II

WHAT IS THE APPROPRIATE
LINGUISTIC ANALYSIS OF
DIFFERENT FORMS OF
REFERRING EXPRESSION?
Issues in the Semantics and Pragmatics of Definite Descriptions in English

BARBARA ABBOTT

1. Introduction

As is well known, Russell assigned indefinite and definite descriptions the interpretations represented schematically in (1) and (2), respectively, where “CNP” stands for “Common Noun Phrase” in the sense used by Montague (1973)—that is, as standing for the constituent which a determiner combines with to form a noun phrase (NP).

(1) a. . . a/an CNP . . .
b. ∃x[CNP(x) & . . . x . . .]

(2) a. . . the CNP . . .
b. ∃x[CNP(x) & ∀y[CNP(y) → y=x] & . . . x . . .]

Examples (3) and (4) are illustrations.

(3) a. Mary bought a car that she liked.
b. ∃x[Car(x) & Liked(m, x) & Bought(m, x)]

(4) a. Mary bought the car that she liked.
b. ∃x[Car(x) & Liked(m, x) & ∀y[Car(y) & Liked(m, y)] → y=x] & Bought(m, x)]
The difference, as is obvious, is the underlined clause expressing uniqueness—exhaustive possession by the entity in question of the property expressed by the CNP.

Szabó (2000) and Ludlow and Segal (2004) (following Kempson (1975), Breheny (1997), and others) defend analyses on which definite descriptions are assigned the same quantificational interpretation as Russell assigned to indefinite descriptions. Thus, on both accounts (3a) as well as (4a) would be given the quantificational analysis in (3b). Both proposals acknowledge that definite descriptions differ from indefinites in their implications—where “implication” is to be understood as neutral between semantic and pragmatic conveyance. One of these implications is what is commonly termed “familiarity”—an assumption that the denotation of the NP1 has already been introduced, as such, to the addressee of the utterance. The other is commonly termed “uniqueness,” in its simplest form as expressed by the underlined clauses in examples (2b) and (4b), but frequently relativized to context and addressee. (We will return to this issue shortly.) Both analyses take familiarity to be more central to the interpretation of definite descriptions, and attempt to derive the uniqueness implication as a conversational implicature.

I agree that, as far as contribution to truth conditions goes, (3b) may suffice for both (3a) and (4a). However, my argument differs from that in both papers in holding that uniqueness is a part of the conventional import of definite descriptions and that familiarity is not conventionally, but only conversationally, associated with them. (Compare also the account in Gundel et al. 1993, 2001.)

Let me be more specific about the view I want to defend. It is, that use of a definite description presupposes the existence and uniqueness clauses of Russell’s analysis (as in (2b) and (4b)), where “presupposes” means roughly ‘conveys as something not needing assertion’ (cf. Abbott 2000). The first of these presuppositions (the existence one) is also an entailment (as encoded in (1b) and (3b)). The second (uniqueness) seems aptly described as a conventional implicature—an implication conveyed semantically, but whose falsehood would not be felt as critically damaging to the truth of the utterance as a whole. Thus, (4a) conveys the information that Mary had liked one and only one car as something not needing assertion, whereas (3a) asserts that Mary had liked a car (and bought it) but says nothing about how many she liked. If Mary had not liked any car, then neither (3a) nor (4a) would be true. However, if Mary had (recently) liked two cars (and bought one of them), then (4a) might be felt to be true although anomalous in conventionally implicating that she had only liked one car.

Before plunging forward, it is necessary to clear up some points surrounding the relevant concept of “uniqueness.” First of all, Russell’s analysis was couched in a logic geared to speak only of discrete entities one at a time. Of course in natural language we do not so confine ourselves, so at the very least if something like this analysis is to be maintained it must be amplified
accordingly. Fortunately, Sharvy (1980) has shown a way to maintain the spirit of Russell’s analysis while extending it to deal with CNPs headed by mass and plural nouns, namely, by restating uniqueness as exhaustiveness (cf. also Hawkins 1978, 1991).

Incidentally, one could hold, with Sharvy, that “the primary use of ‘the’ is not to indicate uniqueness. Rather, it is to indicate totality; implication of uniqueness is a side effect” (Sharvy 1980, 623). Or one could argue that the primary use of the is to indicate uniqueness, and that totality (or exhaustiveness of application of the descriptive content of the relevant CNP) is a side effect of achieving uniqueness with mass and plural CNPs. In any case, I will continue to use the term “uniqueness” for whatever it is we’re talking about.

But there is another, even more pressing issue—the problem of incomplete definite descriptions. Someone can say, boldly, Please put the book on the table, in full knowledge that the world is littered with books and tables. Probably the preferred way to deal with this problem currently is to modify the notion of uniqueness to something like “identifiability in context,” the idea being that use of the signals an assumption that the addressee can individuate the speaker’s intended referent from among the potential referents in that particular context of utterance. (See, e.g., Birner and Ward 1998, 122.) As a statement of the net effect, this is fine, although ultimately I would want to try to argue that it is something Russellian, as elaborated by Sharvy, which is conventionally encoded in definite descriptions, and that the part to do with what the addressee is expected to be able to do can be derived from that plus Grice’s rules of conversation (Grice 1975). I won’t attempt to make the argument here in full, though I will try to sketch something partial along these lines.

The remainder of this paper is organized as follows. In the next two sections I give evidence supporting my claims that familiarity is not conventionally associated with the and that uniqueness is conventionally associated with the. Following that, I look at the case of stressed the, and then look into the derivation of the familiarity implication as a conversational implicature. The penultimate section sketches very briefly the direction of my response to some of the other arguments presented in the papers of Szabó and Ludlow and Segal, and the final section concludes.

2. The Familiarity Implication Is Not Conventional

Ludlow and Segal make the explicit suggestion that familiarity (their term is “givenness”) is a conventional implicature of definite descriptions, but they seem somewhat unclear on what that suggestion entails. For one thing, they seem to believe that conventional implicatures are inferred, and for another, that they can be overridden. Perhaps sensing the inconsistency here, they note in a footnote that they are not “wedded” to the idea of givenness as a conventional implicature, commenting that “all of what we say here can be recast in terms of explication or inference by one’s favorite account of how
the new/given distinction is to be understood.” (Puzzlingly, they continue, “Anne Bezuidenhout has observed that it might be possible to build the new/given information into the semantics of the determiner. In principle we don’t have a problem with that move either. . .” (Ludlow and Segal 2004, 425n5, italics in original; the puzzlement is because conventional implicatures are semantic.)

There is a suggestion here, in the use of the term “inference,” that familiarity, as well as uniqueness, could be derived as a conversational implicature. (Szabó also often sounds as though he had something like this in mind, although he has said elsewhere (Szabó 2003) that that was not his intention.) In any case, there would be a fundamental problem with proposing to derive familiarity or givenness as a pragmatic inference, given that one is also proposing that definite descriptions are assigned the same conventional meaning as indefinites. The problem is that if there is no conventional distinction between the and a, then there is no way to derive familiarity in one case and not the other. Any pragmatic distinction of this type would have to be based on some kind of conventional difference. (This difficulty is discussed in more detail in Abbott 2003.) So let us turn to the idea that familiarity or givenness is a conventional implicature.

It is a characteristic of conventional (as opposed to conversational) implicatures that they cannot be canceled or “overridden.” This follows from their nature: if these implications are linguistically encoded, then there is no way to deny them on the spot without sounding like you are contradicting yourself. Consider (5):

(5) # Even Kim, who is very smart, could solve that problem.

Even conventionally implicates a relatively unexpected instance. This implicature (which is not inferred but semantically encoded in the word) conflicts with the expressed assumption that Kim is very smart, hence would be expected to solve problems. The result is anomaly.

Now contrast the case of the familiarity implication of definites, using (6):

(6) The new curling center at MSU, which you probably haven’t heard of, is the first of its kind.

In (6) any assumption that the addressee is familiar with the curling center is explicitly denied and yet the result is perfectly felicitous. (The denial must be appropriately hedged, of course, since it would be infelicitous for independent reasons to make a bald assertion about the knowledge state of the addressee.) If familiarity or givenness were a part of the conventional meaning of definite descriptions, (6) should be anomalous, but it is not. (See also the research reported in Gundel et al. 2001 and the works cited there, suggesting that in some contexts, as many as a third to a half of occurrences of definite descriptions may introduce unfamiliar referents.)

It should be acknowledged that some definite descriptions are used anaphorically. Standard examples are of the type that Heim gave in presenting
her version of the familiarity theory of definiteness, examples like that in (7) (cf. Heim 1982, 275).

(7) A woman and a man met on the street. The woman said ‘Hi’.

Even if not all definite descriptions convey the familiarity idea, it might be suggested, still some, like the one in (7), do convey that as a part of their conventional meaning. I want to resist this move for several reasons. One is that, given that familiarity is not in general conventionally associated with definite descriptions (as indicated by (6) as well as the research cited in the last paragraph), this move would seem to require that we regard definite descriptions, or the definite article, as ambiguous. This is both contrary to speaker intuitions and methodologically distasteful. Another reason for resisting the claim that definite descriptions at least sometimes convey familiarity conventionally is that it may not be necessary, if we can give a sufficiently convincing account of how the uniqueness approach can account for such examples. We will return to this issue later.

3. The Uniqueness Implication Is Conventional

Both Szabó, and Ludlow and Segal, attempt to treat the uniqueness implication as a conversational implicature, derived pragmatically with the help of the element of familiarity or givenness, however it is obtained. There are both similarities and differences in the two suggested derivations.

Szabó’s approach, which is adapted from the “file card” approach of Heim 1982, involves a potential conflict between familiarity and a pragmatic principle he calls ‘Non-arbitrariness’:

(8) Non-arbitrariness: When filling an utterance, don’t make arbitrary choices.

The idea behind (8) is that it should be clear where the addressee is to look for a referent, in other words that an addressee shouldn’t have to make an arbitrary choice between two potential referents. (Ideally perhaps (8) should be rephrased to apply to speakers, and enjoin them from forcing addressees to make arbitrary choices. It might then be collapsed with Grice’s rules of Manner.) Suppose a speaker has made an assertion using a definite description the F. The addressee, Szabó explains, will reason as follows (41: an F card is a file card representing an entity with the property denoted by the CNP F):

(9) Suppose I had two private F cards with incompatible conditions. Then I could not have filed A’s utterance . . . : either Familiarity or Non-arbitrariness would have been violated.

So the speaker can conclude that the addressee does not have two private incompatible F cards—that is, that he or she knows of at most one F. There are a couple of problems with this line of derivation. One is that it is not so clear that Familiarity and Non-arbitrariness must conflict in this way. Part of the
stipulation for the conflict is that the two potential \( F \) cards have incompatible conditions. That means the two potential \( F \) cards must have different information on them, which in turn suggests that choosing one of them over the other need not be an arbitrary choice after all. Another problem is presented by definite descriptions with plural or mass head nouns, which Szabó does not address in his paper. It is not clear how this derivation will extend to those cases.

The derivation sketched by Ludlow and Segal is also potentially liable to the difficulties presented by plural and mass NPs, but suffers from a more basic problem as well. They use the traditional Gricean format for displaying the calculability of conversational implicatures (p. 427):

(10)  
\[ a. \text{S has expressed the proposition that } ([\exists x: Fx](Gx)). \text{[Recall that this is Ludlow and Segal’s (and Szabó’s) analysis of the truth-conditional contribution of definite descriptions.]} \]

b. There is no reason to suppose that S is not observing the CP and maxims.

c. S could not be doing this unless he thought that \([tx: Fx](Gx)\). Gloss: By invoking the determiner ‘the’, \( S \) intends to communicate that whatever \( F \) or \( F \)s he is talking about is/are given in the conversational context. [This is the conventional implicature of givenness, or familiarity.] By refraining from using a plural noun, \( S \) intends to communicate that just one \( F \) is given in the conversational context. If there were more than one \( F \) given in the context, \( S \) would have used the plural definite description (otherwise \( S \) would flout the maxim of quantity). . .

But it is not a violation of the rule of quantity to speak indiscriminately about one of a number of familiar entities. One can do this without any conversational strain using a phrase like \textit{one of these students}. It would only be a violation of Quantity if we assumed that the descriptive content of \( F \) were sufficient to determine a unique entity, but that would be assuming a uniqueness implication, rather than deriving it.

In any case there is a further problem with any proposal to derive the uniqueness implication pragmatically, as a conversational implicature or some other type of non-conventional inference. The problem is that we would then expect this implication of uniqueness to be cancelable, unlike the case with the implication of unexpectedness with \textit{even} (as in (5) above), but like the case with familiarity (as in (6)). The problem for both sets of authors is that this implication behaves more like a conventional implicature than a conversational implicature, as shown by (11).

(11)  
\[ \# \text{Russell was the author of } Principia Mathematica; \text{ in fact, there were two.} \]

In (11) there has been an attempt to cancel the implication of uniqueness, but the result is anomaly, just as in other cases of attempted cancellation of conventional implicatures. And we get the same effect with a plural definite description, as shown in (12).
(12) # Akmajian and Demers were the authors of Linguistics; in fact, there were three.

The anomaly of (11) and (12) supports the claim that uniqueness/exhaustiveness is conventionally encoded in the definite article, and not derived conversationally.

A major claim of Szabó’s paper is that the implication of uniqueness for definite descriptions and the implication of non-uniqueness for indefinite descriptions are symmetrical and should be treated as such. However this does not seem to be the case. Given that definite descriptions encode uniqueness, the nonuniqueness associated with indefinite descriptions may be readily derived as a conversational implicature: a choice of a rather than the conveys that the is not appropriate, hence that the descriptive content of the NP does not apply uniquely (within the local discourse context). This analysis, which makes a number of subtle predictions concerning contexts in which the implication of nonuniqueness will and will not occur, is defended in some detail in Hawkins 1991; see also Gundel et al. 1993. Note too that the implication of nonuniqueness in the case of indefinites is cancelable:

(13) Russell is an author of Principles of Mathematics, in fact the only one.

Unlike (5), (11), and (12), (13) does not sound self-contradictory, confirming the claim that nonuniqueness is only a conversational implicature of indefinite descriptions.¹

4. Stressed THE

There are other arguments for taking uniqueness rather than familiarity to be the essence of definite descriptions—see e.g. Hawkins 1991, Birner and Ward 1994. I want to mention one argument here that was given in Abbott 1999, and replied to by Ludlow and Segal. That argument involved noting that when the is stressed, it is an implication of uniqueness, rather than familiarity, which is fronted. One example is (14):

(14) That wasn’t a reason I left Pittsburgh, it was the reason.
    [= Abbott 1999, ex. 2]

Concerning this example, Ludlow and Segal remark

We take it that someone may very well utter [14] with stress as indicated despite having several reasons for leaving Pittsburgh. Stressing ‘the’ indicates that the reason in question was not merely one of many reasons for leaving but rather the causally determinate reason—the big reason.

(Ludlow and Segal 2004, 435)

In so remarking, Ludlow and Segal echo comments of Epstein (1996), who described the stressed the in (15) as signaling ‘prominence’ or ‘great importance’.

issues in semantics and pragmatics 67
In other countries, soccer is the sport. If the national team loses, there could be a coup. [Los Angeles Times 6/5/94, p. C9; italics in the original] [= Epstein 1996, ex. 2]

However, as I argued in my 1999 paper in reply to Epstein, it seems both more specific and more accurate,

to describe [15] as conveying this prominence through hyperbole. Obviously soccer isn’t the only sport in any country, but to describe it as such in forceful terms, as is done in [15], is to convey its prominence in a specific way. So on my account the speaker of [15] says literally that soccer is the only sport. This is almost certainly false, invoking standard Gricean mechanisms to arrive at the hyperbolic understanding. In support of this claim notice that one could replace the sport in [15] with the only sport, achieving the same effect although in a slightly more heavy-handed way. . . . [16] below is another naturally occurring example of stressed the where the referent is not actually the unique satisier of the description. . . . In this example the hyperbole is tacitly acknowledged.) (Abbott 1999, 3.)

(16) ‘People say it’s the night that the movers and shakers are going,’ a mischievous-sounding Gehry said when he was asked about his Thursday invitation. ‘I was told, of course, that every night is the most important one,’ Gehry added. ‘But I was told that this was the most important one.’ [The New Yorker 12/22and29/97, p. 50; italics in original] [= Abbott 1999, ex. 8]

Note too the number of other cases where expressions literally expressing uniqueness have come through hyperbole to convey prominence or some other kind of specialness: e.g. the one and only X, X is very unique, they broke the mold . . .

Of course none of these examples seem to have anything to do with the familiarity of the referents involved, and hence are problematic for the familiarity theory. Ludlow and Segal do provide one kind of case which might look at first as though it were familiarity rather than uniqueness which was being emphasized with stressed the:

Ludlow’s third grade teacher had a husband named William Faulkner. On vacation in the South, he was asked, “are you the William Faulkner”? Presumably, the questioner was not asking if he was the unique individual named William Faulkner . . ., but was asking whether this individual was the famous—i.e. given or familiar—William Faulkner. (Ludlow and Segal 2004, 435, italics in original.)

One problem with this analysis is that both William Faulkners are given, in the sense of known to the addressee, in this example. Furthermore we have to assume that to the addressee, he himself is the more given or familiar of the two. Nevertheless, of course, his answer should be No. There is also the fact
that the phrase *the William Faulkner* imparts a sense of luster to the referent. Ludlow and Segal would like to see this as an extrapolation from givenness. But one could just as well see the prominence of the author as granting him a unique salience, and the stressed *the* as expressing that idea. That would have the additional advantage of making this example consistent with those above, where prominence is obviously seen as a species of uniqueness and not givenness.

5. Deriving Familiarity as a Conversational Implicature

I have argued above that familiarity or givenness is not conventionally encoded in definite descriptions, and that uniqueness is. It remains to say something more about where the implication of familiarity comes from, in those cases where it does arise. The prototypical cases are anaphoric uses as in (7) above, repeated here.

(7)  A woman and a man met on the street. *The woman* said ‘Hi’.

As noted at the outset, I am assuming that a definite description will convey as a presupposition that there is one and only one entity meeting the descriptive content of the CNP. In a case such as (7), with a very incomplete definite description like *the woman*, it is clear that this uniqueness must hold within a narrowly circumscribed circumstance and, handily enough, the preceding sentence gives exactly such a circumstance. Furthermore it follows from Grice’s principles that sequences of sentences will be relevant to each other, unless the subject is explicitly changed. An inference that the speaker intends to refer to the female entity introduced in this preceding sentence is almost unavoidable, a fact of which both speaker and addressee must be assumed to be mutually aware.

Despite the naturalness of this derivation let us for a moment consider a claim that familiarity is conventionally encoded in definite descriptions used anaphorically. Could one then try to supplement such a claim by showing that the familiarity is not cancelable? The problem is that in any normal circumstances the utterer of (7) will intend to refer to the woman they have introduced into the conversation. As suggested above, the discourse would be incoherent otherwise. But this does not mean that we need to regard familiarity (or anaphoricity) as encoded in the definite article, any more than we need to regard it as encoded in proper names. The implication of familiarity comes from general principles for coherent discourse, which probably follow from even more general principles as Grice suggested. So although (7′) is definitely anomalous:

(7′)  # A woman and a man met on the street. The woman (not the one I just mentioned) said ‘Hi’.

The anomaly is one of incoherent discourse and not self-contradiction.
6. Brief Replies to Some Other Arguments

Szabó and Ludlow and Segal give two similar kinds of arguments in favor of their approaches. One kind involves citing general facts about determiner interpretation that seem to show that, were the uniqueness implication to be conventionally encoded in an article, English would be out of line, in some sense, with the languages of the world. The other involves the explanation for the definiteness effect in existential sentences. I can only briefly sketch the direction a complete reply to these arguments would take.

On the general claims about possible determiner interpretations, I would suggest that there are a number of quite specific types of information commonly encoded in determiners in some, but not all, of the languages of the world. Some of these refer to aspects of referents such as animacy, shape, visibility, and so on. Of course, it might be claimed that quantificational meanings are special, and there we should expect to find uniformity across languages. However, even here there are notable lacks of uniformity. For example, some languages encode a dual number (as English used to), but clearly not all languages have that category.

The issue of the definiteness effect in existential sentences I have addressed elsewhere (e.g., Abbott 1993, 1997). Briefly, my position is that it is less stipulative, as well as empirically more adequate, to see the awkwardness of indefinites in nonenumerative existentials as a result of a clash between the presupposition of existence grammatically encoded in indefinites and the assertion of existence of an existential sentence. Ludlow and Segal find this unconvincing in view of the fact that “one can perfectly well say ‘The mayor is a mayor’” (p. 433, n. 14). But in this example the conflict is at the level of content morphemes, not grammatical constructions as in the case of existential sentences and definite DPs.

7. Concluding Remarks

In this chapter I have tried to support the uniqueness view of definite descriptions, and to argue that the approaches of Szabó 2000 and Ludlow and Segal 2004, which take familiarity as essential to definite descriptions and uniqueness as derivable, are not correct. There remain plenty of problems with viewing uniqueness as the essence of definiteness, but, in general, they do not support the position that familiarity is the essence instead.

Portions of this paper were read at the annual meeting of the Linguistic Society of America (Atlanta, January 2003) under the title “The Difference between Definite and Indefinite Descriptions.” I am grateful to Larry Horn and Kent Bach for reading drafts of the LSA paper and providing me with extensive and very helpful comments. I am also grateful to Jeanette Gundel and Martin Hahn for organizing a terrific workshop, and for inviting me. I thank the audiences in Atlanta and Vancouver for their comments, and
especially Chris Barker, Michael Israel, Christopher Potts, Ori Simchen, and Gregory Ward. Finally I owe a special debt to Kent Bach for suggesting I think about replying to the papers of Szabó and Ludlow and Segal that are addressed here. Despite all of this high-class help, I must acknowledge the possibility of remaining flaws and accept responsibility for them.

NOTES

1. Strictly speaking, it might be considered to be inconsistent with Russell’s analysis to talk in terms of a denotation for definite descriptions.

2. “… when we implicate that something is given information we are not explicitly saying: ‘this is given information.’ It is simply something that competent users of English can infer from the conventional implicature inherent in ‘the’.” “When the predicate by itself ensures uniqueness, the implicature [of givenness] gets overridden.” (Ludlow and Segal 2004. 425; emphasis added.)

3. The point in this paragraph was also made in Abbott 2003.

4. See Apostolou-Panara 1994 for description of a construction in Greek which may have been influenced by these hyperbolic constructions in English.

REFERENCES


ISSUES IN SEMANTICS AND PRAGMATICS 71


