MODULATING GRAMMAR THROUGH MODALITY: A
DISCOURSE APPROACH

Francisco González García
Universidad de Almería

This paper argues for a dynamic, discourse-based approach to modality which departs from the traditional, monolithic account of modality in a number of interesting respects: (i) modality is characterized as conveying the speaker’s involvement in the propositional content of a given utterance (either in the form of agency or subjectivity) (ii) modality can be said to ramify across the whole lexico-grammatical architecture of the language, (iii) many of the meanings usually ascribed to individual modal verbs are in fact derived either from the verb’s sentential environment or from some wider context of utterance, and (iv) modal meaning in discourse can be said to arise out of the interaction of two closely connected layers of meaning: one embracing the inherent linguistic meaning of the modal verb in conjunction with that of other neighbouring modal devices, and another concerned with principles connected with politeness and face-saving strategies.

1. Introduction

Modality in English and other languages has been the focus of attention of scholars from distinct disciplines and approaches over the last thirty years. Within linguistics, the study of modality has witnessed a gradual shift from a monolithic, static conception to a more dynamic understanding of modality taking into account the relevance of linguistic and extralinguistic contextual factors in the production and interpretation of modal utterances in discourse (Bybee & Fleischman 1995) or the creation of modal textual

* I would like to express my gratitude to Paul J. Thibault (University of Venice) and Judith A. Carini Martinez (University of Almeria) for most helpful discussions on the ideas presented in this paper. I would also like to thank the ELIA editorial board for useful comments and suggestions. All usual disclaimers apply.
coherence (Lundquist 1989). This paper seeks to shed some light on a number of conspicuous issues within this relatively new discourse-oriented program of research into modality, which can be briefly summarized as in (1)-(5) below:

(1) How can modality be aptly and comprehensively defined on both semantic and pragmatic grounds?
(2) What linguistic devices can be considered carriers of modal meaning?
(3) What are the means whereby each modal device modulates the speaker’s involvement which is interpreted from the utterance of that sentence in context?
(4) How can the modal meaning in general and its dynamic relation to the context of utterance in particular, be adequately accounted for from a theoretical and descriptive standpoint?
(5) How can we characterize the enriching (reinforcing or cumulative) effect of contextual factors in the unfolding of “modal” meaning/s in a given piece of discourse?

2. The definition and scope of modality: a question of “attitude”

The extensive bulk of literature on modality shows a clear tendency among linguists1 to accept that semantic areas such as possibility, necessity and prediction (knowledge or epistemic “modality”), on the one hand, and “permission”, “obligation” and “volition” (“deontic” or “root” modality), on the other, constitute the domains of modality. A non-concomitant view with the mainstream conception of modality was already advanced by Halliday (1970:349) in the following terms: “Modality....is the speaker’s assessment of probability and predictability. It is external to the content, being part of the

---

attitude taken up by the speaker”. This “reductionist” view of modality will not be invoked in this paper, since it obscures the fact that deontic modality can nevertheless be regarded as “a form of participation of the speaker in the speech event” (Halliday 1970:335) and plays a significant role in the interpersonal process of negotiation of meaning. Instead, the following definition will be adopted, along the lines suggested by Lyons (1977,1983,1994), where the term “attitude” has been expanded into that of “subjectivity” understood as “subject/speaker’s involvement” so as to embrace both types of modality:

Subjectivity is a matter of speaker’s, or more generally, of the locutionary agent’s involvement of himself in the utterance. In the case of epistemic modality what is involved is his knowledge (or beliefs). In the case of deontic modality it is his will and authority that is involved. But in both cases it is the locutionary agent who is the source of the modality (Lyons

---

2 Therefore, modality in systemic grammar is restricted to what is called “epistemic” modality, and what is often referred to as “root” or “deontic” modality is regarded as forming a different (though obviously connected) category of its own, namely, “modulation”, the reason being that the latter “...are not speaker’s comments on the process referred to...” (Halliday 1970:338).

3 In this respect, two more specific criticisms can be levelled against Halliday’s restrictive formulation of modality: (i) the fact that a clear-cut distinction between both types of modality can be challenged at an ideational level, since cases of indeterminacy (“merger”) may occur (Leech & Coates:1980, Coates:1983, Bald 1990:354, among others); and (ii) both epistemic and deontic modality may be geared towards the interpersonal component of the language (i.e. politeness strategies in general and face-saving in particular). We shall have more to say about this at a later stage in this paper.

4 Palmer (1986:16) also argues that subjectivity is an essential criterion for modality, which he defines as “the grammaticalization of speaker’s (subjective) attitude and opinions.” In our definition of modality, we shall leave out the “grammaticalization” component as it belongs to the domains of mood rather than modality.
Therefore, it can be claimed that modality is concerned with the expression of the subject/speaker’s involvement towards the propositional content of an utterance, whether in the form of agency or subjectivity. An interesting conclusion ensuing from the above definition is that modality need not, and should by no means, be exclusively restricted to modal auxiliary (or semi-auxiliary) verbs.

Over the last fifteen years, linguists have concentrated on surveying a more or less comprehensive inventory of modal elements, a representative sampling of which can be said to include (i) adjectives (e.g. “possible”, “necessary”, etc.), (ii) participles (e.g. “alleged”, “demanded”, etc.), (iii)
nouns (e.g. “necessity”, “consideration”, etc.), \(^8\) (iv) lexical verbs (e.g. “wonder”, “order”, etc.), (v) adverbs (e.g. “perhaps”, “obviously”, etc.), (vi) articles (e.g. “the”/ “a”, “some”/ “any”, etc.), \(^9\) (vii) tense (usually in “preterite” or marked forms) (e.g. “I thought you in Paris”, “In 1492, Columbus discovers America”), \(^10\) (viii) aspect (marked forms) (e.g. “John is having a headache”), \(^11\) (ix) particles (e.g. “if”) and truth-functionally equivalent devices (e.g. subject-verb inversion, putative “should”) (e.g. “If you don’t like it \(^{12}\) should you not like it, that’s your hard luck”), (x) the degree of morphosyntactic compression in the encoding of a sentential complement (finite, non-finite or verbless clause) (e.g. “They consider that

\(^8\) Jacobson’s (1982) study on modal nouns shows, among other things, that there is a much greater variety of lexical items in the field of modality nouns than in that of modal auxiliaries. This finding points to one of the many inconveniences of treating modal verbs as the exclusive or main carrier of modal meaning.

\(^9\) Lyons (1983) provides an interesting outline of the many connections between “demonstrative” and “indicative” and Rauch (1983) also further explores the relationship between modality and deixis in general. Further research will have to address how concepts such as anaphora, generality, individuality (and their grammatical correlates such as “number”) relate to modality.

\(^10\) Lyons (1983:104) justifies the modal treatment of this marked form (ie. historic present) on the grounds that “it conveys an impression of vividness and subjective involvement”. Kress (1977), Hutchinson (1985:1), Fleischman (1989:7) and Kvacanovic (1994:67ff), however, contend that all tense forms imply a given degree of speaker’s distance from and attitude towards a given proposition and can therefore be seen as modalized forms.

\(^11\) Lyons (1983:101) argues that a dynamic, agentive reading of this sentence is feasible given an adequate supporting context (eg., “John is simulating a headache or doing something to bring it on). It must be borne in mind that, for Lyons, agency is superimposed upon the notion of subjectivity. The exact relationship between subjectivity-agency as the main axes of modality remains to be further explored from a linguistic as well as a socio-semiotic viewpoint.

\(^12\) Perkins (1983:110) lumps together all types of “if”-clauses (nominal and adverbial) as modal. Under the view of modality taken here, contextually motivated variants such as indirect speech placed in the past are treated as non-modal. See also Hübler (1983:134) for a position along the same lines.
John is intelligent/John to be intelligent/John intelligent”, etc.),
(hedging expressions (e.g. “She is a great scholar, I think”), (xii) “emphatic”
do (e.g. “Do be getting sleepy!”), (xiii) “get”-passives (e.g. “He got elected
Class President”), (xiv) question tags (e.g. “You liked the movie, didn’t
you?”), (xv) ‘yes’/ ‘no’ questions (e.g. “Have you ever been to
London?”), (xvi) (rising) intonation (e.g. “I feel happy”), and so forth.

The above inventory, by no means complete, is intended to illustrate
how varied and sophisticated the modality system in English is, thus fully
endorsing Halliday’s (1970:331) statement that “…there is (thus) no single
place in the clause where modality is located”. However, the linguistic
system of modal resources is further enriched by their numberless
combinatory possibilities in a given piece of discourse. In the words of
Halliday (1970:331):

Nor do the different non-verbal forms of the same lexical item necessarily
correspond with each other: “obviously” is not the same as “it is obvious
that...”, “surely” as “I am sure that”. But there are discernible groupings, and

---

13 See Ransom (1977) and Gonzálvez (1997) for an account of the embedded complements as
having modal meaning. Differences exist as to where the locus of the modal meaning should
be established. Thus, while for Ransom (1977) the embedded complements have a modal
meaning of their own and independent of that of their embedding predicates, Gonzálvez
(1997) proposes that the modal meaning arises out of a complex, dynamic interaction of the
meaning and form properties of the morphosyntactic encoding of the embedded clause with
those of the embedding predicate.

14 Bailey (1981:168), quoting Bolinger (p.c.) and Granville Hatcher (1949), argues that “get”-
passives are characterized, among other things, by having more saliently affected subjects than
their “be”-counterparts, which motivates their being treated as subjectively modalized.

15 Together with questions, markers of negation are likely candidates for a modal treatment.
See Palmer (1979) for further discussion on the issue of the relationship between “non-
assertion” and modality.

16 See, for further reference, Stevenson (1967), Crystal (1969, 1975), and Roach (1991:164ff),
among others. The latter is particularly insightful for a better and comprehensive
understanding of the attitudinal function of intonation in general.
a clear distinction can be drawn between pairs which are felt to be equivalent, and thus reinforce each other (“as concord”) when both are present, as in “Perhaps he might have built it”, and those which are not equivalent and are thus cumulative in meaning, as in “Certainly he might have built it” (“I insist that it is possible” or “I grant that it is possible”).\textsuperscript{17}

An interesting conclusion ensuing from the above is that modal meaning can be said to resolve itself into a more or less intricate flow of reinforcing or cumulative modal devices which invades the whole lexico-grammatical architecture of the language, always subject to the communicative requirements of the interlocutors in the modalised space-time if dialogic negotiation.\textsuperscript{18}

3. Towards a “dynamic”, context-based view of modal meaning

Traditional (and not so traditional) accounts of modality present in our view two important shortcomings when tackling the issue of modal meaning: (i) they fail to separate the intrinsic linguistic meaning of modal verbs from the linguistic meaning of other neighbouring modal items (of different kinds) in the sentential environment in which they are inserted, and (ii) they also fail to distinguish the meaning of modal verbs from the (endless) number of pragmatic uses or functions these may be put to in specific contexts.

\textsuperscript{17} See Lyons (1977:Chapter 17) for a similar position in the form of a distinction between modally harmonic and non-harmonic meaning.

\textsuperscript{18} In addition, it must be borne in mind that modality may be implicated by the semantic structure of the whole utterance without being signalled by any specific (linguistic) markers. Thus, Costa (1975:119-21) (quoted in Brown & Levinson (1978:274), notes that English statements like “One just doesn’t do things like that” can implicate the deontic modality interpretation (“You must not do things like that”). In our view, the discourse approach to modality invoked in this paper can prove useful to help to determine to what extent covert modality may be implicated with the aid of neighbouring (linguistic) devices in a given discourse scenario.
The consequences to be derived from this cannot possibly be expected to be advantageous from either a theoretical or descriptive view. Thus, as Walton (1991:367) rightly argues, “…the meanings that are ascribed to modal verbs multiply arbitrarily as more and more context is added, and any reference grammar aimed at recording these meanings will come to resemble a lexicon and still fail to cover them all”.

Let us briefly illustrate the relevance of the above statement with particular relevance to the taxonomy of modal meanings for “might” in three well-known works on modality: Leech (1971), Coates (1983) and Hoye (1997). Of the three, the taxonomy proposed by Hoye (1997) is the one coming closer to the line of research into modality invoked here, since it refrains from attributing the meaning of items in the sentential environment of a sentence (i.e. a hypothetical conditional clause, a concessive clause, etc.) to the modal verb itself (“hypothetical might”, “concessive might”), or even lumping together the inherent meaning of the modal with some of the specific pragmatic functions this verb may fulfil in an also specific context of

---

19 Leech (1971) comes up with something like six different meanings for “might”, namely, (i) past tense of “may”, (ii) a “subjunctive substitute” (in clauses of purpose and concession), (iii) hypothetical past tense, (iv) “polite permission”, (v) possibility and, (vi) “tentative suggestion”. Coates (1983:147-8) makes them eight: (i) epistemic possibility, (ii) past of “may” (“epistemic possibility”), (iii) past of “may” (“root possibility”), (iv) past of “may” (“root permission”), (v) remote of “may” (“epistemic possibility”), (vi) remote of “may” (“root possibility) and (vii) remote of “may” (“root permission”). Finally, Hoye (1997:274-5) reduces them to three only: (i) epistemic probability/possibility, (ii) non-epistemic (a blending of possibility and permission) and (iii) collocational combinations with “well” “to convey epistemic probability”, with “just” in “…formulaic expressions of politeness (usually to ask for permission)” and, finally, with “(just) as well” to “make a circumspect or sardonic recommendation” (Hoye 1997:275).
utterance (e.g. “polite permission/suggestion”, etc.).

The question now arises as to how modal meaning can be felicitously systematized in order to account for the dynamic, though nevertheless systematic, relationship between modality, the context of utterance and the interpersonal function of **politeness**. Following Silva Corvalán (1995:73ff), we shall argue for a monosemantic approach to the semantics and pragmatics of modality drawing on intersecting, polysemantic contexts. More specifically, it will be suggested that modal meaning can be shaped and further modulated in terms of (at least) two separate, though closely connected, layers of meaning:

**LAYER 1**: (i) **SYSTEMATIC MEANING** (i.e., the meaning present in all uses of a modal) in conjunction with (ii) **CONTEXTUAL (LINGUISTIC) MEANING** (i.e. the reinforcing or cumulative modal nuances introduced by other satellite elements in the neighbouring sentential (or discourse) environment.

**LAYER 2**: **PRAGMATIC (OR INTERPERSONAL) MEANING** (i.e. the modulation or fine-tuning on the part of interlocutors of the meaning

---

20 However, in our opinion, Hoye’s account is not without problems either, since it fails to account for the relationship between the possibility and permission readings of “might”, on the one hand, and the (metalinguistic) function of politeness, on the other, in a neat, systematic way. More specifically, Hoye’s account misses the important fact that the politeness factor is not a mere appendix of modality, which can be traced in specific formulaic expressions with some collocations, but it should rather be seen as an overarching pragmalinguistic function that can be overtly or covertly marked in deontic, epistemic and dynamic modal utterances, rather than just in the former alone, as Hoye’s taxonomy seems to imply.

21 Politeness is used henceforth in a sense compatible with that of Brown & Levinson (1978), though it is taken to encompass not only linguistic but also those socio-semiotic aspects of communication.

22 The following schema departs from the original presented in Silva Corvalán (1995) in that it replaces the prototypical discourse meaning component with what has been referred to above as pragmatic (or “interpersonal”) meaning.
potential arising out of LAYER 1 basically, though not exclusively, in terms of other non-linguistic features of the context of utterance, with special focus on politeness in general and face-saving in particular).  

The above schema captures, in our opinion, the relationship between the locutionary resources of modality, on the one hand, and the interpersonal function it can perform in the process of negotiation in a dynamic, comprehensive way, on the other, without falling into some of the inconsistencies already discussed.

To round off the discussion being entertained in this paper, we shall proceed to illustrate succinctly the importance of the dynamic, discourse-based view of modality invoked in the preceding pages, with specific reference to some cursory examples featuring “might” from the British component of the International Corpus of English (ICE-GB henceforth) and an extract from the filmscript of “Four Weddings and a Funeral”, reproduced in (6)-(10) below:

(6) “Maybe I might miss something out because we are all human and I’ve overlooked it and you’ve remembered it” (ICE-GB Corpus, S2A-061-36).

Here the linguistic meaning of the modal appears to be that of “remote

---

23 Contextual features also embrace the field of discourse, the mode of discourse, and the style of discourse and should eventually be extended to embrace the social dynamics of negotiation as a whole. A justification is in order here as to why politeness is given special relevance in comparison to other factors such as formality. As Oka (1981:85) has persuasively demonstrated, the dominant factors affecting formality are external variables such as relative status and intimacy, while politeness is more of a psychological state of a speaker towards his addressee. The consideration of contextual factors suggests that formality is more directly dependent upon socio-situational factors of language use, whereas politeness is more dependent upon the speaker’s attitudinal factors within a particular social setting. It is this proximity with speaker’s attitudinal factors that justifies why politeness is seen here as forming part of modal meaning. As for the specific relevance of face-saving within politeness, see Fraser (1990).
possibility”, this being at first sight reinforced by the epistemic modal adverb “maybe”. However, as the message progressively unfolds, one finds that the modal meaning is made to contrast sharply with some non-modal forms (e.g. “are”, “have overlooked” and “have remembered”), thus providing insightful clues as to the intended “factual”-like contextual interpretation of the modal in question by the speaker-addressee. As for the second layer of meaning, one can easily see that the choice of the remote form here is motivated by face-saving factors, or more exactly, the speaker’s plea of sympathy from the audience for any mistake or inaccuracy on his/her part. The point to emphasize, however, is that the choice of the form “might miss out”, unlike any other feasible choice (e.g. “may have missed”, “might have missed” or even the non-modal expression “have missed”) allows the speaker to modulate successfully what is highly likely to be an actual mistake in terms of an eventual (possible) mistake (epistemic modality), while justifying his/her “wrong” course of action in terms of a general weakness of mankind (deontic modality).

(7) “I **might** go back to Cambridge early or something because I’ve got to write an extended essay” (ICE-GB Corpus, S1A-093-26).

As in (6), the use of “might” here could be taken to imply that the possibility of the speaker going back to Cambridge is somewhat remote. However, the neighbouring linguistic environment provides unambiguously “factual” contextual cues which show that the speaker has already made up his mind to go to Cambridge early (e.g. hedging in “or something”, use of a modal phrase implying strong obligation in conjunction with the reinforcing nuance implicated by “extended”). However, the choice of “might” serves the speaker to reveal his intention of leaving early (dynamic modality) without precluding any change of plans (knowledge modality). As for the

---

24 Needless to say, inference strategies are based on the total context of an utterance, which includes not only the socio-physical aspect of an utterance, but also encompasses the mutually shared knowledge between the speaker and the addressee.
interpersonal level of meaning, the high degree of distance of “might” serves the speaker to be conventionally indirect and tactful in communicating the addressee his real intentions.

(8) “(...) Yes, please, don’t bother for a moment because merely I wanted to know whether you disagree, as I think you might do from what you’ve been saying with the passage that I’ve quoted from Dr. Kendall’s evidence (...)” (ICE-GB Corpus, S2A-061-36).

In (8) above, the use of “might”, far from signalling remote possibility, mainly reveals the speaker as adhering to the conventions of politeness in an attempt to beg hearer’s forgiveness for a given transgression in the communicative exchange. In fact, the unambiguously factual evidentiality background on which the speaker’s judgement is based (i.e. “you’ve been saying”, “I’ve quoted”) shows that the speaker is quite sure that the hearer disagrees (epistemic modality) and also that it is the hearer’s will to show his/her disagreement (dynamic modality). However, through the use of “might” (skilfully pre-faced by the hedging expression “I think”), the speaker manages to minimize the categoricalness of his/her assertion, thus saving face.

(9) “You might wish to do this because, while you have it: any Class 1 contribution you pay at a reduced rate do not count for benefit; and you cannot get Home Responsibilities Protection; and you are not allowed to pay voluntary contributions, and if you are a married woman, you cannot get credits, and from October 1989 you could pay more in contribution than colleagues with the same earnings but paying standard rate contributions” (ICE-GB Corpus, W2D-004-70)

Example (9) above, taken from a governmental leaflet on tax payment, provides clear evidence of how the putative surface interpretation of “might” as conveying remote possibility (epistemic modality) or a tentative suggestion (deontic modality) is progressively re-shaped through a modally consistent network with an overwhelmingly factual, categorical tone (e.g. “do not count”, “are not allowed”, “cannot get”). Therefore, the
tentative modal stance encoded in “might wish” becomes more forceful as the text progresses, to end up implying a relatively high degree of possibility (e.g. “You will want to do this”) (epistemic modality) and a forceful recommendation (e.g. “You should do this”) (deontic modality). However, the choice of “might” enables the speaker to mitigate the underlying forceful statement in accordance with politeness strategies in general and the safety claim technique of the discourse of advertising in particular.25

Example (10) above is perhaps more interesting as it furnishes us with a sample of dialogic interaction. Here we know for sure that Charles wants to stay in the pub to meet Carrie, the girl of his dreams. Hence the need to apologize for a sudden change of plans which might be face-threatening. The modal theme initially brought in by “might” (skillfully pre-faced by the apology marker “sorry”) ramifies across an intricate constellation of modal devices (i.e. the italicized elements above) which help him make it clear to Tom that he is definitely staying at the pub (epistemic modality) and minimize the imposition of his will upon that of his friends (dynamic modality).

In the previous pages, we hope to have shown some of the main attractions of a discourse-based approach to modality with a view to gaining a more comprehensive (and possibly better) understanding of its relevance

25 The relationship between politeness and interpersonal meaning, on the one hand, and the specific discourse type in which the interlocutors are engaged still remains to be further explored in future research, as already noted by Brown & Levinson (1978:41), among others.
for the encoding and interpretation of modal utterances in the dialogic process of negotiation.
Bibliographical references


Hermerén, L. 1986. “Modalities in Spoken and Written Language: An
Inventory of Forms” in G. Tottie (ed.) English in Speech and Writing. Uppsala.


