Jesús Navarro-Reyes[*]
University of Seville

SPEECH ACTS, CRITERIA AND INTENTIONS

Abstract

What makes a speech act a speech act? Which are its necessary and sufficient conditions? I claim in this paper that we cannot find an answer to those questions in Austin's doctrine of the infelicities, since some infelicities take place in fully committing speech acts, whereas others prevent the utterance from being considered as a speech act at all. With this qualification in mind, I argue against the idea that intentions—considered as mental states accomplishing a causal role in the performance of the act— should be considered among the necessary conditions of speech acts. I would thus like to deny a merely 'symptomatic' account of intentions, according to which we could never make anything but fallible hypotheses about the effective occurrence of any speech act. I propose an alternative 'criterial' account of the role of intentions in speech acts theory, and analyse Austin's and Searle's approaches in the light of this Wittgensteinian concept. Whether we consider, with Austin, that speech acts 'imply' mental states or, with Searle, that they 'express' them, we could only make sense of this idea if we considered utterances as *criteria* for intentions, and not as alleged behavioural effects of hidden mental causes.

[*] Departamento de Metafísica y Corrientes Actuales de la Filosofía
Universidad de Sevilla
C/ Camilo José Cela, s/n.
41018, Sevilla, Spain
E-mail: jnr@us.es

1. Internalism and externalism in speech acts theory: what makes a speech act a speech act

Robert M. Harnish (2009) has recently tried to classify current speech acts theories between internalist and externalist accounts, considering the relative importance given by each author to the mental states of speakers and hearers in the first case, or to the conditions of the world surrounding them in the second[1]. Despite Harnish's proficient analysis, the result is disconcerting, perhaps because he introduces between the allegedly internalist and externalist positions an intermediate 'mixed' category where, in the end, everybody seems to fit more comfortably. It is intuitively true that, as he points out, approaches inspired by Austin (1962)—in that they try to explain the workings of speech acts by making conventions, rules, and norms explicit—tend to be more akin to postulate external conditions[2]; whereas those who are inspired by Grice (1957)—assuming an approach based on the inferential attribution of intentions to the speaker—tend to make a more constant use of belief-desire psychology, and are thus more prone to internalism. However, as Harnish himself notices, an author like Alston (2000), who is a clear descendent of Austin, seems to be quite internalist; and we find an important inclination to externalism even among new Gricean theoreticians, when they make use of a "Derivative Externalism" in their account of truth (Harnish 2009: 19-20).

The problem is that the relative weight of internal and external conditions within a theory is not something simple to evaluate, and there does not seem to be good reasons to catalogue some theories as 'mixed' and others as 'internalists' if the latter are also 'mixed', since they do not deny the existence of external requirements. On the other side, once the alleged externalism of Sbisà (2002) is discarded by Harnish as a sort of impure position (in a reading that does not seem to be as sympathetic as one could have expected), the externalist extreme of the range seems to be occupied by Gazdar (1981) alone—with the hesitant suggestion that Millikan (2005) might accompany him there. However, *pace* Harnish, Gazdar does not seem to be particularly externalist when he claims that "A speech act is a function from contexts into contexts" (1981: 68). His analysis of the logical structure of speech acts theory could be assumed by an internalist, as far as mental

---

[1] A different way of expressing a similar distinction can be found in Sbisà (2002), where she distinguishes between authors who tend to understand context in terms of "objective conditions" and those who are more predisposed to introduce "cognitive" ones.

[2] Sbisà (2002), (2009); or Moati (2009) are some recent examples.

states have not been banned from the notion of context—and I have not found in Gazdar (1981) any attempt to do this.

In this way, nobody seems to fit properly either in the internalist or in the externalist side of Harnish's arrangement, and the 'mixed' category gets overcrowded. For this reason, I would say that the sharp demarcation line ought to be drawn between those who *do* include mental states as strictly necessary conditions for the performance of speech acts—no matter how tiny their role is in the theory—and those who do not, and explicitly exclude them from the set of necessary and sufficient conditions that *any* utterance has to accomplish in order to be considered as a speech act[3]. Such a demarcation would leave no room for mixed positions, since it is not a matter of relative weight, but of presence or absence of such kind of internal or cognitive requirements.

My intention here is to propose the following question as the crucial one: are unintentional speech acts *a kind* of speech act or, in so far as they have not been intended actions, they are *no* speech acts at all? The answer to our problem will depend on what we mean by "unintentional". And this will in turn depend on what we mean by "intentional" or "intended" and "intention". Most speech act theorists understand these terms in the following way: an act is "intended" if it is caused by an "intention", and an "intention" is a mental state causally efficacious in behaviour. An "unintentional" act would be a piece of behaviour that would not have been *directly* caused by any "intention". An effect that was not expected by the agent—and was only indirectly caused by a different intention of her—would not be an intended action, even if the act itself can be described as intended from a different perspective.[4] Although I do not think that this is the best way to understand intentions, in what follows I will assume this conception, since it seems to be the dominant one among speech acts theoreticians. Thus, if we assume intentions as mental causally efficacious states, the question would thus be: are those states among the necessary and sufficient conditions of a speech act? Or, in

---

[3] Most authors in the tradition of Grice and Strawson talk about "utterances" in general as modifications of the physical environment designed by a communicator to be perceived by an audience and used as evidence of the communicator's intentions. Other authors have preferred the term "stimulus", which does not imply such a bias to linguistic cases of intended communication (Sperber and Wilson 1986:29). I will preserve the original term, but I do not mean by using it that the act must be intentionally performed. In my view, utterances can be done with no intention, and thus with no *communicative* intention, whether linguistic or not.

[4] This would assume two of the main theses of Davidson (1982) on actions: first, that we can quantify over events; and second, that actions can only be intentional under particular descriptions.

other terms: by asserting that a speech act took place, are we *eo ipso* asserting that it is the case that speaker was in quite specific mental states, and that those states have been the cause of her actions?

2. A possible confusion arising from Austin's doctrine of the in-felicities

Austin's (1962) attempt to give an account of performatives was undeniably inspiring: a great example of a philosopher struggling to find his own way in an exciting new field of conceptual analysis. However, his initial effort, in the terms he originally sketched it, could lead to confusion. The reason is that he fails to make clear enough the really crucial distinction: the one between what *is* a performative and what is *not*. This line might be blurred and vague itself and, if this were the case, we should not demand from the analysis a level of exactitude that would be alien to its object. However, whether vague or not, that line ought not to be identified with the famous one initially traced by Austin between felicitous and infelicitous performatives. The reason is very simple: some infelicitous performatives *are* performatives, whereas some others are not. Austin himself was perfectly aware of this, but he induced to mistake when he described unhappy performatives in general as different *kinds of* performatives, when in fact most of them are not performatives *at all*. Austin's approach was tentative, and, at least at that initial point, his aim was just to describe the phenomenon in question, rather than to give an explanation of it, not to say a theory. But the way he approached his subject might have led some of his readers to this kind of conceptual mess.

Let me make this clear with two examples. First: after wandering thirsty in the dessert I think that I see an oasis, just to realise later that it was not an oasis at all, but just a mirage. Although it would make sense to say that 'this one' was a *false* oasis, a false oasis is not *a kind* of oasis: it is not an oasis at all. "False" is not working here as a specification of the *kind* of oasis that I saw. It would make no sense trying to make a list with all the oases that there are in the world, whether they are real or false: such a category would not be correctly defined since, in a way, there are no false oases, but only situations under which somebody deceitfully takes something as an oasis. In contrast to this, it would be perfectly sensible to make a list of all the *real* oasis there are in the word, whether they are "weird" or not: we would just make our requirements less strict. In contrast to "weird", "false" is not an adjective for *that* oasis—since there is no such oasis—but only the effect of a process of categorization: it shows that the process initially failed, or could have failed at a particular moment.

A similar argument would invalidate any pretention to build, for instance, a list with all the false medical doctors that are there in the world. We would have to include in it not only anybody who ever consciously pretended to be a medical doctor but also those cases in which somebody could ever be deceived into taking someone else as such: from fictional characters to quite common misunderstandings, like the pharmacist wearing a white coat that is incorrectly taken for a physician, or the anthropologist that, after being introduced as a "Doctor", is mistakenly considered by somebody as a *medical* doctor.

Those examples show that we cannot build a category of all *P*s, disregarding if they are real or false ones. And it would be equally meaningless to build a category with all the *false P*s. Both categories would be wrongly defined, and assuming them would commit us to look for so confusing things as *really* false oases, or *really* false medical doctors.

Although he did not make this mistake himself, Austin's original intention to approach performatives by describing the conditions that could be wrong in their performance induces the reader to it. An inattentive reading of Austin's work could construct him as trying to guess which are the necessary and sufficient conditions for something to be a performative *tout court*, whether a felicitous or an infelicitous one. But, among the infelicities that he describes, those he called "misfires" (caused by bad invocations or executions of the convention) are not performatives *at all*. They are, in Austin's own words, "purported or empty", and thus "void or without effect" (1962: 16). If we talk about those cases as "performatives", it must be in the same sense we call a false oasis an "oasis": they are not tokens that fall under the type of performatives; it is just a way of speaking about those cases that might induce to failures in the process of categorization. In contrast to this, what Austin called "abuses" (insincerities or inconsistencies with further actions), do give rise to a legitimate kind of performative acts: those whose "professed or hollow" character does not prevent them from being fully committing speech acts, although they might not be "implemented" or "consummated".[5] Considering all of them as different *kinds* of performatives is somehow like trying to make a category of all oases or all medical doctors, disregarding if they are real or false ones.

The situation becomes even more confusing when Austin makes reference to a bunch of other situations that affect the act, but he does not make clear if they could invalidate it (1962: 21-5). Not trying to be exhaustive or systematic, he talks

---

[5] In Bach and Harnish's (1979, 55-7) jargon, "felicity" conditions are not "success" conditions: the act is defective in cases where felicity conditions are not accomplished, but it might be successfully performed nevertheless.

about "extenuating circumstances" (acts done under duress, by accident, owing to mistake or unintentionally), "other kinds of ill which infect *all* utterances" (the possibility of being said on stage, in a poem, or in soliloquy), and "a sort of 'infelicity' […] arising out of 'misunderstanding'" (when the utterance was not properly heard or grasped). It is not clear which of those additional "ills" would turn the utterance into a 'weird' or impure performative, and which of them would force us to exclude the case from the category at all. But one point seems to be plain: taking them all—misinvocations, insincerities, accidents, soliloquies, misunderstandings, and so on—as different species *inside* the category of "performative" would be like saying that the impostor, the fictional character, the pharmacist and the anthropology professor, together with the real medical doctor, are all tokens of the same type of "medical doctors", although some of them are *real*, and some of them are not. And this would be a conceptual mess.

Austin abandoned his initial study of performatives for the broader notion of *speech acts*, which is the one I am here interested in. But the kind of mistake induced by his initial attempt could be easily extended to this second approach, since most of the infelicities that affect performative utterances can also afect speech acts in general.[6]

Returning to the debate on internalism and externalism in speech acts theory, and to the question I proposed in the preceding section, the problem now is which of those ills affect the necessary and sufficient conditions of speech acts up to the point that the speech act would not even take place. I would like to focus on one particular group of infelicities: some of those that Austin classified as "extenuating circumstances"—acts done by accident, owing to mistake or unintentionally—and I will refer to them in general as "unintentional speech acts". Returning thus to the crucial question, I would like to ask if unintended speech acts are a *kind* of speech act, or if they are not speech acts at all. In other words: is "unintentional" here working as an adjective that indicates *the kind* of speech act that took place, or does it mean that, as far as it was not intended, the speech act did not take place at

---

[6] See Sbisà (2009:46). The famous exchange between Searle (1989) and Bach and Harnish (1992) is burdened with the performative/assertive distinction. Despite their important differences, they all seem to be wondering how is it possible to guarantee the performative force of utterances, beyond their merely assertive aspect. In so doing, they preserve assertion as pragmatically unproblematic, and they only find a challenge in the performative side of the utterance. Following Austin in this, I am inclined to consider assertions as just one more kind of speech act, and not as something besides or beyond the pragmatic character of speech acts. It is not clear for me why Searle, who had followed Austin in overcoming his original distinction, did assume Bach and Harnish's terms—probably by influence of Strawson (1964), Urmson (1977) or even Benveniste (1974:271 f).

all? Is "unintentional" like "weird" or like "false"? The position that I will defend is the former one: unintentional speech acts are, in general, *a kind* of speech act, no matter how weird or marginal they are; for this reason, the adjective "unintentional", in most cases, works as a specification of the *type* of act that took place, and not as an invalidation of the act itself.

3. Unintentional speech acts

The aforementioned idea might seem to run counter to most of the alternatives in speech acts theory, which, as Harnish shows, usually claim that speech acts are *constitutively* the result of conscious intentional states. But this general assumption is in clear contradiction with the use of language in ordinary life, where one can easily find the following cases:

(1)   I apologise: I didn't intend to insult you.
(2)   My goodness, I wasn't aware that we were voting *his* proposal!
(3)   I didn't realise then that, by saying so, I was offering them my house.
(4)   But that is what you have just said! —Did I? Oh, I am sorry: I wasn't paying attention.

Those excuses qualify the preceding speech acts as defective in different respects, but neither of them *denies* that the act took place, which would be a different point. In each one of these cases, the change in the normative context is effective: the speakers *did* insult, vote, invite or assert by means of their previous utterances, which belonged to different *kinds* of unintentional speech acts. If that were not the case, excuses would have missed the point, since the preceding utterances would not require 'excuses', but 'explanations'. For instance, along the following lines:

(5)   I'm afraid that you are wrong: I did not insult you.
(6)   It was certainly not his proposal that I voted.
(7)   They are acting as if I had offered them my house, which I never did.
(8)   That is not what I said: I was meaning something else.

The difference between an intended and an unintended speech act is not that the former took place, whereas the latter did not. The latter *did* take place: unintentional speech acts are a kind of speech acts. They might be weird or marginal, but they are not *false* speech acts. Being the causal effect of specific mental states cannot thus be a necessary condition for an utterance to be a speech act.

It is also important to note that the effective performance of the act does not merely rely on the uptake by the audience. Not just *any* uptake is valid or justified:

the hearer might have misunderstood the utterance, or she might have ignored a significant aspect of the context. The question is not that the uptake itself makes the utterance a speech act, but that the uptake might be justified *beyond the speaker's original intention*: in the end, it might be the case that the hearer is correct about the kind of speech act that the speaker performed, despite the speaker's own opinion. In fact, it is an essential aspect of speech acts that their effective performance does not only rely on the alleged intentions of the speaker.

I did not step less strongly on somebody else's foot just because I did it unintentionally, nor did I cause her less pain. By the same token, whether my vote, my acceptance of a proposal, or my invitation were committing speech acts or not cannot strictly depend on their hidden origin in my psychological states. Of course, I could give excuses to try to make this commitment less strong, but those cannot simply rely on the causal role performed by my intention in the coming about of the utterance. Excuses will have to point out to something the hearer could take as a good reason to believe that my action was unintended. If I cannot offer those good reasons, and my act leaves no room for ambiguities, it is impossible for me to excuse it by simply pointing to an alleged lack of intention[7].

I will say something about the nature of those "good reasons" in section 5. But before that, I will try to show that the acceptance of unintentional speech acts as fully committing speech acts is consistent with an explanation of ordinary speech acts as the causal effect of internal mental states. What makes a particular speech act a speech act is one question; what is its causal history is a different one.

4. Usual causes do not have to be necessary conditions

A denial of the role of mental states in the performance of speech acts would commit me to some kind of behaviourism, even to eliminativism, or at least to the abandonment of folk psychological concepts in the explanation of language use. But I am not defending this. It is not my intention to deny the role of internal states in the effective performance of utterances, or that utterances are usually caused by the internal states of the speakers. My point is that such causal connections are not necessary conditions for the utterance to be a fully committing speech act. One

---

[7] There is of course an important difference between excusing an act by saying that it was fully unintended (done with no intention at all), and excusing it by saying that it was the unintended consequence of an intended action. But both cases seem to be equally excluded by those who consider that speech acts must constitutively be the effect of intentions, since in their view the act must be *directly* caused by the specified intention.

thing is to understand how and why are speech acts usually performed (what are the ordinary causes of the utterances that we consider speech acts) and a very different one to explain what makes an utterance a speech act in general.

Let me make this point clear with another example: punts are flat-bottomed shallow boats propelled by means of a long pole thrust. They are usually made of wood, but being made of wood is not a necessary condition for being a punt. A plastic punt would not be *false* punt but a *kind* of punt, although quite a rare, uncommon—and ugly—one. If we wished to propose a 'punt theory', and make explicit the necessary and sufficient conditions for something to be a punt, it would not be advisable to introduce as a requirement that they must be made out of wood. However, by excluding this as a condition we would not be denying (9) or (10):

(9)   Punts are usually made of wood.
(10)   Punts float *because* they are made of wood.

In a similar way, speech acts are usually caused by intentional conscious mental states but, just as any other kind of action, they can also be performed unintentionally, absentmindedly or by mistake. These "extenuating circumstances" make different kinds of excuses pertinent, and those might in the end reduce the normative implications of the speech act, or even make them disappear. But these excuses do not change what happened, nor can they prevent the act from having taken place, just like my unintended stamp does not suddenly disappear from the past just because I beg my pardon to the person who's foot I stepped on. I *did* step on her foot, although I did it unintentionally; I might be excused for that reason, but my excuses do not make the act vanish. Nor do they turn it into a mere event, like the falling of a stone, in which case I would have nothing to apologise for. Both the unintended stamp and the unintended speech act are something I did, although I was having a different intention, or no intention at all. For this reason, unintentional speech acts are a kind of speech acts; they are not something that only appeared to be a speech act, and finally was not, like the false doctor or the false oasis. It is a legitimate category, not the effect of rejection in a process of categorization. Now, just like 'punt theory' ought not make any reference to wood as a necessary condition for something to be a punt, and that would not force the 'punt theoretician' to deny (9) or (10), speech acts theorists ought not consider intentions as necessary conditions, and this would not force them to deny (11) or (12):

(11)   Speech acts are usually the effect of mental states.
(12)   Speech acts usually work *because* they are the effect of mental states.

However, the link between speech acts and mental states is more essential than the one between punts and wood. The latter is a historical and circumstantial relation,

whereas it seems that the former, as we will now see, is somehow constitutive, or even logical.

5. Criteria as logically good—but not unappealable—reasons

In order to defend the possibility of unintentional speech acts I would like to introduce a conceptual distinction inspired in Wittgenstein (2009) between logical entailment, symptoms and criteria, as three different reasons to infer from the patent manifestation of an event p the occurrence of another one, q[8]:

(a) *Logical entailment*, needless to say, is an a priori judgement—i.e. true by definition—that links a particular phenomenon with a set of necessary and sufficient conditions: if p has taken place, and p logically entails q, then q must have taken place. Of course, we can change the conventions that are in the basis of logical entailments, since any definition might be subject to further revisions; but as long as the definition is accepted, logical entailment is strictly necessary.

(b) *Symptomatic induction* is the effect of an *a posteriori* hypothesis by which we take p as an empirically good evidence for the occurrence of q: if p has taken place, and p is a *symptom* of q, then q has probably taken place. This is a piece of inductive evidence, inferred from previous experiences, and usually based on hypothetical causal connections.[9]

(c) *Criterial attribution* is a logically good—but not definitive or unappealable—type of evidence for the occurrence of a phenomenon. In general, when p has taken place, that *means* that q has taken place. Unlike symptoms, criteria are not based on mere hypothetical generalisations from previous experiences: p is not a criterion of q because of empirically observed correlations, but because q is somehow *logically* related to p. It is thus a constitutive connection—or, as Wittgenstein used to say, a *grammatical* question (2009: §353).

---

[8] Some scholars consider this distinction decisive to understand Wittgenstein's work—see (Baker and Hacker 1984:110-5) or (Hacker 1990:545-68)—, whereas others dismiss it as inessential (Hanfling 1989:118-26). In any case, it is not a constant in Wittgenstein's production, since the idea of criterion that he uses in 1929 is very different from the one that we find in the *Blue Book*, and both differ importantly from the version of the *Philosophical Investigations*. It is this final version that I will use here, according to Hacker's (1990) reading of it. However, I am not particularly concerned here with the exegesis of Wittgenstein, and in what follows I will only instrumentally make use of these concepts.

[9] The symptomatic connection does not *have* to be causal, at least not in a straight way. For instance, even if p is a good symptom of q, q does not have to be p's cause. They could usually happen together for another reason—e.g., because they share a common cause, r.

However, unlike logical entailment, criterial attributions are fallible in principle: even if *p* is a *criterion* for *q*, it could be the case that *p*, but not *q*. Like logical entailment, criteria are internal conceptual relationships; but, like symptoms, their relation is contingent, at least in particular occasions.

The limits between these three categories are blurred, even in scientific discourse. In particular, as Wittgenstein himself remarked, "The fluctuation in grammar between criteria and symptoms makes it look as if there were nothing at all but symptoms" (2009: §354). The fact is that, at first sight, both concepts seem to play a similar role: in contrast to logical entailment, neither symptoms nor criteria identify *p* with *q*, and both can fail. However, it should not be neglected that criteria, in contrast to symptoms, are not susceptible to failure because of the fallibility of our empirical generalisations. It is not the fallible character of our empirical inductions what makes criteria occasionally deceitful, since they are simply not based on inductions at all: they are assumed a priori, as meaningfulness conditions of our actions.

The classic example for this is Wittgenstein's account of pain: pain behaviour does not logically entail pain, nor does pain entail pain behaviour—since it is perfectly coherent to find the one in absence of the other—but, at least according to some views, it would be misleading to consider it as a mere symptom of pain, since our use of the word pain and its cognates, and even our understanding of the concept itself is tied to the public situations where pain behaviour is manifest.

There is a peculiarity of criteria that makes the contrast with symptoms quite evident. If *p* is a criterion of *q*, and *p* is accomplished in a particular situation *up to a certain point*, it makes no sense even to ask about the occurrence of *q*. E.g., if somebody had fallen from a second floor and were lying on the ground with an open fractured leg screaming in pain, it would not make *any* sense whatsoever to ask "But is he *really* in pain?". The correct response to such an sceptic inquiry would not be "How could you possibly *doubt* that?", but "What in the world do you *mean*?". Sceptics about other minds would try to claim that we could only see this man's behaviour, but not pain itself, whose existence could only be *inferred* from the observance of behavioural evidence, and thus will always remain hypothetical. But this attitude would only make sense if we were talking about symptoms, which are empirically and hypothetically linked to whatever they are a symptom of. If pain behaviour were just a mere symptom of pain, it would always be possible to remain sceptic about the necessity of the conclusion. But if it is a matter of criteria, such kind of scepticism would only lead to a meaningless use of language, since the sceptic would be trying to make use of an expression completely detached from any social language game that could make it understandable and functional. Pain behaviour is not pain itself, since their relation is not of logical entailment, but under extreme conditions it does not make any

sense whatsoever to doubt about the occurrence of pain. In other words: we *see* pain in pain behaviour, we do not *infer* it from it.

However, the appearance of further contextual evidence could *always* open the way to meaningful scepticism. Imagine for instance that, in the previous example, without realising it, we had got into some scenery where a film were being recorded. The screaming man on the ground would be an actor, and his broken leg a highly lifelike special effect. In that case, the space for sensible scepticism would be reopen. But under those conditions the question "What in the world do you *mean*?" would have quite an easy answer: the sceptic would only have to point to the cameras, the director, or the limits of the stage. However, the fact that a further widening of context would allow us to reconsider criterial attribution does not entail that, in the previous situation, we were making any kind of fallible inference, which would have later been proved to be wrong. There was no such inference on empirical hypothetical grounds, since we were not judging on symptoms, but just acting in accordance to criteria that are constitutive of our concepts.

6. Austin's "implication" as limited entailment—and why the notion of criterion is a better one

It might be helpful to approach Austin's work with those concepts in mind, in particular the moment when he struggles to find out what connection there is between speech acts and mental states. The necessity of this kind of connections is manifested in what Wittgenstein (2009: II, x, §87) had famously called "Moore's paradox": it is somehow odd to say something like "*p* but I do not believe that *p*". However, there is nothing logically incoherent between *p* and somebody not believing that *p* is the case (i.e. somebody being wrong), or between somebody not believing *p* but, at the same time, asserting *p* (i.e. somebody lying). Those are not logical contradictions at all. Moore's paradox does not emerge from the semantic content of the sentence, but from its pragmatic performance, when the speaker both asserts *p* and denies that she has the correlative belief. She is not just not expected to do that, but somehow not *allowed* to do it, because her assertion and her belief *should* be accord.

The scope of this kind of paradoxes is both more extended and more restricted than it was at first thought. On the one hand, it can be extended to other kinds of propositional attitudes and speech acts (whenever a speaker requires something, she is expected to desire it; whenever she promises something, she is expected to have the intention to accomplish it, and so on), and some authors have even extended it to unasserted mental states (an agent could not both believe that *p* is the case and that it is also the case that she does not believe it herself). On the other

hand, as Tsohatzidis (1994) has shown, there can be exceptions to any of those cases.

Moore's paradoxes show that there is some connection between the speech act performed by the speaker and the content of her mental states. Austin calls this kind of inference an "implication" (1962: 47-52). Now the problem is: what exactly is an "implication"? Lacking from the concept of criterion, Austin seems to be equally unsatisfied both by logical deduction and symptomatic induction as models to understand implications. He correctly points out that we are not facing here any kind of logical "entailment", because such an entailment ought to take place between the propositions that composed the assertion, and they are, as we have seen, perfectly compatible with each other. Besides, an implication must be weaker than entailment, since it can be wrong without strict logical contradiction.

While the problem with entailment is that it introduces a too tight requirement, the problem with symptomatic induction is that it is too loose, since such an account of implications would merely point to contingent and probably causal connections between them. When we find somebody asserting *p*, we have a good reason to ascribe to her the believe that *p*; but the reason is not that we have found this correlation in our linguistic community. There is something somehow unexpected in the occurrence of a symptom in absence of the illness it is a symptom of, or of an illness in absence of its alleged symptoms, but that strangeness has nothing to do with the one that we experience when we face Moore's cases. The oddity of Moore's cases does not stem from a break of expectations based on empirical regularities—like that people *usually* believe what they assert, and those beliefs are *usually* the cause of their assertions. Unlike failing symptoms, Moore's cases are the effect of a breach of inner conceptual connections[10].

Furthermore, if speech acts were just an external and causal manifestation of inner mental states—i.e. their symptoms—there would be no way to prevent bad faith speakers from denying the normative implications of their actions. Austin exemplifies this with the case of Euripides' *Hippolytus*, who famously denied to be attached to his promise, excusing himself with the words: *My tongue an oath did take, but not my heart* (I, 612. Trans. Coleridge).

If the causal role accomplished by mental states were a necessary condition for the performance of a speech act, the very existence of the act and its effects would strictly depend on the occurrence of such mental state. In that case, nothing would

---

[10] For similar reasons, Grice (1989:42) refused to give an account of Moore's paradox as an effect of conversational implicatures. As Harnish (1977:370-1) points out, Moore-type sentences would precisely be unsuccessful attempts to cancel such alleged implicature.

prevent the cunning speaker from escaping his duties, since intentions, considered as inner hidden states, cannot in principle be *ever* present in the public scene. An illocutionary act, like a promise or a request, must not rely on the hidden performance of an invisible heart, "or mind or other backstage artist" (Austin 1962: 9-10), and does thus not consist in the outer manifestation of such an inner state. Hippolytus "does promise: the promise here is not even *void*, though it is given in *bad faith*. His utterance is perhaps misleading, probably deceitful and doubtless wrong, but it is not a lie or a misstatement" (1962: 11). *Word is our bound*, not inner commitments, or the effective existence of the implied mental state. The utterance is not a speech act because it is the contingent causal manifestation of an inner act, a mere symptom of thought: the utterance *is* the act itself.[11] By the same token, a speaker's assertion that *p* is not just a good symptom of her belief that *p*: the assertion somehow *means* that the speaker believes what she says. In saying the words, she asserted; in asserting, she manifested her beliefs.

Lacking from the concept of criterion, and equally unsatisfied both by logical entailment and symptomatic induction, Austin had to look somewhere else for an alternative model for implications. He thus proposed a revised notion of logical entailment: it must somehow be the case—*at least under certain circumstances*—that the utterance necessarily *is* a speech act, and the speech act necessarily implies the mental states we expect from it. Austin had thus to appeal to what he called *normal* conditions, restricting the logical entailment of utterances and illocutionary acts to those particular cases. Not just any word utterance by itself entails the performance of illocutionary acts; but word utterance under "normal conditions" *do* entail the intentional states they express, and thus *do* lead to the performance of fully-fledged speech acts.

It is now important to realise that the model proposed by Austin is *not* the one of criterion, as it has been described in the previous section, but the one of a restricted logical entailment whose antecedent has been increased by the addition of further requirements. I.e., Austin substitutes

$$p \rightarrow q$$

for

$$(p \,\&\, r \,\&\, s \,\&\, \ldots) \rightarrow q$$

---

[11] N.B.: when Austin was asked by Chaim Perelman for the need of a meta or extra-linguistic criterion for the performative after a lecture he delivered in France in 1958, he denied that it must be some kind of "sentiment intérieur". See (Béra 1962:271-304), and (Maclean 2004:52).

In this way, the utterance of *p* would be just one of the conditions for the fully effective speech act, and an indefinite number of other contextual conditions (*r*, *s*, and so on) would also be requisites for the entitlement to be strictly necessary. Those conditions would constitute what Austin famously called "the total speech act in the total speech situation" (1962: 53).

Austin himself was aware that the idea of *normality* is extremely vague, since we seem to be forced to consider rehearsals, soliloquies, ironies, metaphors, or quotations as 'abnormal' uses, despite their ubiquity in ordinary language.[12] We do not seem to be able to define what we consider "normal" conditions, but through an enumeration of allegedly 'abnormal' situations. The problem then is that, unless we are given an intensional notion of normality, we are not sure how to follow that extensional enumeration of 'abnormalities'. And, in any case, it cannot be denied that somebody under really weird and abnormal conditions could still try to make a perfectly 'normal', 'serious' and 'literal' speech act.

In an effort to prevent Hippolytus from escaping his duties, Austin tends to heavily rely on the notion of convention: where a clear convention pre-exists the speech act, and the performance fits that convention, the act is fully performed whether the speaker has the implied mental state or not. Conventions would thus link a set of performance conditions with particular illocutionary effects, overcoming the problem of concomitant mental states. E.g., if a judge is dressed in a certain way, in front of a certain audience at a certain moment, and utters the words "I find the defendant guilty of murder in the second degree" he is actually performing the speech act of declaring the accused guilty, quite independently of his alleged intention. However, it has been convincingly shown—at least since Strawson (1964)—that the notion of convention is highly problematic, at least for two reasons: first, it is much more restricted than the idea of speech act, and could thus hardly give a full account of it; in general, conventions as such only seem to play a crucial role in quite rigid, ceremonial or ritual acts (Bach and Harnish 1979: 55). Second, even when a convention can be invoked, that does not seem by itself

---

[12] In Harnish's (2005:21) expression, "None of these activities are non-normal in any normal sense of 'normal'." This is the reason he gives to reject Kemmerling (2002) notion of "*cp*-analytical evidence", which might seem close to the one of criterion proposed here. Kemmerling is searching indeed for a concept that could accomplish the role here attributed to criteria, between mere symptoms and unappealable entailments. However, his appeal to "normality", as Harnish points out, makes the proposal collapse again into "something empirical".

to solve the difficulties it was expected to solve, since conventions are not self interpreting entities (Sperber and Wilson 1986).[13]

In contrast to conventions—whether implicit or explicit—criteria seem to be less limited in its scope. For instance: in the aforementioned example of the person screaming with an open fracture, I really cannot think of any notion of convention that could justify there the alleged 'inference' by the part of the observer as to the pain of the observed person. Being aware of the pain of the other, and being prone to help her in this situation, is simply the immediate and correct way to react. There would be no convention sustaining an inference as to the mental states of the other, and no theoretical premise in any 'theory of mind', but just direct perception of her situation, and response in accordance to that perception. This same idea could be extended from primitive actions and meanings, such as pain behaviour, to complex acts mediated by linguistic structures. As far as the hearer is a competent user of the language, the utterances that she hears are direct expressions of the intentions of the speaker. If somebody asserts p under a particular situation, what we see is that she believes that p is the case: we do not need to appeal to conventions or inferences.[14]

To sum up, the problem with Austin's explanation of Moore's paradox is that he tries to understand the "implication" between speech act and intentional state as a weakened sort of logical entailment, where the antecedent conditions would have been indefinitely extended. But such an entailment is hard to understand, unless we can clearly delimit the conditions that ought to appear in the antecedent. As far as there is an open antecedent, there is always the possibility of a failure in the entailment, and thus, strictly speaking, no such entailment. This would leave the door open to Hippolytus, since it would always be up to him to deny that the conditions were fulfilled, and thus that he should keep his own words. Appealing to conventions would not solve the problem, since conventions—at least conceived as *means* for the action—are much more limited than speech acts and, in any case,

---

[13] Sbisà has tried to defend the original Austinian position by claiming that what is conventional in the performance of speech acts is not their *means*, or at least not necessarily, but their *ends*, that is, what is done in them (Sbisà 2009:43). I am sympathetic with Sbisà's attempt, but I am inclined to believe that criterion is a more effective notion than convention, in order to attain the same goals, namely, making speech acts purely rely on the 'external' conditions of the performance.

[14] There is an obvious resemblance between this idea and the Gricean agenda to get from "natural" to "non-natural" linguistic meaning (Grice 1989:292), but those are not the categories I am using here. It seems to me that Grice's decision to define linguistic meaning in terms of intention to induce beliefs by the recognition of that same intention would also be pushing unintended acts out of the category of communicative actions.

the notion of convention does not properly grasp the immediacy of criterial attributions.

7. Searle's "expression" and the sincerity condition

Searle's (1969) combination of conventions and intentions in the conditions for the performance of speech acts turns his proposal into one of the most 'mixed' cases in Harnish's categorization. With respect to the crucial question proposed at the beginning of this paper, Searle seems to assume that it is a *sine qua non* condition for the performance of the act it to be the result of a quite specific set of intentional states[15]. In particular, the sixth of his conditions, the famous "sincerity condition", involves the conscious entertainment by the part of the speaker of a specific psychological state. For instance, for promises, he claims that:

> 6. *S intends to do A*.[16]

is a condition for the act to be sincerely performed. Similar requirements would be introduced for assertions (S believes p), petitions (S desires the hearer to do A), and so on. Cases like Hippolytus' one, who denies to have had the concomitant intention while his 'tongue' was apparently uttering the oath, are solved by Searle by transforming this condition, assuming that sincerity itself is not a requirement for the act to be performed. However, his transformation still preserves the structure of an intentional attribution:

> 6a. *S intends that the utterance of T will make him responsible for intending to do A*.[17]

---

[15] It is difficult to say if Searle would consider this as a requirement for each and every speech act, since he explicitly confines his theory "to the center of the concept of promising", deliberately "ignoring marginal, fringe, and partially defective promises" (Searle 1969:55). In so doing, he follows Austin's methodological exclusion of infelicities, and exposes himself to the kind of confusions pointed out in section 2. This move was famously denounced by Derrida (1988)—see (Navarro-Reyes 2010). Derrida also defended there the need to reconsider the role of intentions in speech acts, in a way that is not very different from the one here proposed.

[16] Where *S* stands for Speaker and *A* for a future Act.

[17] Where *T* stands for some sentence or similar stretch of discourse.

Even if the speaker does not have the intention to fulfil his promise, his utterance is a committing speech act because he has the intention that his utterance will make him responsible for having that intention. This is, according to Searle, the effect of a law that governs speech acts in general:

> Wherever there is a psychological state specified in the sincerity condition, the performance of the act counts as an expression of that psychological state. This law holds whether the act is sincere or insincere, that is whether the speaker actually has the specified psychological state or not (Searle 1969: 65).

It does not matter for the very performance of the act if the subject is in the alleged psychological state or not: his words count as an expression of that state, and that fact bounds him to them. In contrast to Austin, Searle, following Grice in this, prefers to reserve the word "imply" for what he calls "preparatory conditions", such as that, in a promise, the hearer would prefer the act to be performed—quite the opposite of threads—, or that it is not obvious that the speaker was going to perform it anyway—since that would make the promise pointless, or even counterproductive. What is *implied* in the speech act is what gives it a point, something different in Searle's account from what is *expressed* in the act, that is, the mental state that can be expected from the speaker.

Some authors (Malcolm 1991: 160; Marmor 2009: 124), tend to understand the verb "express" in a factive sense, and that assumption prevents them from correctly understanding Searle's account. A factive use of "express" would imply that the speaker effectively *has* the state she expressed in her words, and this would make insincere speech acts impossible: if having the state were a condition for the act, the lack of intention in a promise (of belief in an assertion, and so on), would prevent the very act from coming to existence. Nevertheless, that is not what Searle has in mind when he uses the concept of "expression": a speaker can perfectly "express" a mental state that she does not have. Effectively having the state is not a condition for the act to take place, but only for its sincere performance.

8. What would Hippolytus deny?

There is one way of understanding Searle's transformation of the sincerity condition that turns it into a pointless move: the problem with the first version of the sincerity condition was that it introduced a 'hidden' condition for the performance of the act, something that never takes place in the social scene of the utterance: the inner intentional attitude. This left the door open to bad faith speakers like Hippolytus to claim that they did not actually perform speech acts at all, since they were not in the concomitant state. Searle's revised version of the sincerity condition is supposed to solve this but, if we understand 6a as establishing

an additional internal commitment, this step clearly fails[18]. Searle's revision of his own sincerity condition would not be successful, since it would not prevent Hippolytus from escaping his duties. The reason is that both 6 and 6a would require an intention by the part of the speaker that could be denied by Hippolytus: just like he refused to be bound by his promise because he did not have the intention to *do* A, he could also refuse to be bound by his promise because he did not have the intention to make himself responsible for intending to do A. It would just be a more complex claim to do, but the possibility of making it would not have changed in any relevant way.

A weak interpretation of Hippolytus' excuse would construe him as denying 6, that is, that he had the intention to perform the promised action. But one can image a stronger bad faith by his part: he could not be denying that he had the intention to *fulfil* his promise, but that he had the intention to *make* the promise. He would not be claiming that his utterance was an insincere promise, but denying that his utterance was a promise at all. There are two senses in which a promise can lack the concomitant intentions:

(13)   'I promise', said when the speaker does not intend to do what she promises.
(14)   'I promise', said with no intention to promise.

By claiming that it was only his tongue, but not his heart, who did that oath, Hippolytus does not seem to mean (13), which is what Searle's revision of the sincerity condition would preclude, but (14). He would be saying that he was not intending to create with his words the alleged bound—not just that he did not intend to respect the bound he was creating. Hippolytus would not be denying 6, but 6a. This might appear as a philosophically sophisticated example, but the case is not so weird as one could expect. In fact, in our ordinary lives we meet almost every day people who fishily deny having talked seriously when they said something, or having said it intentionally, or having been aware of the implications of their words. Hippolytus is not just a speculative case: preventing people of his sort from escaping his duties is one of the most important aims of linguistic practices.

Now, the problem cannot be solved with a merely symptomatic approach to "expression": if by saying that an utterance expresses an intention, we are just making an empirical hypothesis as to its causal psychological precedent, there is no

---

[18] Alston is exposed to the same kind of criticism when he claims that the speech act's committing character depends on the speaker's act of "taking responsibility" for her words (2000: 54).

way to close the escape to Hippolytus, and no way to prevent the performance of the speech act from being, in the end, just a sort of 'internal act' contingently manifested in the social world. If "expression" were a merely symptomatic notion, there would be no reason why the revised sincerity condition would succeed where the original one failed. Hippolytus could still say that his utterance was not binding, because he did not intend to take the responsibility for it being the case that he had that intentional state. Nothing substantial changes if we just replace a first order intention for a second order one in our requisites.

I can only see two options here: whether we preserve intentions such as 6a as necessary conditions (and we must thus give an account of a factive use of "expression") or we simply take intentions out of our requirements. In the first option, by making a promise, the speaker would be "expressing", in a non-factive sense, the intention to fulfil it, but he would be "expressing", in a factive sense, his intention to make the promise, since otherwise there would be no promise at all, and Hippolytus would not be bound by his words. But finding an account of the factive use of "expression" does not seem to be an easy task. Why does that intention *have* to be effectively present? By proposing a factive use of "express" we would be trying to understand the relationship between utterance and intention as logical implication and, as we have seen with Austin, this does not work. We get back to Austin's original problem, since the utterance by itself does not seem logically to entail any mental state, and neither does the utterance plus its context, unless we introduce in the context the intentional state itself—and by doing so we would be begging the question[19].

That leaves us with the second option: taking intentions out of our requirements. And that is the step I propose to take: we will not preclude bad faith speakers from denying their commitments by transforming the sincerity condition into a more complex second order intention, but by clearly formulating the way in which responsibility stems from words and deeds, not from the alleged intentions that generated them. Responsibility does not arise from the actual intentional state of the speaker (whether it is a first or a second order one), but from the fact that her

---

[19] According to Searle (1989: 556), the self-guaranteeing character of performatives would be explained because "as far as illocutionary force is concerned the speaker cannot lie or be mistaken: assuming the other conditions on the speech act are satisfied, *if he intends his utterance to have the force of an order*, then it has that force; because the manifested intention is constitutive of that force" (my emphasis). What I am claiming here is that Hippolytus could simply deny that the subordinate clause of that conditional is a correct description of the situation. Moya's account of "meaningful actions"—as those that necessarily imply the intention to be performed—would be exposed to similar criticism, unless "intention" is assumed in a criterial way (Moya 1990: ch. IV).

utterance *is*—in the sense that it *counts as*—an expression of the state. And that idea of "expression" should clearly not be understood as an act of will, by which the relevant intention *causes* the utterance—that would thus be its symptom—, but by the sort of criterial requirements that the utterance accomplishes. In contrast to the symptomatic sense of "expression", the criterial one does not rely on the purportedly causal connection between utterance and mental state, but on the conceptual connection that allows us to say that a particular pattern of behaviour *counts as* an expression of a specific psychological state. That criterial relation is not up to the speaker, since it is simply not the consequence of any act of his will.

9. Expressibility is not enough

The problem with Searle is that there is some sort of ambiguity in his theory, since in some cases we find the notion of "expression" as pointing to a symptomatic and causal connection, and sometimes we find it as a basis for criterial attribution. In other words: sometimes Searle considers that it is a condition of the speech act that the speaker *intends to express* the mental state indicated in the revised sincerity condition (which would make of the expression a symptom causally related to its cause, the mental state), and sometimes he describes the utterance itself as a fact that *counts as* an expression of the concomitant mental state. But the idea that the utterance itself "counts as" an expression of the mental state does not point to any intended or voluntary action by the part of the speaker (his act of expressing), or to a causal connection between the mental state and the utterance, but to a grammatical or logical relation between the expression and the situation in which the audience attributes mental state predicates.

Searle talks about "expression" indifferently as a voluntary act by the part of the speaker (S expresses the attitude) and as a criterial attribution by the part of the audience (the utterance *counts as* such an attitude). But both ideas are quite different in kind, each one of them leading to different conceptions of the link between mental states and utterances.[20] However, I wouldn't say that Searle's ambiguity in the use of "expression" is the unpremeditated effect of a lack of attention. He makes use of both senses indifferently because he considers that they

---

[20] The symptomatic sense can be found in Bach and Harnish when they claim that "For *S* to *express* an attitude is for *S* to [reflexively] intend the hearer to take *S*'s utterance as a reason to think *S* has that attitude" (1979:15). The criterial sense would point to the utterance itself, and the way in which it gives the hearer good reasons to attribute to the speaker the alleged attitude.

can somehow be, not just compatible, but interchangeable. He explicitly makes that assumption since the beginning of *Speech Acts* (1969: 16), by establishing the *Principle of Expressibility*, according to which "whatever can be meant can be said". The possibilities of any intended expression by the part of the speaker as a voluntary act would be equal to the possibilities of criterial attribution by her target audience. Paraphrased, the *Principle of Expressibility* would guarantee that, at least *de iure*, for any mental estate the speaker might want to *express* (in a symptomatic sense), there is—or could be—a possible utterance that would express such estate (in a criterial sense). That is: for every meaning that the speaker might intend to communicate, there is a sentence whose meaning does the job.

However, even if this Principle were sound—and I have expressed somewhere else my doubts that it could be supported without some substantial qualifications (Navarro-Reyes 2009)—, it would only grant the *possibility* of that accordance between speaker meaning and sentence meaning, but not its necessity. It would only show that the speaker *could* choose ways to express her attitudes that would count as an expression of such attitudes for her audience, but it does not show that whenever she makes use of that ways of expression—ways that are criterially related to the mental attitude—she must effectively intend to perform the alleged speech act. The *Principle of Expressibility* only claims that the symptomatic sense of 'expression' and the criterial sense *can* match, but it does not force the speaker to make them match in any particular situation.[21] If their matching is a matter of inner commitment, that is, of having the alleged mental state as the effective cause of the utterance, Hippolytus is still free to deny it.

That is why Searle's account of speech acts must rely on conventions up to a point that some have considered disproportionate. If the speaker intends to perform a speech act, she has to use means that are previously considered by the community as basis for the attribution of mental states and illocutionary intentions, since the speech act could only work under the assumption that there is some previously shared common knowledge and awareness that such and such piece of behaviour 'count as' an expression of such and such mental state. Every act must somehow be guaranteed by a convention that works as its vehicle, since otherwise the possibilities of the speaker to express her attitudes would not correspond to the expectations of her audience. If the act were not guaranteed by a convention, we

---

[21] Furthermore, neither does it force the *audience* to make them match. It could be the case, for instance, that all criterial reasons to attribute an intentional state were fulfilled, but the audience were justified to deny the occurrence of the corresponding mental state, and would thus not take the alleged expressions as 'counting as' anything at all. A good example is Searle's (1980) Chinese room argument.

could not be sure of what it means to say that the speaker expresses the attitude, since there would be nothing previously considered as a basis for the attitude attribution.

However, we have seen with Austin that conventions do not do the job, since they can only justify indefinitely limited entitlements, not being in any case a guarantee for the factive use of "express". In contrast to this, criterial uses of "express" are definitively not factive—just like criterial attributions are not logical entailments. Neither are they merely symptomatic inductions, since judging on criteria is not making hypothesis as to hidden causes, based on empirical regularities. That is the reason why I claim that the only way to force Hippolytus to keep his word is by attaching his responsibility to his utterance as such, to the criterial conditions that it fulfils, and not to an alleged inner state that ought to be the cause of his utterance. And that is the main thesis of this paper: that it is not mental states that bind, or hidden intentions, but words and deeds. Responsibility does not stem from what we intend to mean, but from what we actually say, quite independently of how our utterance is causally connected to our intentions.[22]

10. Conclusion

I am not claiming that we ought to exclude intentions from our consideration of speech acts. Nor am I denying that beliefs, desires and intentions are generally the causes of those acts, or that they are effectively expressed in them. I am simply claiming that our consideration of an utterance as a speech act is not a symptomatic hypothesis as to the causal role performed by those states. When we claim that somebody made a speech act, we are not making empirical conjectures about the unseen origin of her words: we are describing what she did with them—whether intentionally or not. The price to be paid to force bad faith speakers to keep their

---

[22] Despite the scepticism that some scholars have shown about the convenience to introduce responsibility in an account of speech acts—vg, (Harnish 2005)—, I *do* think that it is a crucial concept, and I agree on this with Searle, and even with Alston. But I do not think that responsibility emerges from the speech act *because* it is the effect of a mental intention. We are made responsible for our words, not for the intentions that we had when we uttered them. We might find understanding in our audience if we allege that our act was unintended—or we might not find it; in any case, the very performance of the act that we are trying to excuse does not strictly depend on the intention that we had when we performed it. And that is precisely what gives our excuse a reason.

words is to detach the performance of the act from its alleged causal history in the mind of the speaker.

I would like to finish with an important qualification: it only makes sense to say that somebody performed an unintentional speech act if that person is able in general to perform intentional speech acts. Unintentional speech acts are necessarily a marginal phenomenon, and could never be the rule. That is the reason why the link between speech acts and intentions is much stronger than the one between punts and wood: it is a constitutive one, not just circumstantial. However, although it is constitutive, it is contingent. That is the contribution of criteria, as logically good—but not unappealable—reasons, which are the basis of our practices: criteria whose conditions must *all* be external, just like the social world in which utterances become speech acts, and acquire normative weight.[23]

References

Alston, William P. *Illocutionary acts and sentence meaning*. Ithaca: Cornell University Press, 2000.
Austin, John L. *How to do things with words*. Oxford: Clarendon Press, 1962.
Bach, Kent and Robert M. Harnish. *Linguistic communication and speech acts*. Cambridge, Mass: MIT Press, 1979.
Bach, Kent and Robert M. Harnish. "How performatives really work: A reply to Searle." *Linguistics and Philosophy* 15 (1992): 93-110, doi: 10.1007/BF00635834.
Baker, Gordon P. and P. M. S. Hacker. *Scepticism, rules and language*. Oxford: Blackwell, 1984.
Benveniste, Émile. *Problèmes de linguistique générale*. Paris: Gallimard, 1974.
Béra, J. (ed.) *La philosophie analytique*. Paris: Les Éditions de Minuit, 1962.
Davidson, Donald. *Essays on actions and events*. Oxford: Clarendon Press, 1982.
Derrida, Jacques. *Limited Inc*. Northwestern University Press, 1988.
Gazdar, Gerald. "Speech act assignment." In *Elements of discourse understanding*, eds. Aravind K. Joshi, Bonnie Lynn Webber, and Ivan A. Sag, 64-83. Cambridge: Cambridge University Press, 1981.
Grice, Herbert Paul. "Meaning." *The Philosophical Review* 66 (1957): 377-388.

—. *Studies in the way of words*. Cambridge, Mass: Harvard University Press, 1989.

Hacker, P. M. S. *Wittgenstein: meaning and mind*. Oxford: Basil Blackwell, 1990.

Hanfling, Oswald. *Wittgenstein's later philosophy*. Albany: State University of New York Press, 1989.

Harnish, Robert M. "Logical Form and Implicature." In Thomas G Bever, Jerrold J Katz and D. Langendoen (eds.) *An Integrated Theory of Linguistic Ability*, 313-391. Brighton: Harvester, 1977.

—. "Commitments and Speech Acts." *Philosophica* 75 (2005): 11-41.

—. "Internalism and externalism in speech act theory". *Lodz Papers in Pragmatics. Special Issue on Speech Actions*, edited by Iwona Witczak-Plisiecka & Maciej Witek. 5.1 (2009): 9-31, doi: 10.2478/v10016-009-0001-2.

Kemmerling, Andreas. "Expressing an intentional state." In Günther Grewendorf and Georg Meggle (eds.) *Speech acts, mind, and social reality*, 83-91. Dordrecht: Kluwer Academic Publishers, 2002.

Maclean, Ian. "Un dialogue de sourds? Some implications of the Austin-Searle-Derrida debate." In Ian Maclachlan (ed.) *Jacques Derrida. Critical Thought*, 49-66. Hampshire: Ashgate, 2004.

Malcolm, Norman. "I believe that p." In Ernest Lepore and Robert Van Gulick (eds.) *John Searle and his* critics, 159-68. Cambridge, MA: Basil Blackwell, 1991.

Marmor, Andrei. *Social conventions: from language to law*. Princeton, N.J.: Princeton University Press, 2009.

Millikan, Ruth Garrett. *Language: a Biological Model*. Oxford: Clarendon Press, 2005.

Moati, Raoul. *Derrida/Searle. Déconstruction et langage ordinaire*. Paris: Presses Universitaires France, 2009.

Moya, Carlos J. *The philosophy of action: an introduction*. Oxford: Polity-Blackwell, 1990.

Navarro-Reyes, Jesús. "Can we say what we mean? Expressibility and background." *Pragmatics & Cognition* 17 (2009): 283-308, doi: 10.1075/p&U38;c.17.2.04nav.

—. *Cómo hacer filosofía con palabras: a propósito del desencuentro entre Searle y Derrida*. Madrid: Fondo de Cultura Económica, 2010.

Sbisà, Marina. "Speech acts in context." *Language and Communication* 22 (2002): 421-436, doi: 10.1016/S0271-5309(02)00018-6.

—. "Uptake and Conventionality in Illocution" *Lodz Papers in Pragmatics. Special Issue on Speech Actions*, edited by Iwona Witczak-Plisiecka & Maciej Witek. 5.1 (2009): 33-52, doi: 10.2478/v10016-009-0003-0.

Searle, John R. "Minds, brains, and programs." *Behavioral and Brain Sciences* 3 (1980): 417-24, doi: 10.1017/S0140525X00005756.

—. *Speech acts : An essay in the philosophy of language*. Cambridge: Cambridge

University Press, 1969.

—. "How performatives work." *Linguistics and Philosophy* 12 (1989): 535-58, doi: 10.1007/BF00627773.

Sperber, Dan, and Deirdre Wilson. *Relevance: communication and cognition*. Oxford: Blackwell, 1986.

Strawson, P. F. "Intention and Convention in Speech Acts." *The Philosophical Review* 73 (1964): 439-460, doi: 10.2307/2183301.

Tsohatzidis, Savas L. 1994. "The gap between speech acts and mental states." In Savas L. Tsohatzidis (ed.) *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, 220-33. London and New York: Routledge, 1994.

Urmson, J. O. "Performative utterances." *Midwest Studies in Philosophy* 2 (1977): 120-127, doi: 10.1111/j.1475-4975.1977.tb00033.x.

Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Wiley-Blackwell, 2009.

About the Author:

Jesús Navarro-Reyes is lecturer of contemporary philosophy at the University of Seville (Spain). He has been Post-Doctoral Fellow at the University of California (Berkeley) and Visiting Scholar at Duke University. He is currently Academic Visitor at Wolfson College, Oxford. He has authored three books and various articles on Montaigne's scepticism, the philosophy of subjectivity, mind and language. The title of his last book is *Cómo hacer filosofía con palabras. A propósito del desencuentro entre Searle y Derrida* [How to Do Philosophy with Words: Reflexions on the Searle-Derrida Debate] (FCE, 2010).