

## Forthcoming in Pragmatics & Cognition

### **From predictions to promises: how to derive deontic commitment\***

#### **Abstract**

*This paper attempts to identify general, cross-cultural cognitive factors that trigger the default commissive interpretation of assertions about one's future action. It is argued that the solution cannot be found at the level of the semantics of the English will, or any other future tense marker, but should be sought in the structure of rational intentions, as combined with the pragmatics of felicitous predictions and with parameters linked to the evolutionary advantage of cooperative behaviour. Some supporting evidence from language development studies is briefly presented.*

Keywords : promises, commitment, futurity, cooperative behaviour

Mikhail Kissine

FNRS

Laboratoire de linguistique textuelle et de pragmatique cognitive

Université Libre de Bruxelles

CP 175

50, avenue F.D. Roosevelt

B-1050 Bruxelles

Belgium

[mkissine@ulb.ac.be](mailto:mkissine@ulb.ac.be)

[homepages.ulb.ac.be/~mkissine/](http://homepages.ulb.ac.be/~mkissine/)

“Cuando no se está demasiado seguro de nada, lo mejor es crearse deberes a manera de flotadores.”

Julio Cortázar, *El perseguidor*

## 1 The problem

Even if it is certainly ill-founded to postulate a one-to-one correspondence between speech act types and sentence-types (e.g. Wilson and Sperber 1988; Vanderveken 1990: 45), there are, without any doubt, systematic links between the declarative mood and assertive acts, between the imperative mood and directive acts, and between the interrogative mood and questions. However, no undisputable evidence has been found that any human language associates a prototypical syntactic mood with speech acts (henceforth, “commissive acts”) such as promises, threats, offers, and the like, whose performance aims at committing oneself to bringing about a certain state of affairs (Sadock and Zwicky 1985).<sup>i</sup> This fact gave rise to overspread doubts about the accuracy of Searle’s (1975) treatment of commissive illocutions as a distinct speech act type, along with assertives, directives, etc. (e.g. Zaefferer 2001; Croft 1994), and led some authors to analyse promises as institutional acts which depend on intracultural social conventions, in the same way as bequeaths, baptisms, and marriages. The following quote from Sperber and Wilson (1995: 245) is a case in point:

Promise and thanking, for example, [...] are institutional acts, which can be performed only in a society with the requisite institutions, and which must be recognised as such in order to be successfully performed.

In a footnote following this statement, Sperber and Wilson are even more assertive about the (allegedly) cultural dependence of the commitment generated by promises:

A promise is a particular, culturally defined form of commitment. [...] We have no doubt that a cross-cultural study of such speech acts would confirm their cultural specificity and institutional nature.

Under such a view, explaining why a successfully performed promise — or, for that matter, any other commissive speech act — imposes an obligation on the speaker requires an independent study of specific cultural institutions.<sup>ii</sup> Yet this line of argument is not very convincing. First, the universality of deontic commitment in human societies (Cummins

1996) makes the dependence of commissive speech acts on a specific, intra-cultural collective convention very unlikely. This universality may be better analysed at a deep conceptual level, so as to explain the existence of cross-culturally valid inferences that are peculiar to the deontic domain (Jackendoff 1999). Second, developmental data in preschool children suggest that the ability to understand deontic commitment and commissive speech acts owes more to innate linguistic and cognitive dispositions than to the immersion into a given social group (Astington 1988b, 1988a; Mant and Perner 1988; Cummins 1996; Bernicot and Laval; Laval 1999; Harris and Núñez 1998).

In order to see clearly the nature of the issues involved once the “culture-dependent” solution is dismissed, Searle’s (1969: 177-182) famous derivation of *ought* from *is* is a good starting point. Searle’s “proof” can be summarised as follows:

1. The speaker (S) uttered the sentence *s*;
2. Under certain conditions *C*, anyone who utters *s* performs a successful promise to bring about the truth of P;
3. Conditions *C* obtain;
4. All promises are acts of placing oneself under an obligation to bring about the propositional content P;

From these four premises we are entitled to conclude (e):

5. *S* is under obligation /ought to bring about the truth of P.

It thus seems that the commitment generated by promises could be accounted for by Searle’s (1995) more general theory of social facts, which are created, according to him, by the formula “In the context *C*, *X* counts as *Y*”, where the clause “*X* counts as *Y*” is potentially recursive. In a certain context *C*, the utterance of *s* counts as a promise — premise (b) —, which, in turn, counts as the undertaking of a certain obligation — premise (d).

Let us begin with premise (b). In Searle’s (1969) derivation of *ought* from *is*, *s* is assumed to be a performative sentence of the form “Hereby, I promise to\_”. It is clear, however, that many, if not most, of our commissive speech acts are realised without using any performative prefix. For instance, in most circumstances, by simply telling you “I’ll proof-read your paper for tomorrow”, I’m placing myself under the obligation to proof-read your paper by tomorrow. Yet, the commissive force is not conventionally attached to this sentence

type.<sup>iii</sup> For instance, it can be explicitly cancelled; compare (1) and (2): while (1) is a typical instance of cancellation of optional pragmatic effects, the second sentence in (2) is pragmatically unacceptable, unless S is understood as retracting her previous utterance.

- (1) I'll proof-read your paper for tomorrow. I can't promise, though.
- (2) I promise that I'll proof-read your paper for tomorrow. ? I can't promise though.

A natural reaction here would be to argue that in the formula “In *C*, the utterance of *s* counts as an act of promising”, *C* stands for appropriate conditions: an utterance counts as a promise only if certain conditions obtain (cf. Searle 1969: 63). But even then, no room is left for the commissive illocutionary force *qua* a cancellable pragmatic component of speaker's meaning. The commissive illocutionary force of the first sentence in (1) can be cancelled because it has been triggered by default, that is, precisely because all preconditions for it to count *ceteris paribus* as a promise obtain in the context of utterance. What is more, it turns out, in the light of subsequent developments of Speech Act Theory, that it is not even possible to tie the illocutionary status of a sentence to a set of context-dependent conditions. Searle and Vanderveken (1985: 75, 78; also Vanderveken 2005: 717-718) claim that, when certain relevant conditions fail to obtain, the illocutionary act, though defective, remains successfully performed. To borrow an example from Tsohatzidis (2007: 211), being mistaken about the existence of Satan, one can promise to kill Satan; the fact that Satan does not exist — which is, in Searle and Vanderveken's (1985) terms, a failure of conditions on the propositional content — does not prevent S from having performed a promise, albeit a defective one. (More on this below.)

The upshot of this discussion is that the default attribution of the commissive force to certain utterances is not to be thought of as a collective agreement about their linguistic meaning. But if so, it is unclear what is really meant by premise (d), which simply states the fact — to be explained — that a promise automatically generates deontic commitment. This deontic status can be traced back, according to Searle, to rules constitutive to the use of every human language (see also Smith 2003): any utterance counting as a promise also has the function of placing S under the obligation to bring about the truth of the propositional content *p*.

Different human languages, to the extent they are inter-translatable, can be regarded as different conventional realisations of the same underlying rule. The fact that in French one can make a promise by saying “je promets” and in English one can make it by saying “I promise” is a matter of convention. But the fact that an utterance of a promising device (under appropriate conditions) counts as the undertaking of an obligation is a matter of [constitutive] rules and not a matter of conventions of French or English. (Searle 1969: 39-40)

In other words, the attribution of the commissive force to a sentence of a certain language would be due to conventions specific to this language, while it would be a cross-linguistic convention that any sentence that receives the status of a promise imposes a deontic commitment on the speaker. But we have just seen that the commissive force is not a feature of the conventional linguistic meaning; therefore, deontic commitment is not an automatic effect of a certain deep-level linguistic feature whose instantiation can be found in every human language. The following quote from Millikan (2005: 149) is an apt statement of the notion of convention that must be avoided here:

[...] the situation with conventional moves is *not* this: that having instanced a certain action shape under certain (difficult to-pin-down) conditions automatically ‘counts as’ x-ing, irrevocably, inexorably, no matter how much you kick and scream. In particular, there can be no such thing as a conventional ‘rule of the form “X counts as Y in context C”’ (Searle 1965:52) where ‘X’ is the description merely of an activity’s shape.

More adapted to the problem in hand is the definition of conventions as behaviours whose pay-offs in a certain type of circumstances have been numerous enough for inducing certain expectations when they are reproduced ; the same expectations account for the fact these behaviours continue to be reproduced. Such a conception allows conventional interpretations to be cancellable and optional (cf. Millikan 2005: esp. chapter 1). In order to unpack premise (d), we thus need to establish this sort of connexion between sentences like “I’ll proof-read your paper for tomorrow”, and the deontic commitment that utterances of such sentences generate by default. This is what this paper sets out to do.

I shall argue that there are deeply anchored cognitive factors that assign a distinctively commissive interpretation to utterances which, structurally, aim at predicting an action of the speaker S; in other words, any commissive speech act is *eo ipso* an assertion but does not

reduce to it. However, before doing so, I should like to address quickly, in Section 2, the question whether the deontic commitment under consideration could not, after all, be traced back to a semantic feature of future-tensed sentences. Even if some readers may feel that I am flogging a dead horse here, this option remains worth considering in that the argument I will rely on to dismiss it will also allow me to emphasise an important constraint bearing on those statements that can acquire a commissive illocutionary force. After outlining, in Section 3, the conditions for a felicitous performance of statements about the future, I will try to show, in Section 4, that when combined with the structure of intentions and the evolutionary impact of cooperation, the pragmatics of predictions explains why such utterances are interpreted by default as commissive speech acts when they concern one's future actions. Threats will be briefly dealt with in the same perspective in Section 5. Finally, I shall mention in Section 6 some supporting evidence from language development studies.

## 2 A semantic solution?

The semantic solution that might come to mind, especially when one looks at the literature on the semantics of English *will*, consists in interpreting the future tense as a kind of deontic necessity. In what follows I shall restrict the discussion to *will*, but the points to be made are general enough to apply to any future tense marker whose semantics would include quantification over possible worlds. Future tenses present well-known similarities with modals (Ludlow 1999: 160; Fleischman 1982; Enç 1996), and since *will* shifts the temporal reference towards the future (Abusch 1998), it seems a very natural step to analyse it as a forward-shifting necessity operator, as proposed by Enç (1996: 354):<sup>iv</sup>

[*will S*] is true at  $\langle w, i \rangle$  iff in every world  $w'$  accessible to  $w$  there is an interval  $i'$  such that  $i < i'$  and  $S$  is true at  $\langle w', i' \rangle$ . [...]  $i$  is the original time of evaluation, i.e. the utterance time. This time is replaced by a future time and the sentence in the scope of the modal is evaluated with respect to this new time.

The accessibility relation that defines the domain (the set of possible worlds) a modal quantifies on — i.e. its modal basis — need not be epistemic (see Lewis 1975; Kratzer 1991); for instance, depending on the context, in (3) *must* may receive an epistemic or a deontic reading:

- (3) John must be in jail.

Nothing prevents us, *a priori*, from assigning a deontic modal basis to *will* in some contexts.<sup>v</sup> Such a solution would mix semantic and pragmatic factors; the commissive reading of statements about the future would then depend on the contextual availability of a deontic modal basis for *will*.

Let us admit, for the sake of the argument, that *will* receives a deontic reading in (4):

(4) I'll attend the meeting tomorrow.

In other words, (4) is true if, and only if, in every possible world that is consistent with the moral/deontological standards in force in the context of utterance, it is true that S attends the meeting the day after.<sup>vi</sup> Since the modal statement in (5) receives the same truth-conditions, the deontic commitment triggered by (4) would be derived at no cost.

(5) I ought to/must attend the meeting tomorrow.

However, if *will* is to be read as a deontic necessity in (4), it is difficult to find any reason why things should be different in (6):

(6) It may be the case/it is possible that I'll attend the meeting.

Indeed, there are contexts where epistemic possibility combines with deontic necessity in that order; for example, the deontic reading of *must* is perfectly acceptable (and, actually, is the only reading available) in (7):

(7) It may be the case/it is possible that you ought to / you must present the paper alone.

Yet the deontic reading is ruled out in (6), as shown by the unacceptability of (8):<sup>vii</sup>

(8) ? It may be the case/it is possible that I'll attend the meeting, I promise.

In fact, epistemic possibility bearing on the propositional content proves simply incompatible with the commissive illocutionary force, a consideration of great importance for the next sections.

(9) ? I promise to possibly attend the meeting.

### **3 The pragmatics of felicitous predictions**

In order to explain why some predictions about the future are interpreted, by default, as commissive speech acts, we need to make two basic assumptions about predictions. According to the first one, any prediction that P will take place at  $i$  is true iff it is true that P does take place at  $i$ .<sup>viii</sup> According to the second assumption, the necessary and sufficient condition, for a prediction that P will take place at  $i$  to be felicitous, is that, at the utterance time, the probability for the proposition P to be true at  $i$  is greater than 0 [ $p(P) > 0$ ] with respect to the common ground. The common ground can be defined as the set of those possible worlds that are compatible with the set of informations S and A mutually accept, i.e. compatible with every proposition A knows that S knows that A knows, ..., that both A and S accept as true (Stalnaker 1978, 2002). Consequently, for a prediction to be successful, the truth of the predicted content must be compatible with at least one possible world belonging to the common ground.

It is perhaps worth emphasising that this does not mean that any prediction that A deems false is unsuccessful. First, if A but not S believes that P is impossible, then neither P nor  $\neg P$  will be mutually accepted, so that P, or for that matter  $\neg P$ , will be compatible with the common ground. Second, suppose that A believes at  $i-1$  that it is mutually accepted that  $\neg P$ , and suppose that S, who looks rational and serious to A, produces at  $i$  an utterance  $u$  that seems to be intended by S as a prediction that P; then, in order to interpret  $u$  as a successful prediction at  $i+1$ , A will have to modify his representation of the common ground so as to eliminate both the proposition that S accepts that  $\neg P$ , and  $\neg P$  itself. This interpretative process is nothing else than presupposition *accommodation* (cf. Lewis 1979; Stalnaker 2002), even if the relevant literature does not usually introduce it in such terms. Note that accommodation cannot take place in every case, for its likelihood depends on the number of modifications it entails in the remaining set of A's beliefs; accommodations that cause, in what A believes to be true, minimal changes only will occur more frequently than those that would provoke a large domino effect resulting in a massive belief revision.

To make all this more palatable, consider (10):

(10) Tomorrow, Mary will grow wings.

Imagine some ordinary context, where A believes that it is mutually accepted by A and S that a human being cannot grow wings. In such a context, (10) will be no successful prediction, for its propositional content is incompatible with the common ground. As a result, A will probably try to figure out some non-literal interpretation or will process (10) symbolically (in the sense of Sperber 1975, 1985). But A may also modify what he takes to be the common ground so as to make it include the proposition that S holds the belief (entirely false, in A's opinion) that some human beings can grow wings, and thus exclude the proposition that (S believes that) human beings cannot grow wings. Imagine that this accommodation takes place or that, prior to the utterance, the context is already such that it is mutually accepted that S believes that some human beings can grow wings: now (10) will be a successful prediction, for this time the propositional content is compatible with the common ground; so that the utterance will be interpreted at its face value. In other words, even if in this second context the content of S's prediction is irrational from A's point of view, the prediction itself remains successful, i.e. rational with respect to S's beliefs. By contrast, in the absence of any accommodation and any mutual belief that humans can grow wings, the prediction is irrational not only from A's point of view, but also with respect to what A believes about S's beliefs (on these matters, see Davidson, 2001: 21-42; 2004: 169-230).

To take a more realistic example, consider utterance (11) in the two following contexts.

(11) This tree will blossom next week.

In the first one, S utters (11) while pointing at a dead tree; in the second one, S points to a budding tree. Imagine, next, that in both cases the tree does not blossom the next week. Clearly, the propositional content of (11) is false in both contexts; but the discursive status of the utterance differs significantly. In case S pointed to a budding tree, she may be said to have made a rational prediction on the basis of the evidence available to her. In case S pointed to a dead tree, A can either (a) assume that S is short-sighted, that S does not know that dead trees cannot blossom or proceed to some similar modification of the common ground, or (b) consider that S was speaking non-literally; if A does not choose among these two options, then he will have no choice but to assume that S is irrational. In more technical terms, the utterance (11), produced at the interval  $i$  in a world  $w$ , is a felicitous prediction if there are at

least one world  $w'$  that belongs to what is the common ground at  $i$ , and an interval  $i'$  in  $w'$ , such that the tree blossoms at  $i'$ , even it happens that, in the real world, the tree does not blossom at  $i'$ .

Let  $W$  be the set of all those worlds that constitute the common ground, and  $W'$  be the set of all those worlds where the propositional content is true. If a prediction is not felicitous,  $W' \cap W = \emptyset$ ; thus  $W' \cap W \neq \emptyset$  for any felicitous prediction.

#### **4 Categorical predictions, intentions and promises**

It is generally agreed that a sincere commissive act implies that  $S$  intends to bring about the truth of the propositional content  $P$  (Searle 1969, 1975; Searle and Vanderveken 1985). According to another dominant view, intentions are formed against a set of beliefs with respect to which the satisfaction of that intention is certain (e.g. Anscombe 1957: 91-93; Davidson 2001: 83-102; Grice 2001: 9-10, 51-57, 101-105; for an empirical confirmation, see Malle and Knobe 2001).<sup>ix</sup> This is not to say that the belief that it is possible that  $\neg P$  is incompatible with the intention to bring about the truth of  $P$ ; but no such belief can belong to the epistemic set that grounds that intention. Therefore, if  $S$  is sincere in uttering (1), from her point of view  $p(P)=1$  with respect to a certain belief set  $\mathbf{E}$ . If  $S$  does not make explicit that  $\mathbf{E}$  is not equivalent to everything she believes to be true,  $p(P)$  will be taken to be 1 with respect to the sum of  $S$ 's beliefs: the unacceptability of (8) and (9) thus stems from the fact that a promise that  $P$  entails the expression of the intention that  $P$ . If  $S$  makes explicit, for instance by using a conditional as in (12), that she intends to bring about the truth of  $P$  with respect to some possible worlds only, then  $p(P)$  is still 1 with respect to these possible worlds — in (12), the possible worlds that are compatible with the antecedent. This is shown by the awkwardness of (13).

(12) If it does not rain, I'll bring you to the movies.

(13) ?I promise that if it does not rain, I'll possibly bring you to the movies. / ?If it does not rain, I promise I'll possibly bring you to the movies.

If  $A$  has no reason to think that  $S$  is not sincere, then by expressing her intention to bring about the truth of  $P$  at  $i$ ,  $S$  will induce in  $A$ 's mind the belief that  $P$  will take place at  $i$  in any possible world compatible with the epistemic set  $\mathbf{E}$  that grounded  $S$ 's intention. In case  $A$  does not entertain contradictory beliefs and has no reason to think that  $\mathbf{E}$  is false, this amounts to causing  $A$  to believe that  $S$  will bring about the truth of  $P$  in any possible world compatible

with what A deems to be true (A's doxastic set **D**) at the utterance time. And if S has no reason to believe that A entertains contradictory beliefs or believes that **E** is false, S may assume that she induced in A the belief that P will take place at *i* in any possible world compatible with what A thinks to be true at the utterance time.<sup>x</sup>

So far, so good. However, the speaker who makes a promise does not simply aim at communicating an intention of hers. As pointed out by Searle (2001: 193-200; also Jackendoff 1999: 75-76), a successful promise is a reason for S to bring about the truth of the propositional content at some time interval *i*, independently of S's effective desires at the utterance time or S's intentions at *i*. In other words, once I have promised to attend the meeting, I have a reason to do so even if, when the day comes, my cognitive environment has changed in such a way that I no longer have the desire and/or intention to attend.

Cooperative behaviour (for instance, our abstention from inducing false beliefs in other minds) is evolutionarily advantageous and forms part of our genetic heritage, because it helps us to reach long-term selfish gains, even when these are in competition with desire-dependent short-term selfish gains (e.g. Axelrod and Hamilton 1981; Cosmides and Tooby 1992; Kitcher 1993; Cummins 1996; Ridley 1996; Nesse 2001; Dennett 2003: 170ff). We have just seen that, given any assertion of S's about a future intentional action of hers, if A takes this assertion as sincere, he will believe that  $p(P)=1$ , unless some fact, unknown to S but known to A, blocks the transition from S's prediction to the belief that S's action will take place. This belief that  $p(P)=1$  may drastically influence A's decisions and actions. To be sure, any belief can be revised, but the revision of one's representation of the world always has a certain cost. Therefore, other things being equal, it is evolutionary advantageous for an individual to avoid interaction with those who repeatedly induced false representations about the future. It is important to realise that two distinct revision mechanisms are at stake here. First, A may understand at *i* that some things that both S and A believed (at the utterance time) to be true are actually false; in other words, the set **E** with respect to which S's intention was formed turns out to be incorrect. Second, A may have to revise her belief that S will see to it that P. In some circumstances, the former revision will entail the latter, because the new set of those propositions A believes to be true is incompatible with S's bringing about the truth of P. But this is not always the case: S and A may have been wrong in their estimate of the world at the utterance time without A's new estimate being incompatible with P. So, it is definitely less costly to proceed to the revision about what is currently true while sticking to the belief that S will make P true.

Since it is clearly disadvantageous to be excluded from interaction, speakers will normally try to make true predictions, that is they will try not to bear any responsibility in the belief revisions carried out by their addressees. In the case of S's predictions about her future intentional actions, the truth of the content P depends on S; therefore, by issuing such a prediction, a cooperative speaker provides herself with a further, desire-independent reason to bring about the truth of P, even if the actual world, while remaining compatible with P, turns out to be incompatible with something she expected at the utterance time. This is why, typically, speakers perform promises by using linguistic devices that aim at showing that the propositional content will be verified (cf. Holtgraves 2005).

Of course, if the truth of P is independent of S's will (for instance, if it is independently obvious that P will be verified at *t*), no commitment is created, for the corresponding intention would be irrational, so that neither A (whose beliefs concerning P do not bear on S) nor a rational S could take the utterance as providing S with an additional reason to act. For instance, a criminal cannot promise to spend five years in jail after receiving a five-year sentence (unless it is mutually accepted that she could easily escape). It is also worth noting that S's mere expression of a rational intention of hers does not suffice to commit S to performing the future act in question. Consider the following exchange, borrowed from Alston (1991: 61):

S said "I'll write those letters tomorrow" in a context that clearly indicates that this would normally be taken as a promise. [A] replies [...] "What makes you think I care whether you write them or not?"

In this dialogue, A makes clear that S's utterance did not induce any modification in the set of A's beliefs that would be relevant for his action planning. Similarly, no commitment is created in the following exchange, where A makes it manifest that he does not believe that  $p(P)=1$  either because he thinks that S is not being sincere or because he thinks that S, though really entertaining the intention to return the book by tomorrow, grounds her intention on an incorrect epistemic support.

(14) S: I promise I'll return the book by tomorrow.

A: Come on, don't be ridiculous! Everyone knows that you're the most unreliable person in town. Given your reputation you are unable to enter in any commitment.

In connection with this last example, it is important to get back to the distinction, mentioned in Section 1, between defective and non-defective promises. It is crucial to distinguish the capacity to make promises from the capacity to make *successful* promises, whose result is to commit S to bring about the truth of the propositional content. A public, manifest inability to keep promises may affect the latter without affecting the former. S's utterance in (14) can be reported by (15).<sup>xi</sup>

(15) S promised to return the book by tomorrow.

To be sure, it can be claimed that by attempting to promise something in a context where it is mutually accepted that A does not believe that S can enter in any kind of commitment, S is attempting to change her social status and to provoke an accommodation effect that would make her promise plainly successful. But it is also the case that after A's answer, S can sensibly claim to be discharged of any obligation to return the book. The same kind of observation applies to Tsohatzidis's example of a speaker who mistakenly believes in Satan and promises to kill him. Tsohatzidis (2007) points out that, even after the speaker's mistaken belief has been corrected, her speech act can be reported as "S promised to kill Satan"; yet, in the same circumstances, an ascription of deontic commitment like "S ought to kill Satan" would be pragmatically infelicitous.<sup>xii</sup>

According to the account presented here, deontic commitment arises by default provided the utterance exerts the relevant influence on A's action planning by inducing the belief that the propositional content will be true. In a context where it is not mutually manifest that such an effect is impossible, the commissive interpretation will be favoured, even if one or both interlocutor(s) entertain(s) the *non-shared* belief that no deontic commitment can take place — for instance, because the promise cannot be kept, because the truth or the falsity of the propositional content has no impact for A's action planning,... In such cases, the promise will be defective but still successful, for nothing in the common ground disallows the commissive interpretation. By contrast, in a context where it is mutually manifest that S's utterance cannot have any effect on A's action planning, no literal commissive force will be assigned to the utterance — the commissive speech act will be plainly unsuccessful. For instance, in a context where it is mutually manifest that S and A know that Satan does not exist, (16) will receive a non-literal interpretation that prevents the report in (17) from being truthful and accurate.

- (16) I promise to kill Satan/ I will kill Satan.  
 (17) S promised to kill Satan.

To sum up, S's expression of a rational intention of hers generates S's commitment *ceteris paribus*, i.e. whenever it is not mutually manifest to S and A either that A does not believe that  $p(P)=1$  or that A does not believe that the truth of P does matter in a certain degree to him. This default inference remains optional and may be blocked or explicitly cancelled, as shown by (1) and (18):

- (1) I'll proof-read your paper for tomorrow. I can't promise, though. [repeated]  
 (18) I intend to come to your party. (But I can't promise.)

Example (18) is particularly interesting.<sup>xiii</sup> I think its analysis should run along Gricean lines. Since any sincere promise entails that S entertains the intention with the same content, but not vice-versa, S's use of "I intend" would be litotic, i.e. literally under-informative if (18) were to be attributed a commissive force. For this reasons, (18) generates an upper-bound implicature: the only interpretation compatible with the Principle of Cooperation is that S is not flouting the first maxim of quantity, and thus makes the strongest claim she can do without being insincere: the effect is to 'downgrade' (Marc Dominicy, p.c.) the utterance so as to reduce it to an assertive ace bearing on S's intention — which also indicates, at the perlocutionary level, that S is reluctant to make the corresponding promise.<sup>xiv</sup>

## **5 A short interlude: threats**

The analysis proposed easily applies to other commissive acts, like pledges, offers and acceptances. Yet the case may seem less clear with threats.

Verbrugge et al. (2004) report an experiment where participants were presented with conditional promises and conditional threats. In the low-credibility setting, the consequent was an excessive reward (promises) or an excessive punishment (threats) with respect to the condition in the antecedent, while in the high-credibility setting, the relationship between the reward or the punishment and the condition in the antecedent was proportional. Participants were then tested on two inference processes, Modus Ponens (MP) and Affirming the Consequent (AC). For instance, after being presented with the conditional threat (addressed by a teacher to a pupil) "If you chatter during the lessons, you will be sent out into the corridor", the participant was asked: "The pupil chattered during the lessons. Do you think

(s)he will be sent to the corridor?” (MP) or “The pupil has been sent out into the corridor. Do you think (s)he chattered during the lesson?” (AC). While, for both high and low-credibility settings, there were more inferences, viz. more positive answers, drawn from promises than from threats, this difference was considerably smaller for AC than for MP.

Contrary to what is claimed by Verbrugge *et al.*, this result does not show that the propensity to generate commitment is lesser for threats than for promises. Recall from the former section that (a) a commissive act is constituted by the expression of an intention, and (b) that an intention is formed against a certain set of propositions with respect to which its probability to be fulfilled is 1. The crucial difference between threats and promises is that threats are almost always used to perlocutionarily prevent an action of the audience; for that reason, the intention that underlies threats, but not promises, is almost always restricted to a very tiny set of possible worlds: threats are (almost) always implicitly or explicitly conditional on some action of A’s. However, as is revealed by the oddity of (19), the probability of the threat to be fulfilled has to be 1 with respect to the situation on which it is conditional.

- (19) ?John has threatened Mary to probably kill her if she doesn’t stop dating Paul.

As convincingly argued by Nesse (2001; see also Krebs and Dawkins 1984; Castelfranchi and Guerini 2007), issuing threats entails no less commitment than making promises, and fulfilling one’s threats is as important for one’s social reputation as is keeping one’s promises. Threats cannot be reduced to mere attempts to prevent or to induce an action of A’s, for if no commitment is generated, a conditional threat would not be a good means of persuasion/dissuasion:<sup>xv</sup>

Let us consider threats. Why after a failure [to persuade A] should [S] waste his resources in harming [A]? It is a manoeuvre suited to other purposes in the future. As a ‘lesson’ it is aimed at teaching [A] that [S]’s threats are credible, that [S] really has the power and will to act according to the threat — hence that he must be taken seriously. This can be further aimed at maintaining the reputation of [S] as a coherent and credible person [...] (Castelfranchi and Guerini 2007: 299)

This being said, in many cases where a threat is fulfilled, the resulting situation is harmful not only for the audience, but also for the speaker. Hence the benefit of fulfilling a

threat can be outweighed, from S's point of view, by the negative effects of such an action. In the experiment carried out by Verburgge *et al.* all threats were instances of what Castelfranchi and Guerini (2007) dub "conditional-influencing threat", where a threat aims at preventing an action of A's that is undesirable from S's point of view. The result that, in the MP setting, people tend to exhibit less expectation for the fulfilment of threats than they do for promises is certainly interesting; but the only thing it provides indisputable evidence for is that, in an experimental situation where no context is given, participants tend to assume that, once the persuasive aim of the threat has failed, honouring the commitment has more negative than positive effects from the speaker's point of view. By contrast, the smaller difference between promises and threats with respect to AC supports the hypothesis that threats generate as much commitment as promises: once the situation corresponding to the fulfilment of a given threat obtains, participants naturally assume that this is so because of the commitment created by performing that threat.

## 6 Evidence from language development studies

Previous sections suggest that promises — at least those that are not produced by the utterance of a performative sentence of the form "I promise to VP" — are constituted by a prediction whose outcome is an action X of S's such that, at the utterance time, S intends to perform X. Provided this action matters to a certain degree to A and A takes the epistemic grounds for S's intention to be correct, such a prediction will be believed to be true by A. In turn, provoking such a belief commits S to the future action in hand — her utterance creates a promise. As we shall see now, this analysis may receive some support from language development studies.

Bernicot and Laval (Bernicot and Laval 1996; Laval and Bernicot 1999; Laval 1999: 116-125) show that, when presented with a scenario where a speaker S makes a promise to some audience A, 3-year children tend to expect the promised outcome to obtain even in those cases where A is not concerned with the fulfilment of S's promise. This suggests that the children evaluate the fulfilment of a promise by taking into account the simple existence of S's action, i.e. by evaluating the truth-value of the propositional content. Bernicot and Laval's results confirm earlier findings by Astington (Astington 1988b), who observed that before the age of 7 children do not distinguish between assertions and promises: for "young children a promise is a true statement, whether it refers to a past or a future state of affairs" (Astington 1988b: 172). This is not surprising, given that before the age of 4 or 5, children do not discriminate between the satisfaction of an intention — which requires a causal link between

the agent and the event —, and the satisfaction of a desire — where no such causal constraint exists (Astington 1993, 2001). The model of interpretation developed above requires, for an utterance to be given a distinctively commissive interpretation, the capacity to view some future states of affairs as causally related to S's cognitive states at the utterance time; in other words, the capacity to attribute the commissive illocutionary force to utterances presupposes the capacity to conceive agents who entertain intentions as presenting causal dispositions to the satisfaction of their intentions.

When presented with unfulfilled promises and non-verified predictions, 7-year children seem to discriminate between those cases where S is responsible for the falsity of the propositional content, and those where she is not (Astington 1988b). However, both Bernicot and Laval's and Astington's studies (also Laval 1999: 116-125) show that, before 9 or 10, children do not evaluate the action with respect to S's commitment; which means that they fail to grasp the fact that the utterance creates a commitment of S's. By contrast, experimental studies with adult subjects have shown that the obligation contracted with respect to the truth of the propositional content is seen as the essential property of promises, be it in production or in reception (Gibbs and Delaney 1987). In fact, the mastery of the concept of deontic commitment does not seem to emerge before the age of 9 or 10; younger children, between 6 and 7, condemn the non-accomplishment of an announced action independently from the impact this has on the addressee: e.g. they do not draw any moral distinction between the situation where S and A agree to meet at the swimming pool, and the situation where S simply mentions that she's going to the swimming pool (Mant and Perner 1988). It is a reasonable guess that conceptualising deontic commitment is related to high-order metarepresentational skills. In order to interpret S's utterance not merely as committing S to the truth of the propositional content, but also as constituting a reason for S to bring about the truth of the propositional content, one needs to attribute to S (a) the belief that her utterance had an impact on A's cognitive environment; (b) the belief that A knows that S entertains the belief that her utterance had an impact on A's cognitive environment.

The two previous paragraphs strongly suggest that two abilities are needed in order to perform and interpret commissive speech acts. First, one should be able to distinguish between the predictions a speaker makes about her own actions, and the predictions she makes about states of affairs that are independent of her actions. Second, one has to be aware of the social commitment S endorses by providing other minds with certain beliefs about her future actions.

## Notes

\* I am grateful to Marc Dominicy for his penetrating comments on earlier versions of this paper. I also benefited from valuable criticisms and suggestions made by Savas Tsohatzidis, Marcelo Dascal and an anonymous referee. A preliminary version of this paper has been delivered at the 9<sup>th</sup> IPRA conference, at Riva del Garda; I am grateful to members of the audience for questions and remarks. My research is funded by a research fellow grant from the Fonds National de la Recherche Scientifique de la Communauté Française de Belgique (F.N.R.S.). The results presented below are also part of the research carried out within the scope of the ARC project 06/11-342 *Culturally modified organisms: "What it means to be human" in the age of culture*, funded by the Ministère de la Communauté française — Direction générale de l'Enseignement non obligatoire et de la Recherche scientifique.

<sup>i</sup> For a possible exception in Korean see Pak *et al.* (2005).

<sup>ii</sup> These institutions can be thought of as a widely shared set of representations that constrain the transmission of further representations within a given group (cf. Sperber 1996).

<sup>iii</sup> An anonymous referee objected that a pragmatic (as opposed to a semantic) theory of promises was already defended by Searle in *Speech Acts*. Quite on the contrary, Searle held — and keeps holding, cf. Searle (2007) — the view that the conventional, literal meaning of a sentence determines the illocutionary act the speaker literally performs when uttering this sentence (see also Recanati 2003; Dascal 2003: chapters 23-25). Cf. his “American soldier” objection against Grice’s account of meaning-*nn* (Searle 1969: 44-45), and his own revised account:

S utters the sentence *T* and means it (i.e., means it literally what he says) =

S utters *T* and

(a) *S* intends (*i*-I) the utterance *U* of *T* to produce in *H* the knowledge (recognition, awareness) that the states of affairs specified by (certain of) the rules of *T* obtain. (Call this effect the illocutionary effect, *IE*.)

(b) *S* intends *U* to produce *IE* by means of the recognition of *i*-I.

(c) *S* intends that *i*-I will be recognised in virtue of (by means of) *H*'s knowledge of (certain of) the rules governing (the elements of) *T*. (Searle 1969: 49-50)

It follows that Searle is compelled to claim that the only promises that are not indirect (i.e. that are literal) are those performed by means of the performative sentence “I (hereby) promise\_”. This view is explicitly endorsed by Searle and Vanderveken (1985: 60).

<sup>iv</sup> In fact, analyses of future tenses in modal terms face insuperable objections (Kissine forthcoming), but this point can be left aside in the present paper.

<sup>v</sup> After all, authors like Palmer (1987) assume that in examples like (i) *will* is a deontic modal (for a criticism of this view, see Kissine forthcoming).

(i) You will leave this room immediately.

<sup>vi</sup> I shall follow the standard practice in conceiving of a possible world as the set of propositions that are true in that world.

<sup>vii</sup> There is a non-commissive reading for (8), irrelevant here, where ‘I promise’ means ‘I affirm, I tell the truth’.

<sup>viii</sup> Some philosophers or linguists are likely to balk at these simple-minded truth-conditions. I have no place here for a proper discussion of the numerous issues involved, so I have to ask them to grant the point. For a defence of a semantics for future tenses along these lines, propounded by Abusch (1998), see Kissine (forthcoming).

<sup>ix</sup> This is so even if the agent cannot foresee the sequence of bodily movements necessary for the satisfaction of her intention: “We can be clear what it is what we intend to do while being in the dark as to the details, and therefore the pitfalls” (Davidson 2001: 94). In other words, those factors that determine the physical means to reach a goal are different from the factors that lead to the decision to reach that goal (see also Dretske 1988: 131-146; Dennett 2003: 237-240), which has the undeniable evolutionary advantage to allow the agents to keep a goal constant across internal or external variations (Talmy 2000: 277-279).

<sup>x</sup> Up to here, my analysis agrees with Armstrong (1971).

---

<sup>xi</sup> Many thanks to Savas Tsohatzidis for pressing me on that point.

<sup>xii</sup> In English, and probably in many other languages, speech act verbs prove singular in that they can be felicitously used for describing illocutionary behaviours that did not lead up to a successful performance of the attempted act. Contrast (i) and (ii):

(i) He married them, but he had no authority to do so.

(ii) ? He killed Mary, but she's still alive.

It is to be noted that examples like (ii) are acceptable in Chinese or in Japanese (e.g. Ikegami 1985; Tsujimura 2003), which suggests that the phenomenon exemplified by (i) depends on more general parameters, whose discussion falls far beyond the scope of this paper.

<sup>xiii</sup> Many thanks to Nausicaa Pouscoulous for drawing my attention to this kind of utterances.

<sup>xiv</sup> Perhaps the same explanation applies to (i), where "I think" causes the utterance to be interpreted as a mere expression of a belief and not as an assertion:

(i) I think that John was at the party. (But you shouldn't rely on me, I was drunk.)

<sup>xv</sup> Castelfranchi and Guerini point out that a conditional threat of the form *if P, Q* entails an implicit promise of the form *if Not-P, Not-Q* which is, of course, a case of "conditional perfection" (e.g. Geis and Zwicky 1971; Ducrot 1984: 13; Cornulier 1985: 83-84; van der Auwera 1997; Horn 2000). Conditional threats conform to the Aristotelian practical syllogism in (i):

- (i) If P, then Q  
 Not-Desirable Q  
 ∴ Not-P

Clearly, in order to (non-demonstratively) draw the conclusion *Not-P*, A must assume that the truth of Q is incompatible with the falsity of P, i.e. that Not-P is an effective means to prevent Q from being true.

---

## References

- Abusch, D. 1998. "Generalizing tense semantics for future contexts". In S. Rothstein, (ed), *Events and Grammar*. Dordrecht: Kluwer, 13-33.
- Alston, W.P. 1991. "Searle on illocutionary acts". In E. Lepore and R. Van Gulick, (eds), *Searle and His Critics*. Oxford: Blackwell, 57-80.
- Anscombe, G.E.M. 1957. *Intention*. Oxford: Blackwell.
- Armstrong, D.M. 1971. "Meaning and communication". *The Philosophical Review* 80(4): 427-447.
- Astington, J.W. 1988a. "Children's production of commissive speech acts". *Journal of Child Language* 15: 411-423.
- Astington, J.W. 1988b. "Children's understanding of the speech act of promising". *Journal of Child Language* 15: 157-173.
- Astington, J.W. 1993. *The Child's Discovery of the Mind*. Cambridge, Mass. Harvard University Press.
- Astington, J.W. 2001. "The paradox of intentions: assessing children's metarepresentational understanding". In B.F. Malle, L.J. Moses and D.A. Baldwin, (eds), *Intentions and Intentionality. Foundations of Social Cognition*. Cambridge, Mass.: MIT Press, 85-103.
- Axelrod, R. and Hamilton, W.D. 1981. "The evolution of cooperation". *Science* 211: 1390-1396.
- Bernicot, J. and Laval, V. 1996. "Promises in French children: Comprehension and metapragmatic knowledge". *Journal of Pragmatics* 25(1): 101-122.
- Castelfranchi, C. and Guerini, M. 2007. "Is it a promise or a threat?" *Pragmatics and Cognition* 15(2): 277-311.

- 
- Cornulier, B. de 1985. *Effets de Sens*. Paris: Minuit.
- Cosmides, L. and Tooby, J. 1992. "Cognitive adaptations for social exchange". In J.H. Barkow, L. Cosmides and J. Tooby, (eds), *The Adapted Mind. Evolutionary psychology and the generation of culture*. Oxford: Oxford University Press, 163-228.
- Croft, W. 1994. "Speech act classification, language typology and cognition". In S. Tsohatzidis, (ed), *Foundations of Speech Act Theory. Philosophical and Linguistic Perspectives*. London: Routledge, 460-477.
- Cummins, D.D. 1996. "Evidence for the innateness of deontic reasoning". *Mind and Language* 11(2): 160-190.
- Dascal, M. 2003. *Interpretation and Understanding*. Amsterdam: John Benjamins.
- Davidson, D. 2001. *Essays on Actions and Events*. Second edition. Oxford: Oxford University Press.
- Dennett, D.C. 2003. *Freedom Evolves*. London: Penguin.
- Dretske, F. 1988. *Explaining Behaviour. Reasons in a World of Causes*. Cambridge, Mass.: MIT Press.
- Ducrot, O. 1984. *Le Dire et le Dit*. Paris: Minuit.
- Enç, M. 1996. "Tense and modality". In S. Lappin, (ed), *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell, 345-358.
- Fleischman, S. 1982. *The Future in Thought and Language: Diachronic Evidence from Romance*. Cambridge: Cambridge University Press.
- Geis, M.L. and Zwicky, A.M. 1971. "On invited inferences". *Linguistic Inquiry* 2: 561-566.
- Gibbs, R.W. and Delaney, S. 1987. "Pragmatic factors in making and understanding promises". *Discourse Processes* 10: 107-126.
- Grice, H.P. 2001. *Aspects of Reason*. Oxford: Clarendon Press.

- 
- Harris, P.L. and Núñez, M. 1998. "Psychological and deontic concepts: separate domains or intimate connection?" *Mind and Language* 13(2): 153-170.
- Holtgraves, T. 2005. "The production and the perception of implicit performatives". *Journal of Pragmatics* 37(12): 20024-22043.
- Horn, L.R. 2000. "From *if* to *iff*: Conditional perfection as pragmatic strengthening". *Journal of Pragmatics* 32: 289-326.
- Ikegami, Y. 1985. "'Activity'-'Accomplishment'-'Achievement'—A language that can't say 'I burned it, but it didn't burn' and one that can". In A. Makkai and A.K. Melby, (eds), *Linguistics and Philosophy: Essays in Honor of Rulon S. Wells*. Amsterdam: John Benjamins, 265-304.
- Jackendoff, R. 1999. "The natural logic of rights and obligations". In R. Jackendoff, P. Bloom and K. Wynn, (eds), *Language, Logic and Concepts. Essays in the Memory of John Macnamara*. Cambridge, Mass.: MIT Press, 67-95.
- Kissine, M. forthcoming. "Why *will* is not a modal". *Natural Language Semantics*.
- Kitcher, P. 1993. "The evolution of human altruism". *Journal of Philosophy* 90(10): 497-516.
- Kratzer, A. 1991. "Modality". In A. von Stechow and D. Wunderlich, (eds), *Semantics: an international handbook of contemporary research*. Berlin: Walter de Gruyter, 639-650.
- Krebs, J.R. and Dawkins, R. 1984. "Animal signals: mind-reading and manipulation". In J.R. Krebs and N.B. Davies, (eds), *Behavioural Ecology. An Evolutionary Approach*. Second edition. Oxford: Blackwell, 380-402.
- Laval, V. 1999. *La promesse chez l'Enfant*. Paris: L'Harmattan.
- Laval, V. and Bernicot, J. 1999. "How French speaking children understand promises: The role of the future tense". *Journal of Psycholinguistic Research* 28(2): 179-195.

- 
- Lewis, D.K. 1975. "Adverbs of quantification". In E.L. Keenan, (ed), *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press, 3-15.
- Lewis, D.K. 1979. "Scorekeeping in a language game". *Journal of Philosophical Logic* 8: 339-359.
- Ludlow, P. 1999. *Semantics, Tense, and Time: An Essay in the Metaphysics of Natural Language*. Cambridge, Mass.: MIT Press.
- Malle, B.F. and Knobe, J. 2001. "The distinction between desire and intention: a folk-conceptual analysis". In B.F. Malle, L.J. Moses and D.A. Baldwin, (eds), *Intentions and Intentionality. Foundations of Social Cognition*. Cambridge, Mass.: MIT Press, 45-67.
- Mant, C.M. and Perner, J. 1988. "The child's understanding of commitment". *Developmental Psychology* 24(3): 343-351.
- Millikan, R.G. 2005. *Language: A Biological Model*. Oxford: Oxford University Press.
- Nesse, R.M. 2001. "Natural Selection and the Capacity for Subjective Commitment". In R.M. Nesse, (ed), *Evolution and the Capacity for Commitment*. New York: Russel Sage Foundation, 1-44.
- Pak, M., Portner, P. and Zanuttini, R. 2005. "What Korean promissives tell us about jussive clause types". *Paper presented at the "Colloque de syntaxe et sémantique", Paris*.
- Palmer, F.R. 1987. *Modality and the English Modals*. London: Longman.
- Recanati, F. 2003. "The limits of expressibility". In B. Smith, (ed), *John Searle*. Cambridge: Cambridge University Press, 189-213.
- Ridley, M. 1996. *The Origins of Virtue*. London: Viking.
- Sadock, J.M. and Zwicky, A.M. 1985. "Speech act distinctions in syntax". In T. Shopen, (ed), *Language Typology and Syntactic Description. Volume I: Clause Structure*. Cambridge: Cambridge University Press, 155-196.

- 
- Searle, J.R. 1969. *Speech acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Searle, J.R. 1975. "A taxonomy of illocutionary acts". In K. Gunderson, (ed), *Language, Mind and Knowledge*. Minneapolis: University of Minnesota Press, 344-369.
- Searle, J.R. 1995. *The construction of Social Reality*. New York: The Free Press.
- Searle, J.R. 2001. *Rationality in Action*. Cambridge, Mass.: MIT Press.
- Searle, J.R. 2007. "What is language: Some preliminary remarks". In I. Kecskes and L.R. Horn, (eds), *Explorations in Pragmatics. Linguistic, Cognitive and Intercultural Aspects*. Berlin: Mouton de Gruyter.
- Searle, J.R. and Vanderveken, D. 1985. *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.
- Smith, B. 2003. "John Searle: from speech acts to social reality". In B. Smith, (ed), *John Searle*. Cambridge: Cambridge University Press, 1-33.
- Sperber, D. 1975. *Rethinking Symbolism*. Cambridge: Cambridge University Press.
- Sperber, D. 1985. *On Anthropological Knowledge: Three Essays*. Cambridge: Cambridge University Press.
- Sperber, D. 1996. *Explaining Culture. A Naturalistic Approach*. Cambridge, Mass.: Blackwell.
- Sperber, D. and Wilson, D. 1995. *Relevance: Communication and Cognition*. Second edition. Oxford: Blackwell.
- Stalnaker, R.C. 1978. "Assertion". In P. Cole, (ed), *Syntax and Semantics 9*. New York: Academic Press, 315-322.
- Stalnaker, R.C. 2002. "Common ground". *Linguistics and Philosophy* 25(5-6): 701-721.
- Talmy, L. 2000. *Toward a Cognitive Semantics. Volume I: Concept Structuring Systems*. Cambridge, Mass. MIT Press.

- 
- Tsohatzidis, S. 2007. "Searle's derivation of promissory obligation". In S. Tsohatzidis, (ed), *Intentional Acts and Institutional Facts*. Dordrecht: Springer, 203-217.
- Tsujimura, N. 2003. "Event cancellation and telicity". *Japanese/Korean Linguistics* 12: 388-399.
- van der Auwera, J. 1997. "Pragmatics in the last quarter century: The case of conditional perfection". *Journal of Pragmatics* 27: 261-274.
- Vanderveken, D. 1990. *Meaning and Speech Acts: Volume I: Principles of Language Use*. Cambridge: Cambridge University Press.
- Vanderveken, D. 2005. "Success, satisfaction and truth in speech acts and formal semantics". In S. Davis and B.S. Gillon, (eds), *Semantics: A Reader*. Oxford: Oxford University Press, 710-734.
- Verbrugge, S., Dieussaert, K., Schaeken, W. and Van Belle, W. 2004. "Promise is debt, threat another matter: the effect of credibility on the interpretation of conditional promises and threats". *Canadian Journal of Experimental Psychology* 58(2): 106-112.
- Wilson, D. and Sperber, D. 1988. "Mood and the analysis of non-declarative sentences." In J. Dancy, J.M. Moravcsik and C.C.W. Taylor, (eds), *Human Agency, Language, Duty and Value. Philosophical Essays in Honour of J.O. Urmson*. Stanford: Stanford University Press, 77-101.
- Zaefferer, D. 2001. "Deconstructing a classical classification: a typological look at Searle's concept of illocution type". *Revue Internationale de Philosophie* 217: 209-225.