nice.

Again, we find utterances of identical truth conditions exhibiting what Lewis would have to diagnose as different capacities for raising the salience of an object. And again we have a systematic phenomenon: Words (including compound words) are always “anaphoric islands” in the sense of Postal (1969).

To sum up my argument, the Grice-Kripke-Lewis approach to reconciling Russell’s quantificational analysis of indefinites with their ability to serve as antecedents to anaphoric pronouns implies that a grammar of English will not only have to assign logical analyses to expressions of English, but in addition, and independently, will have to specify their salience-raising potential. This puts the approach at a prima facie disadvantage in comparison with alternative approaches that manage to get away with assigning logical analyses alone. Certainly this argument does not in any way refute the Gricean type of approach, since the one weakness it points to could turn out to be minor compared to the weaknesses of its competitors.
1.4 Anaphoric pronouns as disguised definite descriptions

There exists a third approach to anaphoric pronouns with definite antecedents that also attempts to maintain Russell’s view that indefinites do not refer, but avoids both Geach’s conclusion that such pronouns are bound variables, and the Gricean suggestion that they receive reference through pragmatic factors to which the utterance of the antecedent contributes only indirectly. This third approach is best represented by Evans.\(^{11}\) Evans acknowledges the existence both of bound-variable pronouns and “pragmatic” pronouns (i.e., those referring by virtue of their referent’s salience), but maintains that the anaphoric pronoun in our example (9) (repeated here) falls under neither of these two types, but under a separate type, which he calls “E-type”.

(9) A dog came in. It lay down under the table.

E-type pronouns always have quantified NPs as their antecedents, but they are not bound by them. Rather, their meaning is that of a definite description which is determined in a certain systematic way by the minimal sentence in which the antecedent occurs. For instance, the sentence “A dog came in,” in which “a dog” occurs, determines the definite description “the dog that came in,” and so that is the meaning of the subsequent “it,” when interpreted as an E-type pronoun. (9) is thus analyzed as meaning the same as:

(9’) A dog came in. The dog that came in lay down under the table.

Here are some more examples of E-type pronouns and the corresponding definite descriptions in terms of which Evans proposes to analyze them (some of the examples are found in Evans (1980)):

(24) Just one man drank champagne. He was ill.
(24’) Just one man drank champagne. The man who drank champagne was ill.

(25) Few congressmen admire Kennedy. They are very junior.
(25’) Few congressmen admire Kennedy. The congressmen that admire Kennedy are very junior.

(26) John gave a book to a student. She didn’t return it.
(26’) John gave a book to a student. The student who John gave a book to didn’t return the book John gave her.

\(^{11}\) Evans (1977; 1980).
It should be clear intuitively from these examples how the appropriate
definite description is constructed from the sentence in which the antecedent
occurs. (It is possible, but not worthwhile for our present purposes, to specify
an exact procedure for this construction.\(^{12}\))

A paraphrase is not yet an analysis, and we cannot assess Evans’ proposal
unless we know how he intends to interpret definite descriptions. He
assumes that a sentence containing a singular definite description with
the predicate “F” implies that there is exactly one object that is F and is
true if that unique F satisfies whatever the sentence predicates of “the F.”
(Whether “imply” is the same as “presuppose,” or more like “entail,” and
whether falsity of the implication results in falsity or in truthvaluelessness
of the assertion, is not specified by Evans and seems indeed to make little
difference here.) As for plural definite descriptions with the predicate “F,” he
assumes the implicature that there are at least two Fs, and that the sentence
containing “the F” (plur.) is true if all Fs satisfy the relevant predicate. These
are commonplace assumptions, of course. Consequently, the predictions
concerning the interpretations of E-type pronouns come out accordingly:
For instance, when “It lay down under the table” is uttered as part of the text
(9), it is predicted to imply that exactly one dog came in; it is furthermore
predicted to be true just in case the unique dog that came in lay down under
the table.

The prediction that E-type pronouns carry uniqueness-implications of
this sort is prima facie controversial, and Evans devotes some remarks to
defending it. First of all, he certainly does not mean to commit himself to the
claim that the object in question should be the only one in the entire universe
that satisfies the predicate restricting the description. Any quantification in
ordinary conversation requires that we tacitly agree on a domain of relevant
instances of evaluation and ignore the countless irrelevant ones that would be
included under a strictly literal reading of the quantifier. In the same way, the
range of objects within which a definite description, or an E-type pronoun,
is presupposed to denote uniquely will always be a pragmatically determined
range of relevant individuals. Take an example like (27):

(27) A wine glass broke last night. It (i.e., the wine glass that broke last
night) had been very expensive.

We must not misunderstand Evans as claiming that (27) carries an implica-
tion that in the entire world only one wine glass broke on the night before
the utterance. Rather, the predicted implication is probably that exactly

\(^{12}\) An exact procedure is spelled out in \textit{Parsons} (1978).
one wine glass broke on that night in the household where the speaker was staying, or in a similarly confined contextually relevant place. (The same restriction would be needed to arrive at a realistic interpretation of an utterance like, “Every wine glass was refilled at least twice last night.”)

Still, Evans I claim that E-type pronouns carry uniqueness-implications is a strong claim, even with the above qualifications. It predicts that the following pairs are by no means paraphrases:

(27) (see above) [A wine glass broke last night. It (i.e., the wine glass that broke last night) had been very expensive.]
   a. A glass which had been very expensive broke last night.

(14) (see section 1.2 above) There is a doctor in London and he is Welsh.
   a. There is a doctor who is Welsh in London.

(Recall that these were pairs of paraphrases under Geach’s analysis.) In the case of (27) vs. (27a), I have no strong intuitions as to whether the meanings differ. Evans predicts that (27a) does not imply that just one glass broke on the night in question (within the relevant range of glasses), whereas (27) does imply this. I can well imagine the following scenario: I dropped a tray with three glasses last night, breaking them all. Two of them had been cheap and I did not mind losing them, but the third one had cost a lot and I was upset at its loss. The next morning, a friend asked me why I was in such a bad mood, and I answered by uttering (27). The utterance was, to me, perfectly appropriate. But I am not sure whether it constitutes a counterexample to Evans’ analysis. He might simply say that, in the context of my friend’s question, the cheap glasses were irrelevant because my mood was not influenced by their loss. This would explain how (27) and (27a), even if interpreted as differently as Evans suggests, end up being interchangeable after all, in the situation described, and probably for similar reasons in most (if not all) other situations as well.

A more perceptible intuitive contrast emerges in (14) vs. (14a), and indeed a contrast that appears to support Evans’ analysis: As we noted earlier (when assessing Geach’s analysis; cf. section 1.2), (14a) is an acceptable way of claiming that at least one person is both Welsh and a London doctor, whereas (14) is not. Evans offers an explanation for this lack of interchangeability: (14) carries the inappropriate implication that there is just one doctor in London, which is absent from (14a). However, it seems to me that this explanation is both unconvincing and replaceable. Imagine the following dialogue, which provides a natural context for (14a):

(28) A: This boy is sick, but he speaks only Welsh and we can’t understand his complaints.
B: There is a doctor who is Welsh in London = (14a)

If we replace (14a) by (14) in (28), the result is odd. But does Evans’ analysis really predict this? Doesn’t the preceding utterance by A establish a context in which only Welsh-speaking doctors qualify as relevant in the first place? If this were assumed, then we would expect (14) in this context to imply merely that there is just one Welsh-speaking doctor in London, and that would not be very awkward at all. So if the awkwardness of (14) as a response to (28) is indeed, as Evans would have it, due to the fact that (14) implicates that London has just one doctor, then we must conclude that domains of quantification do not respond to relevancy criteria quite as flexibly as that. But then this casts doubt on the explanation given above for why (27) is not as inappropriate as we might have thought in a situation where I dropped three glasses. It is hard to see why the cheap glasses should not have counted as relevant in the dialogue described above, whereas the non-Welsh London doctors should count as relevant in (28). Explanations of acceptability judgments that have to rely on mysterious differences like this are not very convincing. Furthermore, there is a different reason to which one might attribute the oddity of (14) in most ordinary contexts: The first conjunct of (14), “There is a doctor in London,” asserts something that is so well known that one can hardly conceive of a context in which there would be any point in asserting it. So it seems plausible that (14) is ruled out by the conversational maxim that one should not assert what is already presupposed. In order to discriminate between (14) and (14a) in the application of this maxim, we presumably must assume that in an assertive utterance of the form “S₁ and S₂,” S₁ and S₂ each count as asserted individually, whereas an utterance of a complex sentence containing a restrictive relative clause is just one assertion, no parts of which must individually conform to the pragmatic appropriateness-criteria for assertions. Note that this line of explanation for the difference in acceptability between (14) and (14a) in no way appeals to a uniqueness implication of the sort that Evans attributes to (14).

I am inclined to deny any systematic validity at all to Evans’ predictions concerning uniqueness-implications. Even with an example like (14), a seemingly insignificant variation makes any suggestion of uniqueness disappear:

(29) There once was a doctor in London. He was Welsh ...

In view of the foregoing considerations, I think that the association of anaphoric interpretations with uniqueness-implications is not well supported by the linguistic facts. This suggests that there is something wrong with Evans’ analysis of the kind of anaphoric pronoun he classifies as “E-type,” but it presently remains unclear what specifically is wrong with it. Re-
call that I have attributed to Evans two assumptions which are independent of each other: (a) the assumption that certain anaphoric pronouns mean the same thing as certain definite descriptions, and (b) the assumption that definite descriptions are to be analyzed in a certain way, which involves predicting uniqueness-implications for singular definite descriptions. As it turns out upon closer investigation of the facts, it is (b) and not (a) that we should question: To the extent that it seems acceptable to utter (27) (“A wine glass broke last night. It had been very expensive.”) without committing oneself to the belief that exactly one wine glass broke last night, it also seems appropriate to utter (27b) without implicating uniqueness:

(27) b. A wine glass broke last night. The wine glass that broke last night had been very expensive.

Likewise, (29a) is equally free from suggestions of uniqueness as (29):

(29) a. There once was a doctor in London. The doctor in London was Welsh.

(Neither of these sound very natural, but the same point can be made by observing the absence of uniqueness-implications with the definite descriptions in (30):

(30) A wine glass broke on Monday night, and a beer glass broke on Tuesday night. The wine glass that broke on Monday night was my favorite wine glass, and the beer glass that broke on Tuesday night had been very expensive.)

So if we replaced Evans’ assumption (b) by a more adequate analysis of definite descriptions, one which does not necessarily predict uniqueness-implications, could we thereby save Evans’ theory of E-type pronouns? Perhaps we could, but the problem is that a more adequate analysis of definite descriptions of the sort needed is not readily available, and the task of developing one faces essentially the same difficulties as the task we have been discussing all along, viz., to account for anaphoric pronouns with indefinite antecedents. Definite descriptions that are anaphoric to indefinites probably work more or less in the same way as pronouns anaphoric to indefinites; and once we have settled on the right theory for the former, it might well carry over to the latter, or vice versa. But while we are in the process of looking for that theory, Evans’ proposal to paraphrase anaphoric pronouns away in favor of anaphoric definite descriptions is not a solution – unless we agree with his view that definite descriptions, including those used anaphorically, have been treated satisfactorily in the tradition to which he subscribes.
1.5 Anaphoric pronouns and the ambiguity hypothesis

While I have so far dealt with various attempts to show how indefinites can be existential quantifiers as Russell suggested, and still serve as antecedents for anaphoric expressions, I will now examine the views of those who have concluded that, to the extent that indefinites participate in anaphoric relations, they simply defy Russell’s analysis. This view goes back to Strawson (1952), where it is suggested that anaphoric relationships like the one in (9) (“A dog came in. It lay down under the table.”) are instances of plain coreference, and that we must therefore acknowledge that indefinites can, at least sometimes, be used to refer.

Among recent defenders of Strawson’s position, Chastain (1975) has been particularly outspoken:

Sentences containing indefinite descriptions are ambiguous. Sometimes ‘A mosquito is in here’ and its stylistic variant ‘There is a mosquito in here’ must be taken as asserting merely that the place is not wholly mosquito-less, but sometimes they involve an intended reference to one particular mosquito. (1975, 212)

Note that the claim is not that indefinites always refer. It is still acknowledged that there is a reading for indefinites under which they conform to Russell’s analysis. In fact, this seems impossible to deny, given the meanings of sentences that contain indefinites inside the scope of negation or quantifiers. One could not possibly claim, for instance, that “a dog” refers to anything when uttered as part of (31), and Russell’s analysis is unrivaled with such examples.

(31) I don’t have a dog.

While examples like (31), where something else takes wider scope than the existentially read indefinite, are the most unmistakable cases of existential readings, Chastain does not think that the existential reading emerges if and only if something else has widest scope. A widest scope existential reading is supposed to exist as well; otherwise a sentence like (32)

(32) There is a mosquito in here,

which contains no other scoped expression, would not be ambiguous, but admit only for the referential reading. But how are the two readings supposed to differ intuitively with examples of this sort? If I understand Chastain correctly, then the difference between a referentially read and an existentially read utterance of (32) is more or less the same difference as the one that Kripke would describe in terms of the presence or absence of

Do Indefinites Refer? | 29
speaker's reference: When (32) is uttered with a particular mosquito in mind, Chastain considers this a case where “a mosquito” refers to that mosquito; when (32) is uttered with no particular mosquito in mind, he considers this a case where “a mosquito” is read as an existential quantifier. The difference between Chastain and Kripke is just that Chastain does not distinguish between semantic reference and speaker's reference, but uses the notion of reference simpliciter in such a way that it is coextensive with Kripke's notion of speaker's reference.

To the extent that Chastain's view differs from Kripke's other than just terminologically, is there anything to be said for it? We are not going to end an answer to this question in any of the considerations that Chastain himself brings to bear on the issue. As far as I can see, he – like Strawson before him – relies crucially on the plausibility that the ambiguity hypothesis derives from its compatibility with a simple explanation of anaphora: An anaphoric pronoun with an indefinite antecedent (where the latter does not bind it) is coreferential with its antecedent. But it is precisely in this respect that Chastain's proposal is almost identical to and no better than Kripke's: just replace “is coreferential with” by “picks up the speaker's reference made in the utterance of,” and you have an equally straightforward explanation for anaphora that combines with Kripke's proposal and covers exactly the same range of examples. (Note that, because of this parallelism, the criticism against Kripke that was based on example-sentences (18)-(19) is equally applicable against Chastain; it can be evaded, though, by recasting Chastain's criteria for reference in such a way that an indefinite counts as referring whenever the utterance it occurs in has the effect of raising a particular object's salience in the sense of Lewis. See above for details.)

So if we stick with Chastain's own reasons – or with the reasons that any other advocate of the ambiguity hypothesis has offered in the philosophical literature I am familiar with — the ambiguity hypothesis buys us nothing, and ceteris paribus, we would of course rank it below a theory that gets around positing an ambiguity. Moreover, there are cases where Chastain's predictions fit poorly with judgments in terms of our pretheoretical notions of truth and falsity. Imagine a situation in which I utter (33):

(33) There is a mosquito in here. I can hear it buzzing.

The fact is there is no mosquito in here, and the buzzing is just an acoustic illusion. Intuitions concerning truth and falsity may be rather uncertain with respect to the second half of this utterance (i.e., the utterance of the second sentence), but they are firm with respect to the first half: My utterance of “There is a mosquito in here” was definitely false, given the facts described. (Recall that this judgment is in line with the predictions of
Kripke's analysis.) Chastain would seem to be committed to saying that this utterance of "There is a mosquito in here" involves an unsuccessful attempt at reference and is consequently neither true nor false. It could be false only if read existentially, but this is against the speaker's intentions, which are clearly signaled by the subsequent use of the anaphoric pronoun. (I suppose that Chastain would not resort to saying that an utterance may have an existential reading even though the speaker means it referentially and is understood to mean it so by the audience.) Short of accepting that the technical distinction between falsity and truthvaluelessness simply fails to match the corresponding intuitive distinction, I see no way for Chastain to respond to this challenge.

But let us not dismiss the ambiguity hypothesis prematurely as a relic from a pre-Gricean past, whose motivation has been undermined as it became conceivable that anaphora might be coreference in a pragmatic, but not a semantic, sense. It might have other advantages, unbeknownst to some of its proponents.

It turns out that a prima facie weakness of Chastain's position, i.e., that he does not attempt to explain the ambiguity he posits in terms of other assumptions that are less ad hoc, serves to render him immune to an objection I have raised against Kripke and Lewis: I noted earlier that anaphoric possibilities are sometimes affected when an indefinite is eliminated in favor of a logically equivalent paraphrase (recall examples (21) through (23) in section 1.3 above). This is problematic for Kripke, who takes the presence of speaker's reference to be a sufficient precondition for a subsequent anaphoric pronoun, and who must therefore concede that it is not only the speaker's intentions and other pragmatic factors which determine whether an utterance with a given semantic interpretation can be accompanied by a speaker's reference, but that the wording of the utterance enters as an additional independent factor. Chastain has no problem here: From his point of view, it is only to be expected that if you paraphrase away a referential expression, such as the referential indefinite "one (of the marbles)" in (21a), you will no longer be able to find a coreferential antecedent for a subsequent pronoun.

Where else could we look for possible arguments in favor of a referential-existential-ambiguity in indefinites? A promising source would seem to be the linguistic literature on the feature ±specific.13 With indefinites, +specific marks the referential reading, -specific the existential reading. (This is how the feature was interpreted wherever it was interpreted at all, as far as I know.)

13 Especially Baker (1966), Karttunen (1968a), Stockwell et al. (1973, Ch. 3).
However, hardly any of the arguments that, at one time or another, were used to motivate the specific/non-specific distinction hold up to closer scrutiny. One reason for this disappointing state of affairs seems to be that linguists were not aware from the outset that many of the contrasts they attributed to specificity vs. non-specificity had perfectly natural explanations in terms of scope-ambiguities. To mention just one example, it was proposed that "only [-SPEC] articles are candidates for undergoing some-any suppletion" (Stockwell et al., 1973, 92), but there was no mention of the alternative hypothesis that being in the scope of negation is a sufficient condition for some-any suppletion. If one disregards this and other contrasts that are equally well or better explained as a result of scope constellations, the linguistic literature offers practically no reason to posit a specific-ambiguity at all. (There are exceptions one of which I will turn to shortly.)

Note that, unlike the philosophers Strawson and Chastain, linguists never really tried to hold specificity responsible for an indefinite's ability to serve as the antecedent to an anaphoric pronoun. Karttunen (1968a; 1976), for instance, devotes much attention to anaphoric relations like the one in (34), where the indefinite antecedent is nonspecific:

(34) I wish she had a car. She would give me a ride in it.

His conclusion is that specificity is sufficient, but not necessary, for the ability to serve as an anaphoric antecedent.

To my knowledge, there is only one family of linguistic phenomena that are better explained in terms of a referential/nonreferential (or specific/nonspecific) ambiguity than in terms of scope-ambiguities. These phe-

---

14 It is of course not self-evident that what can be analyzed as a scope ambiguity should be so analyzed. Natural language indefinites could, in principle, be subject to a two-way ambiguity specific vs. non-specific, rather than be scoped elements. However, there are plenty of arguments to be found in the literature which show a specific-nonspecific-ambiguity to be by itself insufficient to account for the full range of readings that sentences with indefinites exhibit. See Stockwell et al. (1973) for examples with an indefinite in the environment of more than one (other) scoped element and the unexpected (from the point of view of a two-way ambiguity hypothesis) multiplicity of readings in such examples. While Stockwell et al. (1973) point out that a two-way ambiguity in indefinites cannot account for these examples, they fail to observe that a theory according to which indefinites are unambiguous, but admit of variable scope, could have predicted not only these examples, but all other ambiguities that they cite as manifestations of the specific-nonspecific contrast as well. Only later was this conclusion put forward. See Karttunen (1976), who proposes that the notions of specificity and nonspecificity should, if not discarded altogether, be understood as relative to a given scoped element; in other words, that to be specific should be taken to mean precisely to have wider scope than...
Do Indefinites Refer?

nomena and their relevance for the hypothesis that indefinites are ambiguous were discussed most recently and most sophisticatedly by Fodor and Sag (1982). The point of their argument is that the ambiguity hypothesis is compatible with a simpler and more explanatory theory of scope islands than the alternative hypothesis that all indefinites are existential quantifiers. It is based on data like the following:\textsuperscript{15}

\begin{enumerate}
\item[(35)] John overheard the rumor that each of my students had been called before the dean.
\item[(36)] John overheard the rumor that a student of mine had been called before the dean.
\end{enumerate}

As (35) illustrates, a “that”-complement to a noun like “rumor” is a scope island which not even the quantifier “each,” which tends to prefer wide scope, can escape: (35) does not have a reading according to which for each student \(x\), John overheard the rumor that \(x\) was called before the dean. We would therefore expect that, if “a student of mine” is an existential quantifier, then its scope in (36) should likewise be confined to the “that”-complement. But this is not borne out by the facts: (36) seems quite naturally understood in such a way as to be judged true in a situation where Mary is a student of mine and John overheard the rumor that Mary has been called before the dean.

Note that this situation is one which verifies (36) under a reading where “a student of mine” is given widest scope, but does not verify it under the reading with “a student of mine” taking narrow scope. Yet the latter reading is predicted to be the only one possible if indefinites are existential quantifiers and respect scope islands like other quantifiers do. So the price of defending an analysis of indefinites as unambiguously quantificational is that indefinites must be considered exceptional in their ability to violate the principles which account for scope islands. On the other hand, no such price has to be paid by the proponent of the ambiguity hypothesis: If indefinites have a referential reading in addition to a quantificational reading, then (36) may be interpreted in such a way that “a student of mine” refers to Mary. Under this interpretation, it will be true in the situation described above, i.e., where Mary is a student of mine and John overheard a rumor that Mary has been called before the dean. There is of course no violation of scope islands involved in such an interpretation, since referring expressions have no scope at all.

Other examples (Fodor and Sag give a variety of them) lead to the same conclusion: As far as simplicity and generality of scope-restricting principles

\textsuperscript{15} Examples from Fodor and Sag (1982).
go, the ambiguity hypothesis compares favorably with the hypothesis that indefinites are unambiguously quantificational.

1.6 Summary

Before I restate the results of the preceding five sections, let me say a few words about a reading of indefinites that was obviously not under consideration in any of the discussion so far: “generic” indefinites, as they are found in the following sentences, under their most natural readings.

(37) A dog has four legs.
(38) An up-to-date encyclopedia is expensive.

All of the authors I have cited would consider the generic interpretation as a separate reading of the indefinite article, distinct from the existential-quantifier reading as well as from the referential reading, if the latter exists. None of the reasoning that I am about to summarize was intended to apply to generic indefinites. This said, let us continue to disregard them.

We have considered Russell’s analysis of the indefinite as a means of existential quantification, as well as an opposing view, which holds indefinites to be ambiguous between a quantificational and a referential use. The Russellian position was seen to face a prima facie problem in the fact that indefinites can serve as the antecedents of anaphoric pronouns. We looked at three approaches to this problem within the limits of the Russellian position, and found some (possibly minor) defect in each of them:

Geach’s proposal to analyze the problematic anaphoric pronouns as bound variables forces him to assign truth-conditions primarily to discourses (i.e., utterances of sequences of sentences), not to monosentential utterances, and thereby falls short of reconstructing our intuitive notion of truth and falsity, which often does apply to each uttered sentence separately. Furthermore, Geach’s analysis seems to force him to assume that indefinites as existential quantifiers show an exceptional behavior’ with respect to the scope island constraints that otherwise govern the interpretation of quantifiers in natural language. (Besides, Geach’s analysis did not seem to carryover to texts with a plural indefinite followed by an anaphoric plural pronoun without practicing inadequately weak truth conditions.)

The Grice-Kripke-Lewis family of proposals has no problems with any of the above challenges, but because of their commitment to a purely pragmatic account of the relation between antecedent and anaphor, they leave unexplained how two structurally different expressions of the same meaning could ever differ in their capacity of making an object available for subsequent anaphoric reference.
Problems with Donkey Sentences

Then there was Evans’ proposal to analyze the problematic anaphoric pronouns as paraphrases of certain definite descriptions that can be constructed from the antecedent and the sentence it occurs in in a systematic way. This proposal avoids both the objections against Geach and against the Griceans. However, it has little substance if the proper analysis of (anaphoric) definite descriptions remains itself unclear, and it is subject to criticism if coupled with the standard analysis of definite descriptions that Evans presupposes. The problem is that definite descriptions under the standard analysis carry an implication that their referent satisfies the descriptive predicate uniquely, and it is not true, or at the very least doubtful, that this uniqueness is required for an anaphoric use of a pronoun or definite description to be felicitous.

As for the anti-Russellian position, according to which indefinites have a reading as referring expressions, it turned out to be immune to any of the objections that have been raised against the Russellian approaches. On the other hand, it is at a prima facie disadvantage by positing an irreducible ambiguity where its competitors purport to predict all uses of indefinites from a unique lexical meaning by independent principles. Moreover, its predictions contradict pretheoretical semantic judgments on the question of whether certain utterances with indefinites are false or else truthvalueless. Still, the ambiguity hypothesis has the advantage of being compatible with a simpler and more explanatory theory of scope island constraints than any version of the view that indefinites are unambiguously quantificational.

If we were to choose among these four alternatives at this point, we would not be able to settle on any one without some misgivings. And we have not even yet looked at a phenomenon that presents all four of them with difficulties that outweigh most of the weaknesses noted so far. The phenomenon in question is the type of anaphoric relation that indefinite antecedents and pronouns (or definite descriptions) exhibit in the so-called “donkey sentences.” This is the subject of section 2.

2 Problems with Donkey Sentences

Donkey sentences are sentences that contain an indefinite NP which is inside an if-clause or relative clause, and a pronoun which is outside that if-clause or relative clause, but is related anaphorically to the indefinite NP. The following are examples:

(1) If someone
    anyone
is in Athens, he is not in Rhodes.

(2) If a man owns a donkey, he beats it.