


---

**Modality and polarity**

Sjie Barbiers
Royal Netherlands Academy of Sciences, Meertens Instituut, Amsterdam

In this paper it is argued that the possible interpretations of a modal verb primarily depend on properties of its complement. On the basis of non-verbal complementation of modals in Dutch it is shown that modals always require a complement that denotes a bounded scale. The nature of this scale determines whether the modal gets a root or an epistemic interpretation.

1. **Introduction**

This paper investigates the relation between the interpretations of modal verbs and the nature of their complements, starting out with a cross-linguistically rare construction in which modals take non-verbal complements:

```
  (1) a. Het licht moet aan.
      the light must on
      'The light must be switched on.'

  b. De fles moet leeg.
      the bottle must empty
      'The bottle must be emptied.'
```

The ensuing analysis of this construction leads to the following claims:

i. The non-verbal complements in (1) cannot be analysed as involving a silent infinitive (Section 2; contra Geerts et al. 1984; Vanden Wyngaard 1994).

ii. The complement of a modal must denote a value on a bounded scale, regardless of the interpretation of the modal. If the complement does not denote such a value, the sentence is ungrammatical (Section 2).

iii. The data in (1) are problematical for analyses according to which root interpretations always involve control structures while epistemic interpretat-
tions always involve subject raising structures (e.g. Hofmann 1966; Ross 1969; Perlmutter 1970). While the modals in (1) must be analysed as raising verbs, they can only have root interpretations, not epistemic interpretations.

iv. In root interpretations, a modal can be subject-oriented (dyadic) or non-subject-oriented (monadic). A modal is always monadic in epistemic interpretations (Section 3).

v. Modals with a root interpretation always involve a potential change of the value on the scale introduced by the modal’s complement, i.e. a potential polarity transition. When such a change is impossible, e.g. when the complement of a modal denotes an invariable property of the subject, an epistemic interpretation is forced (Section 4).

vi. The semantic ambiguity of modals is determined by two parameters: (1) the subject-orientation of the modal, and (2) the availability of a potential change of value on a bounded scale. This yields four classes of modal interpretations (Section 5).

vii. The syntactic complement of the modal determines its interpretations. There are two types of complements that cannot be interpreted as involving a potential change of value. In the case of a definite nominal complement, this gives rise to sympathy/antipathy interpretations. In the case of a verbal complement, this yields an epistemic interpretation. These interpretations are possible only if negation or (abstract) affirmation is present to provide the required value on a bounded scale (Section 6).

viii. Epistemic interpretations are non-subject-oriented. This follows from the assumption that the semantic relation between a verb and the subject is established by an (abstract) verbal head with D-features. In the epistemic interpretation, the projection of this verbal head including the subject is the single argument of the modal, making it impossible to establish a modal relation between subject and verbal complement (Section 7).

ix. The negative polarity of moeten ‘must’ in (2a) and the bipolarity of mogen ‘may’ in (2b) follow directly from the general lexical argument structure of modals. There is no need to assume that this polarity is an accidental lexical idiosyncrasy or that polar and non-polar instances of moeten and mogen involve different lexical entries (Section 8).

(2) a. Marie moet die jongen *(niet).
   Marie must that boy not
   ‘Marie does not like that boy.’

b. Marie mag die jongen (wel)/niet.
   Marie may that boy affirm/negate
   ‘Marie likes/not like that boy.’

2. Modals with non-verbal complements

Traditionally, Dutch modals with a non-verbal complement have been analysed as involving an infinitive that is phonologically silent or deleted at PF (Geerts et al. 1984; Vanden Wyngaerd 1994). An advantage of such an analysis is that it allows the cross-linguistic generalisation that modals always take a verbal complement, as in English.

(3) a. Jan moet weg (gaan).
    Jan must away go
    ‘Jan must go away.’

b. Jan wil een pizza (hebben).
    Jan want a pizza (have)
    ‘Jan wants to have a pizza.’

However, there are a number of strong counterarguments. The first argument against PF-deletion of the infinitive in the case of modals with a non-verbal complement is that certain cases would require deletion of a passive complement. This incorrectly predicts that such cases should allow a BY-phrase or an agent-oriented adverb. The crucial observation on which this argument rests is that it is not always sufficient to assume that GO, HAVE, BE or some other basic verb is deleted: other, more complex verbal expressions must be assumed to be deletable as well.

(4) a. Deze lampen moeten uit (≠ gaan/zie).
    These lights must out (go/be)
    ‘These lights must be switched off.’

b. De boeken mogen weg (≠ gaan/zie)
    These books may away (go/be)
    ‘These books can be thrown away.’

The sentences in (4) with audible GO/BE yield an interpretation that is completely different from the interpretation with GO/BE absent. With GO/BE absent, the interpretation must be as given in the translation, whereas audible GO forces an interpretation that the lights go out by themselves and the books go away by themselves, while audible BE forces an interpretation that the lights
must be out and the books must be away at utterance time or some point in the future. As the translations in (4) indicate, the interpretation of (4a, b) without an infinitive is as with an embedded passive. We could assume that (4a, b) is the result of PF-deletion of the embedded passive:

(5) a. Deze lampen moeten uit worden gedaan.
   These lights must out (be done)
   'These lights must be switched off.'

   b. Deze lampen moeten uit worden gedaan.
   These books may away (be done)
   'These books can be thrown away.'

If (5a, b) were the correct analysis, it should be possible to add a BY-phrase or an agent-oriented adverb, since passives allow such an addition. However, BY-phrases and agent-oriented adverbs are impossible when there is no overt verbal complement (6a, b), even though deletion of infinitive and participle normally does not force deletion of BY-phrase or agent-oriented adverb (6c, d).

(6) a. Deze lampen moeten door jou uit worden gedaan.
   these lights must by you out (be done)
   'These lights must be switched off.'

   b. Deze lampen moeten door jou uit worden gedaan.
   these lights must carefully off (be done)
   'These lights must be switched off carefully.'

   c. De lamp moet door P. aan worden gedaan en door J. uit worden gedaan.
   the light must by P. on be done and by J. off be done
   'The light must be switched on by Peter and switched off by Jan.'

   d. De lamp moet door P. aan worden gedaan en uit worden gedaan.
   the light must carefully on be done and off be done
   'The light must be switched on and switched off carefully.'

The second argument against a PF-deletion analysis involves the reduced ambiguity of modals with a non-verbal complement. Since the alleged deletion of verbal material would take place at PF, this should not have any consequences for the interpretation possibilities. However, whereas a modal with a verbal complement can get an epistemic interpretation, a modal with a non-verbal complement can never be interpreted epistemically. This is illustrated in (7).^4

(7) a. Jan mag weggaan, (ooit zal hij terugkeren).
   Jan may away-go (someday will he return)

A third argument against the PF-deletion analysis involves cases in which there is no suitable candidate for deletion:

(8) a. Jan kan niet tegen katten.
   Jan cannot against cats
   'Jan is allergic to cats.'

   b. Jan kan twee keer in die jas.
   Jan can twice in that coat
   'That coat is a size too big for John.'

   c. Jan kan zijn werk niet aan.
   Jan can his work not on
   'Jan cannot cope with his work.'

   d. Jan moet Marie niet.
   Jan must Marie not
   'Jan does not like Marie.'

   e. Jan mag Marie wel.
   Jan may Marie aff
   'Jan likes Marie.'

The final and perhaps strongest argument against the PF-deletion analysis is the following observation. Only constituents that denote a value on a bounded scale can be the non-verbal complement of a modal. The facts in (9) can be stated directly in terms of selectional restrictions imposed by the modal if we assume that there is no silent infinitive. If there were a silent infinitive, we would be forced to say that the modal imposes selectional restrictions on the complement of its complement, a rather unusual state of affairs.^5 (Note by the way that the stage-level-individual-level distinction is irrelevant here, as the

---

^4 Sjef Barbiers

^5}
examples (9d, e) show. It is also not sufficient for the complement to denote a scale that is unbounded, as (9f) shows.6,7)

9a. De fles moet leeg.
the bottle must empty
b. Het raam kan open.
the window can open
c. Het licht mag uit.
the light may off
d. *Het konijn kan ziek.
the rabbit can sick
e. *Jan moet intelligent.
Jan must intelligent
f. *De speler moet lang.
the player must long

I will assume on the basis of these four arguments that Dutch modals can have a non-verbal complement, as long as this complement denotes a value on a bounded scale.

3. Against a raising-control analysis of the epistemic-root distinction

On the basis of the conclusion that modals in Dutch can have non-verbal complements, it is possible to show that the semantic epistemic-root distinction cannot always be expressed in terms of the syntactic distinction between raising and control verbs.8 An example of this traditional syntactic analysis of modal ambiguity (cf. Hofmann 1966; Ross 1969; Perlmutter 1970) is given in (10).

10. John must be kind.
   i. *It must be true that John is kind.'
      epistemic
      John, must be [t, kind]
      raising
   ii. *John has the obligation to be kind.'
      root
      John, must be [PRO, kind]
      control

In (10i), the modal is a monadic predicate, taking the proposition as its sole argument: must (John be kind). The subject does not receive a Theta-role from the modal. In (10ii), the modal is a dyadic predicate, a relation between the subject and a predicate: must (John, be kind). The subject John receives a Theta-role from the modal, expressing that the obligation is attributed to the subject in ii. but not in i.

Feldman (1986) and Brennan (1993) have already observed that sentences such as the ones in (11) clearly can have deontic interpretations but cannot be analysed as involving control:

11a. De brief moet van Jan voor vijf uur in Amsterdam zijn.
the letter must of Jan before five hour in Amsterdam be
   *Jan requires that the letter be in Amsterdam before five o'clock.'
   b. De nieuwe machine mag geen problemen geven.
new the machine may no problems cause
   *It is unacceptable if the new machine causes problems.'

The sentences in (11) cannot be cases of control since the obligation and permission are not attributed to a subject; rather, the obligation/permission applies to an entire state of affairs here: the state of affairs 'the letter be in Amsterdam before five o'clock' is obligatory, and the state of affairs 'the new machine causing problems' is not permitted. The modals in (11) must therefore be analysed as monadic predicates.

Modals with non-verbal complements provide an even stronger argument against the raising-control analysis of the epistemic-root distinction. As is well-known, small clause PP-complements in Dutch cannot be extrapoated (12a; Hoekstra 1984), whereas all other types of PPs, such as PP-arguments (12b) and PP-adjuncts (12c), can:  

12a. dat Jan <naar huis> gaat <naar huis>
that Jan to home goes to home
   *that Jan is going home'
   b. dat Jan <naar Marie> luistert <naar Marie>
that Jan to Marie listens to Marie
   *that Jan is listening to Marie'
   c. dat Jan <in de tuin> zit te lezen <in de tuin>
that Jan in the garden sits to read in the garden
   *that Jan is reading in the garden'

PP-complements of modals behave like small clause complements in this respect:

13a. dat Jan morgen <naar Amsterdam> moet <naar Amsterdam>
that Jan tomorrow to Amsterdam must to Amsterdam
   b. dat de brief morgen <in de prullenbak> mag <in de prullenbak>
that the letter tomorrow into the trashcan can into the trashcan
If the PPs [naar Amsterdam] 'to Amsterdam' and [in de prullenbak] 'into the trashcan' are small clause complements, then the DPs [Jan] 'Jan' and [de brief] 'the letter' must be their subjects. Their subjects cannot be PRO, since small clause complements cannot have PRO as their subject (14a). This in turn implies that the subjects in (13) must have raised from a position inside the small clause, as is illustrated in (14b).

(14) a. Jan drinkt [sc zich/*zro ziek]  
   Jan drinks se sick  
   b. Jan moet [pp t naar Amsterdam]  
   Jan must to Amsterdam

We have now established that (13a, b) must be analysed as subject raising constructions. If subject raising corresponded to an epistemic interpretation, the prediction would be that sentences like (13a, b) can only have an epistemic interpretation. But this prediction is wrong: sentences like (13a, b) cannot have an epistemic interpretation; they can only have root interpretations:

(15) Jan moet naar Amsterdam.  
   i. *Jan has the obligation to go to Amsterdam.' direct deontic  
   ii. 'It is required that Jan goes to Amsterdam.' indirect deontic  
   iii. *'It must be true that Jan is going to Amsterdam.' epistemic

We can conclude that the raising-control analysis of the epistemic-root distinction cannot be entirely correct.

4. Potential polarity transition as a crucial property of root interpretations

As discussed in the previous section, modals with a root interpretation can be subject-oriented and non-subject-oriented (cf. Feldman 1986; Brennan 1993). Subject-orientation therefore cannot be used as a defining property of root interpretations. Non-subject-orientation, however, still counts as one of the defining properties of the epistemic interpretation. In this section it is shown that there is another factor playing a role in the distinction between the two types of interpretations: the availability of a potential polarity transition.

One of the arguments mentioned in Section 3 in favour of the existence of constructions with a modal and a non-verbal complement is that this allows us to express selectional restrictions on the complement of the modal straightforwardly: the non-verbal complement of a modal must denote a value on a bounded scale. If it does not, the sentence is ungrammatical. It is clear what role this value on a bounded scale is playing in the semantic interpretation of a sentence with a modal. For example, a sentence like (16) may be paraphrased as: 'the value that the bottle has on the scale of empty (0) to full (1) is not 1 when (16) is uttered and should become 1 at some point in the future.' Notice that the sentence does not say anything about the question whether this value will ever be reached. We may say that the interpretation of (16) involves a potential, not an actual polarity transition from [not 1] to 1.

(16) Deze fles moet vol.  
   this bottle must full

My claim is that the above paraphrase of (16) reveals two essential properties of root interpretations: (i) the complement of the modal must denote a value on a bounded scale, and (ii) this value is not the actual value at the moment 't denoted by the modal. The latter requirement makes sense intuitively: (16) cannot be uttered in a situation in which the bottle is already full.

It can be shown that a potential polarity transition is also crucial for the availability of root interpretations when modals have a verbal complement. When the complement of the modal denotes a fixed property of the subject, an epistemic interpretation is forced:

(17) a. De aarde moet een planeet zijn.  
   the earth must a planet be  
   i. *'It must be true that the earth is a planet.' epistemic  
   ii. 'It is required that the earth is a planet.' root
   b. Jan moet een moedertaalspreker van het Fins zijn  
   Jan must a native speaker of the Finnish be  
   i. *'It must be true that Jan is a native speaker of Finnish.' epistemic  
   ii. 'It is required to be a native speaker of Finnish.' root
   c. Jan moet zijn kamer gisteren voor vijf uur hebben opgeruimd.  
   Jan must his room yesterday before five hour have cleaned  
   i. *'It must be true that Jan has cleaned his room before five.'  
   o'clock yesterday' epistemic  
   ii. 'Jan is required to have cleaned his room yesterday.' root

Non-verbal complements and infinitival complements of modals thus have in common that they only allow root interpretations if they denote a variable property. For non-verbal complements, this is not yet sufficient, as we have seen: the potential polarity transition applies to a value on a bounded scale provided by the complement. Many stage-level predicates cannot be the non-
verbal complement of a modal, since they do not denote a value on a bounded scale (cf. 9d, f).

Now that it has been established that a potential polarity transition is also crucial for the availability of root interpretations with infinitival complements, the question must be asked whether infinitival complements involve a value on a bounded scale. The null hypothesis clearly is that they do. If we were to hypothesise that they do not, we would in fact claim that the requirement of a bounded scale in the case of non-verbal complements is completely accidental and that non-verbal and verbal complements are fundamentally different, despite the fact that they both must denote a variable property.

The next question is which constituent provides the bounded scale in the case of infinitival complements. I would like to suggest that it is the embedded verb itself that does that. The behaviour of the adverbial modifier half 'halfway' supports this view. It typically modifies the end point of a bounded scale (18a) and therefore does not occur with constituents that do not denote an endpoint (18b). As expected, it can modify the infinitival complement of a modal (18c, d).

(18) a. De fles is half leeg.
   the bottle is halfway empty
b. ‘Jan is half oud.
   Jan is halfway old
c. Jan moest half huilen.
   Jan must-past halfway cry
   ‘Jan was torn between laughing and crying.
d. Ed mag maar half beseffen wat de bedoeling is.
   Ed may only half realise what the intention is
   ‘Ed should realise only partly what the intention is.’

The sentence in (18c) means that there was a half, not a full crying event. Aspect is not involved here, as (18c) is not about the completion of the crying event. What (18c) means is that Jan was halfway between not crying and crying. Similarly, (18d) means that Ed should be halfway between not understanding and understanding what the intention is. The scale involved is simply a numerical scale from 0 to 1, from no event to one event. As I argue in Section 6, this numerical scale should not be confused with the scale involved in epistemic interpretations.

A striking difference between verbal and non-verbal complements of modals is that modals do not impose any further selectional restrictions on the embedded verb: any verb can be the complement of a modal. This can be explained if we assume that a verb is always dominated by a functional projection, the head of which denotes the value 1 (cf. Section 6 and Barbiers 1995). The presence of the verb then guarantees the presence of this functional projection, and the functional projection satisfies the selectional restriction imposed by modals, making the choice of verb free.13

As the notion of potential polarity transition may easily give rise to misunderstandings, four qualifications of the observations in (17) are called for (cf. Barbiers 1995 and the Introduction to this volume):

(19) i. The aspectual class of the embedded verb is irrelevant for modal interpretation. More specifically, stative complements allow root interpretations. The crucial property of root interpretations is whether the situation denoted by the complement can switch from non-existent to existent; whether this situation is dynamic or stative internally is irrelevant.

   ii. Individual-level complements only block the root interpretations when the subject is a name, not when the reference of the subject may vary with different situations or worlds (20b–c). This is because the property denoted by the predicate within the complement can only be a permanent, invariable property relative to a particular referent.

   iii. Perfective complements force an epistemic interpretation only when they refer to the completion of an event at some point in the past, not when they refer to the completion of an event at some point in the future (20d).

   iv. It is possible to find contexts in which fixed properties such as the one in (17a) can be interpreted as variable properties. This actually confirms our point. In a possible world in which the earth is yet to be created, the property of being a planet or a star has not yet been fixed and is still variable. Therefore root interpretations are possible (20e).

(20) a. Jan moet een auto hebben.
   Jan must a car have
   i. ‘It must be true that Jan has a car.’
   ii. ‘Jan is required to have a car.’

b. The new professor must be a native speaker of Finnish.
   i. ‘It must be true that the new professor is a native speaker of Finnish.’
   ii. ‘It is required that the new professor be a native speaker of Finnish’

c. Jan must be a native speaker of Finnish.
   i. ‘It must be true that Jan is a native speaker of F.’
5. A new classification of modal interpretations

The two parameters for modal interpretations isolated in the previous section, [± subject-oriented] and [± potential polarity transition], are independent, yielding the four different types of modal interpretation in Table 1. An example of each interpretation is given in (21). The most interesting result is that the two parameters correctly predict that modal interpretations of type d. should exist. This type of interpretation is discussed in Section 8.

Table 1. A new classification of interpretations of modal verbs in Dutch

<table>
<thead>
<tr>
<th>[± potential polarity transition]</th>
<th>[+ subject-oriented]</th>
<th>[- subject-oriented]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[± potential polarity transition]</td>
<td>a. Dispositional</td>
<td>c. Indirect deontic</td>
</tr>
<tr>
<td></td>
<td>b. Direct deontic</td>
<td></td>
</tr>
<tr>
<td>[- potential polarity transition]</td>
<td>d. Sympathy/antipathy</td>
<td>e. Epistemic</td>
</tr>
</tbody>
</table>

(21) a. Jan mag graag de baas spelen.
Jan may eagerly the boss play
'Jan likes to be the boss.'

b. Jij mag van mij de kamer uitzetten.
you may from me the room out
'I allow you to leave the room.'

c. De kamer moet schoon zijn.
the room must clean be
'It is required that the room be clean.'

d. Jan moet Marie niet.
Jan must Marie not
'Jan does not like Marie.'
e. Jan moet z'n kamer gisteren hebben opgeruimd.
Jan must his room yesterday have cleaned
'Jan must have cleaned his room yesterday.'

As will be clear from Table 1, the two parameters cannot distinguish between dispositional and direct deontic interpretations. Perhaps this is simply a lexical ambiguity that cannot be reduced to the syntactic context of the modal. I leave this issue for future research.

6. How the complement determines modal interpretation

Table 1 captures the conclusion drawn in Section 3 that the epistemic-root distinction does not depend on the subject-orientation of the modal: a modal can be a monadic predicate and at the same time have a root interpretation. Under the assumption that the ambiguity of modals is not lexical, the complement of the modal is the only possible source of ambiguity. There are a number of differences between complements of epistemic modals and complements of root modals:

Table 2. Differences between epistemic and root interpretations in Dutch

<table>
<thead>
<tr>
<th></th>
<th>Epistemic</th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential polarity transition</td>
<td>not required</td>
<td>required</td>
</tr>
<tr>
<td>Scale of complement</td>
<td>negative--affirmative</td>
<td>0–1</td>
</tr>
<tr>
<td>(or: no–yes)</td>
<td>(number of events)</td>
<td></td>
</tr>
<tr>
<td>Category of complement</td>
<td>verbal only</td>
<td>all categories</td>
</tr>
<tr>
<td>Definite complement</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

The first three differences have been illustrated above. The fourth difference, the definiteness of the complement, is the key to understanding the influence of the complement on the interpretation of the modal. 'Definite complement' in this table means that the complement of an epistemic modal is a definite verbal argument, not a nominal argument, since it arguably is not. The complement of an epistemic modal behaves like a definite constituent on a number
of tests, whereas the infinitival complement of a modal in the root interpretations behaves like an indefinite constituent on the same tests.15 For example, indefinites but not definites can occur in the exclamative constructions in (22) and (23). These exclamative constructions disambiguate the modals. Modals can only have root interpretations here, so the infinitival complement of an epistemic modal patterns with definite constituents.

(22) a. Een auto dat Jan heeft!
    a car that Jan has
    'Jan has a very nice car!' (exclamative)

b. *Die auto dat Jan heeft!
     that car that Jan has

c. Werken dat Jan kan!
    work that Jan can
    i. 'Jan is able to work very hard.' root
    ii. '*It is possible that Jan is working very hard.' epistemic

(23) a. Wat heeft Jan een auto!
    what has Jan a car
    'Jan has a very nice car.'

b. *Wat heeft Jan die auto!
    what has Jan that car

c. Wat kan Jan eten!
    what can Jan eat
    i. 'Jan can eat very much!' (exclamative) root
    ii. '*It is possible that Jan eats very much!' epistemic

I propose to characterise this difference between infinitival complements of epistemic and root modals syntactically as in (24). Root modals take an indefinite verbal complement IndP, whereas epistemic modals take a definite verbal complement D,P. This D,P complement must be specified by Negation/Affirmation, as will be argued below.16

(24) a. root: [Dv Dv [ModP Modal [IndP [Ind ONE [vp V]]]]

b. epistemic: [ModP Modal [Dv NEG/AFF [Dv Dv [IndP [IndONE [vp V]]]]]

The IndP (Individuator Phrase) with (abstract) head ONE is the verbal counterpart of an indefinite noun phrase (or NumP), and provides the value 1 on the numerical 0–1 scale. In root interpretations with an infinitival complement, modals take an IndP, thus satisfying the requirement that the complement of a modal denote a value on a bounded scale, leaving the choice of the verb free.

I take an (abstract) verbal Dv to establish the semantic relation between subject and event, identifying the subject as the source, starting point, agent, possessor of the event (cf. Barbiers 1995). In traditional terms, Dv is responsible for assigning an external Theta-role to the subject in its Spec when the verb is transitive or unergative. However, I take it to be present with unaccusative verbs as well. Dv is the equivalent of abstract or little v (Larson 1988; Hale and Keyser 1993; Chomsky 1995; see Collins 1997 for the D-like properties of little v).

Due to the presence of Dv, the complement of the modal is definite in the case of the epistemic interpretation. If it is correct that Dv has to be present in the complement of an epistemic modal it follows that only verbal complements allow an epistemic interpretation. Verbs will always be dominated by verbal Dv. When Dv dominates a predicate of category A, P, N, the resulting constituent will be verbal too (cf. Hale and Keyser 1993). Only when Dv is absent can the resulting complement be non-verbal, and then epistemic interpretations are impossible.

Maximally generalising the lexical argument structure of modals, I assume that also in the case of the epistemic interpretation the complement should denote a value on a bounded scale. The presence of Negation/Affirmation in the complement of an epistemic modal is necessary to provide such a value. Since a definite constituent denotes an entity or an event/state, not a scalar value, it cannot be the complement of a modal when negation/affirmation are absent. Epistemic and root interpretations involve different scales, then, because in epistemic interpretations Negation/Affirmation provides the scale and in root interpretations IndP.

The claim that definite constituents without Negation/Affirmation cannot be the complement of a modal is supported by the behaviour of nominal complements of modals. Whereas indefinites and other constituents that provide a value on a bounded scale can be the complement of a modal, NPs with a definite determiner cannot.17

     Jan can all/thing
     'Jan is able to do everything/thing.'

b. *Jan kan het werk.
     Jan can the work

c. Jan moet een/*het vriendinnetje.
     Jan must a/thing
     'Jan wants to find a girl friend.'
As in the case of epistemic interpretations, the modal can take an NP with a definite determiner when negation or affirmation provides the required value:

   Jan can the work not/affirm
   'Jan is (not) able to do the work.'

b. Jan moet het vriendinnetje *(niet).
   Jan must the girl friend not
   'Jan does not like the girl friend.'

In this view, then, epistemic modals and the modals in the construction in (26) have in common that they take a definite complement that cannot provide a value on a bounded scale. This forces the presence of negation/affirmation.

The absence of a polarity transition both in the epistemic interpretation and in the definite nominal constructions in (26) (cf. Tables 1 and 2) can now be related to the nature of the complement. When the complement is definite, a polarity transition is impossible, whereas when the complement is indefinite, adjectival, prepositional or verbal, a polarity transition is possible. It is plausible that the impossibility of a transition is related to the fact that definite constituents have a fixed reference, while the possibility of a transition derives from (among other things) the variable reference of indefinite constituents.

7. The syntactic representation of subject-orientation

An epistemic modal is never subject-oriented (it is always monadic), whereas a root modal may or may not be subject-oriented. This follows straightforwardly from the structures proposed in (24). If it is correct that $D_v$ establishes the semantic relation between a subject and a verb, identifying the subject as the source, starting point, possessor, agent of what is denoted by the verb, it depends entirely on the base position of the modal whether a semantic relation will be established between the subject and modal or not.\(^{18}\)

(27) a. root: \[
[D_v \cap [\text{ModP} \cap [\text{ModP} \cap [\text{IndP} \cap [\text{IndP} \cap [\text{VP Verb}]]]]]]
\]

b. epistemic: \[
[\text{ModP} \cap [\text{ModP} \cap [\text{IndP} \cap [\text{IndP} \cap [\text{VP Verb}]]]]]]
\]

According to standard compositional semantics (cf. Heim and Kratzer 1998), in (27a) $D_v$ is a binary relation between subject and ModP, so the subject will be interpreted as the source of the modality (dispositional interpretation) or the possessor of the modality (direct deontic interpretation). In (27b), $D_v$ cannot be a binary relation between subject and ModP, because ModP dominates $D_v$. The fact that an epistemic modal is never subject-oriented is thus a direct consequence of the assumption that the modal selects $D_v$ (including the subject) in such cases.

Indirect root interpretations, in which the modal is a monadic predicate, may now arise in two ways: (i) $D_v$ is entirely absent and the modal selects NumP, PP or AP; (ii) $D_v$ is present, but the subject reconstructs at LF into a position within the scope of the modal, [SpecVP]. I will not discuss these possibilities any further here (cf. Barbiers 1995 for discussion).

The argument structure of modals at LF is summarised in (28). The lexical relational information associated with modals, whatever its precise content, must fully match the semantic relations determined by syntactic structure. A modal can be dyadic or monadic. In both cases, one argument must provide a value on a bounded scale. Class V gives rise to polarity effects and is discussed in the next section.

(28) I. Dispositional
   Relation: modal (entity, value on bounded scale)
   Example: mag ([Jan], [IndP schaatsen])

   may (Jan, skate)
   Jan mag graag schaatsen.
   Jan may eagerly skate
   'Jan likes skating.'

II. Direct deontic
   Relation: modal (entity, value on bounded scale)
   Example: mag ([Jan], [IndP schaatsen])

   may (Jan, skate)
   Jan mag van mij schaatsen.
   Jan may of me skate
   'I allow Jan to skate.'

III. Indirect deontic
   Relation: modal (value on bounded scale)
   Example: moet ([IndP de brief op tijd in Amsterdam zijn])

   must (the letter be in Amsterdam on time)
   De brief moet op tijd in Amsterdam zijn.
   the letter must on time in Amsterdam be
   'The letter must be in Amsterdam on time.'

IV. Epistemic
   Relation: modal (value on bounded scale)
Example: \[\text{mag } ([\text{NegP} \ [\text{DP} \text{ Jan net zijn kamer opgeruimd hebben}]])\]
may (Jan just his room cleaned have)
Jan mag net zijn kamer opgeruimd hebben, het is
Jan mag just his room cleaned have, it is
al weer een roombel.
already again a mess
'It may be true that Jan has cleaned his room,
but it's a mess again.'

V. Negative/Positive
Relation: modal (entity, value on a bounded scale)
Example: \[\text{mag } ([\text{Marie}, \ [\text{NegP} \text{ niet/wel}])\]
may (Marie, not/AFFIRMATIVE)
Jan mag Marie niet/wel
Jan may Marie not/AFFIRMATIVE
'Jan likes Marie/Jan does not like Marie'

8. Negative polar moeten and bipolar mogen

As we have seen in the previous sections, the modals moeten 'must' en mogen 'may' are polarity items when they occur with definite nominal complements (cf. 29), while they are not polarity items in their other uses. In this section it is argued that this polarity is not an accidental lexical idiosyncrasy, but follows from the general requirement that one of the arguments of the modal denote a value on a bounded scale. In the unmarked case, the scale is provided by the complement of the modal, and then moeten and mogen are not polarity items. A definite nominal complement cannot provide the required scale and therefore negation or affirmation must be present.

\[(29)\]
(a) Jan mag Marie * (niet/wel)^20
Jan mag Marie not/AFFIRMATIVE
'Jan does not like/likes Marie.'
(b) Jan moet Marie * (niet)/('wel).
Jan must Marie not
'Jan does not like Marie.'

This explanation entails that Negation/Affirmation can be one of the arguments of a modal and that the modals in (29) should not be analysed as simple transitive verbs. Maintaining the hypothesis that the relation between a verb and its 'external' argument is established by abstract D, (cf. Section 6 and 7), I propose the structure in (30) for this construction. According to standard compositional semantics, the modal in (30) is a binary relation between Negation/Affirmation and DP2 Marie.

\[(30)\] \[\text{[DWP } [\text{DP} \text{ Jan}] \ [\text{DVP } \Theta \ [\text{ModP niet/wel} \ [\text{Mod} \text{ mag/moet } ([\text{DP2 Marie} ] ) ] ] ] ]\]
Jan not/AFFIRMATIVE may/must Marie

The idea that Negation/Affirmation can be an argument of a verb is supported by the behavior of epistemic verbs (cf. Barbiers (2000) for a detailed analysis and more arguments). Note that the construction in (31a) cannot be analysed as a case of ellipsis since a full complement with a verbal core is impossible (31c). Put differently, van niet/wel and full CP-complements are in complementary distribution.

\[(31)\]
(a) Jan denkt/hoopt/gelooft van niet/wel.
Jan thinks/hopes/believes not/so.
(b) Jan denkt/hoopt/geloof dat Marie komt.
Jan thinks/hopes/believes of not/so that Marie comes.

Finally, the fact that moeten 'must' is negative polar while mogen 'may' is bipolar with definite nominal complements follows from an independent difference between the two modals: moeten 'must' is a universal quantifier, whereas mogen 'may' is an existential quantifier ([not may not] = must, just like [not one not] = all). It is a general property of universal quantifiers that they cannot be specified by wel (AFFIRMATIVE), whereas they can be specified by niet 'not'. This may be a tautological effect: wel indicates that a given value is the highest on a scale, which is redundant in the case of universal quantifiers, which denote the highest value on a scale themselves. Existential quantifiers can be specified both by affirmative and negative morphemes, as the following contrasts show.

\[(32)\]
(a) Niet/*Wel allemaal gingen we naar het feest.
not/AFFIRMATIVE all went we to the party
'We did not all go to the party.'
(b) Niet/*Wel alles heeft Jan gedaan.
not/AFFIRMATIVE everything has Jan done
'Jan did not do everything.'
(c) Niet/*Wel altijd regende het
not/AFFIRMATIVE always rained it
'It was not always raining.'
(d) Niemand/Wel iemand heeft Jan gesproken (maar niet veel gasten).
no one/AFFIRMATIVE someone has Jan spoken (but not many guests)
Thus, *moeten* is unipolar because it is a universal quantifier and universal quantifiers cannot be specified by an affirmative morpheme.

In sum, the negative polarity of *moeten* and the bipolarity of *mogen* are the result of the interaction between two properties: (i) the universal versus existential nature of *moeten* and *mogen* and (ii) the general property of modals that they require an argument providing a value on a bounded scale.

9. Conclusion

This investigation of the complements of modals leads to the conclusion that the ambiguity of modal verbs is primarily determined by syntactic and semantic properties of the complement of the modal.

Notes

* I thank Jack Hoeksema and Virginia Brennan for comments on earlier versions of this paper. The usual disclaimers apply.
1. I discuss Dutch data only. The construction also exists in Norwegian, which has the same restrictions on the type of complement as Dutch. German and Afrikaans also have the construction, but with more restrictions; these languages allow only a subset of the complements that can occur with a modal in Dutch and Norwegian.
2. In the epistemic interpretation the modal qualifies the truth of a proposition. The root interpretations involve obligation, permission, ability and will. See Lyons (1977), Palmer (1986) and Barbiers (1995) for discussion and classification of modal interpretations.
3. The ‘*’ sign means that the interpretation of the clause without an infinitive is not equivalent to the interpretation of the clause with an infinitive.

4. For reasons that I do fully not understand, the epistemic interpretation of *mogen* ‘may’ is only felicitous in a concessive context, which requires the presence of a consecutive clause.
5. Cf. Barbiers (1995) for an elaboration of this argument and for other arguments.
6. Interestingly, complements that yield ungrammaticality because they do not denote a scale become grammatical in the comparative

   (i) a. *Deze wijn moet oud
t this wine must old
   b. *Deze wijn moet (nog een jaar) ouder
t this wine must one more year older

In view of our generalisation, this must mean that the comparative denotes a value on a bounded scale. It is clear that a comparative introduces a scale with a lower bound, namely the degree of comparison. The upper bound must then be taken to be implicitly given, as in (1, b). The difference with the ungrammatical non-comparative cases is that here even implicit specification is impossible.

7. As is the case with selectional restrictions in other contexts, type-coercion arises when the complement does not satisfy the restriction that it denote a bounded scale. The effect of this is that the hearer will try to interpret as bounded any complement that denotes an unbounded scale in the unmarked case. Provided the proper context, this is possible for many unbounded predicates. Therefore, it is impossible to give a complete list of predicates that can be the complement of a modal.

8. *open* is an adjective here; a verb would have the infinitival form: *openen*.
9. For an overview of arguments against the raising—control analysis of the epistemic—root ambiguity, see the Introduction of this volume.
10. See below for a refutation of potential counterexamples.
11. The ‘*’ indicates that an interpretation is not available.
12. When the perfective complement denotes the completion of an event at some point in the past, this becomes a fixed property: once it has been established that John has cleaned his room at five o’clock, January 10, 1998, this remains true forever and cannot change anymore.
13. Contrary to what I claim in Barbiers (1995), it cannot be the infinitival form (the morpheme -en) that denotes the value 1. Finite verbs introduce the value 1 as well, if the diagnostic of modification by *half* is reliable:

   (j) Jan lachte half
   John laughed half

   (ii) Jan besefte maar half dat hij gekozen was
   John realised only half that he chosen was

14. For example, infinitival complements of modals may follow the finite verb in embedded clauses, unlike nominal complements.
15. The non-nominal nature of the complement may be determined by the verbal head or by functional elements.
16. Cf. Barbiers (1995) for an elaborated version of this idea within a compositional syntactic theory, and for more arguments.
17. Cases like Jan kan dat 'John can that' do not count as counterexamples because dat is not a definite determiner here but a predicate, as in Jan is dokter en Piet is dat ook, lit. 'John is doctor and Pete is that too'.
18. For reasons of space we have omitted Negation/Affirmation in this structure.
19. Obviously, it has to be specified in the lexicon that Dutch modals allow nominal complements. This may be considered an idiosyncracy given that in many languages modals cannot have a nominal complement. The claim in the main text, however, is restricted to the polarity of modals with a definite nominal complement. This is not a lexical idiosyncracy but follows from general properties of modals.
20. Since affirmative is the unmarked value of clauses, it may be implicit.
21. The structure abstracts away from linear word order.
22. The continuations in this and the following examples are necessary to make the variants with an affirmative morpheme acceptable. Such a continuation does not help in the negative polar cases.

References
