Teorema
Vol. XVIII/1, 1999, pp.73-83

A Critical Evaluation of Searle's Connection Principle

Robbert van Baaren

RESUMEN

Este trabajo evalúa críticamente el Principio de Conexión de Searle (PC). En primer lugar se presentará PC en una de sus versiones publicadas, y se comentarán brevemente los pasos que conducen al mismo. A continuación, se plantearán dos cuestiones. (1) La noción crucial de forma aspectual resultará ser ambigua: puede entenderse de un modo estricto o amplio. PC exige la concepción amplia, pero no hay necesidad alguna de adoptarla. (2) Se mostrará que existen candidatos plausibles a ser considerados como estados intencionales inconscientes, entidades cuya posible existencia pretende excluir PC.

ABSTRACT

This paper critically evaluates Searle's Connection Principle (CP). First, the CP will be presented in one of its published forms and its numbered steps will be briefly commented upon. Then two issues will be raised. (1) The crucial notion of aspectual shape will turn out to be ambiguous: it allows for a narrow conception and a broad conception. The CP requires the broad conception, but there is no necessity to adopt it. (2) It will be shown that there are plausible candidates of occurrent unconscious intentional states, entities the possible existence of which the CP aims to rule out.

I. INTRODUCTION

On several occasions Searle [1989, 1990, 1992] has presented a thesis he labels the Connection Principle, henceforth called CP. The CP asserts that unconscious mental states are dispositions to cause conscious mental states. This paper aims to evaluate that claim and the argument Searle brings forward to support it. In particular, two points will be raised. First, could the notion of aspectual shape, which plays a crucial role in the argument, as will presently be shown, not be a hybrid notion, i.e., could it not be the case that aspectual shape is in fact a complex property of intentional states consisting of a strict aspectual and a phenomenal constituent? Second, Searle invites critics to refute the CP by coming up with counterexamples to it, viz., unconscious mental states that are not dispositions to cause conscious states. A class of possible counterexamples will be presented, after which one example, about which it will be argued that it poses a problem if not a counterex-

ample to the CP, will be highlighted. Finally, it will be shown that the two points raised hang together in that one plausible explanation for the possible counterexample assumes the disconnection of the phenomenal and aspectual properties of intentional states. And if phenomenal and aspectual properties of intentional states are in fact disconnected, the principled connection between intentionality and consciousness, which the CP aims to establish, cannot be maintained either.

II. PRESENTATION OF THE CP

Of the three occurrences of the CP in Searle's writings the most recent version, viz., that of *The Rediscovery of the Mind*, Chapter 7, "The Unconscious and Its Relation to Consciousness", will be presented here. All versions of the argument for the CP consist of numbered steps, which gives the impression that we are dealing with a deductive argument. Unfortunately, this is not the case, as was pointed out by Rey [Searle (1990), p. 620] and acknowledged by Searle [Searle (1992), p. 156]. The version under consideration here has one extra step (#2) — which points out the relevance of the first premise for the rest of the argument — as compared with the previous presentation, viz., that in *Behavioral and Brain Sciences* (BBS), entitled "Consciousness, Explanatory Inversion and Cognitive Science", but is otherwise identical. In accordance with the format of BBS, the presentation of the CP in a target article evoked numerous commentaries, to which the author responded. Some of the commentaries will be mentioned briefly in the next section. It is, however, not my aim to go through them extensively, but they are included to give an idea of the directions in which the criticism of the CP go. The first presentation of the CP, in an article entitled "Consciousness, Unconsciousness and Intentionality", is the most extensive one. It consists of 10 steps, but as Searle remarks in a footnote in the BBS version: "The argument here is a condensed version of a much longer development in Searle (1989). I tried to keep its basic structure intact, and I apologize for a certain amount of repetition". [Searle (1990), p. 596, n 2]. Only in the context of the discussion of the notion of aspectual shape shall I fall back on the longer version, for clarification.

III. THE ARGUMENT FOR THE CP IN SEARLE (1992), PP. 156-60

1. There is a distinction between intrinsic and as-if intentionality; only intrinsic intentionality is genuinely mental.
2. Unconscious intentional states are intrinsic.

3. Intrinsic intentional states, whether conscious or unconscious, always have aspectual shapes.

4. The aspectual feature cannot be exhaustively or completely characterized solely in terms of third-person, behavioral or even neurophysiological predicates. None of these is sufficient to give an exhaustive account of aspectual shape.

5. But the ontology of unconscious mental states, at the time they are unconscious, consists entirely in the existence of purely neurophysiological phenomena.

6. The notion of an unconscious intentional state is the notion of a state that is a possible conscious thought or experience.

7. The ontology of the unconscious consists in objective features of the brain capable of causing subjective conscious thoughts.

IV. Preliminary Comments on the CP

re 1. For Searle, the distinction between intrinsic and as-if intentionality is crucial and basic. Intrinsic intentionality is real and original, as-if intentionality is no intentionality at all. There is another form of intentionality, viz., derived intentionality as of, for example, speech acts, which is real but not original. The distinction between intrinsic and derived intentionality is a special case of observer-independent and observer-relative properties. For example, the mass of physical objects and their chemical composition are observer-independent properties, whereas their functional identities, their being a knife or being a piece of money, are observer-dependent. There are thus two pairs of predicates of intentionality: real vs. as-if and original vs. derived. Also, there are two kinds of real intentionality, original (or intrinsic) and derived [Searle (1997), pp. 9-12].

The distinction between intrinsic and as-if intentionality has been disputed, in particular by Dennett. To borrow Searle's terminology, Dennett's position could be characterized by saying that he takes all intentionality to be observer-relative. Dennett takes the idea of original intentionality to be a myth and suggests that our intentionality is derived from the intentionality of our "selfish" genes [Dennett (1990), p. 59]. Two very strong intuitions seem to be clashing here. One is that our intentionality is real and original, that our intentional states are caused by our brains and that therefore intentionality must be an emergent property of brains. The other is that intentionality is a higher-order property of brains, that intentionality must be reducible to lower-order properties of brains and that therefore intentionality is derived from these lower-order properties. The premise implicit in the second strong intuition is that causal reduction implies ontological reduction. This, Searle

admits, is generally a valid inference, but when we come to consciousness, it is invalid. The reason it is valid in most cases, as in the reduction of water to H_2O, is that there is an objective physical reality on the one hand and a subjective appearance on the other. In the case of consciousness and intentionality, the appearance is the reality. The latter claim would be denied by Dennett. The gap between them is deep and I will not either try to bridge it or take sides here. For the sake of the argument, I will grant Searle this premise.

re 2. Having granted the distinction between intrinsic and as-if intentionality, I will grant this premise as well. Searle presumably would not object to the claim that unconscious intentional states are as-if intentional, since in his view anything can be as-if intentional, even "thirsty" lawns. If, however, unconscious intentional states are to have any explanatory power — and this is the very ground for assuming their existence —, they need to be intrinsically intentional, or so Searle argues. He does not here consider the possibility that unconscious intentional states have real, but derived intentionality. If one believes, as Dennett seems to do, that intentionality is derived from the genes, then one should face up to this possibility. Searle, however, thinks of derived intentionality as being derived from intrinsic or original intentionality and original intentionality is tied to consciousness, as he tries to establish. There is a threat of circularity here. Without the CP I do not see why unconscious intentional states could not be derived.

re 3. Assuming that besides conscious and unconscious intentional states, there are no other (intrinsic) intentional states, (3) becomes: All intentional states have aspectual shape. In order to evaluate this claim, one needs to know what Searle means by "aspectual shape". My inclination is to think that aspectual shape relates to conditions of satisfaction in a way that is similar to the relation between sense and reference. The (Fregean) examples Searle offers to illustrate aspectual differences — Morning Star/Evening Star; water/H_2O — give rise to this supposition. In one commentary, Searle offers "mode of presentation" as an alternative term for aspectual shape, which further supports the supposition. For the sake of the argument I will grant this premise.

re 4. In its formulation this premise is redundant. Since behavioral and neurophysiological predicates are expressed in third-person terms, (4) in fact says: Aspectual shape cannot be fully characterized in third-person terms. Rey [Searle (1990), p. 620] called this claim a negative conceivable, and negative conceivables are notoriously hard to establish. I suspect this to be the most controversial claim. Searle argues that aspectual shape cannot be fully characterized in behavioral terms, because identical behavior can be motivated by aspectually different intentional states. For example, water-seeking behavior could be motivated by a want for water or for a want for H_2O. In addition, aspectual shape cannot be fully characterized in neuro-

physiological terms, because even if we had complete neurophysiological knowledge, it would still take an inference from the specification of aspectual shape in neurophysiological terms to the specification of aspectual shape in intentional terms. Notice that Searle does not evaluate other candidate third-person characterizations of aspectual shape, such as functional or representational predicates. Although all representational and functional proposals that have been suggested so far suffer from some problem or other, the issue of the possibility of the characterization of aspectual shape has not been settled yet. I conclude, therefore, that (4) is not beyond dispute.

re 5. Since intentional states are mental states, (5) implies: unconscious intentional states are neurophysiological. This claim is not controversial to the naturalist. Those who deny that intentional states are brain states, such as the practical realist Lynne Rudder Baker, would have to deny this premise.

re 6. For "thought or experience" I will substitute "intentional state", hence: The notion of an unconscious intentional state is the notion of a possible conscious intentional state. Many have criticized the notion of "possible", but I will not go into that issue.

re 7. This is not about states, but about the unconscious as a whole. My argument will be directed at the argument up to "the first main conclusion" (6), so for my purposes (7) can be left out of consideration.

V. The Crucial Notion of Aspectual Shape

Searle posits that "This aspectual feature must matter to the agent. It must exist from his/her point of view" [Searle (1989), p. 199]. To cite one of his examples, one may want water or one may want H_2O; these are different intentional states that may nevertheless result in the same behavior. The idea is that there are differences in aspectual shape that are concealed from a third person point of view, but not from the first person point of view. And behavioral predicates seem to be too coarse-grained to catch this difference.

Suppose, then, that H_2O and water are coreferential from the point of view of some subject. And suppose no behavior of his would ever reveal the difference between wanting H_2O and wanting water. Then the aspect under which the water/H_2O is represented does not make a difference to the behavior of the subject. It is not the case that he wants water iff he wants to drink water and he wants H_2O iff he wants to do chemical experiments. If he wants water/H_2O in order to drink it, he wants either water or H_2O and only he knows. That is the situation under consideration.

Searle remarks that it does not help to ask the subject whether he wants water or H_2O, because his answers would not fix the aspectual shapes.

> There is no way just from the behavior to determine whether the person means "H_2O" what I mean by "H_2O" and whether the person means "water" what I mean by "water" [Searle (1990), p. 587].

(The argument Searle uses here seems to be a variant of the inverted-spectrum argument. This already gives a clue about the direction in which the argument is going: intentional states have qualitative or phenomenal properties that cannot be fully expressed.)

This does not convince me yet. If I were to ask "Do you want it under the aspect of being a chemical or under the common sense aspect?", I can find out which of the two notions of water the subject has in mind. So, the distinction Searle has in mind is more fine-grained than that. Let me push a little further. If, on a second occasion, the subject shows water seeking behavior, we can ask him whether the same intentional state was involved as the first time. If he (sincerely) affirms this, then there is no distinction between the two states from his point of view. If he denies it, then we know that there is a distinction between the two states from his point of view, but also from ours, because of the content of his verbal response. In other words, even the most fine-grained differences between the aspectual shapes of intentional states can still be communicated at least in principle.

But Searle is moving in another direction. It is not the difference, but the properties making the difference that cannot be communicated. It is not *that* there are differences, but *what* these differences are that cannot be communicated. The 1989 presentation of the CP reveals what Searle has in mind:

> None of these [third-person accounts] is sufficient to give an exhaustive account of *the way it seems to the agent* [Searle (1989), p. 199; emphasis in the original; "it" refers to "the aspectual feature"].

Searle seems to claim that in conscious intentional states the aspectual feature is connected with a phenomenal feature. There are at least two models of how conscious states should be understood. One assumes that a conscious state is a state one is conscious of. This implies that there is a so-called (by Rosenthal for one) higher-order thought the object of which is the state one is conscious of. One consequence of this view is that there must be unconscious thoughts, because the higher-order thought need not itself be conscious. Either it is or it is not conscious. If it is not conscious, the conclusion that there must be unconscious thoughts is already reached. If the higher-order thought is conscious, there is a thought of an even higher order, which need not be conscious, etc. At some level, this recursion has to end as otherwise it leads to an infinite regress. In other words, there are, in principle, unconscious thoughts. Since thoughts are intentional states, the higher-order

thought conception of consciousness leads to a denial of the CP, according to which there are no in-principle unconscious intentional states. Thus, this cannot be the way consciousness is conceived of by Searle, and it isn't.

According to Searle the "of" of consciousness is not the "of" of intentionality. "[W]hen I have a conscious experience of anxiety [...], the experience of anxiety and the anxiety are identical [...]" [Searle (1983), p. 2]. If the experience of anxiety is identical with the anxiety itself, then one cannot be anxious without being aware of it. In other words, one cannot have unconscious phenomenal states. But can one have unconscious *intentional* states? Here a distinction should be made between dispositional and occurrent states. Dispositional unconscious intentional states are not problematic: Searle acknowledges their existence. Indeed, they are the paradigm cases for which the CP works very well. Dispositional unconscious intentional states are possible conscious intentional states or dispositions to cause conscious intentional states. The problem is occurrent unconscious intentional states, since they are ruled out by the CP.

Why are they ruled out? Because being intentional states, they have aspectual shape, and having aspectual shape, they should exist from the point of view of the subject. But being unconscious, they do not exist from the point of view of the subject, and being occurrent rather than dispositional, they cannot ever become conscious. The notion of an occurrent unconscious intentional state is unintelligible, because it is a contradiction in terms. One should object, I think, against the claim that aspectual shape requires existence or even possible existence from the point of view of the subject, so-called subjective ontology. So I am not denying that no exhaustive third-person account can be given of the way the aspectual feature appears to the agent; I am agnostic on this account. What I *am* denying is that the way the aspectual feature appears to the agent is an integral part of the intentional state.

By way of comparison, consider the following example: Is the way Clinton appears to me a property of Clinton? Or rather of me? I think we should say that the way Clinton appears to me, and the way the aspectual feature of some intentional state appears to me, are relational properties. Relational properties are properties of a relation, not of either or both of the *relata*. Thus if the aspectual feature of an intentional state includes the way it appears to me, then the aspectual feature is not an intrinsic property of that intentional state. And if it is not an intrinsic property, we should not worry that it "[...] cannot be exhaustively or completely characterized solely in terms of third-person, behavioral, or even neurophysiological predicates" [Searle (1989), p. 199]. If, on the other hand, the aspectual feature does not include the way it appears to me, the worry that no exhaustive third-person account can be given of the aspectual feature cannot be justified on the same grounds.

Our first conclusion reads as follows. Aspectual shape seems to be a hybrid notion consisting of two components: a strictly aspectual one and a phenomenal one. Searle offers no reason why the purely aspectual component could not be characterized in third-person terms, but I am willing to grant that the phenomenal component cannot be thus characterized. And it *could* be that intentional states have their phenomenality in virtue of being conscious, not in virtue of being intentional. There are two indications supporting the idea that phenomenality and aspectual shape are dissociable. (1) There are conscious states that are phenomenal, but not intentional. They are phenomenal not in virtue of having aspectual shape, because they don't have aspectual shape, so that they must be phenomenal in virtue of something else. (2) On the other hand, there seem to be intentional states that are not conscious and thus not phenomenal, so even though they have aspectual shape they lack phenomenality. The point is that aspectual shape seems neither necessary nor sufficient for phenomenality. We now turn to examples of this second category.

VI. Possible Counterexamples to the CP

Searle invites people to refute the CP by coming up with counterexamples to it. A counterexample would be an unconscious intentional state that is not accessible to consciousness or is not a disposition to cause a conscious intentional state.

> I would be convinced that I was wrong if someone could give me both a clear sense and clear evidence for the claim that there are inner qualitative subjective mental states going on in me which are totally unconscious, that consciousness could be pealed off from mental states, *leaving everything else intact* [Searle (1995), p. 231].

This is the challenge I will face up to, albeit that I will try to show that there are such states going on in certain people, rather than in Searle, hoping that that will be sufficient to convince him. One problem is that it is by definition impossible to give direct evidence of such states. Being mental states, they are inaccessible to third persons, and being unconscious states, they are inaccessible to the first person. Therefore, the best I or anyone can do is to provide a clear sense of and *indirect* evidence for the claim that there are such states.

The type of cases one is bound to come up with are cases of occurrent unconscious intentional states, rather than dispositional unconscious intentional states. The controversy, as I see it, is not about the latter type of cases. They

seem to be paradigm cases for which the CP works very well. My knowledge that Clinton is the president of the USA is somehow and somewhere stored in my brain and this brain state can very well be characterized as the capability of causing the conscious thought that Clinton is the president of the USA The problematic cases are those where there seems to be something going on in the brain, without there being consciousness of it, and where this something going on in the brain is most properly described as a mental state. Let us call these cases mental events as opposed to the former type of cases, which are most properly called states.

The challenge is thus to come up with cases of unconscious mental events that make clear sense and for which there is clear evidence, albeit indirect evidence. Although proponents of unconscious mental events will say that they are ubiquitous, in order to convince Searle of their existence we will have to concentrate on very specific cases. The type of cases I am thinking of have been brought into the discussion before and have not convinced Searle, but I think that if anything could convince him, it is these types of cases. So here they are: blindsight, priming, subliminal perception, Freudian unconsciousness and the like. They have in common that they occur in situations in which the subject acts without being conscious of the reason or cause of his actions. In some cases this is due to a lesion, in other cases this has to do with normal functioning.

Let us concentrate on the case of blindsight. The subject, in this case a patient, is able to act on visual information, although he has to be externally motivated to do so. If we accept the very general scheme of a belief and a desire causing an action, as in the case of the belief that drinking water will quench my thirst plus the desire to quench my thirst (plus my being thirsty, plus my awareness of being thirsty) will normally make me form an intention to drink water, which, in turn, will normally make me drink water. Suppose a blindsight patient will be presented with a cup of water in his blind field. He will not act, because he has no awareness of the cup of water. But when motivated to grab the object in his blind field, he will make a movement with his arm and hand appropriate to grab the cup, even though he claims to have no knowledge or awareness of what is in front of him.

One could describe to him the belief that there is a cup-like object in front of him, because he acts accordingly. But does he really have that belief? At least not consciously, or so he claims. One could suggest that he has the unconscious belief that there is a cup in front of him, but that this unconscious belief does not cause a conscious belief. What is missing besides the consciousness of the belief is the motivation to act. If we accept the Dretskean distinction between structuring cause and triggering cause, it seems that in the blindsight case the structuring cause for the action is there, but not the triggering cause. That has to come from outside. The triggering cause

supplies the motivation to act. It motivates the subject to grab the object in front of him. But it does not tell him how to act. Yet his action is appropriate, so the triggering cause, i.e., the instruction to act, is not a sufficient explanation of the successful action.

The subject is grabbing the object in front of him in the way in which he would grab the cup were it in his visual field and within reach. There is therefore reason to suppose that the object in front of him is represented in a way similar to the way the cup would be represented if he did see it. And a representation represents its object under certain aspects. Thus here we seem to have a case of unconscious aspectual shape.

VII. Conclusion

My conclusion is that the CP has not been established. The crucial notion of aspectual shape turned out to be hybrid consisting of a phenomenal and a purely aspectual component. Phenomenality is not sufficient for aspectuality, because — as Searle acknowledges — there are conscious states that are not intentional and thus not aspectual. On the other hand, phenomenality is not necessary for aspectuality, because there are strong indications that there must be intentional states that are not conscious and thus not phenomenal, as in the blindsight cases. If phenomenality is neither sufficient nor necessary for aspectuality, then the two are dissociable and the plausibility of the CP is disputed.

Katholieke Universiteit Brabant
Department of Philosophy
Warandelaan 2, 5000 Tilburg (The Netherlands)
E-mail: R.vanBaaren@kub.ul

References

Dennett, D. (1990), "The Myth of Original Intentionality", in Mohyeldin Said, K. A., Newton-Smith, W. H., Viale, R. and Wilkes, K. V. (eds.) (1990), *Modelling the Mind*, Oxford, Clarendon Press, pp. 43-62.

Nelkin, N. (1993), "The Connection between Intentionality and Consciousness", in Davies, M. and Humphreys, G. W., *Consciousness. Psychological and Philosophical Essays*, Oxford, Blackwell, pp. 224-39.

Searle, J. (1969), *Speech Acts. An Essay in the Philosophy of Language*, Cambridge, Cambridge University Press.

—— (1983), *Intentionality. An Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press.

—— (1989), "Consciousness, Unconsciousness and Intentionality", *Philosophical Topics*, vol. XVII, no. 1, pp. 193-209.
—— (1990), "Consciousness, Explanatory Inversion and Cognitive Science", *Behavioral and Brain Sciences*, vol. 13, pp. 585-642.
—— (1992), *The Rediscovery of the Mind*, Cambridge Mass., The MIT Press.
—— (1995), "Consciousness, the Brain and the Connection Principle: A Reply", *Philosophy and Phenomenological Research*, vol. LV, nr. 1, pp. 217-32.
VAN GULICK, R. (1995), "Why the Connection Argument Doesn't Work", *Philosophy and Phenomenological Research*, vol. LV, nr. 1, pp. 201-7.