182
Notre Dame Journal of Formal Logic
Volume XXI, Number 1, January 1980
NDJFAM

REFERENTIAL OCCURRENCE

DAVID E. COOPER

In his recent book *Meaning and Modality* (Cambridge University Press, 1976), Casimir Lewy presents an objection to Quine's criterion for referential occurrence, and offers a criterion of his own. First I shall argue that Lewy's own criterion cannot be correct, and second, that the counterexample on which his objection to Quine rests is illegitimate.

Quine's criterion is as follows:

'a' has referential occurrence in Fa iff $(e)(e \text{ COREF } a \supset\ !\ Fa/e)$;

which is to be read as:

'a' has referential occurrence in Fa iff, for the substitution of any expression coreferential with 'a' for 'a', the truth-value of Fa is preserved.

Hence, 'cat' occurs referentially in "This cat is brown", but not in " 'cat' is an English word". (Actually Quine sometimes says that, in a sentence like the second, it is not 'cat' which occurs, but ' 'cat' '. Following Lewy, however, "I shall talk. . . as if it were correct to speak of the occurrence of the interior of a quotation in a sentence which contains that quotation" (p. 24)).

Lewy proposes the following counterexample to Quine's criterion: it is possible that all expressions coreferential with 'cat' should have three letters, in which case the sentence " 'cat' has three letters" would remain true whichever of these expressions replaced 'cat'. Hence, by Quine's criterion, 'cat' would be occurring referentially in that sentence. But clearly it would not. As Lewy says, that sentence "still wouldn't be about cats" (p. 25).

Lewy's own proposal is that for an occurrence to be referential it is required that "substitution (of coreferentials) be *necessarily* truth-preserving—that it be *logically impossible* for it to lead from truth to falsehood" (p. 27). Since it would be a merely contingent fact that substitution of any expression coreferential with 'cat' should preserve the truth of " 'cat' has three letters", then by Lewy's criterion, 'cat' would not be occurring referentially in that sentence. Formally, Lewy's criterion is:

Received March 9, 1977

'a' occurs referentially in Fa iff $\Box(e)(e \text{ COREF } a \supset \,!\, Fa/e)$.

For the time being, I ignore whether Lewy's alleged counterexample is genuine. I want, first, to show that Lewy's own criterion is plainly unsatisfactory. Consider the sentences:

(A) Either 'cat' has three letters or 'cat' has not three letters.
(B) 'cat' has at least one letter.

Both (A) and (B) are necessarily true, and remain necessarily true whatever expression coreferential with 'cat' is substituted in them. Since it is logically impossible for such a substitution to lead from truth to falsehood then, by Lewy's criterion, 'cat' has referential occurrence in (A) and (B). Clearly, though, it does not. Neither sentence is about cats.

One might expect Lewy to repair matters by requiring that substitution of coreferentials should *entail*, and not simply *strictly imply*, truth-preservation (in some sense of 'entail' that is narrower than that of 'strictly imply'). The right-hand side of Lewy's criterion would then become:

$$\ldots(e)(e \text{ COREF } a \text{ ent. } !\, Fa/e).$$

Such a repair may seem attractive for the following reason: plainly the fact that 'e' is coreferential with 'cat' is *irrelevant* to the fact that the truth of (A) and (B) is preserved when 'e' replaces 'cat'. Substitution of *any* expression will necessarily preserve the truth of (A) and (B).

There are two comments to make on such a proposal. First, it is not open to Lewy himself to make since, in later chapters of his book, he argues that "the best decision is to identify entailment with strict implication" (p. 135). (He is particularly hostile to introducing the notion of relevance into that of entailment.) Second, the proposal should only be adopted as a last resort, for despite the efforts of Anderson and Belnap, there is as yet no standard and generally agreed notion of entailment narrower than that of strict implication.

A suggestion for overcoming the problem posed by (A) for Lewy might be this: where 'a' occurs in a molecular sentence, it occurs referentially iff it occurs referentially, by Lewy's criterion, in each component atomic sentence. So, since 'cat' occurs non-referentially, by that criterion, in each disjunct of (A), it occurs non-referentially in (A) itself. There are three comments to make on this suggestion:

(i) it is *ad hoc*

(ii) it is not clear that (A) cannot be reformulated so as to resist the suggested treatment (e.g., as "'cat' has the property of being three-or-not-three-lettered")

(iii) (most important) no similar treatment can be given to (B); yet it would be most counterintuitive to suppose that the problems posed by (A) and (B) are different ones, requiring different solutions.

The correct solution, I think, requires an addition to Quine's original criterion—and a purely extensional one at that. (I am still shelving the question of whether Lewy's alleged counterexample is genuine.) The new criterion is:

$$\text{`}a\text{' occurs referentially in } Fa \text{ iff } (1)\ (e)(e \text{ COREF } a \supset\ !\ Fa/e) \text{ and } (2)\ (\exists e)(-!\ Fa/e)$$

(with (2) to be read as: There is at least one expression, substitution of which for '*a*' in *Fa*, does *not* preserve the truth-value of *Fa*.) By this criterion, 'cat' does not occur referentially in (A) and (B), since (2) is not satisfied. Substitution of *any* expression for 'cat', and not simply of those coreferential with it, preserves their truth.

Someone—Lewy perhaps—will ask how one guarantees that no sentence can be constructed such that: (a) an expression clearly occurs non-referentially in it, (b) substitution of any expression coreferential with it preserves the truth of the sentence, but (c) substitution of some other expression does not preserve truth. I am not sure that one can guarantee this, but I think one could by a generalization of the argument (to be offered later) against the possibility of a counterexample like Lewy's "'cat' has three letters". Moreover, the following consideration provides a strong presumption against the constructibility of a sentence meeting the conditions (a)-(c). There is no natural limit—orthographic, phonetic, morphological, etc.—to the variety of expressions coreferential with a given one. This is a consequence of the conventionality of the relation between expressions and what they stand for, of phonetic/semantic, orthographic/semantic, etc., pairings. All one could expect to be in common to members of a class of coreferential expressions is that they have the same reference. With a certain exception to be discussed below, it becomes difficult to conceive of a statement about those members which is true of each of them but not true of every expression—because, considered apart from their reference, they constitute an arbitrary class.

The exceptional case is illustrated by the following sentence:

(C) 'cat' is coreferential with 'cat'.

If an expression is substituted for the first occurrence of 'cat', the truth of (C) is preserved iff that expression is coreferential with 'cat'. It would not be plausible to handle (C) in the way that some philosophers (Quine ?) might wish to handle

(C′) 'cat' denotes cat;

namely, by holding that 'cat' *does* occur referentially at the beginning of the sentence, since the context "'. . . .' denotes ___" serves to *disquote* the first expression. For from (C), certainly, one can infer nothing about cats—not even that there are, or might be, any.

I find it difficult, however, to be troubled by the problem allegedly posed by (C). It might be possible to exclude cases like (C) by adding extra clauses to my criterion of referential occurrence—but this would wear an *ad hoc* appearance. It is better, I think, to stress that my clauses (1) and (2) are meant as a genuine *criterion*, i.e., *test*, for referential occurrence and thereby for coreferentiality. ('*e*' and '*a*' will be coreferential iff they are substitutable *salva veritate* in all sentences in which they occur referentially. There is an inevitable, and hence tolerable, element of circularity here. But it is surely perfectly respectable to employ previous evidence that an occurrence is referential to help determine when expressions are coreferential, and vice versa.) When two expressions, '*e*' and '*a*', pass the test, we are entitled to assert:

(C″) '*e*' is coreferential with '*a*'.

((C) is an instance of the limiting case where, in (C″), '*e*' = '*a*'.) Clearly it would be absurd to employ sentences of the type (C″) as ones against which to test for referential occurrence and coreferentiality. If one had to have reason for thinking '*e*' and '*a*' are interchangable in (C″) *before* thinking they are coreferential, one could not begin to test for their coreferentiality. I suggest, then, that in the statement of my criterion, there are implicit restrictions on the range of sentences which might replace *Fa*. In particular, sentences are excluded which serve to state the results of expressions passing or not passing the test laid down by the criterion. These restrictions are ones which any genuinely usable criterion for referential occurrence and coreferentiality *must* respect.

My criterion, like Quine's, is only acceptable if Lewy's alleged counterexample can be warded off. If he is right, 'cat' in "'cat' has three letters" could satisfy clauses (1) and (2), but would not thereby be occurring referentially. But *could* it be the case that all expressions coreferential with 'cat' have three letters? I think not.

I try to establish the categorical conclusion that not all such expressions could have three letters only after trying to establish the conditional conclusions that this must be so *if* the problem of 'cat's referential occurrence can be stated, and that this must be so *if* 'cat' can have non-referential occurrence. My arguments are formally valid, although the modal premises can only be established, if at all, by informal philosophical reflection.

Consider the following four propositions:

(α) (Given that '*a*' is a meaningful expression in a language $\mathcal{L}$) the expression '*a*' can have non-referential occurrence in sentences intelligible to users of $\mathcal{L}$.[1]

(β) The question of whether '*a*' is occurring referentially or not in a sentence can be intelligibly raised by users of $\mathcal{L}$.

(γ) The semantics of '*a*' can be expressed by users of $\mathcal{L}$.

(δ) Not every expression coreferential with '*a*' contains the same number of letters as '*a*'.

My first argument is:

$$\text{(I) } (\beta) \rightarrow (\gamma)$$
$$\text{(II) } (\gamma) \rightarrow (\delta)$$
$$\therefore \text{(III) } (\beta) \rightarrow (\delta).$$

('$\rightarrow$' signifies strict implication.) If the argument is correct it shows, I think, that Lewy's position is self-defeating—for by even raising the question of whether 'cat' is occurring referentially, he is necessarily committed to employing language in which not every expression coreferential with 'cat' has three letters. Hence it is not possible, as he imagines it is, for him or anyone else to truly *say* "I can ask if 'cat' occurs referentially in certain sentences, and every expression coreferential with 'cat' has three letters". I think this is already a serious objection to Lewy although he, I think, would not. For he would want to distinguish between (i) the possibility of truly uttering the above sentence, and (ii) the possibility of what is expressed by that sentence being true. And he would argue that what is expressed might be true even if it would not be true if anyone has the means of expressing it. (Compare the sentence "No one possesses a language".)

I shall not, here, challenge Lewy's distinction, but proceed instead to my second argument, which is:

$$\begin{array}{rl} (\text{I})' & (\alpha) \rightarrow (\beta) \\ (\text{I}) & (\beta) \rightarrow (\gamma) \\ (\text{II}) & (\gamma) \rightarrow (\delta) \\ \therefore (\text{IV}) & (\alpha) \rightarrow (\delta). \end{array}$$

That is: if 'cat' does or can have non-referential occurrence in sentences intelligible to speakers of English, if follows that not every expression coreferential with 'cat' has three letters. So it is not merely the case that one cannot truly utter the proposition that 'cat' may have non-referential occurrence and that all coreferential expressions have three letters; that proposition itself cannot be true.

Lewy might still remain relatively undaunted, for I have not shown that (δ) must be true—only that its truth is entailed by the truth of (α). So I now proceed to my third argument, which is:

$$\begin{array}{rl} (\text{I}'') & \Box(\alpha) \\ (\text{I}') & (\alpha) \rightarrow (\beta) \\ (\text{I}) & (\beta) \rightarrow (\gamma) \\ (\text{II}) & (\gamma) \rightarrow (\delta) \\ \therefore (\text{V}) & \Box(\delta). \end{array}$$

How do I establish the premises (I), (II), (I′), and (I″)? I can do no more than outline the very deep and complicated considerations which, I believe, establish them.

(I) ((β) $\rightarrow$ (γ)): To intelligibly raise the question of whether '*a*' is occurring referentially in a sentence one must understand the notion of the reference of '*a*'. But to understand this notion it is necessary that one can semantically specify the reference of '*a*', if only in an uninformative way such as "the reference of '*a*' ". For it is a condition of understanding the notion of '*a*'s reference that one implicitly recognizes as true such sentences as " '*a*' refers to *a*", "The reference of '*a*' is *a*", etc.

(II) ((γ) $\rightarrow$ (δ)): If there are means for semantically specifying the reference of '*a*', it follows that not all expressions coreferential with it have the same number of letters as it. For example 'the reference of 'cat' ', 'what 'cat' refers to', etc., do not contain the same number of letters as 'cat'. Even if, *per accidens*, some linguistic device for semantic specification of the reference of 'cat' produced an expression having three letters, the recursive applicability of the device to its own former product will guarantee that expressions coreferential with 'cat' can be generated which do not have three letters. Someone may argue that since no token of an expression coreferential with 'cat', but not having three letters, might ever have been produced, then all *actual* expressions coreferential with 'cat' might have been three-lettered. But it would be wrong to thus equate the existence of an expression with its actually having received phonetic or orthographic realization. It is better to identify the existence of expressions with the existence of certain devices for producing tokens. Otherwise one would have to deny, *inter alia*, that there are English sentences which have never been spoken or written.

(I′) $((\alpha) \rightarrow (\beta))$: Suppose speakers of $\mathcal{L}$ cannot raise the question of whether '*a*' might be occurring non-referentially in sentences. Then, it would seem, they are unable to raise the question of whether a sentence is about '*a*' or about what '*a*' stands for. But, in that case, they could not understand sentences, if there are any, in which '*a*' occurs non-referentially. For to understand a sentence about '*a*', they must at least understand *that* it is about '*a*'. And this they could not do unless they were able to raise the question of *whether* it is about '*a*' or not. Hence: if the speakers cannot raise the question of whether '*a*' is occurring referentially, they cannot understand sentences in which '*a*' occurs non-referentially. Hence: if there can be sentences intelligible to them in which 'a' occurs non-referentially (i.e., if (α) is true), then it is possible for them to raise the question of whether an occurrence of '*a*' is referential or not (i.e., then (β) is true).

(I″) $(\Box(\alpha))$: In order for speakers to understand sentences in which 'cat' occurs referentially, it is at least necessary that they (i) recognize the word 'cat' as the word 'cat', (ii) distinguish between well-formed and ill-formed sequences containing 'cat', and (iii) understand the meaning of 'cat'. It is not, I think, conceivable that speakers should acquire the abilities illustrated by (i)-(iii)—that, in other words, they should learn their language—unless means were available for saying things which serve to (i) identify words, (ii) state well-formedness restrictions, and (iii) elucidate meanings. (Things like: " 'cat' is a three-lettered noun", "You can't say 'cat am up' ", "That is not the sort of creature properly called 'a cat' ".) I believe this is more than a matter of empirical necessity. It is a conceptual or logical point, in which case I *am* entitled to use Lewis' '$\Box$' operator in front of (α) and not merely some other modal operator. For unless means of the kind mentioned were available, there could be no way of even debating, much less settling, whether different speakers were employing the same words, the same grammar, or the same system of meanings. But then there could be no way of debating, much less settling, whether the different speakers were using the same language—whether, that is, there was *an* $\mathcal{L}$ of which they were all speakers.

NOTE

1. It should be noted that (α) does not claim that 'a' can occur non-referentially *in* $\mathcal{L}$. So it would be no objection to (α) if $\mathcal{L}$ were an object-language so characterized that, in it, 'a' cannot occur non-referentially. Analogous remarks may be made concerning what is *not* claimed by (β) and (γ).

University of Surrey
Surrey, England