

An early version of this paper was published in Savas L. Tsohatzidis, ed., *Intentional Acts and Institutional Facts: Essays on John Searle's Social Ontology* (Berlin: Springer 2007). The present, revised, version appears in Savas L. Tsohatzidis, *Truth, Force, and Knowledge in Language: Essays on Semantic and Pragmatic Topics* (Berlin & Boston: De Gruyter 2020).

Searle's derivation of promissory obligation

Savas L. Tsohatzidis

1. Introduction

In “How to derive ‘ought’ from ‘is’” (Searle 1964), perhaps the most famous among his early articles, John Searle set out to show that what is sometimes called “the naturalistic fallacy”—the fallacy that is allegedly committed by those who affirm that it is possible to deduce evaluative conclusions from wholly non-evaluative (“descriptive”) premises—is not at all a fallacy, contrary to what David Hume and the long list of philosophers who follow Hume on this matter have supposed. The reason that it is not a fallacy, Searle claimed, is that there are certain kinds of clearly evaluative statements (specifically, statements about what a particular person *ought* to do or *has the obligation* to do) that *are* deducible from certain sets of wholly non-evaluative (“descriptive”) statements (specifically, from sets of statements that include a statement about what that person has *promised* to do). That being so, Searle contended, not only is there no such thing as “the naturalistic fallacy” but those claiming that there is commit themselves to the denial of the validity of a series of logically impeccable inferences (and so become victims of what might be called “the naturalistic fallacy fallacy”).

Searle presented an improved version of his proposed derivation of “ought” from “is” (in effect, of statements about obligations from statements about promises) in the final chapter of his first major book, *Speech Acts* (Searle 1969), arguing that all criticisms of the derivation that in the meantime had been produced have been unsuccessful, and claiming that the derivation is one among the many philosophically interesting by-products of the general account of meaning and speech acts that that book proposes. In outline, Searle's improved version of his derivation proceeds as follows. From a *descriptive* statement to the effect that a speaker X has uttered, at time *t*, a sentence of the form “I (hereby) promise to *p*”, where *p* refers to a future action of X, we can infer, Searle says, the equally *descriptive* statement that X has promised, at *t*, to *p*, simply by making the empirical assumption that X's utterance satisfies the conditions that the theory of speech acts specifies as necessary and sufficient for a speaker's utterance to constitute a promise on his part to perform a future action. But once these two purely *descriptive* statements are in place, Searle contends, they logically necessitate, in conjunction with three analytic (and so, non-evaluative) statements, a series of clearly *evaluative* statements, and so contradict the thesis this is impossible to logically deduce evaluative conclusions from non-evaluative premises. Specifically: (a) in conjunction with the analytically true statement that *whoever promises, at a given time, to do something, undertakes, at that time, the obligation to do it*, the statement that X promised, at *t*, to *p* entails the statement that X undertook, at *t*, the obligation to *p* (e.g., the statement that Tom promised, at *t*, to give Mary five dollars entails the statement that Tom undertook, at *t*, the obligation to give

Mary five dollars); (b) in conjunction with the analytically true statement that *whoever undertakes, at a given time, the obligation to do something, has, at that time, the obligation to do it*, the statement that X undertook, at *t*, the obligation to *p* entails the statement that X had, at *t*, the obligation to *p* (e.g., the statement that Tom undertook, at *t*, the obligation to give Mary five dollars entails the statement that Tom had, at *t*, the obligation to give Mary five dollars); and finally, (c) in conjunction with the analytically true statement that *whenever one has the obligation to do something, one ought, as regards that obligation, to do it*, the statement that it was true of X at *t* that he had the obligation to *p* entails the statement that it was true of X at *t* that he ought, as regards that obligation, to *p* (e.g., the statement that it was true of Tom at *t* that he had the obligation to give Mary five dollars entails the statement that it was true of Tom at *t* that he ought, as regards that obligation, to give Mary five dollars). But since statements about what a particular person at a particular time *ought* to do, or *has the obligation* to do, are clearly evaluative statements, it follows, Searle concludes, that such statements *can* be logically deduced from appropriate sets of wholly *non-evaluative* statements, contrary to the Humean claim that it is never possible to logically deduce evaluative statements from wholly non-evaluative ones.

Searle's early thesis about the deducibility of statements about obligations from statements about promises deserves attention not only on account of its intrinsic interest, but also on account of the special significance it has acquired in the context of the ambitious philosophical projects that he has launched in *The Construction of Social Reality* (Searle 1995) and in *Rationality in Action* (Searle 2001). One of the main aims of *The Construction of Social Reality* was to explain, compatibly with a naturalistic world-view, how institutional facts are possible. A key element of the proposed explanation is that institutional facts are facts that only exist because groups of agents collectively acknowledge their existence by virtue of accepting constitutive rules of the form "X counts as Y in context C", whereby status functions are assigned to entities. And since, according to Searle, the assignment of a status function to an entity amounts to the creation of a *deontic power* (that is, of a right or of an obligation, or, more generally, of an entitlement or of a requirement) involving that entity (as Searle succinctly puts it [1997: 451], "the imposition of a status-function is the imposition of a deontology"), the fundamental kinds of things that, in the domain of institutional reality, agents bring into existence simply by acknowledging their existence are the deontic powers (e.g., the rights or the obligations) that follow from the collective assignment of status functions to entities. But since, according to Searle, deontic powers (as opposed to brute powers) are not determinable on the basis of the physical constitution of the entities to which status functions are assigned, it is often necessary, in order for a deontic power to be brought into existence, to be publicly and conventionally *represented* as being brought into existence. And since the commonest means of *public* and *conventional* representation is language, language plays, according to Searle, a key facilitating role in the creation and, especially, in the proliferation of institutional facts. Now, if institutional facts are, essentially, impositions of deontologies, and if language plays the indicated key facilitating role in the creation and proliferation of such facts, it should not be surprising to find that an obligation, which is a particular kind of deontic power, is brought into existence by a promise, which is a particular kind of linguistic act. Viewed in the context of *The Construction of Social Reality*, therefore, Searle's early derivation of statements about

obligations from statements about promises does not merely represent, if correct, an isolated instance of the dependence of an extra-linguistic institutional fact on a linguistic fact, but rather a paradigmatically clear, though by no means unique, instance of the pervasive dependence that, according to Searle, obtains between facts of the former kind and facts of the latter kind.

Searle's early derivation has an even more important connection with another major theme of his recent work, namely, the critique of the classical model of practical reason that he develops in *Rationality in Action*.¹ One of the fundamental defects of the classical model of practical reason, Searle contends, is that, according to it, the only reasons for action that agents could possibly recognise and have are their desires or other states essentially dependent on their desires. And the fundamental move that, according to Searle, is required in order for that defect to be overcome, consists in acknowledging that, among the reasons that can motivate agents to act, there are not only the *desire-dependent reasons* that agents may recognise, but also the *desire-independent reasons* that they may recognise, and especially those desire-independent reasons that they cannot *fail* to recognise because they intentionally *create* those reasons for themselves. But understanding the significance of such desire-independent reasons for action amounts, in Searle's view, to understanding the significance of linguistically generated deontic powers, and in particular the significance of promissory obligations. For, according to Searle, the prototypical instance of an intentionally created desire-independent reason for action is the *obligation* that an agent creates for himself when he *promises* to perform some future action: that obligation is clearly a reason that the agent has to perform the future action; and it is, furthermore, a reason that, on the one hand, holds quite independently of the agent's desires (even if an agent does not *want* to do what he has promised to do, the fact that, through his promise, he has placed himself under the *obligation* to do it, provides him with a *reason* for doing it) and, on the other hand, cannot fail to be recognised by the agent, since it was not dictated by others but was purposively brought into existence by the agent himself through his act of promising. Indeed, not only are the obligations created by acts of promising paradigm cases of intentionally created desire-independent reasons for action, but it is also the case, Searle suggests, that virtually all other linguistic acts create, in ways that are perhaps less obvious but no less certain, many other obligations of a similar sort (e.g., an assertion creates for its author the obligation to provide, if challenged, evidence for the truth of its content, a request creates for its author the obligation not to obstruct the realisation, by its addressee, of the state of affairs that the addressee has been requested to realise, etc.); and since each of these obligations constitutes a reason for action that holds quite independently of the agent's desires, it follows that virtually every normal use of language constitutes a counterexample to the classical model's assumption that the only reasons for action that an agent might have are, or are essentially dependent on, the agent's desires. Obviously, Searle's early thesis that statements about promises entail statements about obligations acquires deep significance in this context, and that makes it understandable why it is frequently invoked by him throughout *Rationality in Action*. For although, as Searle now

¹ See also Searle (1999) for a preliminary statement of that critique, explicitly connected with the topic of Searle's early derivation.

stresses, promises are not, in his view, unique among speech acts in creating obligations (and therefore desire-independent reasons for action), their conceptual ties with the creation of obligations are so strong and evident that they provide a model on which all linguistically generated obligations (and, therefore, all linguistically generated desire-independent reasons for action) should be analysed.

The main purpose of the present essay is to argue that Searle's thesis that promises necessarily create obligations is open to a family of decisive counterexamples. The reason why these counterexamples have so far remained unnoticed is, in my view, that most of the critical literature that the thesis has provoked (and which I do not propose to review here) does not concern itself with Searle's central contention that statements about promises entail statements about obligations (does not address, that is, what I will henceforth call the *deducibility claim* that Searle makes), but rather effectively grants to Searle that statements of the former kind entail statements of the latter kind and only disputes Searle's *separate* claim that statements of the former kind are purely 'descriptive' whereas statements of the latter kind are purely 'evaluative' (it only addresses, that is, what I shall henceforth call the *categorisation claim* that Searle also makes). My counterexamples concern specifically the deducibility claim to which Searle is centrally committed—the claim, that is, that, necessarily, whenever a statement of the form “X promised to *p*” is true, a corresponding statement of the form “X ought/has the obligation to *p*” is also true—and, if valid, they make it strictly speaking unnecessary to further discuss the categorisation claim: if, contrary to Searle's contention, “X promised to *p*” does *not* entail “X ought/has the obligation to *p*”, then Searle would presumably not expect any momentous philosophical conclusions to follow *merely* from the assumption that “X promised to *p*” is 'descriptive' whereas “X ought/has the obligation to *p*” is 'evaluative', even if that assumption turned out to be correct. I believe, however, that it is itself doubtful whether Searle would currently be entitled to that last assumption, given some independent theses that he has recently propounded; and I consequently believe that it is doubtful whether he would *now* be entitled to all the conclusions he was originally aiming to draw from his discussion of the “is”–“ought” question, even assuming, contrary to fact, that statements of the form “X promised to *p*” do entail statements of the forms “X ought to *p*” and “X has the obligation to *p*”. I will therefore use the present occasion to first offer some critical reflections on the status of categorisation claim, and will then proceed to the primary task of determining the truth or falsity of the deducibility claim.

2. The instability of the categorisation claim

Assume for the sake of argument that statements of the form “X promised to *p*” do entail corresponding statements of the forms “X ought to *p*” and “X has the obligation to *p*”. As Searle recognises, in order for the existence of such entailments to imply anything determinate about the “naturalistic fallacy”, it should be pre-theoretically quite obvious that statements about promises cannot be anything else than purely 'descriptive' and that statements about obligations cannot be anything else than purely 'evaluative'. And Searle, in all presentations of his derivation, does take both of these things to be

pre-theoretically quite obvious—so obvious, in fact, that he spends no time at all *arguing* in their favour. One might doubt, however, that pre-theoretical obviousness could be reasonably assumed in this area (after all, ‘descriptive’ and ‘evaluative’ are not terms like ‘polysyllabic’ and ‘monosyllabic’), and one might accordingly expect that, given different theoretical frameworks within which the labels ‘descriptive’ and ‘evaluative’ are being deployed, different decisions as to how these labels should be applied to statements about promises or to statements about obligations would be called for. I suspect that Searle would simply deny that the conditions of application of the labels ‘descriptive’ and ‘evaluative’ are theory-sensitive in that sense, and that they lack the degree of stability that his categorisation claim presupposes. I propose to show in this section, however, that such a denial would be unwarranted, since some of Searle’s *own* recent theoretical positions commit him to the thesis that statements about promises are no less *evaluative* than statements about obligations, contrary to his explicit early claim that such statements, unlike statement about obligations, are obviously not evaluative.

The source of the instability is Searle’s thesis, frequently asserted in *The Construction of Social Reality* and in subsequent works, that *linguistic facts are institutional facts*—a thesis by which Searle do not mean, of course, that the human capacity for producing speech sounds is itself an institutional phenomenon, but rather that the particular systems of symbolising conventions that particular human communities create by exercising that capacity are institutional phenomena. The fact that a member of a human community has produced a sequence of sounds is, Searle admits, not in itself an institutional fact. But the fact that the very same sequence of sounds is a word, or a sentence, in one human community and neither a word nor a sentence in another human community, or the fact that what are phonetically the same words, or the same sentences, in two human communities may have totally different semantic contents in these two communities are, Searle insists, clearly institutional facts—as a passage in *The Construction of Social Reality* (Searle 1995: 99) explains, such terms as “word”, “sentence” or “semantic content” do not refer to intrinsic properties of the sound structures that are produced by members of a linguistic community, but rather to “statuses” that have been collectively “imposed” on those sound structures by members of that community. Indeed, not only are linguistic facts, in that sense, institutional facts, but they are, according to Searle, *paradigmatic* institutional facts. Thus, in giving what he takes to be especially clear examples of institutional facts that are created out of brute facts through the imposition of “status functions on entities” in accordance with “the general formula ‘X counts as Y in C’”, Searle writes:

I make noises through my mouth. So far, that is a brute fact: there is nothing institutional about it. But, as I am a speaker of English addressing other English speakers, those noises *count as* the utterance of an English sentence; they are instances of the formula ‘X counts as Y in C’. But now, in an utterance of that English sentence, the Y term from the previous level now functions as the X term at the next level. The utterance of that English sentence with those intentions in that context *counts as*, for example, making a promise. (Searle 1998: 128–129; original emphasis)

On Searle's current view, then, both the fact that a sequence of sounds in a *sentence* of a particular human language and the fact that the utterance of that sequence of sounds constitutes the performance of a particular *illocutionary act* in that language, are *institutional* facts holding in the community of humans that are speakers of that language.

But now recall that, according to Searle, *whenever* an institutional fact is brought into existence through the assignment of a status-function to an entity certain *deontic powers* are necessarily created—that is, certain rights and obligations (or more generally, certain entitlements and requirements) that did not previously exist come to exist within the human group that collectively acknowledges the fact's existence. If that is so (if, to put it in Searle's own words, "everything turns out to be deontic" (1995: 109) in the domain of institutional reality), then it follows that the *mere* fact that a sequence of sounds is the utterance of a *sentence* of a particular language, or the *mere* fact that the utterance of that sentence is the performance of a particular *illocutionary act* in a particular language, are already *deontic* facts of particular kinds, each associated with its proprietary set of entitlements and requirements. But if that is true, and if it is also true, as Searle is independently committed to holding, that *every* statement of a deontic fact is an *evaluative* statement, then it follows that the mere statement that a sequence of sounds is the utterance of a *sentence* of a particular language, or the mere statement that the utterance of that sentence constitutes the performance of a particular *illocutionary act* in that language, are already *evaluative* statements of particular kinds.

The upshot of all this is, of course, that *every* singular statement figuring in the derivation of "ought" from "is" that Searle had originally proposed turns out to be, by his own stipulations, and contrary to his original intentions, ultimately evaluative. Searle's derivation, it will be recalled, was supposed to show that from certain purely *non-evaluative* singular statements such as (1) and (2)

- (1) Tom uttered the sentence "I hereby promise to give Mary five dollars".
- (2) Tom promised to give Mary five dollars.

one can validly deduce, through the mediation of appropriate analytic truths, certain deontic—and therefore, according to him, unquestionably *evaluative*—singular statements such as those in (3) and (4):

- (3) Tom has the obligation to give Mary five dollars.
- (4) Tom ought to give Mary five dollars.

But, in view of Searle's current conception of linguistic reality as a particular kind of institutional reality, it now turns out that (1) and (2) are *themselves*, in their own subtle way, no less *evaluative* statements than (3) and (4) are, since (1) and (2) report particular *institutional* facts of the linguistic variety, since every institutional fact is ultimately, according to Searle, a *deontic* fact, and since every statement of a deontic fact is, according to Searle, an *evaluative* statement: As it should be clear by now, a sequence of sounds cannot, for Searle, be Tom's utterance of the *sentence* "I hereby promise to give Mary five dollars" unless Tom is a member of a linguistic population where that sequence of sounds is collectively *accepted* (or *counted*) as the utterance of a sentence; and a sequence of

sounds cannot, for Searle, be Tom's *promise* to give Mary five dollars unless Tom is a member of a linguistic population where the utterance of that sequence of sounds is collectively *accepted* (or *counted*) as the making of a promise. But if (1) and (2) report facts that are institutional, and therefore, in their own subtle way, no less deontic than those reported in (3) and (4), then, given Searle's non-negotiable assumption that deontic statements are necessarily evaluative, it follows that (1) and (2) are no less evaluative statements (though, of course, less *blatantly* evaluative statements) than (3) and (4) are. Consequently, even if it should turn out, as Searle claims, that, from (1) and (2), one can validly deduce (3) and (4), it would *not* follow, as Searle *also* wanted to claim, that this shows that certain evaluative statements can be deduced from certain *non*-evaluative ones. In short, Searle's early claim that "uttering certain words in certain conditions just *is* promising, and the description of these conditions needs no evaluative element" (Searle 1964: 50; second emphasis added) is simply inconsistent with his current view that the conditions under which a series of sounds constitutes a promise are conditions whose satisfaction transforms the occurrence of that series of sounds into an institutional fact by imposing on it a particular *deontology*, and are, therefore, as full of 'evaluative elements' as any deontological condition is.²

Of course, it would be open to Searle to try to avoid this result by radically modifying his theory of institutional facts in a way that would allow him to withdraw his current commitment to regard statements like (1) and (2) as subtly evaluative, and so to restore his early claim that such statements are definitely not evaluative. But not only do I not see any basis on which such a move could be motivated within Searle's theory as it is now developed, but it is clear to me that, if attempted, it would only reinforce the point with which the discussion in this section has begun, namely, that the question of the applicability of the labels 'evaluative' or 'non-evaluative' to various kinds of statement cannot be answered in the theory-neutral way that Searle was taking to be possible when he originally proposed the derivation of statements about obligations from statements about promises. If that is so, the only sufficiently well-defined issue that remains to be discussed with respect to Searle's derivation does not concern the categorisation claim that it embodies (nor, consequently, the implications or non-implications of that claim on the "naturalistic fallacy"), but rather the deducibility claim that it embodies. It is to the discussion of that claim that I now turn.

3. The falsity of the deducibility claim

In order to show that, contrary to Searle's deducibility claim, statements about promises do not entail statements about obligations, it would be sufficient to show that there exist conditions under which it is possible for a statement of the form "X promised to *p*" to be true of an individual at a given time but impossible for corresponding statements of

² It is worth noting that, as the second sentence in the above quotation from Searle makes clear, Searle's original position was not merely that "promise" is not an 'evaluative' word, but that none of *the conditions under which* a series of sounds constitutes a promise involves any evaluative element.

the forms “X has the obligation to p ” or “X ought to p ” to be true of that individual at that time.³

Such conditions clearly exist. Suppose that Tom has just heard a story about Othello and Desdemona, but does not know that, unlike himself, Othello and Desdemona are merely creatures of an author’s imagination. If, outraged by what the story represents as Othello’s mistreatment of Desdemona, Tom says, at t , “I hereby promise to kill Othello”, or “I hereby promise to protect Desdemona”, the fact that neither Othello nor Desdemona exist has no power to make it false that Tom has *promised*, at t , to kill Othello and to protect Desdemona. But the fact that neither Othello nor Desdemona exist certainly makes it false that Tom *has the obligation* to kill Othello, or that Tom *has the obligation* to protect Desdemona. An imaginary thing cannot be a thing that one *has the obligation* to kill, or that one *has the obligation* to protect, though it *is* a thing that one can, in one’s ignorance, *promise* to kill or *promise* to protect (and, incidentally, it is also a thing that one can, in one’s ignorance, *believe* to be worth killing or *believe* to be worth protecting). For that reason, neither a statement like (5) nor a statement like (6) is contradictory:

- (5) You did promise to kill Othello, Tom, but, since Othello doesn’t exist, there is, so far, no one that you have the obligation to kill.
- (6) You did promise to protect Desdemona, Tom, but, since Desdemona doesn’t exist, there is, so far, no one that you have the obligation to protect.

And the fact that neither a statement like (5) nor a statement (6) is contradictory suffices for showing that, contrary to Searle’s deducibility claim, it is *not* the case either that “Tom promised to kill Othello” entails “Tom has the obligation to kill Othello” or that “Tom promised to protect Desdemona” entails “Tom has the obligation to protect Desdemona”.

Vacuous names (that is, names without referents, such as “Othello” and “Desdemona”) present an obvious problem for Searle’s derivation. Since Atlantis is a non-existent city, one can be under no *obligation* to liberate Atlantis; but this does not make it impossible for one to *promise* to liberate Atlantis, if one mistakenly believes both that such a city exists and that it is unjustly occupied. Since Sherlock Holmes is only a fictional detective, one can be under no *obligation* to hire Sherlock Holmes for the investigation of a robbery; but this does not make it impossible for someone to *promise* to the victims of a robbery to hire Sherlock Holmes for the investigation of that robbery, if one mistakenly believes, or simply wants to lead *others* to mistakenly believe, that Sherlock Holmes is a real detective. Beliefs in the existence of non-existent entities may sometimes be sufficient in order for promises involving those entities to be possible, but they are never sufficient in order for obligations involving those entities to be possible.

Vacuous names, however, are not the only, or the commonest, kinds of expression that show Searle’s deducibility claim to be untenable. What I shall here call *vacuous descriptions* (that is, definite descriptions that, as a matter of contingent fact, nothing satisfies) give rise to a far wider, indeed in principle unlimited, range of

³ I will be assuming with Searle that “ought”-statements have at least one sense in which they clearly ascribe obligations, and that this is the sense of their occurrence in Searle’s derivation.

counterexamples.⁴ Suppose that Tom, confusing me with another person, incorrectly believes that I lost my wallet, even though the description “my lost wallet” is vacuous, since, in fact, I never *had* a wallet to lose. Now, if Tom says to me, “I hereby promise to find your lost wallet”, I can without contradiction say,

- (7) Tom promised to find my lost wallet, but I never lost, since I never had, a wallet, so he must be confusing me with somebody else.

What I *cannot* say without contradiction, however, is,

- (8) Tom ought to find my lost wallet, but I never lost, since I never had, a wallet, so he must be confusing me with somebody else.

If I never had, and so I never lost, a wallet, then it cannot be the case that Tom *ought* to find my lost wallet; but it *can* be the case that Tom, confused as he was, has *promised* me to find my lost wallet. And since, in that circumstance, “Tom promised to find my lost wallet” will be true even though “Tom ought to find my lost wallet” is not true, the former statement cannot entail the latter, contrary to what Searle’s deducibility claim requires. Similarly, suppose that Tom, confusing his own exam paper with somebody else’s exam paper, and incorrectly believing that his own exam paper contains a spelling mistake, says, “I hereby promise to correct the spelling mistake in my exam paper”, even though the description “the spelling mistake in Tom’s exam paper” is vacuous, since, in fact, Tom’s exam paper contains no spelling mistakes at all. I could then without contradiction say,

- (9) Tom promised to correct the spelling mistake in his exam paper, but he must be confusing his own exam paper with somebody else’s, since his exam paper contains no spelling mistakes at all.

What I could *not* say without contradiction, however, is,

- (10) Tom ought to correct the spelling mistake in his exam paper, but he must be confusing his own exam paper with somebody else’s, since his exam paper contains no spelling mistakes at all.

If Tom has made no spelling mistake, then it cannot be the case that Tom *ought* to correct his spelling mistake, but this does not prevent Tom from incorrectly *believing* that he has made a spelling mistake, and so from *promising* to correct it. And since, in these circumstances, “Tom promised to correct his spelling mistake” can be true even though “Tom ought to correct his spelling mistake” cannot be true, the former of these statements cannot entail the latter, contrary to what Searle’s deducibility claim requires. In short, the existence of vacuous descriptions, no less than the existence of vacuous

⁴ In what follows, I am assuming the ‘attributive’ reading of definite descriptions, whose existence no one, as far as I know, doubts. I note in passing that Searle belongs to those who believe, in addition, that the ‘attributive’ reading of descriptions is the only semantically (as opposed to pragmatically) available one. See Searle (1979: 137–161).

names, makes the inference from “X promised to *p*” to “X ought to *p*” invalid, and thus shows Searle’s fundamental claim that promises necessarily create obligations to be untenable. Indeed, since the vast majority of linguistically possible definite descriptions are such that merely knowing what they mean does *not* amount to knowing whether or not they are vacuous, I suspect that the number of promises that have been exchanged in the course of human history without *any* obligations having thereby been generated must be staggering.

The pertinent generalisation could be formulated as follows. Call “content-clause” the clause *p* that, in statements of the form “X promised to *p*” and “X ought/has the obligation to *p*”, specifies the content of the promise or of the obligation that such statements ascribe. Then, if the content-clause *p* contains a proper name or a definite description and is such that, when it occurs *unembedded*, it is false whenever that name or that description are vacuous, the effect of *embedding* clause *p* in matrix sentences of the forms “X promised to *p*” and “X ought/has the obligation to *p*” is necessarily the falsity of the latter matrix sentence but not necessarily the falsity of the former matrix sentence. Assuming *one* of the senses of the distinction between extensionality and non-extensionality that Searle himself recognises⁵ (the sense in which a sentence is extensional with respect to an item it contains if and only if it entails that the purported referent of that item exists), this can be re-expressed as follows: if the content-clause of a promise-ascription or of an obligation-ascription is independently certifiable as extensional with respect to the names or descriptions it contains, then the promise-ascription is *non-extensional* with respect to those names or descriptions, even though the obligation-ascription *is* extensional with respect to those names or descriptions. In short, the difference between “X promised to *p*[. . . *r* . . .]” and “X ought/has the obligation to *p*[. . . *r* . . .]”, where *p* is a content-clause and *r* a name or description inside that content-clause, is that, when *p* is extensional with respect to *r*, “X ought/has the obligation to *p*[. . . *r* . . .]” is *also* extensional with respect to *r*, even though “X promised to *p*[. . . *r* . . .]” is *not* extensional with respect to *r*.⁶ The claim that promises necessarily create obligations results, then, from the failure to recognise that statements about obligations are extensional in a way in which statements about promises are non-extensional. For it is precisely this difference that makes it possible for statements of the latter kind to be true even though statements of the former kind are false, and it is, in turn, that possibility that prevents statements of the latter kind from entailing statements of the former kind.

It is strictly speaking unnecessary, but it may nevertheless be useful, to point out that the kind of problem exposed above could not be avoided by deploying the sort of strategy that Searle had used in rebutting certain early objections to his derivation that were appealing to the notion of “*prima facie* obligation”. Those objections were aiming to dispute Searle’s claim that promises entail obligations by relying on the observation that the obligation a person assumes in making a promise may sometimes be in conflict with *other* obligations that that person independently has, and may consequently have to be discharged in order for such a conflict between obligations to be resolved. As Searle has pointed out, that observation, though correct in itself, cannot be construed as an *objection*

⁵ See Searle (1983: 181ff; 2004: 174ff).

⁶ I will not examine here the interesting question whether promise-ascriptions or obligation-ascriptions are extensional or non-extensional in senses of that distinction other than the one explained above.

to his derivation: two obligations cannot be in conflict unless they both *exist*, and an obligation cannot, logically, be discharged unless it is already in existence; so, the fact that an obligation incurred by an act of promising may have to be subsequently discharged in order for a conflict between obligations to be resolved has no tendency to show that it was non-existent *at the time* of promising; and the bringing into existence of an obligation by virtue of an act of promising *at the time* of that act of promising is all that Searle's derivation, properly understood, requires.⁷ Notice, however, that this sort of move would be entirely inapplicable to the family of counterexamples considered above. When a vacuous name or a vacuous description figures in the content-clause of a promise, *no obligation whatsoever* is created at the time of promising regarding anything mentioned in the content-clause of the promise (if, for example, I never lost, since I never had, a wallet, and Tom says to me, "I hereby promise to find your lost wallet," no obligation of Tom's to find my non-existent lost wallet is thereby created, no matter what Tom might imagine); but an obligation that has never existed cannot, logically, be discharged, and the question whether the person making such a promise has or does not have any *other* obligations with which his *non-existent* promissory obligation might be in conflict makes no sense. Consequently, the cases considered above cannot be assimilated to cases where a promissory obligation does exist but is *subsequently* discharged because it conflicts with independently assumed obligations; and so, Searle's way of deflating objections to his derivation that rely on situations where conflicts between actually existing obligations are possible is of no use in solving the problem posed by the cases presented above.

Granting the obvious point that, in the cases presented above, no obligations are generated, someone (though, I presume, not Searle) might make a desperate final attempt to save Searle's derivation by claiming that, in the same cases, *no promises* are made either—in other words, that if, for example, Tom says to me, "I hereby promise to find your lost wallet," and it turns out that I never lost, since I never had, a wallet, it not only follows that Tom has not thereby acquired any *obligation*, but also that he has not even made any *promise*. That this is false is evident not only from the already noted fact that, in the described circumstances, Tom's utterance could without contradiction be reported by a statement such as "Tom promised to find my lost wallet, but I never lost, since I never had, a wallet, so he must be confusing me with somebody else", but also from the fact that, in the same circumstances, Tom's utterance would be subject to a variety of forms of appraisal that would be unintelligible *except* on the assumption that it *was* the utterance of a promise. For example, knowing better than Tom whether or not I ever owned a wallet, I might respond to his utterance of "I hereby promise to find your lost wallet" by truly and intelligibly saying to him, "Well, Tom, the promise you just made cannot be fulfilled, since I never had, and so I never lost, a wallet." But that perfectly true and intelligible remark would have to be counted either as just unintelligible or as blatantly false if Tom had *not* made the promise to which I would be referring, and if, in particular, the proposition that I never had, and never lost, a wallet had the power to entail the proposition that Tom did not *promise* to find my lost wallet. What the proposition that I never had, and never lost, a wallet does entail is, precisely, that Tom cannot *fulfil* the

⁷ Searle's more general misgivings about the standard interpretation of the notion of "*prima facie* obligation" are presented in detail in Searle (1978).

promise he has made, not that he cannot *make* the promise he has made. But fulfilling a promise and making it are two quite different things, and the fact that a promise cannot be fulfilled does not imply that it is not a promise, any more than the fact that an assertion cannot be true implies that it is not an assertion, or the fact that a request cannot be complied with implies that it is not a request. Rather, a promise can be fulfilled *or* unfulfilled only if it is already in existence, just as an assertion can be true or false only if it is already in existence, and a request can be complied or not complied with only if it is already in existence. Indeed, Searle's familiar and fundamental distinction between the *success* of an illocutionary act and the *satisfaction* of its propositional content (see, e.g., Searle 1991) was designed precisely in order to acknowledge these elementary facts; but none of these elementary facts could be acknowledged if the desperate strategy now under consideration was to be systematically applied, since its application would have the effect of systematically misdescribing what are, in fact, conditions of propositional content satisfaction as if they were conditions of illocutionary act success. It would require claiming, for example, that the existence of a King of Germany is not only a condition on the *fulfilment* of the promise to arrest the King of Germany, but also a condition on the *existence* of the promise to arrest the King of Germany; that it is not only a condition on the *truth* of the assertion that the King of Germany has been arrested, but also a condition on the *existence* of the assertion that the King of Germany has been arrested; that it is not only a condition on *complying with the request* that the King of Germany be arrested, but also a condition on *making the request* that the King of Germany be arrested; and so on. Since no one (and least of all Searle) would presumably be prepared to accept any of these claims, it should readily be granted that the presence of vacuous names or vacuous descriptions in the content-clause of a statement ascribing a promise does *not* necessitate the falsity of that statement. And since, on the other hand, the presence of vacuous names or vacuous descriptions in the content-clause of a corresponding statement ascribing an obligation *does* necessitate the falsity of that statement, there is, as far as I can see, no escape from the conclusion that ascriptions of promises do not *entail* ascriptions of obligations.

4. Conclusion

I have first argued that, even if it were true, as Searle claims, that statements about promises entail statements about obligations, it would not follow, as he also claims, that non-evaluative statements entail evaluative ones (and that, therefore, the “naturalistic fallacy” is not a fallacy), since, as the instability of his own categorisations indicates, the conditions of application of the labels ‘evaluative’ and ‘non-evaluative’ to statements about promises or to statements about obligations are far more theory-sensitive than they would have to be if the latter claim was to follow from the former one. Concentrating, then, on the former claim alone, I have argued that it is not in fact the case that statements about promises entail statements about obligations, since statements about promises can be true whereas corresponding statements about obligations are false when vacuous names or vacuous descriptions occur in the clauses that specify the content of the promises or of the obligations that such statements ascribe—since, in

other words, statements about obligations are extensional in a way in which statements about promises are non-extensional with respect to vacuous names or vacuous descriptions appearing in their respective content-clauses.

As indicated in the introductory section, the claim that statements of the form “X promised to *p*” entail corresponding statements of the forms “X ought to *p*” and “X has the obligation to *p*”, apart from its early involvement in Searle’s discussion of the “naturalistic fallacy”, plays a key role in the accounts of sociality and rationality recently developed in Searle’s *The Construction of Social Reality and Rationality in Action*. In arguing that Searle’s claim cannot, at least as formulated, be sustained, I do not, of course, mean to imply that the various innovative theses put forward in these works should simply be discarded. On the contrary, I believe that they deserve sustained examination, and regard the present discussion as a preliminary step in seeking better ways of articulating their content and implications. If one abandons the dogmatic policy of accepting as worthy of one’s theoretical attention only those inferential relations that can be *deductively* validated, one will have no difficulty in acknowledging the obvious fact that people do have the tendency to *defeasibly* infer “X ought to *p*” or “X has the obligation to *p*” from “X promised to *p*”, in a way in which, for example, they do not have the tendency to infer, even defeasibly, “X ought to *p*” or “X has the obligation to *p*” from “X expressed the wish to *p*”. The problem, then, is to explain, in a naturalistic spirit, what it is that warrants inferences of the former sort without warranting inferences of the latter sort, *even though* neither the former nor the latter conform to deductive canons. And it may well be that, once this naturalistic explanation is at hand, it can be made to cohere with many of Searle’s ideas about sociality and rationality. For, as Searle would probably agree, sociality and rationality are not entirely, or even largely, deductive affairs.

References

- Searle, John R. 1964. How to derive ‘ought’ from ‘is’. *The Philosophical Review* 73. 43–58.
- Searle, John R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, John R. 1978. *Prima facie* obligations. In Joseph Raz (ed.), *Practical reasoning*, 81–90. Oxford: Oxford University Press.
- Searle, John R. 1979. *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.
- Searle, John R. 1983. *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Searle, John R. 1991. Response: meaning, intentionality, and speech acts. In Ernest Lepore & Robert Van Gulick (eds.), *John Searle and his critics*, 81–102. Oxford: Blackwell.
- Searle, John R. 1995. *The construction of social reality*. New York: Free Press.
- Searle, John R. 1997. Responses to critics of *The Construction of Social Reality*. *Philosophy and Phenomenological Research* 57. 449–458.
- Searle, John R. 1998. *Mind, language and society: Philosophy in the real world*. New York: Basic Books.

- Searle, John R. 1999. Fact and value, 'is' and 'ought', and reasons for action. In G. O. Mazur (ed.), *Twenty-five year commemoration to the life of Hans Kelsen (1898–1973)*, 20–34. New York: Semenenko Foundation.
- Searle, John R. 2001. *Rationality in action*. Cambridge, Mass.: The MIT Press.
- Searle, John R. 2004. *Mind: A brief introduction*. Oxford: Oxford University Press.