SENSE, REFERENCE AND PURPORTED REFERENCE (*)

Howard G. CALLAWAY

I

Frege employed his distinction between sense and reference attempting solution to a puzzle concerning identity. ([1]) The relation obtains between objects, but only between an object and itself. Frege seeks to explain how sentences of the forms '$a = a$' and '$a = b$' differ in cognitive significance. If '$a = b$' is true then 'a' and 'b' have the same reference; the same predicate appears in each sentence. Yet there is an interesting difference between the two sorts of sentences. He concludes that a difference in sense between 'a' and 'b' is required to explain how it is that the one sort of sentence is informative while the other is not.

Frege's distinction is often emphasized, even by those such as Quine who will not countenance Fregean senses. ([2]) Frege is credited with originating the modern distinction between meaning and reference. ([3]) I will show that solution to Frege's puzzle does not require introduction of Frege's special ontology of senses.

The puzzle is best explained entirely within the theory of reference. Following Quine and others, the notion of ontological commitment belongs to the theory of reference. ([4]) The following account of Frege's puzzle makes use of 'ontological commitment' and other vocabulary of the theory of reference to provide an extensionalistic simulation of Frege's distinction. This solution serves to emphasize the importance of the distinction between reference and purported reference.

We have '$a = b$' by hypothesis and thus 'a' and 'b' name the same object. The same predicate '=' occurs in each of the two sorts of sentences. Still the ontological commitments of the two sorts of sentences diverge. The existentials '$(\exists x)\ a = x$', '$(\exists x)\ x = b$' and importantly, '$(\exists x)(a = x \,\&\, x = b)$' follow from '$a = b$', but the latter two do not follow from '$a = a$'. This begins to explain the intuitive claim on which Frege insists. Sentences of the form '$a = b$' are more informative since they force wider existential commitments. This will be regarded as genuine empirical information depending upon how well-established '$a = b$' has become and how it was established. But whether well-established or not the existential claim is evident. Similarly, a theory which logically implies that there are F's and there are G's is more informative than one which merely implies that there are F's. The former is more informative since it has wider existential commitments. Nothing is changed if in fact all and only F's are G's. Any other conclusion confuses the ontology or commitments of a theory with the objects of a theory. ([5])

Endorsing Frege's intuition does not require adoption of the notion of cognitive significance (or meaning). Using only referential semantics, the differences between sentences of the two forms may be explained by noting their logical implications. As will be shown shortly, this difference in the logical implications of the two sorts of sentences is of substantial importance in understanding referential semantics.

If '$a = b$' is more informative this does not require that it is more informative in every context of discussion. If someone already accepts '$a = b$', the sentence will not be informative to that person. It will not convey new information to him. Likewise, if a person accepts a theory T which already includes many logical implications of '$a = b$', the sentence will not convey as much information to that person. What is primarily of interest here is the informative character of sentences rather than the information a sentence conveys to particular people. The information conveyed by a sentence to someone in particular is merely indicative of the informative character of the sentence.

Further important differences between the two sorts of sentences become evident on consideration of the logical consequences of augmenting a theory T (which does not imply '$a = b$') with '$a = b$'. Suppose that T implies various existentials such as '$(\exists x)\ Fxa$' and let T′ be the conjunction of T and '$a = b$'. The existential commitments of T and T′ diverge. For example T′ logically implies '$(\exists x)\ Fxb$' whether T does or not. Whatever T tells us by use of 'a' we may obtain an additional bit of new information in T′ by means of the law of substitutivity. Moreover, it is not merely additional logical implications in a given theory which are of interest here. For there may be a range of theories under consideration as alternatives to T, and '$a = b$' as compared with '$a = a$' will force wider existential commitments in these as well.

Whether we consider the two sorts of sentences in isolation or in combination with various relevant theories, the existential consequences are different. Sentences of the form '$a = b$' are more informative than sentences of the form '$a = a$' (whether or not the latter is regarded as a logical truth) in view of the more extensive existential implications of '$a = b$'. It is the sum of these existential implications within various relevant theories which constitutes the informative character of a sentence.

Various psychological questions arise if we attempt to say what information is conveyed to a particular person by a sentence. We must first find a theory which approximates the person's relevant belief system, and then consider the additional logical implications which arise by augmenting that theory with the sentence in question. Complications arise since a person may not have considered even all the interesting logical consequences of his beliefs and may not see the logical consequences of augmenting his beliefs. But these are sorts of questions which need not be treated here.

Otherwise, in considering the information conveyed by a sentence of the form '$a = b$', three sorts of cases arise. In the first (1), the person in question has not previously accepted any sentence making use of 'a' or 'b'. In this case the information conveyed will correspond to the existential implications of '$a = b$'. In case (2) the person has accepted a theory T which included some but not all of the logical implications which may be obtained, within the augmented theory, by substitution once '$a = b$' is accepted. In this case the information conveyed corresponds to the additional implications of T augmented by '$a = b$'. In the last case (3) the person accepts a theory T which implies all the sentences which may be obtained by augmenting T with '$a = b$' and substituting 'a' for 'b' or 'b' for 'a'. Thus in such a case the person already holds that a exists and that b exists, and a and b are not discernible within his theory; he accepts an existential sentence involving a corresponding to every existential sentence with b for a and vice versa. This case is of interest because it may seem to provide counter-examples, where '$a = b$' does convey information, but this cannot be accounted for by additional existential implications in the augmented theory. However, there are still new existential implications in the augmented theory. '$(\exists x)(a = x \ \& \ x = b)$' is logically equivalent to '$a = b$'. This new existential in the augmented theory is sufficient to guarentee that case (3) provides no counter-examples. Thus, the information conveyed to someone by a sentence of the form '$a = b$' can be accounted for in terms of additional existential implications in the augmented theory.

Whether we consider the two sorts of sentences alone or as augmenting various theories, the differences in the existential implications serve to account for Frege's intuition that sentences such as '$a = b$' are informative in principle. Since this account only appeals to referential semantics, Frege's argument for senses is undercut. Indeed, Frege's puzzle retains its problematic character only so long as we concentrate on sameness or difference of reference (designation), ignoring the larger resources of referential semantics.

II

We find Frege's argument plausible while neglecting a point which is otherwise understood. Factors other than reference affect deductive relationships. Here this point will be developed, making use of the notion of purported reference. Though obvious enough once attention has been focused on it, the point has been consistently ignored in certain contexts of discussion. Before examining one such context it will be important to clarify 'purported reference'.

This notion, as applied to individual constants, is best understood in terms of its relationship to 'ontological commitment'. Since 'ontological commitment' belongs to the theory of reference, the following schematic definition serves to establish that 'purported reference' also belongs to the theory of reference. Although I will be primarily concerned with individual constants, schemata similar to those below serve to apply the notion of purported reference to general terms.

(I) ᵗP(a)'a' iff C(a)t

A constant 'a' in a theory T purports reference to an object if and only if T is committed to the existence of a. The positions within parenthesis are not referential. Rather there are distinct monadic predicates such as '^{t}P(a)—' and 'C(a)—' in the metalanguage for each individual constant appearing in the theory of interest. A theory may, of course, have certain commitments when, in fact, there are no objects as called for by the theory and the theory is false. This point makes it evident that 'is ontologically committed to' is not a relation. A relation obtains only if the relata exist. Seeing this much, one also sees the need for distinct predicates in the metalanguage to state the ontological commitments of a theory. Similar points establish that 'purports reference to' is not a relational expression, but rather a part of longer monadic predicates. These distinct monadic predicates are true of an individual constant, relative to a theory, depending on the logical relations of sentences of the theory T in which the constant occurs. But more on this shortly.

Ontological commitment to individuals may be explicated by the following schema:

(II) C(a)t iff t logically implies '$(\exists x)\ x = a$'

A theory T is committed to a if and only if T logically implies that there is an object identical with a.(6) A given theory may employ 'a' without commitment to an object designated. For example the theory may tell us that there is no such object, implying '$-(\exists x)\ x = a$'.(7) It is not the mere occurrence of the constant which determines whether it purports reference. In view of (I) and (II) we get the following which makes this point explicitly.

(III) ^{t}P(a)'a' iff t logically implies '$(\exists x)\ x = a$'

The conditions under which two individual constants in a theory T have the same or different purported reference are as follows:

(IV) (^{t}P(a)'a' & ^{t}P(a)'b') iff t logically implies '$a = b$'

Two constants 'a' and 'b' have the same purported reference in a theory T if and only if T logically implies the corresponding identity. Just as reference (which is relational) must be distinguished from purported reference (which is not), so co-reference must be distinguished from sameness of purported reference. Two expressions 'a' and 'b' may be co-referential without the theory of our interest taking notice of this. On the other hand if 'a' and 'b' do have the same purported reference in T they may not be co-referential since T may be false.

The relationship between reference and purported reference is illustrated by this final schema:

(V) If 'P(a)'*a*' & *t* is true, then '*a*' designates *a*

If the constants of a given theory purport reference to objects of certain descriptions and the theory T is true, then these constants refer to objects so described. ([8])

Purported reference, following ontological commitments, will differ from theory to theory since one theory may deny the existential commitments of another. This point serves to distinguish 'purported reference' from Frege's 'sense' since the sense of a constant does not vary depending upon whether it is employed in existential claims and identity claims. Still variations in purported reference from theory to theory do not prevent such theories from having referents in common. Questions concerning co-reference are not settled simply by taking note of the purported reference of constants in particular theories.

With this fragment of a theory of purported reference at hand, the next step is to make use of the notion to explain how factors other than reference affect deductive relations. Israel Scheffler overlooks this point in a much discussed criticism of the thesis of incommensurability between alternative theories. Citing Frege on the distinction between meaning and reference, Scheffler emphasizes reference. But his claim involves a mistake:

> As for deduction within scientific systems it should be especially noted that it requires stability of meaning only in the sense of reference in order to proceed without mishap. ([9])

But stability of reference is not sufficient for deduction without mishap. This is illustrated by the following pair of arguments.

(1) Fa
$a = b$
———
Fb (valid)

(2) Fa
$a = a$
———
Fb (invalid)

The difference between the two is that in (2) we find 'a' in place of 'b' in the second premise. The contrast between (1) which is valid and (2) which is not cannot be accounted for merely in terms of a difference in reference between 'a' and 'b'. Whether 'a' and 'b' are co-referential or not, this only effects the truth of '$a = b$' and not the validity of either argument. Exchanging coreferential constants here turns a valid argument into an invalid argument, so long as the identity is not added as an extra premise.

Sameness and difference of purported reference are relevant to the contrast between (1) and (2). Whether or not the premises of an argument suffice to settle questions regarding sameness or difference of purported reference influences the range of interpretations to be taken into account in deciding the validity of the argument. An argument is valid, on the standard account, if there is no interpretation in any non-empty domain which renders the premises true and the conclusion false. The premises of (1) provide that 'a' and 'b' have the same purported reference in any theory where these premises are available. Hence, on each interpretation 'a' and 'b' are assigned to the same object. Thus if the premises are true on an interpretation so is the conclusion.

In contrast, the premises of (2) do not determine whether 'a' and 'b' have the same purported reference. These premises will be available in theories where the two constants differ in purported reference. We must therefore allow for interpretations on which the premises are true but '$a = b$' and 'Fb' are false. Thus, even if the corresponding conditional '(Fa & $a = a$) only if Fb' is true on the preferred interpretation (as it is on every interpretation of a theory including '$a = b$'), still it is not a logical truth and (2) is invalid.

Though refraining from the sense/reference distinction, we need not restrict ourselves to the relational vocabulary of referential semantics. Referential semantics, as here augmented, provides resources sufficient to give an account of the logical differences between the two forms of identities without recourse to non-extensional notions.

Constructing set-theoretic models for theories involves much concern for purported reference of expressions, the ontological commitments of the theory, and little concern with reference. We need not accept a theory or its ontological commitments in order to construct models for it. Giving interpretations of a theory on which it is true does not require us to know whether or not the theory is in fact true. Moreover, we need not make non-hypothetical claims regarding which elements and subsets of a domain are identical or distinct. An interpretation does presuppose a non-empty domain. But the validity or satisfiability of a formula in a domain D depends only on the number of elements in D and not upon their specific identities. ([10]) Thus we need not make specific assumptions regarding the elements. It is important to avoid the assumption that there is no domain with more than n elements, since a formula satisfied on all interpretations in all domains of n elements may be invalid considering domains with n + 1 elements. But this is not to assume (except hypothetically) that there are domains with n + 1 elements, rather the point is that if there are such domains (and there may be for all the logician is required to know), the relevant formulas are not satisfied on all interpretations. Thus specific existential assumptions are required to decide the truth-value of such a formula.

The point I am emphasizing is that there are important hypothetical sentences, making use of the relational vocabulary of the theory of reference. These hypothetical claims, e.g., «If T is true, then '*a*' designates *a*», do not involve specific ontological commitments. Moreover, the argument above strongly suggests that in constructing models one makes use of such hypothetical claims for the purpose of elucidating the purported reference rather than the reference of constant expressions of a theory. We have also seen that a valid argument may be converted into an invalid argument by exchanging co-referential constants. Although further investigation is required, these two points do suggest that it is purported reference rather than reference which is crucial to validity.

III

Having explained the distinction between reference and purported reference, we may now return to Frege's puzzle with clearer insight into the difficulty. John Searle has stated the problem as follows:

> ... if proper names simply stand for objects and nothing more, how could such identity statements ever convey factual information? If we construe such statements as solely about the referents, then it seems they must be trivial, since, if true they say only that an object is identical with itself. If on the other hand, we construe the statements as giving information about the names, then it seems they must be arbitrary, since we can assign any name we wish to an object. (11)

Attention to the distinction between reference and purported reference allows one to grasp either horn of the dilemma.

. A sentence of the form '$a = b$' does convey factual information not conveyed by sentences of the form '$a = a$'. We can understand this in terms of the differences in their logical implications. Both sorts of sentences do state that an object is identical with itself. Yet how this is said makes all the difference. The differences in the logical implications of the two sorts of sentences are easily understood once we recognize that two names may differ in purported reference in a theory, even if they are (in fact) co-referential. On discovery that $a = b$ we may obtain a wealth of new information, the logical and particularly the existential implications of '$a = b$' within our theory. Sentences of the form '$a = a$' do not have the same interesting implications.

If we assumed (for the sake of argument) that '$a = b$' is about the names, does it follow that such claims must be arbitrary, «... since we can assign any name we wish to an object» ? The mistake here comes of thinking that semantic claims must be trivial in view of their relation to linguistic conventions. Certain semantic claims do have a trivial character. For example, consider the following which may account for the temptation to think of '$a = b$' as about the names.

$$a = b \text{ iff } (\exists x)(\text{'}a\text{' designates } x \;\&\; \text{'}b\text{' designates } x)$$

But this biconditional stands in sharp contrast with other semantic claims which are by no means so obvious. Take the claim that 'a' and 'b' designate the same object, for example. There is no more reason to think such a claim trivial than there is to think that 'renate' and 'cordate' are true of only the same objects is a trivial claim. In this case it is obvious that the semantic claim is true if an only if all renates are cordates, and this is clearly a matter open to empirical investigation. A similar point holds regarding the co-referentiality of '*a*' and '*b*'. Even if we set out to give two names to one object, minimal skepticism creeps in regarding success. After giving the name '*a*' to an object one might mistake some other object for *a* and give the name '*b*' to that object instead. The conventionality of naming does not guarantee the triviality of semantic claims regarding sameness and difference of reference.

Being able to distinguish two objects depends upon the truth of our beliefs concerning them. Even if all our beliefs are true, this does not guarantee ability to distinguish them. As our beliefs are continually augmented, we may find criteria to distinguish *a* from *b*, just as we may find criteria to distinguish F's from G's.

Solution to Frege's puzzle does not require departure from referential semantics. Rather it requires that we make use of resources readily at hand. In particular it is important to recognize the distinction between reference and purported reference, and emphasize how claims regarding reference and coreference are dependent upon empirical and existential considerations.

Drexel University
University of Ibadan

H.G. CALLAWAY

(*) An earlier version of this paper was read at the Western Division meeting of the American Philosophical Association in April, 1979, and on some other occasions as well. I want to thank Hugues Leblanc, Henry Hiz, Dennis Stampe and LeVerne Shelton for criticisms and helpful suggestions.

(1) G. FREGE, «On Sense and Reference,» in Frege, *Philosophical Writings* (Peter Geach and Max Black eds.) Oxford: Blackwell, 1952.

(2) See, for instance, W.V. QUINE, «Two Dogmas of Empiricism,» in *From a Logical Point of View*, Cambridge: Harvard University Press, 1953, p. 21, and *Word and Object*, Cambridge: MIT Press, 1960, p. 201.

(3) See, QUINE, «On What There Is,» in *From a Logical Point of View*, p. 9 and R. CARNAP, *Meaning and Necessity*, Chicago: University of Chicago Press, 1947, p. 118 ff. The distinction between theory of reference and theory of meaning has been questioned, however, for instance by Hintikka. See the opening pages of his «Semantics for Propositional Attitudes,» in *Models of Modalities*, Dordrecht: Reidel, 1969, p. 87 ff.

(4) QUINE, «Notes on the Theory of Reference,» in *From a Logical Point of View*, pp. 130-131.

(5) To speak of the ontological commitments of a theory or sentence is to indicate which existential sentences or existence claims follow from the theory or sentence. These may, of course, be true or false, and still the commitments of the theory. But if they are false then there are no such objects.

(6) Cf. QUINE, «Existence and Quantification,» in *Ontological Relativity and Other Essays*, New York: Columbia University Press, 1969, p. 93 ff.

(7) Names or individual constants are to be thought of here as introduced via definitions of the form '*a*' for '(ɿx)Ax' following Quine's assimilation of names and definite descriptions. Thus '*a*' need not purport reference to an object in a theory T since 'A' may denote the null set. The logic I assume here drops the presupposition that every name names an object.

(8) We may view designation as involving description in view of note 7 above.

(9) Israel SCHEFFLER, *Science and Subjectivity*, New York: Bobbs-Merrill, 1967, p. 58.

(10) Cf., for instance, Albert Blumberg's discussion of validity in «Logic, Modern,» *The Encyclopedia of Philosophy*, New York: Mac-Millan and The Free Press, 1967, Vol. 5, p. 31.

(11) John SEARLE, «Proper Names and Descriptions,» in *The Encyclopedia of Philosophy*, Vol. 6, p. 488.