

**AN ANALYSIS OF JOHN R. SEARLE'S *HOW TO  
DERIVE "OUGHT" FROM "IS,"*<sup>1</sup>**

**Doç. Dr. TEO GRÜNBERG**

Searle sets himself the task of proving, by means of giving a counterexample, the falsity of the classical thesis that "no set of statements of fact by themselves entail any statement of value" (p. 43). The counterexample presented is qualified in the following way: "We have thus derived (in as strict a sense of "derive" as natural languages will admit of) an "ought" from an "is". And the extra premises which were needed to make the derivation work were in no case moral or evaluative in nature. They consisted of empirical assumptions, tautologies, and descriptions of word usage" (p. 48).

The counterexample in question consists in the following derivation:

(1) Jones uttered the words "I hereby promise to pay you, Smith, five dollars".

(1a) Under certain conditions C anyone who utters the words (sentence) "I hereby promise to pay you. Smith, five dollars" promises to pay Smith five dollars.

(1b) Conditions C obtain.

(2) Jones promised to pay Smith five dollars.

(2a) All promises are acts of placing oneself under (undertaking) an obligation to do the thing promised.

(3) Jones placed himself under (undertook) an obligation to pay Smith five dollars.

(3a) Other things are equal.

(3b) All those who place themselves under an obligation are, other things being equal, under an obligation.

<sup>1</sup> Published in *The Philosophical Review*, Vol. LXXIII, No. 1, January 1964, pp. 43-58. (All page numbers in the main text refer to this article.)

(4) Jones is under an obligation to pay Smith five dollars.

(4a) Other things are equal.

(4b) Other things being equal, one ought to do what one is under an obligation to do.

(5) Jones ought to pay Smith five dollars.

Searle points to the fact that the *ceteris paribus* clauses (3a) and (4a) frequently involve evaluative considerations, but that "it is not logically necessary that they should in every case" (pp. 46, 48). He remarks further that even if these clauses were involving necessarily evaluations "we can always rewrite... steps (4) and (5) so that they include the *ceteris paribus* clause as part of the conclusion. Thus... we would then have derived "Other things being equal Jones ought to pay Smith five dollars", and that would still be sufficient to refute the tradition, for we would still have shown a relation of entailment between descriptive and evaluative statements" (p. 48).

(1) is obviously a descriptive statement (one about the linguistic behavior of Jones), (1a) is considered as stating "a fact about English usage" (p. 44), or as "a description of word usage" (p. 48). So it seems that Searle construes (1a) as an *empirical* statement (and not at all as an analytic one). Concerning (1b), he points that "the conditions under which a man who utters "I hereby promise" can correctly be said to have made a promise, are straightforwardly empirical conditions" (p. 45). He considers further the premisses (2a), (3b) and (4b) as *tautological*. The remaining two premisses (3a) and (4a) can be made irrelevant by incorporating them respectively in the conclusions (4) and (5) in the manner indicated above. On the other hand, it is obvious that the derivation (1)-(5) (or (1)-(5') -(5')) being the statement '*other things being equal Jones ought to pay Smith five dollars*- is *formally correct*. Hence, if the premisses had really the characters attributed to them by Searle, an *evaluational statement* (viz., (5) or (5')) would have been derived from *empirical assumptions* (viz., (1) and (1b); *tautologies* (viz., (2a), (3b) and (4b)); and a *description of word usage* (viz., (1a)).

Now I think that there is no problem about the character of premisses (1) and (1b). They are both empirical (descriptive) statements. We could also admit without discussion that (3b) and (4b) are, if anything, *analytic*.<sup>2</sup> There

<sup>2</sup> From now on I shall use the term '*analytic*' in place of 'tautological', since the latter is applied by most contemporary logicians exclusively to truth-functional truths.

remain only two premisses, namely (1a) and (2a) which give way to dispute. Searle contends that:

- (i) Premiss (1a) is a *descriptive* statement (an *empirical* statement).
- (ii) Premiss (2a) is *analytic* (and not at all an evaluational statement).

We see that Searle's argument can be refuted only by refuting one at least of these two contentions (i) and (ii). We can therefore confine our analysis to them.

Let us first consider (ii). Searle defends it by saying that "promising is, by definition, an act of placing oneself under an obligation" (p. 45). Also, referring to the derivation (1)-(5), he says that "the whole proof rests on an appeal to the constitutive rule that to make a promise is to undertake an obligation" (p. 56). Indeed, Searle considers (2a) as expressing a *constitutive rule* of the institution of promising. I would rather explain the analyticity of (2a) by pointing to the (presumed) fact that nobody would consider a certain utterance as a promise unless that utterance is taken as entailing an obligation.<sup>3</sup> Furthermore we could agree with Searle that a promise *is* a certain speech act. The situation is similar to the following one: To have a banknote of five dollars *is* to have a certain rectangular bit of paper with green ink on it (p. 54). But although my having a certain bit of paper does not logically entail that I can exchange it against certain goods, my having a banknote of five dollars does (other things being equal) logically entail that I can make such an exchange. Just in the same way, though my uttering certain words does not logically entail that I undertake any obligation, my making a promise does logically entail my undertaking an obligation. But making a promise is nothing else but my uttering certain words (in certain determinate conditions). In short, I am inclined to agree with Searle regarding his making the assumption (ii).

So we remain only with assumption (i). Searle's attempt of deriving "ought" from "is" is successful if and only if either (i) is true, *i. e.*, premiss (1a) is in its literal sense an *empirical* statement, or else (1a) is *analytic*. Now there are four possible alternatives:

- (A) (1a) is an empirical statement,
- (B) (1a) is an analytic statement,

<sup>3</sup> This could be established by means of tests such as those proposed by Arne Naess. (Cf. Arne Naess, "Toward a Theory of Interpretation and Preciseness", reprinted in Linski, L., *Semantics and the Philosophy of Language*.)

(C) (1a) is an evaluational statement,

(D) (1a) is not a statement at all (emotive theory of ethics).

Searle's attempt is succesful if and only if (A) or (B) is the case, *i. e.*, if and only if neither (C) nor (D) is the case. I shall now proceed to the analysis of these four alternatives.

*Alternative (A): (1a) is an empirical statement.*

Then (1a) is either an empirical generalization or a "theoretical" statement.

*First case : (1a) is an empirical generalization.*

Let us use 'U' as short for the predicate 'uttered the words "I hereby promise to pay you, Smith, five dolars"'; 'C' as short for a description of the conditions in which Jones uttered the above mentioned words; and 'P' as short for the predicate 'promised to pay Smith five dolars'. Then, sentence (1a) can be expressed as

$$\forall x[Cx \rightarrow (Ux \rightarrow Px)]$$

This sentence can be transformed into

$$\forall x(Cx \wedge Ux \rightarrow Px)$$

Using 'Q' in such a way that 'Qx' is short for 'Cx  $\wedge$  Ux', we get finally

$$(1a') \quad \forall x(Qx \rightarrow Px)$$

We can consider 'Q' as an observation predicate; while 'P' if anything is not an observation predicate, unless 'P' is construed as synonymous with 'Q', or better to some disjunction of observation predicates (*e. g.*, 'Px =<sub>DF</sub> Qx  $\vee$  Q<sub>1</sub>x  $\vee$  Q<sub>2</sub>x  $\vee$  ...'). Now to believe in such a synonymy is to commit the "naturalistic fallacy". But even if such a definition were granted (1a') would not be empirical but rather *analytic* (alternative (B)). So let us consider only the case that 'P' is *not* synonymous with 'Q' (nor with any other observation predicate).

Now (1a') is an empirical generalization only if it is confirmable by means of such singular statements as 'Qa  $\wedge$  Pa', 'Qb  $\wedge$  Pb', 'Qc  $\wedge$  Pc', . . . But since 'P' is *not* an observation predicate, these singular statements are themselves not confirmable by direct observation. Hence (1a') (*i. e.*, (1a) also) is not an empirical generalization.

*Second case: (1a) is a theoretical<sup>4</sup> statement*

In this case 'P' is a "theoretical term" and so is *a fortiori* the predicate 'undertakes an obligation'. We can then consider the statements (1a') and (2a) as the postulates determining the meaning of these two theoretical terms. (1a') would be a C- postulate (correspondence rule) and (2a) a T-postulate (theoretical postulate).<sup>5</sup>

Since (3b) and (4b) are analytic, we can consider as a T- postulate, instead of (2a) the statement

(2b) Anyone who promised to pay Smith five dollars ought, other things being equal, to pay Smith five dollars.

which follows from (2a), (3b) and (4b). Let us use 'O' as short for the predicate 'ought, other things being equal, to pay Smith five dollars'. We can then express (2b) in the form

$$(2b') \quad \forall x(Px \leftrightarrow Ox)$$

We consider 'P' and 'O' as the theoretical terms and 'Q' as the observation term of our "theory", while (1a') is the C-postulate and (2b') the T-postulate. Then the *Ramsey-sentence*<sup>6</sup> of the theory will consist in

$$\exists F \exists G [\forall x(Qx \leftrightarrow Fx) \wedge \forall x(Fx \leftrightarrow Gx)]$$

The latter statement is analytic, since it follows —by means of a double application of the rule of existential generalization— from the logical truth

$$\forall x(Qx \leftrightarrow Qx) \wedge \forall x(Qx \leftrightarrow Qx)$$

Now generally the empirical content of a "theory" is determined by its Ramsey-sentence. But in our case the Ramsey-sentence is itself an analytic

4 Cf. R. Carnap, "The Methodological Character of Theoretical Concepts", in Feigl, H. and Scriven, M., *Minnesota Studies in the Philosophy of Science*, Vol. I.

5 Cf. *ibid.*, esp. p. 43 and pp. 47 f.

6 T being the conjunction of the T-postulates, C the conjunction of the C-postulates and TC the conjunction of T and C, the *Ramsey-sentence* is the statement R obtained from TC by substituting in the latter suitable variable for all its theoretical terms and by binding existentially these variables. Cf. R. Carnap, "Beobachtungssprache und Theoretische Sprache", in *Logica: Studia Paul Bernays Dedicata*, (Editions du Griffon, Neuchatel-Suisse, 1959). (Cf. also C.G. Hempel, "The 'Theoreticians' Dilemma", in *Minnesota Studies in the Philosophy of Science*, Vol. II, pp. 80-81.)

statement. Hence the theory, and so each postulate, can be considered as analytic.<sup>7</sup> Thus (1a') (*i. e.*, (1a) also) is *analytic*.

There is a way of considering (1a) as empirical while still construing 'P' as a theoretical term. Namely, we can enlarge our "theory" by introducing some additional C-postulates. In particular we can construe an "ought" such as to entail certain observable events, say the occurrence of a *sui generis* "feeling" of obligation. Let 'R' be short for the predicate 'has a feeling of obligation'. Then we can add the C-postulate (correspondence rule)

$$(5a) \quad \forall x (Ox \leftrightarrow Rx)$$

Such a theory has clearly an empirical (observational) consequence. Indeed, from (1a'), (2b') and (5a) we can deduce

$$\text{a} \quad \forall x (Qx \leftrightarrow Rx)$$

where both 'Q' and 'R' are observation terms.

However (5a) is inadmissible. Indeed if (5a) were true, then 'O(a)' being any ought-statement, one could refute 'O(a)' merely by observing that R(a) is not the case. But the word 'ought' is used in such a way that no ought-statement can be falsified by sense experience. Consequently we see that if 'P' is a *theoretical* term, then (1a) must be considered as *analytic*.

Now, so far as I can see, Searle is himself construing the term 'P' neither as an observation term nor as a theoretical term. In the contrary case, 'P' would be descriptive. In Searle's terminology, *that a is P* would be a brute fact, not an institutional one. Indeed, Searle says that (1a) states a fact about English usage, or is a description of word usage. Hence according to him, (1a) consists in a description of the usage of the word 'promise' in English. But in its literal sense, (1a) is not at all a description of the use of the word 'promise', unless it be interpreted as an elliptical form of

(1a") According to the Anglo-saxons, uttering in certain conditions C the words "I hereby promise..." is called a promise.

But then instead of (2) one gets

<sup>7</sup> An *analytic* statement is explicated as one which follows (logically) from the conditional statement  $R \rightarrow TC$ . Now if R is logically true, TC is analytic, since it follows then from  $R \rightarrow TC$ . TC consists in the conjunction of all (T-and-C-) postulates of the theory, Thus each one of these postulates is analytic, since it follows from TC and hence from  $R \rightarrow TC$ . (Cf. Carnap, *op. cit.*, pp. 41-42.)

(2') Jones did what the Anglo-saxons call promising,  
and finally instead of (5') one gets

(5'') According to them, Jones ought, other things being equal, to pay Smith five dollars.

But as stated by Searle himself (pp. 51-52), statement (5'') is a statement of fact. Therefore I can say that Searle must renounce to construe (1a) as "a description of word usage" (since that would annihilate his argument). But then he is forced to consider it as a non-empirical statement, *i. e.*, he must renounce to make his assumption (i).

*Alternative (B): (1a) is an analytic statement.*

Since (2a), (3b) and (4b) are analytic, the statement

(1c)  $\forall x(Qx \rightarrow Ox)$

which follows from (1a), (2a), (3b) and (4b) will be *analytic* in case (1a) is analytic. So if we could show that (1c) is *not* analytic, we would have refuted (B).

Now (1c) is clearly not analytic. I think that Searle himself would deny that (1c) is a "tautology". But even if he were inclined to consider (1c) as analytic, he could not, for that would be question-begging for his argument. Indeed his aim consists in deriving an evaluative statement of form 'O(a)' from a descriptive statement of form 'Q(a)'. But then, to consider ' $\forall x(Qx \rightarrow Ox)$ ' and hence ' $Q(a) \rightarrow O(a)$ ' as analytic, is to trivialize the proof. In any case, it cannot be said that any person who does not accept the truth of (1c) is misusing the English word 'ought'. To maintain that a statement like (1c) is analytic, is to maintain that the whole of morals is merely a grammatical affair, so that a moral fault is nothing but a grammatical mistake. That is too ridiculous a view. So we can state safely that (1a) is *not analytic*. But we have seen that it is also not an empirical statement either. Hence it must be either an evaluational statement or else no statement at all.

*Alternative (C): (1a) is an evaluational statement:*

(2a), (3b) and (4b) being analytic, the statement (1c) which follows from (1a) by means of these three analytic statements will be itself an evaluational statement in case (1a) is itself an evaluational one. Unless we adopt the emotive theory according to which (1c) is no statement at all, it would conform to usa-

ge to consider (1a) as an evaluational (i. e, neither empirical nor analytic) one. So Searle's argument collapses.<sup>8</sup>

*Alternative:* (D): (1a) *is not a statement:*

In this case the derivation (1)-(5) becomes meaningless.

### CONCLUSION

From the four possible alternatives (A), (B), (C) and (D) about the nature of sentence (1a), I have shown that the first two, (A) and (B), are untenable. As to (C) and (D), they cause both the collapse of Searle's argument, (C) being incompatible with his thesis and (D) making his argument meaningless.

<sup>8</sup> In case (following W.V. Quine and Morton White) we renounce to the *analytic-synthetic* distinction, Searle's argument loses at once all its force, since it would then reduce merely to a derivation of an ought-statement from premisses containing the word 'ought'.