Relevance and Modality

José Luis Berbeira Gardón
University of Cádiz

ABSTRACT

This paper sets out to reasess some relevance-theoretic analyses of the English modals, in particular Walton (1988), Groefsema (1995), Klinge (1993) and Berbeira Gardón (1996a), and show how a unified treatment of modal verbs is possible within the framework of Sperber and Wilson's Relevance Theory. Before going through each of these proposals one by one, a summary is offered of the main theoretical assumptions common to all of them.

In the last two sections, it is argued that a unified account of the English modals can be given if we assume that they encode that the state of affairs described in the proposition expressed is located in a potential world. This proposal draws heavily on Wilson and Sperber's (1988) account of non-declarative sentences.

1. Introduction

The literature on the modals can usually be divided into those which assign a meaning to the modals in isolation from a specific context of use (the so-called 'monosemantic' approach), and those which regard the meanings of the modals as being largely, if not entirely, dependent upon a specific context of use (the 'polysemantic' approach). Both these approaches have their advocates (cf. Ehrman, 1966; Kratzer, 1981; and Perkins, 1983; for the first view; Palmer, 1979; Coates, 1983; Leech, 1971; and Marino, 1973; for the second).[1] It is not a striking fact that in the past few years the monosemantic, or unitary meaning view, has been largely strengthened since the publication of a number of

relevance-theoretic analyses which, in chronological order, include the following titles: Walton (1988) *The Pragmatics of English Modal Verbs*, Walton (1991), "The semantics and pragmatics of CAN," Groefsema (1992) "*Can, may, must* and *should*: unitary semantics, diverse pragmatics," Berbeira Gardón (1993) *La Pragmática de los Verbos Modales Ingleses*, Klinge (1993) "The English modal auxiliaries: from lexical semantics to utterance interpretation," Groefsema (1995) "*Can, may, must* and *should*: a Relevance-theoretic account," Klinge (1995) "On the linguistic interpretation of contractual modalities," and Berbeira Gardón (1996a) *Los Verbos Modales Ingleses. Estudio Semántico-pragmático.*[2] This paper sets out to reassess some of these analyses (in particular Walton, 1988; Groefsema, 1995; Klinge, 1993; and Berbeira Gardón, 1996a), and show how a unified treatment of modal verbs is possible within the framework of Sperber and Wilson's Relevance Theory.[3]

Within Relevance Theory, people discuss the role of the modals from rather different angles, although there are as well some very important similarities among these accounts. What brings these works together is a number of background theoretical assumptions:

1. The modals are not polysemous but have a single unitary meaning, which gives rise to different interpretations depending on contextual factors. Dissatisfaction is shown towards their treatment in much of the preceding literature, where they are said to be ambiguous between different meanings, the ambiguity being a result of the interpretation of an utterance containing the modal in a particular context, without further explanation about how the resolution of this ambiguity is actually carried out in the interpretation process.[4] In the semantic level, the modals have a basic meaning common to all their instances, the distinction between different modalities (epistemic, deontic, dynamic) being a matter of utterance interpretation which depends on the contexts in which the modals occur, thus belonging to pragmatics.

2. Not any 'monosemantic' approach will do. According to Walton (1991: 325), previous attempts to specify a "core" or "basic" meaning are incomplete. Groefsema (1995: 60) states that the problem with earlier proposals of unitary meanings for the modals (Walton, 1988 included) is that our intuitions about these verbs meaning different things in different utterances have not been accounted for in any precise way.[5] For Klinge (1993: 318), "the recurring problem for linguistic analyses of the modals has been the lack of a principled account of *how* we arrive at an explicit interpretation of an utterance of a sentence containing a modal."

3. The different interpretations of the modals are one more instance of the underdetermination of the propositional content by the semantic input of sentences (what Kempson has called the 'underdeterminacy thesis'): the logical form of utterances is incomplete, and must be developed in a process of inferential enrichment which uses information from context in order to produce a propositional form from which propositions can be constructed.[6] The inferences so made are called explicatures. These are constructed in accordance with the principle of relevance, that is, the utterance must be processed in such a way that its contextual effects are greatest and compensate for the effort involved in

the processing. Relevance theory is then, as stated by Klinge (1993: 320), "a theory which allows an explicit place for inferential enrichment of linguistic semantic input before it can be mentally represented in a propositional form."

2. Walton's proposals (CAN, MAY, MUST, SHOULD, WILL)[7]

Walton (1988) sets out to give an account of the modals which is based on the interaction of linguistic semantics and pragmatics as advocated by relevance theory. He proposes that the modals have a unitary linguistic content and uses paraphrases to express the modals' propositional operator function:

MODAL	UNITARY MEANING	PROPOSITIONAL OPERATOR
CAN	potential	the potential exists that p
MAY	possibility	it is possible that p
MUST	necessity	it is necessary that p
SHOULD	expectation	it is expected that p
WILL	anticipation	it is anticipated that p

(where p is the proposition expressed by the sentence in which the relevant modal stands as first auxiliary in the verb phrase) (Walton, 1988: 50).

Walton says about his paraphrases that they are putative and that their usefulness will be judged by their ability to express the sense of each modal. He notes that "it may be objected that labels such as *anticipation* hide real semantic distinctions, and that the precise meaning of *anticipation* is far from clear," but "the explanatory power of these paraphrases will be shown to be semantically adequate in clearing the undergrowth of 'contextual meaning' which surrounds the modals, and in dealing with a wide range of issues in modal pragmatics" (1988: 51). The question is then how these paraphrases interact with assumptions in the context of utterance to yield the different interpretations of the modals.

As an example of CAN Walton gives:

(1)   He can recite a poem now.

This sentence can mean different things in different contexts. Walton gives as examples (among others):

(1.1)   A: Has Freddie learnt anything new in his English class recently?
         B: He can recite a poem now.
(1.2)   A: When are you finally going to let Freddie contribute to the party?
         B: He can recite a poem now.

Following Sperber and Wilson (1986), Walton argues that in interpreting B's answer the addressee first recovers the logical form of the utterance:

(1.3)   The potential exists that Freddie recite a poem now.

This logical form has to be enriched to yield a fully propositional form. In the first place reference has to be assigned to "he." Because of the preceding utterance, "Freddie" is the most accessible referent. However, this is not the only thing that has to be established. In (1.1), "a poem" could mean "a specific poem" or "any poem he is given to learn to recite," and "now" has to be interpreted as having present and future time reference. In (1.2), however, "a poem" must be interpreted as "any poem he chooses/knows/is capable of reciting," and "now" must be interpreted as "immediately/in the next few minutes." Walton says that given the combination of the linguistic content of CAN as *potential* together with the syntactic information that CAN as a modal operator constrains the scope of the predicate, the addressee can establish the nature of the potential expressed in (1.1) and (1.2). For (1.1) this means that given the context established by A's question, whether Freddie has learnt anything new, the only enrichment consistent with the principle of relevance is to assume that the potential expressed is that of "ability to do something." For (1.2) Walton says that in the context of A's question the interpretation of CAN as expressing "ability" does not achieve sufficient relevance, and that the only enrichment of CAN consistent with the principle of relevance is that the potential expresses "permission to recite a poem."

Walton's analysis, however, raises a few questions. In the first place, his account of (1.2) implies that the interpretation of CAN as expressing "ability" is tried first, before the "permission" interpretation. This account misses the fact that the principle of relevance does not warrant the selection of more than one interpretation for a single ostensive stimulus: the interpretation whose selection it warrants is the first one tested and found to be consistent with the principle. In the second place, if "permission" is a direct development of CAN's sense of *potential*, then the question arises of how this relate to the "permission" interpretation that we get from MAY's sense of *possibility*. According to Walton, both CAN's *potential* and MAY's *possibility* have as a direct development a "permission" interpretation, without there being any difference between them. However, an account of these modals should give an explanation of why MAY is felt to be more formal or polite than CAN on a "permission" interpretation (cf. Coates, 1983). Moreover, taking *permission* as a direct development of *possibility* (in the case of MAY) suggests that *possibility* is a more basic notion than *permission*, which is not supported by the data (cf. Klinge, 1993: 323).

Another problem with Walton's account is the fact that he does not account for the role of sentential context in the interpretation of modals. Walton (1988: 46) points out that "it is not the sentences with CAN or CAN itself which select the meanings of 'ability', 'possibility' and 'permission' or any other meaning of CAN. These are interpretations of the

modal sentence in a particular context." This implies that when we encounter a sentence such as (2):

(2) John can speak fourteen languages.

in isolation, we are not able to interpret it as (2'):

(2') John has the ability to speak fourteen languages.

However, when processed in isolation, we do interpret (2) as (2'). Although it is true that context plays a very important role in the interpretation of sentences containing modals, I agree with Klinge's (1993: 318) claim that sentential context is also a determining factor in the interpretation process, and requires an explicit place in the theoretical framework.

3. Groefsema's account (CAN, MAY, MUST and SHOULD)

Groefsema (1995) proposes unitary meanings which, she claims, are rich enough to account for the range of interpretations these modals can have, but which are specific enough to account for why they get these interpretations and not others.

She proposes that the basic meanings of CAN, MAY, MUST and SHOULD express relations between the proposition expressed by the rest of an utterance containing them and a set of background assumptions, while putting constraints on what sets of assumptions are recovered during the interpretation process. These basic meanings are characterized as follows:

CAN: p is compatible with the set of all propositions which have a bearing on p.
MAY: There is at least some set of propositions such that p is compatible with it.
MUST: p is entailed by the set of all propositions which have a bearing on p.
SHOULD: There is at least some set of propositions such that p is entailed by it.
(where p is the proposition expressed by the rest of the utterance). (1995: 62).

The notion of 'bearing' in the basic meanings of the modals is a technical one, which has been defined formally:

"A proposition P can have a bearing on a proposition Q in different ways:
P positively has a bearing on Q iff Q or -Q follows P, for example, the proposition 'It's Sunday afternoon and I'm working' (P) has a bearing on the proposition 'I'm working' (Q), because we can conclude Q from P. Or Q or -Q follows from a set of propositions X together with P (where X = {x_i, ..., x_n}), and Q or -Q does not follow from X alone, and Q or -Q does not follow from any set of propositions X' together with P (where X' refers to set X without the sentence x_i, for any i, $1 \leq i \leq n$). For example, the proposition 'I'm working' (R) has a bearing on the proposition 'I put a sign on my door' (S), because

we can conclude S from R together with the two propositions (the set X) 'If I'm working, I put a sign on my door'. S does not follow if we leave either proposition out of the set X (i.e. the set X'). If we can conclude Q in this way, we can say that P positively has a positive bearing on Q. If we can conclude -Q in this way we can say that P positively has a negative bearing on Q.

P negatively has a bearing on Q, iff Q or -Q follows from -P, or Q or -Q follows from X together with -P, and Q or -Q does not follow from X alone, and Q or -Q does not follow from any X' together with -P. If we can conclude Q in this way, we can say that P negatively has a positive bearing on Q. If we can conclude -Q in this way, we can say that P negatively has a negative bearing on Q." (Groefsema, 1995: 62).

In informal terms, this notion focuses the addressee's attention on all the evidence for the proposition expressed by the rest of the utterance. This evidence, she says, can be of whatever nature, epistemic or otherwise.

She then sets out to explain how the basic meanings proposed for CAN, MAY, MUST and SHOULD give rise to the wide variety of interpretations that utterances containing these modals can have.

As an example of *possibility* CAN she uses (3) in a context in which a group of workers building a house are discussing the schedule for the day:

(3) (a)   A: Who is doing what?
     (b)   B: The painters can paint the doors.

According to Groefsema, B's answer would yield the incomplete logical form:

     (c)   [$_p$ the painters paint the doors] is compatible with the set of all propositions which have a bearing on p.

Once this logical form has been inferentially enriched (reference assignment to "painters" and "doors" and assignment of the temporal reference "today"), it focuses the addressee's attention on all the propositions which have a bearing on p, such as that the painters have the ability to paint, that the doors are ready for painting, that painting the doors will not interfere with the other jobs that have to be done, that paint and brushes are available, etc. This then gives us the intuitive interpretation of (3) as expressing that it is *possible* for the painters to paint the doors today.

Groefsema illustrates the *ability* interpretation of CAN with the following example:

(4) She can speak fourteen languages.

In interpreting (4), the addressee first recovers the logical form of the utterance:

(4'   [$_p$ Ann speak fourteen languages] is compatible with the set of all propositions which have a bearing on p.

Groefsema argues that when we process the incomplete logical form (4'), the only proposition which has a bearing on the proposition expressed is the proposition that Ann has the ability to speak fourteen languages, because we know in general that if someone performs an action/activity, they have the ability to perform that action/activity, which means that someone performing an action/activity entails that that person has the ability to perform that action/activity. That's why (4) gives us an ability interpretation.

The intuition that CAN can express *permission* also is a direct development of the basic meaning ascribed to CAN by Groefsema. As an example of this kind of interpretation, she uses (5) in a context in which John is visiting Ann and Bruce for the first time. He pulls a packet of cigarettes and asks:

(5)   Can I smoke in here?

This yields the logical form in (5'):

(5') John is asking whether ($[_p$ John smoke in room (5) at time $t]$ is compatible with the set of all propositions which have a bearing on p).

According to Groefsema, (5') focuses the attention of Bruce and Ann on all the propositions which have a bearing on p, such as whether John has the ability to smoke, and whether he has cigarettes. However, since (5') involves the instantiation of the property of smoking at a particular time, an ability interpretation is not enough. Moreover, all three can see that John has a packet of cigarettes, so that the only evidence that will make the question relevant, is whether or not Bruce and Ann give John permission to smoke in their room. This means that although all the evidence supports the proposition, the only evidence that yields contextual effects is Bruce and Ann's permission, and therefore our intuition that (5) expresses permission.

Although I sympathize with Groefsema's analysis, her account does not seem precise enough to explain the contribution of modals to the interpretation of utterances containing them. The main problem with her four-way division is that the notions of "compatibility" (in the basic meanings of CAN and MAY) and "entailment" (in the meanings posited for MUST and SHOULD) are not easily distinguishable from the more traditional notions of "possibility" and "necessity." As pointed out by Kiefer (1987: 69):

> "The proposition p is epistemically necessary iff p is entailed by what the speaker knows about the world and the proposition p is epistemically possible iff p is compatible with what the speaker knows about the world. The speaker's knowledge is representable as a set of propositions."

In this respect, Klinge (1993: 319) argues that "the final, more acute, problem is that POSSIBILITY and/or NECESSITY as primitives in the description of the modals (...) will

clearly result in a defective description. (...) there is simply no obvious place for WILL and SHALL in a framework based on POSSIBILITY and NECESSITY."

The question is then how Groefsema would propose to go on with WILL without leaving the uncontroversial assumption that the semantic behaviour of modals is homogeneous. A possible way out would be to claim that there is a semantic field shared by the modals.[8] This solution will be pursued in the last two sections of this study.

4. Klinge's analysis (CAN, MAY, MUST, WILL, and SHALL)

Klinge also draws on Relevance Theory in his analysis of the modals CAN, MAY, MUST, WILL and SHALL. According to him, the linguistic semantic representation of a sentence is at the level of logical form. Through a process of inferential enrichment, the addressee can arrive at a fully propositional form, what he calls a SITUATION REPRESENTATION. He then distinguishes between the SITUATION REPRESENTATION arrived at on the basis of the linguistic semantics of the PROPOSITIONAL CONTENT of a sentence, and assumptions about the actual state of affairs it is taken to describe.

Klinge's concern is with the semantic contribution the modals make in the OPERATOR position.[9] He looks for a semantic field shared by the modals and adopts the terms POTENTIAL and POTENTIALITY to stand for that semantic field.[10]

The cognitive elements of POTENTIALITY are taken to be the following:

1. The object of the POTENTIAL is an assumption about a WORLD SITUATION;
2. The WORLD SITUATION is not verified to be the case;
3. One resolution of the POTENTIAL is that the WORLD SITUATION turns out to be the case;
4. One resolution of the POTENTIAL is that the WORLD SITUATION turns out not to be the case.

These elements are the basis on which Klinge defines POTENTIALITY. This essentially conceptual model is applied in order to make assumptions about any unverified WORLD SITUATION, and it can be applied to unverified WORLD SITUATIONS irrespective of whether they are temporally located in past time, present time or future time.

If we let MODAL stand for any modal, then we can use the following string as an illustration:

John MODAL buy Mary's ticket.

The MODAL represents POTENTIALITY. The first element of POTENTIALITY is that the SITUATION REPRESENTATION is not asserted to be true of a WORLD SITUATION (i.e., that the proposition expressed is not asserted to be a true state of affairs in the actual world). So when the sentence is uttered, it only signals that there is a

POTENTIAL that the SITUATION REPRESENTATION turns out to be a true description of a WORLD SITUATION. This POTENTIAL can become resolved in two ways: one way is that it turns out that the SITUATION REPRESENTATION is a true description of a WORLD SITUATION; another way is that it turns out that the SITUATION REPRESENTATION is not a true description of a WORLD SITUATION.

It is this potential correspondence between the SITUATION REPRESENTATION and a WORLD SITUATION that is the shared semantic field of the modals as OPERATORS. Klinge then goes through the modals one by one and shows how they share their common semantic field of POTENTIALITY and how each serves to produce a different constellation of potential correspondence between a SITUATION REPRESENTATION and the WORLD SITUATION it is taken to describe in a given context of utterance. Klinge, then, proposes that the modals contribute to the interpretation of an utterance by providing a relation between the proposition expressed and an actual state of affairs. It is in this respect that his analysis differs from Groefsema (1992, 1995) and Berbeira Gardón (1993, 1996a), who suggest that the modals contribute a relation between the proposition expressed and a set of assumptions.

He proposes the following meanings for the modals:

CAN: The SITUATION REPRESENTATION turns out to be a true description of a
          WORLD SITUATION and the SITUATION REPRESENTATION turns out not to be
          a true description of a WORLD SITUATION.

MAY: The SITUATION REPRESENTATION turns out to be a true description of a
          WORLD SITUATION or the SITUATION REPRESENTATION turns out not to be
          a true description of a WORLD SITUATION.

WILL: The SITUATION REPRESENTATION turns out to be a true description of a
          WORLD SITUATION.

SHALL: The SITUATION REPRESENTATION turns out to be a true description of a
          WORLD SITUATION.

MUST: The SITUATION REPRESENTATION does not turn out not to be a true description
          of a WORLD SITUATION.

Before proposing these meanings, however, Klinge outlines some theoretical assumptions about the notion of WORLD SITUATION. He draws on Relevance Theory and Durst-Andersen's (1986, 1992) Mental Grammar.[11]

The temporal index assigned to a SITUATION REPRESENTATION leads to a fundamental distinction in Klinge's model between what he calls WORLD-EVENTS and AGENT-EVENTS. The former are WORLD SITUATIONS that merely happen; AGENT-EVENTS are WORLD SITUATIONS that only happen if an agent chooses intentionally

to bring them about. In the case of AGENT-EVENTS, we assign to an agent the ability and the intention to bring about the WORLD SITUATION.

Klinge (1993) draws on this distinction between WORLD-EVENTS and AGENT-EVENTS to account for the differences between different modalities. When an utterance containing a modal is understood as what is traditionally called deontic modality, we interpret it in terms of a WORLD SITUATION that is an AGENT-EVENT not motivated by the agent himself, and when an utterance containing a modal is understood as what is traditionally called dynamic modality, we interpret it in terms of a WORLD SITUATION that is an AGENT-EVENT motivated by the agent himself: "In interpretation of utterances in terms of dynamic and deontic modality we add the SITUATION REPRESENTATION to our assumptions about agents in the world, and we focus on the intentional activity element of the AGENT-EVENT, because, by definition, wilful agents control the activity that brings about a WORLD SITUATION." In the case of epistemic modality, we add the SITUATION REPRESENTATION to our assumptions about the world at past or present time or about a future WORLD-EVENT with no agent control inferred.

The inferential process through which the addressee goes when faced with an utterance containing a modal auxiliary is represented as follows:

> "1. Is my assumption that the WORLD SITUATION described by the SITUATION REPRESENTATION is a past or present WORLD SITUATION, or that it turns out as a future WORLD-EVENT? If so, as a result of the modal operator, I will not add to my assumptions about the world that there is, or that there has been a WORLD SITUATION corresponding with the SITUATION REPRESENTATION, or, in the case of a future WORLD-EVENT, I will add to my assumptions about the world that the course of the world will determine whether a WORLD SITUATION occurs corresponding with the SITUATION REPRESENTATION.

> 2. Is my assumption that the WORLD SITUATION described by the SITUATION REPRESENTATION is a future WORLD SITUATION that only turns out if someone wilfully produces an activity to produce an AGENT-EVENT? If so, I will add to my assumptions about the world that the intentional activity of an agent in the world will determine whether a WORLD SITUATION occurs corresponding with the SITUATION REPRESENTATION. Is the WORLD SITUATION desired or not desired by the agent and is the AGENT-EVENT motivated or not motivated by the agent himself?" (Klinge, 1993: 329).

If the process of inference stops at 1, that is, if the WORLD SITUATION is taken to be a past or present WORLD SITUATION or a future WORLD-EVENT, the addressee will identify the utterance as epistemic, and add an unverified SITUATION REPRESENTATION to her assumptions about the world. If the preferred interpretation is 2, the utterance will be identified as dynamic or deontic: the WORLD SITUATION described is interpreted in terms of an activity by an AGENT with the intention to produce an AGENT-EVENT. In this case, the addressee will add the SITUATION REPRESENTATION to her assumptions about agents in the world.

The interpretation process can be illustrated with the following example:

(6) Yes, we can send you a map, if you wish.

Klinge analyses (6) in a situation in which the speaker has maps at his disposal which he sends as he chooses; however, he does not wish to impose on the addressee one of the outcomes of sending and not sending, and therefore he uses the sentence containing CAN, which is neutral, thus leaving it to the addressee to decide whether correspondence or non correspondence between the SITUATION REPRESENTATION and WORLD SITUATION turns out, because the addressee has the option of saying 'Yes, please' or 'No, thank you'. The contribution of CAN in this example is to communicate that the proposition expressed is not asserted to be an actual state of affairs, but rather a potential state of affairs. According to Klinge's proposal, CAN in this example signals that the SITUATION REPRESENTATION 'WE SEND YOU A MAP' turns out to be a true description of a WORLD SITUATION and that the SITUATION REPRESENTATION 'WE SEND YOU A MAP' turns out not to be a true description of a WORLD SITUATION.

However, this analysis does not explain examples like (7)

(7) Oh, you can leave me out, thank you very much.[12]

which are interpreted as conveying a command. This example would be interpreted in Klinge's model as follows:

> The SITUATION REPRESENTATION 'YOU CAN LEAVE ME OUT' turns out to be a true description of a WORLD SITUATION and the SITUATION REPRESENTATION 'YOU CAN LEAVE ME OUT' turns out not to be a true description of a WORLD SITUATION.

This implies that the addressee has a choice about whether he leaves the speaker out or not: according to this interpretation the speaker is merely not stopping the addressee from leaving him out. However, it is doubtful whether the addressee would interpret the utterance as offering him a choice. According to the scenario, the speaker is ordering the addressee to leave him out, and the addressee would interpret (7) as such. This example then expresses more than just a potential correspondence/non correspondence between a SITUATION REPRESENTATION and a WORLD SITUATION.

The view of the modals proposed by Klinge also runs into problems when trying to account for another puzzling phenomena, namely the interplay between modals and negation. Consider the following example:

(8) John cannot come to the party tonight because he's in France now.

which yields the logical form (8'):

(8') The SITUATION REPRESENTATION 'JOHN COME TO THE PARTY TONIGHT' turns out not to be a true description of a WORLD SITUATION and the SITUATION REPRESENTATION 'JOHN COME TO THE PARTY TONIGHT' does not turn out not to be a true description of a WORLD SITUATION.

The second resolution of the potential in this example turns to be counterintuitive, if not absurd.

These are only some of the problems which find no solution in Klinge's model.

5. The semantics of CAN, MAY, MUST, SHOULD and WILL

A number of people have proposed that modality can be best handled by a possible-worlds semantic framework, Kratzer's (1981) model being the most cited. However, a few problems arise when the analysis of modality involves the introduction of possible worlds into the framework. Such an approach would be unsatisfactory for different reasons. On the one hand, this kind of approach does not reflect psychological reality.[13] On the other, there are many pragmatic factors which come into play in the interpretation of modals, but which are generally ignored by the formal approaches. Kratzer (1981), for example, proposes basic meanings for the modals which are complemented by the phrase 'in view of'. Her theory includes a 'conversational background' which involves a set of assumptions against which the modal relation can be understood. She distinguishes different sorts of conversational backgrounds, such as an epistemic background and a deontic background. These can be exemplified in the uses of MUST in (8a-b) respectively:[14]

8a. You must be very careful, since you are able to paint such delicate pictures. ("In view of what I know, you must be very careful")
8b. You must be very careless so that we can scare the guests off once and for all. ("In view of what is necessary, you must be very careful")

However, as Groefsema (1995: 60) points out, "it turns to be as difficult to classify all the different possible conversational backgrounds as it is to classify all the possible interpretations of the modals." Another problem with Kratzer's theory is that the idea of a fixed conversational background which favours one interpretation over others presupposes the existence of a unique context fixed independently of the utterance. But, as Sperber and Wilson (1986) argue, no unique context is given in a discourse; rather, the context for understanding an utterance is part of the interpretation process.[15]

In the analysis I propose,[16] it is held that a satisfactory account of modals can be given within the spirit of Relevance Theory and without invoking the notion of possible worlds, hence dispensing with the necessity of relying on a classical framework. Instead, I use Wilson and Sperber's (1988) notion of 'potential worlds'. Potential worlds are defined in

Wilson and Sperber (1988: 85) as "worlds compatible with the individual's assumptions about the actual world which can be, or become, actual themselves."

The use of this notion in the analysis of the modals, without invoking possible worlds, seems to be have been supported very recently by Papafragou (1996: 188-189), who claims that interlocutors expect each other to remain within a relatively limited range of inter-world similarity:

> "(...) why should a rational being entertain alternative possibilities if they had absolutely nothing to do with the way the world actually is? The issue relates directly to the way representations of alternative realities achieve relevance: by introducing a set of separate assumptions (deontic, stereotype-based, or other) into a body of assumptions about the real world which is held maximally constant, humans are typically able to derive cognitive gains which they could not have produced by manipulating solely representations of actuality. So in most cases, in order to achieve relevance, representations of a given world should be descriptions of a world not simply possible but also potential, that is, similar enough to the actual world to be able to be, become or have been actual itself."

A potential world, according to Papafragou (1996: 189), is similar to the actual world both quantitatively and qualitatively. On the one hand, the two worlds have to share a number of propositions; on the other, the sort of propositions shared by both worlds has to be important: any departure from the current construal of the world demands considerable cognitive effort and therefore has to be offset by adequate cognitive effects: these effects represent a different picture of what the *actual* world could be (become/have been), if a set of conditions had been met. Apart from the cognitive side, there is also a communicative side to modality. When inviting the addressee to entertain a representation of an alternative state of affairs, the speaker should try to convey a representation which the addressee can reconstruct from the evidence available to him, in other words, from his own representation of actuality. Since the actual world is (by definition) the only world which is realised, it is going to be the one which furnishes a vast range of assumptions mutually manifest to the interlocutors; it would be communicatively infelicitous to represent all aspects of possibilia. If the addressee cannot see how to construct a representation of a possible world while preserving the structure of the actual world as much as possible, he will probably give up trying and refuse to process the utterance altogether.

In section 2, I pointed out that Groefsema's four-way division was not fully explanatory. In what follows, a solution will be attempted by introducing the notion of 'potential worlds' into the basic meanings posited by Groefsema. I will claim that a unified account of the modals can be given if we assume that they encode that the described state of affairs is located in a potential world. The basic meanings of the modals would then be as follows:

CAN: p is compatible with the set of all propositions which have a bearing on p, and the world type is potential.

MAY: There is at least some set of propositions such that p is compatible with it, and the world type is potential.

MUST: p is entailed by the set of all propositions which have a bearing on p, and the world type is potential.

SHOULD: There is at least some set of propositions such that p is entailed by it, and the world type is potential.

WILL: p is true, and the world type is potential.[17]

In the next section, I will outline how the different interpretations of the modals are derived, assuming that they semantically encode that the speaker regards the world represented as potential. My analysis draws heavily on Wilson and Sperber's (1988) account of non-declarative utterances.

6. The interpretation process

Following the semantic analysis proposed in the previous section, I will show that the information semantically encoded by CAN, MAY, MUST, SHOULD and WILL contributes to the construction of a higher-level explicature carried by the utterance and that the different interpretations of a modal utterance involve inferentially enriching the basic meaning of the modal auxiliary.

Sperber and Wilson draw a very important distinction between the implicatures and the explicatures of an utterance. The latter are developments of a logical form and are generated through a process of inferential enrichment which uses the immediately accessible contextual information in order to reach a complete propositional form capable of bearing truth conditions. The explicatures of an utterance have two properties: (1) they are explicitly communicated assumptions and (2) the speaker communicates them to the addressee so that she can entertain them as true. The propositional form of an utterance can also be embedded in a speech-act or attitudinal description in order to generate the higher-level explicatures of the utterance. These are conceptual representations capable of entailing and contradicting each other and representing determinate states of affairs: though true or false in their own right, they do not generally contribute to the truth conditions of the utterances which carry them.

So, on the explicit side of communication, an utterance can communicate different propositions: the propositional form of the utterance (roughly similar to what Grice calls 'what is said'), and one or more higher-level propositions, resulting from the embedding of the propositional form in a speech-act or attitudinal description.

According to the semantic analysis proposed in the last section, the modalized utterance in (10):

(10) You can wash the dishes for a change.

encodes that the state of affairs described by the proposition expressed, namely "You wash the dishes for a change," is located in a potential world. The presence of CAN also indicates that the occurrence of the state of affairs described is compatible with all the evidence the speaker has available. The output of linguistic decoding is then:

(10.1) [p You wash the dishes for a change] is compatible with the set of all propositions which have a bearing on p, and the world type is potential.

This is of course information about the speaker's attitude towards the proposition expressed but is much weaker than what the speaker intends to communicate. Intuitively, (10) is interpreted as a covert imperative.

On the Relevance Theory view, the schematic output of linguistic decoding is inferentially enriched to the point where it expresses the minimal proposition capable of giving rise to enough contextual effects for no unjustifiable effort. The processes of enrichment typically consist of the context dependent inferential processes of fixing referents, resolving ambiguities and specifying the contribution of vague terms in accordance with the criterion of consistency with the principle of relevance. So, for example, the pronoun "you" will be set a referent, namely the addressee, on the basis of the available contextual information and considerations of relevance. A referent will be also set for "dishes." A temporal index will also have to be assigned to the proposition (for example, in the next few minutes, after dinner, etc.), thus producing:

(10.2) [p Addressee washes the dishes after dinner for a change] is compatible with the set of all propositions which have a bearing on p, and the world type is potential.

However, (10.2) is still schematic. It cannot be what the speaker intended to communicate. In other words, it does not give rise to an adequate level of contextual effects in this context and therefore does not satisfy the assumption of optimal relevance. Thus, the addressee has good reasons to assume that this is not what the speaker intended to communicate and that more inferential work on his part is needed.

In order to interpret (10) the addressee, guided by the Principle of Relevance will form the hypothesis that what the speaker intends to communicate is not only the potentiality of the state of affairs described, but also that this state of affairs is desirable to the speaker, because according to the scenario, if the addressee washes the dishes, the speaker will feel better.[18] So the speaker is also communicating:

(10.3) [p Addressee washes the dishes after dinner for a change] is compatible with the set of all propositions which have a bearing on p, and the world type is potential and desirable.

In Relevance Theory terms, (10.3) is a higher-level explicature carried by the utterance in (10). According to the proposed analysis, the information semantically encoded by CAN contributes to the truth conditions of this higher-level explicature and hence to one of the thoughts communicated by the utterance but not to what is standardly taken to be the truth conditions of the proposition expressed by the utterance.

However, given that this example is interpreted as a covert imperative, the question arises of what the relation is between (10) and an imperative such as (11):

(11) Wash the dishes.

As Sperber and Wilson's (1988) analysis predicts, (11), as an imperative, encodes that the proposition expressed is entertained as a true description of a world that the speaker regards as both potential and desirable. According to our analysis, (10) encodes simply that the state of affairs described is located in a potential world; this then is enriched into "potential and desirable" as the result of inference imposed by the context and considerations of relevance. Since (10) is less explicit than (11) and therefore requires more processing effort, one would expect on the basis of the principle of relevance that it would yield some extra contextual effects. And indeed it does. (10), when compared to (11), is less direct, this overtone following naturally from the semantics of CAN. In the case of the modalized utterance, (10) semantically encodes the notion of potential world, and as a result of contextual enrichment, the notion of desirable world is inferred. The combination of these two types of world, the former semantically encoded and the latter pragmatically inferred makes the interpretation of (10) similar to that of the imperative.

I have just shown that Relevance Theory offers a means of accounting for the contribution of modals to the interpretation of utterances by demonstrating the operation of a single pragmatic principle. The basic meaning of a modal is one input to the process of utterance interpretation: it is the contribution of natural language semantics. Given that the semantic content of an utterance underdetermines the proposition expressed, the features of context which are necessary to determination of the proposition expressed are the pragmatic input to that process. From the basic meaning of the modals, contextual features and the Principle of Relevance, it has been shown that a broad array of interpretations of modal utterances is obtainable.[19]

7. Conclusions

These relevance-theoretic proposals have a clear theoretical interest with respect to the nature of the modals. If a satisfactory account can be given to deal with modality without invoking possible worlds, we have strengthened the basis of relevance theory as a viable alternative to traditionally conceived semantic frameworks. Furthermore, I also hope to have produced some insight into the division of labour between the various factors that play a role in the human linguistic capacity, in particular the extent to which context, pragmatic

knowledge and common sense reasoning can be assigned responsibility in modelling the understanding of modality.

Some areas for further research remain, such as indirectness, politeness, the interplay between modals and negation, and the semantic relationship between the English modals and their 'periphrastic' counterparts (MUST/HAVE (GOT) TO; WILL/BE GOING TO; CAN/BE ABLE TO, etc.). We have clearly not solved all these problems. I believe, however, that we have been able to use the insights of Relevance Theory to cast a little more light on the nature of the English modals and their role in utterance interpretation while maintaining the simplest possible semantics.

Notes

1. An intermediate position is taken by Sweetser (1990), who proposes that the relation between the different meanings of the modals is more systematic than the 'polysemy' view allows for, but rejects a 'monosemantic' approach. She argues that the root meanings of the modals are basic, and that these meanings are metaphorically mapped onto the epistemic domain to yield epistemic interpretations: "My proposal is that root-modal meanings are extended to the epistemic domain precisely because we generally use the language of the external world to apply to the internal mental world, which is metaphorically structured as parallel to that external world." (Sweetser, 1990: 50)).

2. To this it should be added that Anna Papafragou (University of London) is now doing some very interesting work on the modals also from the perspective of Relevance Theory.

3. For a more detailed study of all these proposals, cf. Berbeira Gardón (1996a).

4. Groefsema (1995: 55) states that a problem with the polysemy view is that, "even when more, and more fine-grained categories of modality are distinguished, there seems to be no one-to-one correspondence between the kinds and degrees of modality that one can distinguish, and the different meanings of the modal verbs." According to Klinge (1993: 319), "within a framework based on the three categories of modality, epistemic, deontic and dynamic, each modal is assigned several semantic interpretations, which gives rise to the question whether the interpretations are really of the modals."

5. "The intuition that is shared by the different proposals for unitary meanings of the modals is that the full interpretation of utterances containing modals depends on the interaction between these unitary meanings and assumptions in the context of the utterance. The question then is how we can account for this interaction." (Groefsema, 1995: 60-61).

6. "The sense of a sentence is often an incomplete logical form. (...) when a natural language sentence is uttered, the linguistic input system automatically decodes it into its logical form (or in the case of an ambiguous sentence a set of logical forms), which the hearer is normally expected to complete into the fully propositional form that the speaker intended to convey." (Sperber and Wilson, 1986: 73).

7. Unfortunately, evaluating Walton's account of the past tense forms COULD, MIGHT and WOULD falls outside the scope of this paper.

8. In this respect, Klinge (1993: 323) points out that "it is a drawback to existing monosemantic accounts that paraphrases for the individual modals appear to be essentially unrelated, and accordingly do not suggest that the modals share a semantic field."

9. Modal sentences can be arranged according to two different types of linguistic semantic information:

OPERATOR [PROPOSITIONAL CONTENT]

as in the following example (taken from Klinge, 1993: 321):

OPERATOR [JOHN BUY MARY'S TICKET]

The linguistic semantic content of the modal should be understood as a form of modifier of the proposition expressed by the sentence (cf. Rescher, 1968: 26).

10. The notion of 'potentiality' in the analysis of the modals had previously been used by Marino (1973: 312) and Bouma (1973). In the next section, it will be argued that Sperber and Wilson's (1988) notion of 'potential worlds' is more adequate than Klinge's notion of 'potentiality' when trying to account for the interpretation of utterances containing a modal (cf. Berbeira Gardón, 1993).

11. Durst-Andersen (1992) develops some ideas about universal mental models of verbal aspect. According to him, aspect is a linguistic manifestation of mental models which must be formed prior to the acquisition of tense systems.

12. Example taken from Palmer (1979: 60).

13. As Johnson-Laird (1982: 33) points out, "the relation between psychological performance and a model-theoretic semantics remains a remote one. When you evaluate an assertion about a possibility as true, then you will either have succeeded in constructing a state of affairs in which it is realized, or at the very least failed to have shown that there cannot be any such state of affairs. These two outcomes are psychologically very distinct. Only if it is conceded that they are *logically* equivalent, may the outcome of your proceedings be characterized in model-theoretic terms: there is at least one accessible possible world in which the putative event occurs."

14. Examples taken from Klinge (1993: 317).

15. "It is not that first the context is determined, and then relevance assessed. On the contrary, people hope that the assumption being processed is relevant (or else they would not bother to process it at all) and they try to select a context which will justify that hope (...) In verbal comprehension in particular, it is relevance which is treated as given, and context which is treated as variable." (Sperber and Wilson, 1986: 124).

16. Cf. Berbeira Gardón (1996).

17. The basic meaning of WILL, suggested first in Berbeira Gardón (1996b) remains, however, a provisional one, and will be evaluated in the future.

18. Any overtones of desirability in modals are not semantically encoded but pragmatically inferred. Let us compare the infinitival imperative in French:

Ne pas se pencher en dehors.

This could be analysed in terms of (pragmatically inferred) higher-level explicatures:

[Passengers are requested] not to lean outside.

or perhaps there is some way of introducing desirability as an implicated contextual assumption -as is the case of (10), where the speaker concedes *p*, which has already been established as desirable to the addressee. This seems to me a good way of analysing the deontic interpretation of modals, and covert imperatives such as (10), which look as if they are giving permission, and hence communicate that washing the dishes is desirable to the addressee.

19. For a detailed study of the distinction among epistemic, dynamic and deontic interpretations of modals, cf. Berbeira Gardón (1996a), chp. 4.

Works cited

Berbeira Gardón, José L. *La Pragmática de los Verbos Modales Ingleses.* Microfiche. Cádiz: Servicio de Publicaciones, 1993.

_____ . *Los Verbos Modales Ingleses. Estudio Semántico-Pragmático.* Cádiz: Servicio de Publicaciones, 1996a.

_____ . "Epistemic modality and discourse connectivity." *Pragmalingüística* 3-4 (1996b): 223-240.

Bouma, L. *The Semantics of the Modal Auxiliaries in Contemporary German.* The Hague: Mouton, 1973.

Coates, Jennifer. *The Semantics of the Modal Auxiliaries.* London: Croom Helm, 1983.

Durst-Andersen, Per. "Russian qualitative aspect: referential and propositional semantics." *Aspects of Aspect.* Ed. Sørensen, F. Copenhagen: Cebal, 1986. 27-57.

_____ . *Mental Grammar.* Columbus, OH: Slavica, 1992.

Ehrman, Madeline. E. *The Meanings of the Modals in Present-Day American English. Janua Linguarum, Series Practica XLV.* The Hague: Mouton, 1966.

Groefsema, Marjolein. "CAN, MAY, MUST and SHOULD: Unitary semantics, diverse pragmatics." *Cognitive Science Memo* 37 (1992), Colchester: University of Essex.

_____ . "*Can, may, must* and *should*: a Relevance-theoretic approach." *Journal of Linguistics* 31 (1995): 53-79.

Klinge, Alex. "The English modal auxiliaries: from lexical semantics to utterance interpretation." *Journal of Linguistics* 29 (1993): 315-357.

_____ . "On the linguistic interpretation of contractual modalities." *Journal of Pragmatics* 23 (1995): 649-675.

Kratzer, Angelika. "The notional category of modality." *Worlds, Words, and Contexts.* Eds. Eikmeyer, H.-J. and H. Rieser. Berlin: de Gruyter, 1981. 38-74.

Leech, Geoffrey N. *Meaning and the English Verb.* London: Longman, 1971.

Marino, Matthew. "A feature analysis of the modal system of English." *Lingua* 32 (1973): 309-323.

Palmer, Frank R. *Modality and the English Modals.* London: Longman, 1979.

Papafragou, Anna. "On generics." *UCL Working Papers in Linguistics* 8 (1996): 165-198.

Perkins, Michael. "The core meanings of the English modals." *Journal of Linguistics* 18 (1982): 245-273.

Rescher, Nicholas. *Topics in Philosophical Logic.* Dordrecht: Reidel, 1968.

Sperber, Dan and Wilson, Deidre. *Relevance: Communication and Cognition.* Oxford: Basil Blackwell, 1986.

Sweetser, Eve E. *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press, 1990.

Walton, Alan L. *The Pragmatics of English Modal Verbs*. Ph.D. thesis. London: University of
    London, 1988.
_____. "The semantics and pragmatics of CAN." *Linguistische Berichte* 135 (1991): 325-345.
Wilson, Deirdre and Sperber, Dan. "Mood and the analysis of non-declarative sentences." *Human
    Agency: Language, Duty and Value*. Eds. Dancy, J., Moravcsic, J. and C. Taylor. Stanford:
    Stanford University Press, 1988, 77-101.