

A COMMENT ON THREATS AND COMMUNICATIVE RATIONALITY

Cristina CORREDOR*

Received: 1999.6.18

* Departamento de Filosofía, Universidad de Valladolid, Pza. del Campus s/n, 47011 Valladolid. E-mail: corredor@fyl.uva.es

BIBLID [0495-4548 (2001) 16: 40; p. 147-166]

ABSTRACT: The article studies two specific forms of social interaction, linguistically mediated: promises and threats. Two pregnant theoretical accounts are to be considered here. Firstly, the analysis propounded within the framework of Game Theory, assuming an intentionalist account of human agency and an instrumentalist concept of rationality; and secondly, the attempt carried out by Speech Acts theorists. In the first case, it can be shown that the theoretical premisses are insufficient to offer a proper account of such basic forms of social interchange. This result gives indirect support, so it is argued, to the second theoretical framework considered. Yet some of the solutions offered seem to be also unsatisfactory. Reasons are given of why it is so, and an attempt of solution, within the same theoretical framework, is suggested.

Keywords: Speech acts theory, Game theory, indirect speech acts, communicative rationality, threats, Formal pragmatics.

I

In a pair of illuminating articles, J. Heath has studied the shortcomings and difficulties of game theory for a satisfactory account of linguistically mediated interaction and, in particular, of communicative rationality. Taking a point of departure in the works by other specialists, he convincingly shows that "standard non-cooperative game theory is unable, in principle, to model speech acts, and is therefore unable to specify what is rational about linguistically mediated interactions"¹. In searching for a theory of social action, he considers game theory to be bound to model instrumental rationality, and defends a necessity to account for a non-instrumental, norm-governed form of rational action, which would be analizable as a generalization of J. Habermas' notion of communicative rationality. Nevertheless, in what seems to be a further development of this programmatic suggestion, he has tried to show that certain forms of linguistically mediated interaction, usually considered to be strategic manifestations of instrumental rationality (namely, threats), can be reconstructed in terms of this non-instrumental, norm-governed rationality.

THEORIA - Segunda Época
Vol. 16/1, 2001, 147-166

Aim of the present paper is to study the validity of the two thesis given support to in Heath (1995). Here the author shows, firstly, that threats, analysed as a linguistically mediated, coordinative interaction, cannot be adequately accounted for from within standard game theory and therefore are not instrumentally rational actions -in the precise technical sense of the theory. Secondly, he claims threats to be pragmatically dependent on norm-governed actions and thus susceptible of integration within the frame of the formal-pragmatic meaning theory developed by J. Habermas in the first formulation of his seminal theory of communicative action (hereafter *TkH*)². Although Heath's first thesis is a well-established result within the modified theoretic model he proposes, I see the second more problematic and difficult to sustain. The following reflection is an attempt to show why it is so.

II

A core intuition in Heath's reconstruction and propounded model is that, for a threat to be successful, the threatening action must be something the agent would rather prefer not to perform, namely a "mutually damaging action for both participants in the interaction". For, as he argues, only then could the hearer (here, the receiver of the threat) consider the threat rational and thus credible; that is, only then could the hearer consider reliable the agent's "promise", that he will not carry out his threat -hence maximizing, together with the hearer's, his own utility. For in such a case he receives, as a counterpart, another utility he prefers. But as a result, "in order to threaten credibly, players have to be willing to engage in non-utility maximizing actions" (Heath 1995, p. 232). This means that the agent's performing the threat would contradict the supposition that he is rational, given the instrumental concept embodied by standard game theory.

This very difficulty, in the general case of any cooperative interaction or interaction requiring coordination, has been made manifest by the proponents of rational choice theory themselves. Following J. Elster here, this theory assumes a principle of methodological individualism; and, in order to explain and justify behaviour in a choice situation, three elements are taken into account: the feasible set of all courses of actions (occasionally, with constraints); the causal structure of the situation, which determines which course of action leads to what outcomes; and a subjective ranking ("preferences") of the feasible alternatives, these numbers being called the *utilities* of the corresponding outcomes or consequences. Acting rationally means choosing the highest-ranked element on the feasible set; furthermore, the normatively proper decision cri-

terion is to choose that element which maximizes expected utility. The *expected utility* of an action is defined as the weighted average of the utilities that the action will yield under different states of the world, the weights being the probabilities of the states³.

In his paper, Heath recalls a further main distinction applying to rational-choice situations, namely that between *parametric* and *strategic* decisions. Whereas in a parametric decision the agent faces constraints that are in some sense external or parametric, strategic situations are characterized by interdependence of decisions. The latter are specifically studied by game theory, which considers rational choice in situations where the outcome depends on the choices of more than one agent and the agents are aware of one another's rationality. The choice takes place between two or more strategies. Moreover, games in which agents do not communicate with one another, *except through their actions in the game* itself, are termed *non-cooperative*. In strategic situations in general, an equilibrium point is a set of choices, one for each person involved in the interaction, that are optimal against each other. A *Nash equilibrium* is reached whenever each player's strategy maximizes his expected utility, given the other player's strategy. This requires *mutual consistency* and hence that each player holds true beliefs about what the other will do. Moreover, each player's choice of strategy is constrained by knowledge that the other is rational and so similarly constrained -according to the concept of rationality as individual utility maximization embodied by the theory⁴. As a result, premisses of game theory make it difficult to account for coordination or to explain elementary facts of social life, like commitments and, paradigmatically, promises. The assumption that communication exists leads the way either to an infinite regress of anticipations that undermine determining a maximizing strategy, or to the necessary postulate of a mutual knowledge of rationality or *perfect knowledge*, which is to the effect that communication becomes "cheap talk" and is in fact restricted to what is manifested through the performed actions themselves. This is to the effect that accounting for commitments evidences "unmistakeably a deficiency in the model", given that "covenants of mutual trust are not possible for rational players"⁵.

This conclusion, seemingly paradoxical, comes from the very concept of rationality into play. As Hollis/Sudgen show, the utility-maximizing strategy for the agent in the 'Promising Game', archotypically in the 'Prisoner's Dilemma Game', would be to say the words, thus inducing the hearer to perform, but then not to perform in return. But if agents generally acted in this way, the speaking of the words would not be a reliable indicator of future performance. For this reason "there is a deep-seated belief among game theorists

that words like 'I promise' are *cheap talk* and convey nothing" (Hollis/Sudgen 1993, pp. 15, 17-18, 25). As R.B. Myerson observes, this is due to the fact that non-cooperative game theory does not permit us to suppose that communication can determine the effective equilibrium actually played, because the meanings of all statements are supposed to be determined by the effective equilibrium itself; non-cooperative games with signalling and communication derive the meanings of all statements and signals from the equilibrium in which they are used⁶. In such contexts it has been proven that, whenever a pre-existing common language is taken into account whose meanings are not endogenously determined by the equilibrium in which they are used, the degree to which the players' interests coincide limits the equilibrium effectiveness of cheap talk⁷. Heath notices that, according to this view, a speech act turns out to be nothing more than a "move" in the game and any action transmits information by virtue of what the other player is able to infer about the speaker's intentions, on the basis of the move made. Hence communication cannot determine the equilibrium reached, because the *meaning* of all statements is determined endogenously by the equilibrium itself⁸.

To account for common phenomena of social life, specialists assume a necessity to extend the model, mainly with the purpose of integrating motivation -psychological, social (conventions and norms), or other- as a determining factor on the decision. Different procedures have been advanced to sever the conceptual gap between utility and choice, e.g. a concept of constrained maximization, in order to credit rational agents with the powers of strategic reflection. Nevertheless, in such cases parametric environments have to be abandoned, since it is no longer possible to justify the idea that players link subjective probabilities to the others' decisions. Within certain games there is a necessity to communicate the strategic reflection outcome to the others, what drives back to the problems of credibility and cheap talk mentioned. For these cases, *cooperative models* are proposed, in which strategies are defined as a set specifying, for each move in the game, an action and a message; thus a strategy profile would include both actions and messages. Here again, however, "introducing knowledge of language as a new *source* of beliefs" raises new problems, particularly because the cooperative framework fails to provide the standard solution equilibria for non-cooperative games and the models, once constructed, do not guarantee the existence of a solution (cf. Heath 1996, pp. 21-25).

Alternatively, so-called "evolutionary perspectives" have been applied to game theory, in an approach which views equilibria not as the consequences of ideal rationality, but as historically contingent conventions. This explanation

is considered to suggest that a Wittgensteinian perspective would be possible, when accounting for the function of *social norms* and *normative expectations*: to the extent that meaning assigned to behaviour is considered to be a matter of rule following in an institutional or social context, *interaction would be analytically prior to action*. Hence games as the 'Promising Game' would not be "context-neutral", and normative traits should be constitutive of the players' own conception of their strategies (cf. Hollis/Sudgen 1993, p. 31).

III

This very necessity of taking into account a normative background is present in Heath's analysis. He proposes a model for threats and promises following the Prisoner's Dilemma extensive form, but modifies it to face the challenge "to find a different game to model the situations in which promises are made" (Hollis/Sudgen 1993, p. 17). According to the author, this should result in "an account of threats as norm-governed actions" (Heath 1995, p. 225). On this reconstruction, threats differentiate themselves from promises because, in the former case, performing the commitment would be "mutually damaging" to both the threatened person and the agent, whereas in the latter the action "would only hurt" the agent (Heath 1995, p. 233). Furthermore, Heath assumes R. Selten's notion of *subgame-perfect equilibrium*, according to which not only should a reasonable equilibrium strategy profile be Nash in the large game, but the relevant subset of it should be also Nash in any proper subgame, even if this subgame is never reached in the course of the game.

Heath's extensive form game for threats would be the following⁹:

1	2	
	D	R
o	o	(0,0)
U	L	
(2,2)	(3,1)	

It is possible to reconstruct what the author represents here. To account for covenants of mutual trust -archetypically, in the 'Promising Game'-, a situation is figured where one party performs some service for the other, in return for a promise that the other will perform some service later -in Heath's represented situation, not to perform the threat¹⁰. Here, player 1 is the threatened person and player 2 is the agent; U is a situation where 1 performs what 2 requires

(e.g. a threatened bank clerk hands the money over to a bank robber); in D, on the contrary, 1 does not perform; then it is 2 who has to "move", performing (situation R) or not performing (situation L) the threat (e.g. shooting). The utilities for both players at each possible move are represented by the ordered pairs, where the first number is the utility for 1 and the second the utility for 2. According to Heath's reconstruction, the utility of not performing the threat is for the agent higher than that of performing it.

To this analysis it can be objected that in real situations, however, there is no perfect knowledge of the other's preferences, so that 1, the threatened person, cannot be certain as to which utility does 2, the agent, assign to L. It seems possible to figure a different model, alternative to Heath's, where the implausible assumption that performing the threat is *always* damaging for the agent be suspended -and thus where the theoretical postulate of perfect knowledge of rationality do not include this assumption¹¹.

In fact, Heath's final proposal is intended to be of wide import. For, from the double observation that "[t]here are some clear instances in which threats are directly norm-governed actions" -instances that, as it will be argued here, should perhaps be considered warnings- and that "the imperative contains an obvious reference to the normative background that licenses the action", together with the structural similarity in game theory between promises and threats, this author draws the conclusion that "[i]llegitimate threats are therefore parasitic upon legitimate threats, in exactly the same way that *insincere* promises are parasitic upon sincere ones". Therefore he considers justified to conclude that the basic position in his reconstruction is "recognizably Habermas" (Heath 1995, pp. 234, 239) [m.e., C.C.].

Nevertheless, the tacit nivelation here effected of two different validity criteria, that of sincerity (referred to the expressive domain) and that of legitimacy (belonging to the social domain of values and norms), seems to point out to some lack of concordance between Heath's proposal and the *TkH* framework. Hence a more detailed analysis is called for. From the study to be carried out it will result that, from a formal pragmatics perspective, a very important triple distinction is urged: that among, firstly, legitimate warnings (what Heath terms "legitimate threats, secondly paradigmatic regulative speech acts as promises or commitments, and finally *coercive* interchanges (on the present account, threats in a strict sense).

VI

Heath's remark that there is an insufficient distinction between parametric and strategic rationality within *TkH* is undoubtedly accurate, if these notions are

taken in strictly game-theoretical terms. The author aims at giving support to what seems to be Habermas' own assumption, namely, that instrumental rationality in the sense of the *TkH* -as opposed to communicative- effectively corresponds to the notion introduced by game theory. Yet what seems to follow from Heath's analysis is that they are two different notions, in part determined by the general theoretical frameworks to which they belong. On the one hand, as Heath recalls, "standard game theory models specifically exclude communication among partners" (Heath 1995, p. 235). On the other hand, *TkH* explicitly finds a point of departure in G.H. Mead's notion of symbolically (linguistically) mediated interaction; its reconstruction of the rational basis of speech underlying processes of human understanding is therefore grounded on this form of analytically prior social interaction.

According to the formal pragmatics embedded by *TkH*, what characterizes the nature of an illocutive act is the aspect under which the speaker claims validity, hence that susceptible of a critical positioning from the part of the hearer. In the case of what, following Heath's account, would represent a threat not conforming to norms (or a threat *simpliciter* on the present account), is this very possibility what results nullified in the interaction. This fact allows assigning to threats a different conceptual status from that conferred on those cases, like promises or warnings of sanction, where conformity to norms makes the illocution susceptible of criticism. This requisite of *dialogical fallibility*¹², which is what distinguishes communicative from non-communicative (strategic-instrumental) speech acts, is absent from Heath's considerations.

Habermas' original thesis asserted the conceptual preeminence of norm-conformed illocutions over any other uses of language, these latter being then termed non-communicative. This statement turned out to be problematic in the particular case of *manifestly strategic speech* acts -such as negotiations, threats and offerings of advantage, and even bare imperatives-, to the extent that some critical studies have suggested restricting formal pragmatics analysis to the communicative uses of language¹³. Inversely, in the case of *hiddently strategic speech acts* -paradigmatically, insincere promises-, their dependence on sincere, norm-conformed ones is considered well-founded and indisputable. Here the validity claim of sincerity, referred to the expressive dimension of speech, has in any case to be distinguished from the validity claim of correctness or legitimacy, this relative to the normative dimension where interpersonal relationships are established. Finally, a third validity claim of truth would be linked to the propositional component of the speech act, and would become susceptible of criticism in the paradigmatic case of an epistemic use of language -resting as a tacit presupposition in other cases. There-

fore it is a generalized notion of validity, in the sense of *rational acceptability*, that which comes to constitute the core of this pragmatist meaning theory. The strong underlying thesis: "man versteht einen illokutionären Akt, wenn man weiß, was ihn akzeptabel macht", asserts that an orientation to the possible validity of utterances belongs to the pragmatic conditions not only of human agreement, but also -and with priority- to those of language understanding in general.

In its most recent revision, however, the theory has made it explicit that not every language use is communicative, and that not every linguistic communication is oriented to an agreement based on validity claims intersubjectively recognised. Thus in an epistemic language use applied to the expression of objective world knowledge, and in a teleological use applied to calculations of success in human action, the orientation towards strictly illocutionary ends is not essential. In the former case, it suffices that the speaker knows the statement truth conditions and that the statement expresses his own belief on its truth when he utters it. In the latter, it suffices that the speaker knows the success conditions of his utterance, whenever this utterance is an intentional statement employed monologically and directed towards a plan of action; here, the success conditions are identical with the conditions under which the statement would become true. None of the two language uses needs an orientation towards an intersubjective recognisance, of the statement truth in the former case or of the intention seriousness in the latter, and are then termed *non-communicative* (cf. Habermas 1996, pp. 73-76). For, according to the revised theory, these language uses are possible through an abstraction process which suspends the "always given" reference of epistemic statements to truth and intentional statements to seriousness. Therefore no claim to rightness underlies the speaker's speech act. Since any demand to intersubjective recognisance is abstracted away, the speech act justificatory grounds become relative to the viewpoint of the speaker (*aktorrelative Gründe*) and are rational only relatively to him (cf. Habermas 1996, p. 77). This trait seems to be, in the revised version of the theory, what definitely distinguishes the non-communicative uses of language from the communicative one.

Thus any expression of will addressed by the speaker to the hearer reduces to a non-communicative language use whenever the virtual justificatory grounds giving support to the exigence are barely relative to the speaker. This happens, e.g., when the speaker resorts to sanctions or rewards, thus attributing these to the hearer as grounds-for-him, in order to have his exigence successfully performed. And it is easy to see that this would be the case of any coercive threat.

It is essential to this formal-pragmatic analysis to take account of the double structure of speech, which articulates an illocutionary component conveying the mode of communication and a propositional component conveying what is communicated. The validity claims offered with any speech act are internally connected with the illocutionary component. Correctness is thus referred to the performance of the act and its accord with a set of *constitutive* rules and *normative* presuppositions, susceptible of reflexive reconstruction. In the case of a promise, the *intrinsic* validity claim could result violated, e.g., whenever the agent is unable to accomplish the commitment, even if the promise was sincere. But this regulative speech act is also susceptible of criticism, depending on whether the promised action is morally, or legally, legitimate. This latter criticism could be considered, in principle, an *external* one, to the extent that it evaluates the (propositional) content of the regulative speech act, to be distinguished from the intrinsic correctness dimension relative to the performance of the act itself. Nevertheless, in consonance to the formal-pragmatic reconstruction, this legitimacy or moral rightness of the promised action -of what is stated in the propositional component- should be seen as a normative presupposition *constitutive* for the illocution, and internally connected to the validity claim intrinsic to the act. For only then would the promise be susceptible of criticism in the reflexive domain of *discourse*, in reason of the legitimacy of the promised action. And, in such a case, this condition should be included into the set of *essential conditions* in the formal-pragmatic sense. Yet the question arises whether or not all regulative speech acts, as commitments in general, advises, warnings, requirements, offerings, etc., are *constituted* by this normative presupposition concerning the legitimacy of the action stated in the propositional component.

This distinction, when applied to Heath's examples of "threats" -a policeman orders to stop under threat of shooting, a mother demands a correct behaviour under threat of punishment-, allows for a difference between the normative background which constitutes these speech acts, on the one hand, and the potential of reasons allowing for a justification of the demands themselves, on the other. In the policeman's case, the agent is not engaging in a personal commitment, or putting himself under a moral obligation with respect to the hearer, but making the hearer take notice of his professional assignment in the circumstances. It is against a specific institutional background that the policeman's speech act can thus take the illocutionary force it has -that of a warning, more than that of a *coercive* imperative or "threat" in the strict sense. Only so understood could the legitimacy of the embedded demand be justified or questioned. In the case of the second example, it is a social and cultural back-

ground of recognised roles and educative values and norms which turns the mother's warning in something different from a mere personal commitment and would eventually allow for a justification. In both cases, Austin's criterium to differentiate illocutions from perlocutions could be of help: the illocutionary mode of the speech act could be made explicit through the locution "I advise/warn you...", but more difficultly so with something like "I threaten you of...", or "I coerce you by..."¹⁴.

In *TkH*, Habermas accounted for the binding force of illocutions through the notion of *acceptability conditions*, which he termed also *illocutionary success conditions*. He subdivided it and spoke of applying the notion in a *broader* and a *narrower* sense (cf. Habermas 1986, pp. 359, 362-363). Acceptability conditions in the narrower sense would correspond, so do I think, to the conditions of illocutionary success in Searle/Vanderveken's last speech act theory¹⁵. They refer to the hearer's understanding of the speaker's utterance, and include the propositional and the pragmatical force component, the latter in terms of "speech act X counts as Y for the interlocutors". Acceptability conditions in the broader sense refer to the position taken by the hearer, i.e. to his accepting the offering embedded in the speech act and his recognizing its validity. They would correspond to conditions of non-defective performance as introduced by Searle/Vanderveken's theory. For both notions include the set of rules constitutive for the type of speech act considered; namely, preparatory, propositional content, and sincerity conditions. But, finally, speech act theory in Searle/Vanderveken's proposal includes a set of satisfaction conditions (of fit between words and world); and to these conditions does it belong the subsequent action in correspondence to the speech act -the accomplishment of an order, the fulfilment of a promise, etc. According to the formal pragmatics reconstruction, in contrast, any subsequent action different from the hearer's acceptance is seen as part of the illocution perlocutionary effect. Therefore the "hearer's acceptance" has to be understood as a critical "Yes" to the speaker's offering, grounded on the hearer's recognizing or accepting it as valid¹⁶. This positioning is not a mere change of "belief" -as it could be considered from a Gricean perspective, or from a D. Lewis' one, to which Heath seems to subscribe-, but it is rather to be seen as inseparable from the possibly explicit (linguistic) answer and its relevance for the subsequent interaction -even if this consists just of a prosecution of the understanding process, linguistically carried out.

Hence, in the original formal pragmatics theory, anything beyond the fulfilment of the acceptability conditions -i.e. of the *satisfaction conditions* in the *TkH* sense, or pragmatical understanding, together with the *validation condi-*

tions, or the hearer's recognizing and accepting the validity of the embedded claims- was conceptualized as *perlocutionary*. This provided us with a criterium to delimit strategic from communicative linguistic interaction. For, as emphasized above, instrumental rationality turns words into "cheap talk", and meaning emerges from actions effectively performed and their outcomes, thereof communication adds nothing within this game-theoretical framework. Still the last revision of the theory defends the dependence of any perlocutionary success on the manifest success of an illocutionary act. Although this dominance of illocutionary ends seems to disappear in the case of a particular language use, namely *perlocutions*, these do need the vehicle of illocutionary acts. And, according to the theory, any strategic language use works following the model of perlocutions; furthermore, threats belong in particular to this type (cf. Habermas 1996, pp. 82-84). From these theoretic statements, it seems legitime to conclude that the epistemic and expressive non-communicative uses of language have turned out to be a sort of "intermediate" categories between the communicative one, which claims rightness through a demand of intersubjective recognition, and perlocutions. It is relativity to the speaker's viewpoint what distinguishes in general non-communicative from communicative interaction linguistically mediated; and perlocutions are to be differentiated from other non-communicative uses by their exclusive perlocutionary aim, *versus* the virtual illocutionary object of epistemic or expressive language uses¹⁷.

What happens in the case of a threat, that is to say, of a coercive interaction? Heath proposes to see threats as a kind of commitment, close to a promise: "I promise you that, if you do *x* [hand the money over], I won't do *y* [shoot]". Adapting Searle's original analysis in his first theory of speech acts¹⁸, constitutive rules for *commitments* -here, "not to do *y*"- would include the following: that it not be obvious, both for hearer and speaker, that the latter will not (or would prefer not to) do *y* in the circumstances, and that the hearer does prefer the speaker's not performing *y*. Contrary to this, on Heath's account the particular commitment embedded in a threat equals that of a promise because the speaker would, in any circumstance, prefer not doing *y* and thus would not do *y* -this is what makes him "instrumentally irrational", whenever the threat is sincere and therefore the speaker sincerely intends to do *y*. From this theoretic viewpoint, an "irrational threat" would not satisfy one of the constitutive conditions for the type of speech acts to which it is supposed to belong. But what kind of reasons would allow the hearer critically take a position and answer with "No" to the speaker's "offering", in such a way that this answer be re-

spected and accepted by the speaker as an attitude entailed by his illocutionary obligation or commitment?

On Heath's analysis, an "illegitimate threat" is said to be similar to an insincere promise because, for the speech act to obtain perlocutionary success, the hearer has to be deceived by it: either he thinks the speech act to be a normatively correct illocution, or he thinks the speaker sincere though instrumentally irrational, when this one decides to stick to the commitment. Nevertheless, in contrast with this analysis and according to the sketched *TkH* framework, the formal-pragmatic reconstruction of communicative action differentiated, among illocutionary success conditions (i.e. acceptability conditions in the broader sense), the sincerity conditions from the essential ones. In the case of paradigmatic regulative speech acts, as argued above, the speaker's demand of recognition is relative to the correctness and legitimacy of his promise or commitment; but this takes place under the assumption that the sincerity rule is kept. This demand of recognition does rest on a potential of reasons that would allow the hearer to question it. In coercive imperatives, however, the rational basis of speech is displaced and substituted by a request of interchange to which no rational objection can be opposed. In conformity to the original formal pragmatics theory, an illegitimate commitment of any sort -the status claimed for "illegitimate threats"- should be compared to an illegitimate promise, which is not the same as an insincere one, the latter undoubtedly parasitic upon sincere.

This observation seems to be supported by the theory in its last revision. According to it (cf. Habermas 1996, pp. 83-84), the illocutionary act which consists of a specific warning of sanction becomes a threat when the speaker explicitly refers to the perlocutionary effect of intimidation or fear which he on purpose seeks to inflict on the hearer. Therefore, not only can the speech act be critically questioned [a] because of its lack of seriousness as an intentional claim, or [b] because of its falsity as a statement whose realisation is not objectively possible, but also [c] due to the speaker's incapacity to bring about the announced perlocutionary effect. In this last case [c], the speech act is questioned through its inefficacy. Only illocutionary acts are susceptible of being questioned in terms of their validity.

But the last revision of the theory has introduced also a new conceptual distinction, which would apply in the particular case of perlocutions characterised as threats; the distinction depends on the type of contexts within which they appear. Namely, the subclass of perlocutions taking place in a *normative* context. This would be the case of reproaches, legal imputations, and moral reprobations; and also that of official warnings of sanction (cf. Habermas 1996,

p. 84). The latter would correspond to Heath's legitimate threats. It becomes necessary, in all these cases, to make appeal to a normative background, which in the latter case should have been brought about by a legitimate political and juridical order, assumed to be based in an original consensus or social contract and seen in its turn legitimizing for the sanctioning norms invoked.

Nevertheless, it is important to notice here that "normative threats of sanction" are thematised by the theory as perlocutions that, only "secondary", can turn out to be embodied in a normative context. Contrary to this, and as stated before, in strategic contexts of action language exclusively works following the model of perlocutions. Here communication is subordinated to end-oriented action, and interlocutors try to influence each other in benefit of their own plans of action. Illocutionary ends become pertinent solely as conditions of perlocutionary success. As a result, speech acts are devoid of their illocutionary force; the presuppositions of communicative action are left suspended, together with the agent's orientation towards validity claims aiming at intersubjective recognition -claims thus susceptible of discursive resolution. Communication turns out to be indirect and parasitic with respect to a shared language knowledge. Each interlocutor has to suppose that the other decides rationally, and has to infer from individual convictions and his own preferences and designed ends (cf. Habermas 1996, p. 85). This allows us to conclude, contrary to Heath, that threats so considered in normative contexts do not result in a breakening of the internal link between rightness of the propositional content component and correctness of the illocutionary one.

In its turn, a regulative speech act, paradigmatically a promise, is questionable whenever some of its constitutive rules -in Searle's original sense- are violated. This happens, e.g., when the speaker is not in situation of realising his commitment, or the hearer does not desire what is offered to him. But another reason to question it is that the promised action be morally reproachable, or that it be so the action demanded from the hearer. In such cases, it is not the formal or procedural correctness of the illocutionary act what fails, but the legitimacy of what the propositional content states and therefore the moral character of the tie or obligation that the illocution would constitute between the interlocutors. As defended above, there seems to be an internal link between the formal correctness of regulative speech acts -whereof the interlocutors put themselves under an obligation or establish a moral tie between them- and the legitimacy of their propositional contents. Finally, there is another case in which a promise or commitment is violated, as noticed by J. Bohmann: "A promise, say of equality, is not ideological when it is simply violated, as

much as when it is left standing and yet does not bind those with power in their subsequent interaction"¹⁹.

The previous observations give support to the significance of making explicit and keeping in mind the triple distinction indicated above. Firstly, "legitimate threats" in Heath's sense receive their legitimacy from that of an institutional and social background of laws and institutionalised values and norms, but not directly from the rational basis of speech and a moral, interpersonal tie established on it. This allows concluding that they are cases of *legitimate warnings*. Secondly, in contrast to these legitimate warnings, whoever utters either a *coercive* imperative, or a *threat* in the strict sense on the present account, means his speech act not as an illocutionary demand of recognition, to be based on the interlocutor's rational acceptance of the illocution -that is to say, based on the potential of reasons that make it correct and legitimate. He rather pretends that the coercion perlocutionary prompts from the hearer the performance of an action. And thirdly, in the case of *regulative speech acts* in the strict sense, the speaker's demand for the interlocutor's rational acceptance is constitutive of the illocution itself.

In this last case, furthermore, the critical acceptance or positioning from the part of the hearer is not to be considered either, as noted above, an induced belief. Yet this seems to underly Heath's reconstruction, when he argues that, since "*in principle* beliefs cannot be outcomes, and instrumentally rational action is by definition directed toward outcomes", one can conclude that then "actions aimed at producing beliefs cannot be modelled game-theoretically". From that he draws his final proposal: "An action that has another's *belief* as its goal (...) is an understanding-oriented action" (Heath 1995, pp. 236, 238). Bohmann's quotation is here intended to make it evident that, in *TkH*, beliefs are taken into account because they are relevant for the subsequent interaction: they cannot be severed from actions, for it is this very fact what denounces their status of distorted communication.

Carried to its limit, Heath's argument is intended to reach the conclusion that any *linguistic*, or *linguistically mediated* interaction whatsoever cannot possibly be a manifestly strategic action, where the agent has an exclusively individual goal and aims at an outcome. For any communication process would have the structure of a coordination problem and would presuppose a common set of standardized intracommunicative objectives. Nevertheless, a different possible conclusion is that this "weak" coordination, based on a set of standardized structural elements, is to be referred back to what constitutes illocutionary success or acceptability in the narrower sense (understanding). These structural elements do not include necessarily "the [counterfactual] assumption

that agents always operate with a mutual ascribed interest in open and honest communication", for this very regulative presupposition is contradicted by a coercive exchange, where no demand of intersubjective recognisance underlies communication. Acceptability in the broader sense depends on the satisfaction of conditions different from mere sincerity from the part of the speaker; and the normative background from which regulative speech acts obtain their correctness and legitimacy is not to be identified with factual social institutions, but rather with formal, procedimental rules clearly violated in the case of a coercive imperative or threat.

Here, R. Alexy's set of constitutive rules for practical discourse, arguably presupposed in any other form of *communicative* interaction, can be appelled to in order to realise that even in the case of "sincere threats", such rules result necessarily violated. In particular, the *rational rules of justifiability* express "demands for equality, universality, and *lack of constraint*" by means of a *general rule of justification*: "Every speaker must justify what he or she asserts upon request, unless he or she can provide grounds which justify avoiding giving a justification". This general statement is further developed through a set of rules that correspond to the conditions proposed in *TkH* for the "ideal speech situation"; to the extent that they define "the most typical preconditions for discourse theory's concept of rationality", are termed by Alexy *rules of reason*. The main point here is that whoever justifies something pretends to accept the other person as equal partner in justification and "neither to exercise constraint nor to support constraint exercised by others"²⁰. It is difficult to see how, from an interaction based in a coercive imperative or a threat, the step to practical discourse could be taken.

V

The conclusion to be drawn is that the "normative background" invoqued for "legitimate threats" is not to be seen as inherent or constitutive of their illocutionary force as such. On the one hand, as argued above, Heath's examples should be seen more as warnings than as coercive threats. When a reconstruction is searched on *how* these speech acts *count* for speaker and hearer, the normative background that permits legitimizing the speech act includes a factual, extra-linguistic social institution.

On the other hand, either a coercive imperative or a threat in the proper sense could be considered successfully (instrumentally) performed, in the narrower sense, whenever the general context conditions are of a kind that they impose an unequal distribution of forces, thus making the hypothesised action

stated by the speaker possible: 'if those antecedent conditions are (not) satisfied / if such-and-such state of things does not take place [through the hearer's action], then this state of things will take place [through the agent's action]'. Here, the peculiar conditional structure of the speech act ("if (not)..., then...), which approaches it to constatives, is essential. The hearer's acceptance does not go back, nor does it depend on, a potential of reasons that would allow the speaker to justify his statement. Actions are counterfactually presented here objectualised (reified) and devoid of any moral dimension. Even if the hearer believes the speaker sincere and acting "irrationally" from the instrumental, game-theoretical viewpoint, the "binding force" of a coercion does not come from constitutive rules inherent to the rational basis of speech -as in the case of commitments that can legitimately be considered such-, but from extra-linguistic circumstances and conditions. The grammars of the concept of *coercion*, to speak along with Wittgenstein, does exclude an appeal to normative presuppositions or legitimizing rules.

All things considered, Heath's theoretical option seems to resort to a "weakening" of the notion of *understanding-oriented action*, to the effect that only acceptability conditions in the narrower sense -and the correlative notion of illocutionary success- are taken into account. Hence, the notion turns out to include actions typified in Habermas' original theory as instrumental, and in the last revision of the theory as non-communicative.

Yet the underlying difficulty here could be deeper. The conceptual pair instrumental/communicative in *TkH* goes back to the Frankfurter Schule and the distinction between an *objective* and a *subjective* (or *instrumental*) rationality, and to M. Weber's theory of rationality. Habermas seems to have been confident on the following: that strategic-instrumental action in this sense might be analysable in terms of rational choice theory, thus tacitly assuming that the theoretical notion of *instrumental rationality* embedded here would coincide with his own. Nevertheless, it could happen that they be, in fact, different notions, only partially overlapping and to be necessarily situated within their respective conceptual frames. Nevertheless, the last revision of the theory allows for a conceptual distinction that restricts strategic rationality to perlocutions and not to every non-communicative language uses. It seems possible then a new effort to integrate game-theoretic analysis within this restricted domain.

Notes

- ¹ Heath: 'Is language a game?', *Canadian Journal of Philosophy* 26/1, 1-28, here p. 1. The other work referred to is Heath: 1995, 'Threats, promises and communicative action', *European Journal of Philosophy* 3/3, 225-41. The former article, though later published, contains the general theoretical elements presupposed in Heath (1995) and is therefore to be seen as a previous development.
- ² Particular reference is here made to Habermas, J.: 1981, *Theorie des kommunikativen Handelns*, 2 vols., Frankfurt, vol. 1, cap. 3, "Erste Zwischen Betrachtung: Soziales Handeln, Zweckrätigkeit und Kommunikation"; and id.: 1992, *Nachmetaphysisches Denken*, Frankfurt, cap. 2, "Pragmatische Wende". -A last development concerning aspects of particular relevance for the present discussion is to be found in Habermas: 1996, 'Sprechakttheoretische Erläuterungen zum Begriff der kommunikativen Rationalität', *Zeitschrift für philosophische Forschung* 50, 1/2, 65-91.
- ³ Cf. Elster, J.: 1986, 'Introduction', in id. (ed): *Rational Choice*, Oxford, pp. 1-33; Sudgen, R.: 1991, 'Rational choice: a survey from economics and philosophy', *The Economic Journal* 101, 751-85; Hollis, M., Sugden, R.: 1993, 'Rationality in action', *Mind* 102/405, 1-35.
- ⁴ Cf. Nash, J.: 1951, 'Noncooperative games', *Annals of Mathematics* 54, 279-95. Hollis/Sudgen observe that then the equilibrium condition is stronger than the common-knowledge-of-rationality condition in the sense of D. Lewis (Lewis: 1969, *Convention*, Cambridge, Mass., p. 56; ref. in Hollis/Sudgen 1993, p. 8).
- ⁵ Hollis/Sudgen (1993, p. 25); cf. also pp. 17, 21-26, and Sudgen (1991, pp. 752, 769, 775).
- ⁶ Cf. Myerson: 1989, 'Credible negotiation statements and coherent plans', *Journal of Economic Theory* 48, 264-303, here p. 265.
- ⁷ Cf. Heath (1995, p. 237), and id. (1996, p. 18); he refers to results by J. Farrell and V.P. Crawford/J. Sobel.
- ⁸ Cf. Heath (1996, pp. 4, 10). -In order to account for the incidence of information exchange in equilibria, some alternative models have been advanced from which Heath gives an illuminating critical sketch; he shows how these models are bound to similar insufficiencies. Thus in *sequential equilibria* -equilibria consisting of a set of assessments and where the 'meaning' of sentences is inferable using Bayes's rule-, as well as in *games of incomplete information* -where the 'meaning' of the messages is the information content they reveal in the equilibrium-, the same intrinsic limitations of the non-cooperative model become manifest. (Cf. Heath 1996, pp. 12-18).
- ⁹ Heath (1995, p. 231); for an explicit explanation of the formal and theoretical elements into play cf. Heath (1996, pp. 10-12).
- ¹⁰ In *Leviathan*, ch. xiv, Hobbes gives the example of a prisoner of war who is released on the promise that he will pay a ransom. Cf. Hollis/sudgen (1993, p. 14).
- ¹¹ From an applied point of view, Heath's reconstruction seems slightly "ad hoc"; it would be more adequate to account for certain situations -e.g. a person who threatens his/her lover with committing suicide, in the case he/she is quitted- than others -e.g. the indiscriminate, non-conditional threat of expulsion or slaughter against Juives in Nazis' programs. -Two possible alternative models for threats would be the following:

(i)

1	D	2	
			R
o		o	(0,0)
U		L	
(2,2)		(3,0)	

(ii)

1	D	2	
			R
o		o	(0,1)
U		L	
(2,2)		(3,0)	

The first model, (i), would reflect a situation where there is not a relevant equilibrium for all relevant subgames -thus Selten's postulate is not satisfied-, but the agent's decision would be prompted by some psychological (or another) motivation. Here the situation for 2, the agent, is one of indifference, consonant with the notion of rationality considered -and with reality too. The second model, (ii), would account for an agent's preference to accomplish the threat, on some psychological (or another) motivation, whenever the threatened person does not perform first. Here, the mutual damaging action would be suboptimal for the hearer.

¹² The term is due to M. Cooke; cf. id.: 1994, *Language and reason*, Cambridge, Mass. and London.

¹³ On the problem of manifestly strategic use of language, cf.: Apel, K.-O.: 1987, 'Sprachliche Bedeutung, Wahrheit und normative Gültigkeit', *Archivio di Filosofia* 55, 51-88; id.: 1992, 'Illokutionäre Bedeutung und normative Gültigkeit', *Protosoziologie* 2, 2-15; Habermas, J.: 1981, *ibid.*; id.: 1986, 'Entgegnung', in A. Honeth/H. Joas (eds): *Kommunikatives Handeln*, Frankfurt, pp. 327-405; id.: 1989, *ibid.*; Wellmer, A.: 1989, 'Was ist eine pragmatische Bedeutungstheorie?', in Honeth et al. (eds): *Zwischenbetrachtungen. Zum Prozeß der Aufklärung*, Frankfurt, pp. 318-370. On the last critical suggestion, cf. Cooke (1994, *ibid.*).

¹⁴ But it would be equally inappropriate something like "I commit myself to, or put myself under the moral obligation of... [shooting against you, or punishing you]". This last remark permits distinguishing between advices and warnings, on the one hand, and archetypical regulative acts as promises and commitments, on the other. In the first case, as Austin noticed, the speech act is close to constative illocutions; it takes the form of a *counterfactual statement* relative to a hypothesised state of things, to be actualised given certain antecedent conditions. What is called for is the hearer's recognition of the epistemic validity in the counterfactual prognosis -namely, that the prognosticated state of things is likely to take place. Moreover, in the first of the examples here considered, the particular institutional background that makes a warning norm-governed allows the hearer to critically consider if the corresponding demand or suggestion is rightly integrated in such an institutional context, or if the agent is legitimated to invoke his membership to it. In such a

case, however, no personal commitment or tie is established between agent and hearer, but an institutional one. In contrast, archetypical regulative speech acts constitute a legitimate or moral, *personal* relationship between agent and hearer, which is to be differentiated in its contents from the *procedural* correctness concerning the speech act itself -in the sense previously pointed to.

- 15 Cf. Searle/Vanderveken: 1985, *Foundations of illocutionary logic*, 2 vols., Cambridge; Vanderveken: 1996, 'Illocutionary forces', in M. Dascal et al. (eds): *Sprachphilosophie/Philosophy of Language/La Philosophie du Langage*, Berlin and New York, vol. 2, pp. 1359-1371.
- 16 Originally in his (1981), Habermas spoke uniquely of "acceptability conditions in the narrower sense", without any reference whatsoever to the correlative broader notion, which is strongly suggested by his expression mode. He nevertheless distinguished two components in the former notion: *satisfaction conditions*, and *validation conditions* (cf. Habermas 1981, vol. 1, pp. 402-403, 406). -The explicitation of these two notions can be interpreted in terms of the pragmatic distinction existent between the hearer's understanding of the speech act, including the validity claims the speaker could offer to support by reasons if required, and the hearer's acceptance of the speech act, rationally motivated, this latter without requiring that the warranty be effected -or once the speaker has effected it. Hence it seems justified to assume that the original distinction between satisfaction and validation conditions in *TkH* does coincide with that between acceptability conditions (or illocutionary success conditions) in the narrower and the broader sense, resp., as later introduced (in Habermas 1986, op. cit.). In what follows this conceptual correlation will be assumed to be correct.
- 17 This statement arises the question of when epistemic or expressive uses of language could be considered communicative, i.e., as demanding intersubjective recognition to a claim to rightness. It seems that only in a *discursive* domain would it be so. But then no pregnant difference appears to persist between the critical examination of one or another of the two components which the speech act consists of, namely the illocutionary and the propositional.
- 18 In particular, for promises. Cf. Searle: 1969, *Speech acts*, Cambridge, 1992, pp. 58-59. Searle makes the interesting point that, because "I promise" is among the strongest illocutionary force indicating devices for *commitment* provided by English language -and undoubtedly so in many other natural languages as well. This would explain that the expression is used in the performance of speech acts that are not strictly speaking promises.
- 19 Bohmann: 1992, 'Critique of ideologies', in M. Dascal et al. (eds): *Sprachphilosophie/Philosophy of Language/La Philosophie du Langage*, Berlin and New York, pp. 689-704, here p. 700. -Bohmann makes this reflection in his searching for a "systematic account of the pragmatic mechanisms by which relations of domination and power are constituted and maintained through 'distortions' of the basic structure of speech acts and reciprocal communicative interaction" (ibid., p. 699). Cf. also Bohmann: 1986, 'Formal pragmatics and social criticism', *Philosophy and Social Criticism* 12, 332-352.
- 20 Cf. Alexy: 1990, 'A theory of practical discourse', in S. Benhabib/F. Dallmayr (eds.), *The communicative ethics controversy*, Cambridge, Mass., pp. 151-190, here pp. 165-67. Cf. also Alexy: 1978, *Theorie der juristischen Argumentation*, Frankfurt 1983.

Cristina Corredor is University teacher at the Departamento de Filosofía of the Universidad de Valladolid, where she teaches Philosophy of language and a doctoral course on Argumentation theory. She is a member of the SLMFC and of the SEFA. Her initial interest and research dealt with Formal logic, and afterwards her work has focused on Philosophy of language. She has published several papers on these matters in national and international journals, and it is going shortly to appear the work *Filosofía del lenguaje: una aproximación a las teorías del significado del siglo XX* (Madrid, Visor).