70 The Syntax of Modal Auxiliaries

SJEF BARBIERS

---

1 Introduction

Modality is a semantic concept and covers, *inter alia*, notions such as possibility, necessity, probability, obligation, permission, ability, and volition. Modality can be expressed syntactically by modal verbs, imperatives, verbal inflection, modal adverbs, and modal particles. This chapter is restricted to modality expressed by modal auxiliaries such as CAN, MUST, MAY, WILL, WANT, and NEED.[1]

A great deal of research on the syntax of modal auxiliaries has concentrated on the distinction between epistemic and root interpretations. Roughly, epistemic interpretations are a class of interpretations involving a speaker-oriented or, in the case of embedded clauses, matrix-subject oriented qualification or modification of the truth of a proposition, while root interpretations involve the will, ability, permission, or obligation to perform some action or bring about some state of affairs (cf. Lyons 1977 and Palmer 1986 for a fine-grained classification and description of modal interpretations). Sentences are often ambiguous between the two readings. In many Germanic and Romance languages the same set of modals can have both the epistemic and the root interpretations. An example is given in (1):

(1)

John must be home at six o'clock.	
Epistemic:	'(Given what I, the speaker, know), I conclude that John will be home at six o'clock.'
Root:	'John is obliged to be home at six o'clock.'

An important question is whether this systematic semantic ambiguity entails the existence of two different syntactic structures. One of the first answers to this question (Ross 1969a) is that modals with an epistemic interpretation are one-place predicates, taking the entire proposition as their complement, whereas root

modals are two-place predicates, i.e., they involve relations between the subject and the rest of the clause:

(2)

Epistemic:	must (John be home at six o'clock)
Root:	must (John, be home at six o'clock)

I take this idea as a starting point, as a great deal of work has been invested to translate it into syntactic terms.

There are different ways to represent the alternation between monadic and dyadic modality syntactically. A first option would be to try to reduce the modal alternation to similar alternations with verbs like *break* and *smoke*. However, it will be shown in that these are alternations of a fundamentally different type.

A second way to represent the alternation in terms of argument structure is to analyze root modals as control structures and epistemic modals as subject raising structures. This is the most common generative analysis of the root–epistemic ambiguity (Hofmann 1966 ; Ross 1969a ; Perlmutter 1970b and numerous studies since then). However, as shown in , there are a number of serious problems for this analysis.

Third, the modal alternation could be the result of generating modals in different base positions, a higher, possibly functional, position for epistemic modals and a lower, possibly lexical, position for root modals. This analysis is not entirely without problems either, as show.

A fourth way to analyze the modal alternation is to assume that epistemic modals, but not root modals, undergo movement at LF. This possibility is discussed in . On this approach, the monadic nature of the modal verb in epistemic interpretations is not established until LF.

A different perspective on the possibility of generating modals in different head positions is to look at how different types of complement of modals influence modal interpretation. This perspective is chosen in .

2 Argument structure of modals

1 Introduction 3 The structural position of modals 4 The complement of modals 5

Conclusion NOTES REFERENCES

2.1 Transitive–intransitive alternations

As is well known, English modals have a defective inflection paradigm and a distribution different from that of main verbs (see). Therefore, the question whether the dyadic–monadic alternation is reducible to other transitive–intransitive alternations does not seem to be relevant for English. It is a relevant question, however, for languages in which modals have many properties in common with main verbs, such as Dutch and German. This section shows that in such languages the modal alternation is not reducible to transitive–intransitive alternations either.

Transitive–intransitive alternations are found cross-linguistically with main verbs like *break* and *smoke* (3a–d). A common analysis of such alternations is that each of these verbs has only one lexical entry, which corresponds to the transitive variant. The intransitive variant is derived by a lexical operation that reduces transitive predicates to intransitive predicates by removing the external argument.

(3)

a.	John broke the window.
b.	The window broke.
c.	John smoked the eel.
d.	The chimney smoked.

In fact, *break* and *smoke* represent two different types of alternation: *break* is a transitive–unaccusative alternation and *smoke* a transitive–unergative alternation (cf. Abraham 2000). The two can be distinguished by using classic tests described in Burzio (1986) and Hoekstra (1984b). In Dutch, the unaccusative alternant selects the auxiliary BE in the perfect tense (4b) and its adjectival participle can modify a noun that corresponds to the surface subject of the unaccusative (4c). The unergative alternant selects the auxiliary HAVE (4e) and its participle cannot modify a noun that corresponds to the surface subject of the unergative (4f).

(4)

Transitive–unaccusative				
a.	Jan	breekt	de	ruit.
	John	breaks	the	window

(4)

b.	De	ruit	is/	*heeft	gebroken.
	the	window	is/	has	broken

c.	de	gebroken	ruit
	the	broken	window

Transitive–unergative

d.	Jan	rookt	een	paling.
	John	smokes	an	eel

e.	De	kachel	heeft/	*is	gerookt.[2]
	the	stove	has/	is	smoked

f.	*de	gerookte	kachel
	the	smoked	stove

If the root–epistemic ambiguity of modals were to involve reduction of dyadic predicates to monadic ones, then these tests would show that modals belong to the same class as *roken*'smoke', as (5a–c) illustrate. In the perfect tense HAVE is selected and the participle cannot modify a noun corresponding to the surface subject of the intransitive alternant:

(5)

Root–epistemic

a.	Jan	mag	eten.
	John	may	eat
	'John is allowed to eat.'		

b.	Eten	heeft/	*is	altijd	gemogen.
	eat	has/	is	always	may-Part
	'It has always been allowed to eat'				

c.	*het	gemogen	eten
	the	allow-Part	eating

However, the monadic alternant of the modal (5b) does not have an epistemic interpretation, which shows that the epistemic–root ambiguity is not an instance of the transitive–unergative alternation either. Moreover, for the epistemic

interpretation to be possible, the external argument does not need to disappear, as (1) shows. We can thus conclude that the epistemic–root ambiguity can neither be reduced to the transitive–unaccusative alternation nor to the transitive–unergative alternation.

2.2 The raising–control analysis of the epistemic–root ambiguity

A consequence of treating root modals as dyadic predicates is that they involve a control structure, if something like the theta-criterion holds. If a root modal assigns a theta-role to the surface subject DP and if a DP must receive exactly one theta role, it cannot be the case that root modals involve raising structures. If they did, the raised subject would receive a theta-role from both the embedded predicate and the modal. Therefore, root modals can be analyzed only as dyadic predicates if they involve control structures, which implies that PRO in the complement of the modal receives the external theta-role of the embedded verb, while the matrix subject receives its theta-role from the modal directly.

(6)

a.	[[$_{DP}$ John] must [$_{DP}$ John] work from nine to five]].	Raising & epistemic
b.	[[$_{DP}$ John]$_i$ must [PRO$_i$ work from nine to five]].	Control & root

The control–raising analysis rather elegantly captures the idea that root modals are dyadic predicates while epistemic modals are monadic predicates. However, many problems for this analysis have been noted in the literature. First, assuming a PRO-subject in the infinitival complement of the modal is problematic because infinitival complements without *te*'to' usually cannot have PRO as their subject, as the Dutch example in (7a) illustrates.[3]

(7)

a.	Jan	voelde	[zich/	*PRO	wegzakken	in de modder].
	John	felt	himself/	PRO	sink	in the mire
	'John felt that he was sinking in the mire.'					
b.	Jan	voelde	dat	hij	wegzonk	in de modder.
	'John felt that he was sinking in the mire.'					

(7)

c.	Jan	at [$_{sc}$	zich/	*PRO	ziek].
	John	ate	himself/	PRO	sick
	‘John ate so much that he got sick.’				

The infinitival complement in (7a) requires a reflexive as its subject if it is to express the same meaning as (7b). If this infinitival complement can be analyzed as a small clause (SC), the impossibility of PRO in (7a) is an instance of the broader generalization that SC-complements cannot have a PRO subject, as (7c) illustrates (cf. Stowell 1981).

A possible way out would be to argue that modals do not take small clauses as their complements, but larger constituents, e.g., CP or TP. For the Germanic languages there does not seem to be much evidence, however, that this is the case. C-elements such as complementizers are impossible in the complement of a modal, as is embedded *wh*. If infinitival *to* is a T-element, its obligatory absence in most modal contexts shows that the complement of a modal is not a TP either. Higher adjuncts such as evidential and modal adverbs cannot be part of the complement of a modal, which also suggests that complements of modals are smaller than CP.

However, it cannot be concluded from the situation in Germanic that it is generally impossible for modals to have a CP or TP complement. As Tomić (2002) argues, Macedonian has a set of modals that take subjunctive CP-complements. Incidentally, these Macedonian constructions should not be confused with cases in which modals in Germanic seemingly take a CP-complement, as in (8a). There is reason to think that CP in this construction is not a complement but an adjunct (cf. Bennis 1986). As with adjunct clauses, CP in this construction is opaque for extraction, as the Dutch example in (8b) illustrates.

(8)

(8)

	which	team	can	it	that	Netherlands	team	tonight	beats

A second problem for the analysis of root modals as dyadic control predicates is provided by sentences in which the matrix subject cannot be interpreted as the person being obliged or permitted to do something (Feldman 1986 ; Klooster 1986 ; Brennan 1993 ; Barbiers 1995):

(9)

a.	This	letter	must be in London	before five	o'clock.
	'It is necessary/obligatory that this letter be in London before five o'clock.'				
b.	Al	deze	kranten	mogen	weg.[4]
	all	these	newspapers	may	away
	'It is permitted to throw away all of these newspapers.'				

In these sentences, the modal clearly is a monadic predicate taking the entire clause as its argument, but (9a) can and (9b) must have a root interpretation. Obviously, it is the addressee of utterances like (9a, b) that has the obligation or permission to do something, but the crucial point is that this addressee cannot be expressed syntactically here. In addition, it has often been observed that root modals (in fact, modals in general) cannot be passivized.[5] This is unexpected if they are dyadic verbs, since dyadic verbs which assign an external theta role satisfy an important condition for passivizability (Burzio 1986). However, it should be noted that modals share this inability to occur in the passive with other stative transitive verbs; it may be that the crucial factor blocking passivization here is stativity.[6]

A third problem for the control analysis of root modals involves the behavior of expletive subjects (Thráinsson and Vikner 1995 ; Warner 1993). In some languages, e.g., English, expletive subjects occur with raising verbs but not with control verbs (10a, b), the explanation being that expletives cannot receive a theta-role but would get one in a control structure. If root interpretations always involved control structures, expletive subjects are predicted to be incompatible with root interpretations, contrary to fact, as (10c) illustrates.[7]

(10)

(10)

a.	There seems to be a man in the room.
b.	*There tried someone to call you.
c.	There must be a solution to this problem on my table this morning.

A fourth problem emerges when we look at languages with morphological case. Wurmbrand and Bobaljik (1999) claim that in Icelandic modals with a root interpretation behave like raising predicates with respect to quirky case. When a verb requiring quirky case on its subject is embedded under a raising verb, the surface subject may appear with quirky case (11a). When such a verb is embedded under a control verb, the surface subject must appear with nominative case (11b). Regardless of the interpretation, the subject of modal clauses always occurs with quirky case when the embedded verb requires this (11c–d).[8]

(11)

a.	Harald/%Haraldur virðist ekki vanta peninga.	Icelandic
	Harold-Acc/Harold-Nom seems not lack money	
	'Harold seems not to lack money.'	
b.	Haraldur/*Harald vonast til að vanta ekki peninga.	
	Harold-Nom/*Harold-Acc hopes for to lack not money	
	'Harold hopes not to lack money.'	
c.	Haraldi/*Haraldur ætlar að líka vel í Stuttgart.	
	Harold-Dat/*Harald-Nom intends to like well in Stuttgart	
	'It looks like Harold will like it in Stuttgart.'	
d.	Umsaekjandann verður að vanta peninga.	
	the-applicant-Acc must to lack money	
	'The applicant must lack money.' (In order to apply for this grant)	

A fifth problem involves modals taking a non-verbal complement. Barbiers (1995, 2002 ; cf. also) shows that Dutch modals can take a non-verbal complement, e.g., a PP or AP complement.[9] Modals with a PP or AP complement are subject raising constructions. Yet, they can only get non-epistemic interpretations.

Finally, Wurmbrand and Bobaljik (1999) show that A-reconstruction for scope, i.e., the subject taking scope under the modal, is possible for both root and epistemic modals in German. As is well known, A-reconstruction is possible with raising verbs but not with control verbs (cf. May 1985 , among many others). The crucial example is given in (12a), which shows that a root interpretation is compatible with wide scope of the modal.[10] Butler (2003) makes the same point for English (12b).

(12)

a.	Ein Österreicher muß das nächste Rennen gewinnen (damit Österreich	
	die Führung im Weltcup übernimmt).	German
	an Austrian must win the next race (in order for Austria to have the most gold medals)	
	i. 'It is necessary that an Austrian (whoever it is) win the next race.'	
	Scope: modal > subject	
	ii. #'There is an Austrian and it is necessary that he win the next race.'	
	Scope: subject > modal	
b.	Some philosophers must go to those seminars.	
	i. 'There are some philosophers which are required to go to those seminars.'	
	Scope: subject > modal	
	ii. 'It is required that some philosophers go to those seminars (as a condition of our being given money to run them).'	
	Scope: modal > subject	

In sum, the observations presented in this section pose serious problems for a control analysis of modals with a root interpretation, but not for a raising analysis. We have also seen that root modality comes in two types. In the first type, subject-oriented root modality, the surface subject has the permission, obligation, ability to perform some action, or achieve a certain state. In the second type, non-subject oriented modality, the event or state denoted by the embedded verb is allowed, permitted, or made possible, but the permission, obligation, or ability is not attributed to the surface subject. Consequently, it is not possible to distinguish root and epistemic modality in terms of theta role assignment to the surface subject.

Moreover, the evidence presented in the previous section suggests that modal auxiliaries are raising verbs, which do not assign a theta role to the surface subject. The question whether the two types of root interpretation can and should be distinguished syntactically is still open. If all modals are raising verbs and if in the subject-oriented interpretation the surface subject gets a theta role from the modal, subject-oriented root interpretations pose a serious problem for the theta criterion.

2.3 Prepositional arguments of modals

To conclude this section, I add some well-known observations on the argument structure of modals that any theory should capture. In Dutch, the source of permission or obligation can be realized as a PP introduced by the preposition *van*'of' (13a–c). The imperative clause in (13d) shows that it is really the modal that licenses the presence of a *van* PP.

(13)

a.	Jan moet van zijn vader naar huis gaan.
	John must of his father to home go
	'His father forces John to go home.'
b.	Jan mag van zijn vader naar huis gaan.
	John may of his father to home go
	'His father permits John to go home.'
c.	Jan hoeft van mij niet naar huis te gaan.
	John needs of me not to home go
	'I do not force John to go home.'
d.	Ga (*van mij) maar naar huis.
	go (of me) just to home
	'Just go home.'

Although these modals apparently have a theta-role to assign to the source argument, they are not able to license their case, witness the occurrence of *van*'of', the same preposition that shows up to license case on arguments of nouns.

The source argument cannot be licensed by modals expressing ability or will:

(14)

a.	??Jan kan van mij naar huis gaan.
	John can of me to home go
	'I allow John to go home.'
b.	*Jan wil van mij naar huis gaan.
	John wants of me to house go

The status of (14a) is not so clear. The impossibility of a *van* PP with *willen* 'want' in (14b) is presumably due to the fact that the source argument is already expressed by the subject. Whereas the source of permission and obligation is distinct from the possessor of permission and obligation, the source and possessor of will coincide. This also suggests that the modal auxiliary *willen* is not always a raising verb, a claim which is supported by the minimal pair in (15a, b). In addition, passivization and use as a prenominal modifier is possible with *willen* (15c, d).

(15)

a.	Jan/*Het wil dat Piet gaat.
	John/It wants that Pete goes
	'John wants Pete to leave.'
b.	Het/*Jan schijnt/kan dat hij gaat.
	it/John seems/can that he goes
	'It seems/is possible that he is leaving.'
c.	Dat wordt door veel mensen gewild.
	that is by many people wanted
	'Many people want that.'
d.	een gewild object
	a wanted object

Another intriguing case of PP arguments licensed by modality is found in so-called modal passives. When a PP with *voor* 'for' is combined with a modal passive, the modal passive can have only an ability interpretation. On the other hand, when a PP with *door* 'by' occurs, only an obligation interpretation is available (see 16a, b) (cf. Bennis 1990).

(16)

(16)

a.	een voor Jan op te lossen probleem
	a for John up to solve problem
	‘a problem that can be solved by John’
b.	een door Jan op te lossen probleem
	a by John up to solve problem
	‘a problem that must be solved by John’

In addition to raising questions about the relation between argument structure and modality, the modal passive construction in (16) shows that syntactic constructions may yield modal interpretations in the absence of any modal verb. This seems quite generally to be the case cross-linguistically: BE TO INFINITIVE and HAVE TO INFINITIVE combinations have a modal meaning in many languages.

3 The structural position of modals

1 Introduction 2 Argument structure of modals 4 The complement of modals 5 Conclusion NOTES REFERENCES

An alternative way to shed some light on the epistemic–root distinction is to investigate the structural position of modals. It turns out, however, that their position is rather difficult to establish; a simple and universal picture seems virtually impossible given the empirical facts. At least four types of distinction should be taken into account:

(i) the distinction between modal auxiliaries and modal main verbs;
(ii) the distinction between epistemic and root modality;
(iii) fine-grained distinctions within the class of epistemic interpretations and within the class of root interpretations;
(iv) the distinction between the position where a modal is interpreted and the position where a modal is pronounced.

The evidence currently available strongly suggests that cross-linguistic parameterization involves all four distinctions.

3.1 Different positions for modal auxiliaries and modal main verbs

The distinction between modals behaving like auxiliaries and modals behaving like main verbs has engendered a great deal of discussion. A well-known view is that modals in English are auxiliaries generated in Infl or Aux (cf. Roberts 1985) because they have a defective inflectional paradigm and behave like other auxiliaries in that they can undergo residual V2 and do not trigger DO-support in negative clauses. On this view, modals in languages like German and Dutch are main verbs projecting their own VPs, as their properties are largely (though not completely) identical to those of main verbs.

However, the distribution and properties of modals in Dutch suggest that Dutch modals can be used both as main verbs and as auxiliaries. Evidence that they can be used as main verbs was provided in the previous section. It was shown there that the volition modal *willen*'want' behaves like a transitive verb when it selects a CP complement. Obligation and permission modals optionally select a PP argument that expresses the source of permission or obligation, but ability and volition modals do not. Non-verbal complements of modals in languages such as Dutch, German, Afrikaans, Frisian, and Norwegian provide a further piece of evidence for the claim that modals can have their own argument structure (see).

The main-verb use of modals can also be distinguished from their use as auxiliaries by looking at inflection. Unlike English modals, Dutch modals have a participial form. When used as a main verb embedded under a perfective auxiliary, the modal occurs as a participle (17a). When used as an auxiliary, the modal appears as an infinitive instead of a participle (17b). This so-called Infinitivus-pro-Participio effect (IPP) is typical for auxiliaries, witness the contrasts in (17c, d) and (17e, f).

(17)

a.	Jan had dat best gekund/*kunnen.
	John had that best could-PCP/can-INF
	'John would very well have been able to do that.'
b.	Jan had dat best kunnen/*gekund doen.
	John had that best can-INF/could-PCP do-INF
	'John would very well have been able to do that.'
c.	Jan is thuis gebleven/*blijven.
	John is home stayed-PCP/stay-INF

(17)

	'John stayed home.'
	Jan is blijven/*gebleven werken.
	John is keep-INF/kept-PCP work-INF
d.	'John kept on working.'
	Jan had niet hoeven/*gehoeven te komen.
	John had not need-INF/needed-PCP to come
e.	'It was not necessary for John to come.'
f.	Jan had gezegd/*zeggen te komen.
	John has said-PCP/say-INF to come
	'John had said that he would come.'

Facts such as (17a) are a problem for the hypothesis that modals always have a verbal complement which may be a phonologically silent or a lexically empty verb (Vanden Wyngaerd 1994 ; van Riemsdijk 2002), as this hypothesis predicts the IPP effect to occur in such cases as well. The absence of the IPP effect cannot be ascribed to the fact that the verbal complement is not pronounced, given cases such as (18) where the IPP effect is obligatory despite the absence of an audible verbal complement.

(18)

Jan had de kamer mogen/*gemogen maar niet hoeven opruimen.
John had the room may-INF/may-PCP but not need-INF clean
'John was allowed to clean the room but he did not have to do it.'

It therefore seems to be necessary to specify in the lexicon that in languages like Dutch, modals are syntactically ambiguous between main and auxiliary verbs. This does not complicate the grammar in any way, as such specification is independently needed for other verbs. For example, the verbs *beloven* 'promise' and *dreigen* 'threaten' can be used both as main verbs (19a, c) and as epistemic auxiliaries (19b, d). Interestingly, as auxiliaries they do not just lack a participial form, they cannot occur in the perfect tense at all (19a, b).[11] In addition, in its auxiliary use *beloven* 'promise' cannot license an indirect object (19a, b). The possibility to be used as an epistemic auxiliary seems to be idiosyncratic, i.e., there

does not seem to be a way to predict that *beloven*'promise' and *dreigen*'threaten' can be used as auxiliaries while *zeggen*'say' en *verraden*'betray' cannot (19e).

(19)

a.	Jan heeft (ons) beloofd/*beloven te komen.
	John has (us) promised-PCP/promise-INF to come
	'John promised (us) to come.'
b.	Het belooft/*heeft beloofd/*heeft beloven (*ons) een mooie dag te worden.
	it promises/has promised-PCP/has promise-INF (us) a fine day to become
	'It promises (*us) to be a fine day.'
c.	Jan heeft gedreigd/*dreigen op te stappen.
	John has threatened-PCP/threaten-INF to leave
d.	Het dreigt/*heeft gedreigd/*heeft dreigen te gaan regenen.
	it threatens/has threatened-PCP/has threaten-INF to go rain
	'There's a threat of rain.'
e.	*Het zegt/verraadt een mooie dag te worden.
	it says/betrays a beautiful day to become

To conclude this section, there are two reasons to distinguish main verb and auxiliary modals. The first is the difference between English on the one hand and languages such as Dutch and German on the other. The second is the language-internal behavior of modals in languages like Dutch. In English, the syntactic distribution and behavior of modals patterns with other auxiliaries. In Dutch, modals can both be used as main verbs and as auxiliaries. Since English, German, and Dutch modals are all ambiguous between epistemic and root interpretations, the auxiliary–main verb distinction does not correspond to this ambiguity.

Language internally, the main verb–auxiliary distinction does not correspond to the root–epistemic distinction either. In (20a), the modal is used as a main verb but it has an epistemic interpretation. In (20b), it is used as a main verb with a root interpretation.

(20)

	Dat Jan te laat gekomen is, dat kan haast

(20)

a.	niet.
	that John too late come is, that can almost not
	'It is almost impossible that John was too late.'
b.	Jan had best naar huis gemogen.
	John had best to home may-PCP
	'John would have been allowed to go home.'

3.2 Different positions for epistemic and root modals

Although modals across and even within languages may differ with respect to their categorial status, it is clear that this distinction alone does not determine the structural position of modals. A complication is that languages may have two base positions for modals. Picallo (1990) claims that this is the case in Catalan. There is a higher IP position above clausal negation that corresponds to epistemic interpretations, and there is a lower VP position below clausal negation corresponding to root interpretations. Evidence that Catalan has two different base positions for modals, one for epistemic and one for root modals derives, among other things, from the interpretation of adjacent modals and the interaction of modal interpretations with the position of perfective auxiliaries. According to Picallo, in the Catalan sentence in (21) only the first modal can have an epistemic interpretation.

(21)

	En Pere due poder tocar el piano.	Catalan
	the Pere must can play the piano	
	i. 'It must be the case that Pere is able/allowed to play the piano.'	
	ii. *'It must be the case that it is possible that Pere would play the piano.'	

Although this may be true for Catalan, the restriction does not hold universally, as the Dutch sentence in (22a) (from Barbiers 1995 : 194) and the Icelandic sentence in (22b) (from Thráinsson and Vikner 1995 : 76) show:

(22)

a.	Het moet hem nog wel eens willen tegenzitten.	

(22)

	it must him once in a while want against-sit	
	'It is said that things go against him once in a while.'	
b.	það mun vilja rigna meðan þið eruðþar.	Icelandic
	it will want rain while you are there	
	'It will tend to rain while you are there.'	

There's a more robust generalization that seems to support the idea that epistemic modals are generated higher than root modals, namely the fact that root modals can be embedded under epistemic modals but not vice versa (but see the next section for a different interpretation of these facts). This is illustrated for Catalan in (23) (from Picallo 1990 : 294). The modal *deure* 'must' can only have an epistemic interpretation and therefore causes ungrammaticality when embedded under a root modal. The generalization also holds for Dutch (Barbiers 1995 : 194), Danish, and Icelandic (Thráinsson and Vikner 1995 : 98).

(23)

*En	Joan	pot	duere	tocar	el	piano.	Catalan
the	Joan	may	must	play	the	piano	

Picallo's second piece of evidence involves the interaction between modal and perfective auxiliaries. Perfective auxiliaries in Catalan can precede root modals but not epistemic modals, while both root and epistemic modals can precede perfective auxiliaries (Picallo 1990 : 293). This follows if root modals are generated lower than epistemic modals. The restriction also holds for Dutch (Barbiers 1995 : 197–198).[12]

(24)

a.	*En	Pere	havia	degut	venir.		Catalan
	the	Pere	had	must	come		
b.	En Joan	ha	pogut	anar	a	Banyoles.	
	'Joan	has been	allowed to	go	to	Banyoles.'	

3.3 Different positions for necessity and possibility modals, and further distinctions

Cormack and Smith (2002) argue that necessity/obligation modals are generated

in a position higher than negation and possibility/ability modals in a position lower than negation. This captures the facts in (25).[13]

(25)

a.	Alfred shouldn't eat nuts.	(deontic)
	'It is advisable for Alfred not to eat nuts.'	should > not
	*'It is not the case that it is advisable for A. to eat nuts	*not > should
b.	Bob shouldn't be late.	(epistemic)
	'It is predictable that Bob will not be late.'	should > not
	*'It is not the case that it is predictable that Bob will be late.'	*not > should
c.	Edwin can't climb trees.	(deontic)
	*'Edwin is able not to climb trees.'	*can > not
	'It is not the case that Edwin can climb trees.'	not > can
d.	Jean can't have left.	(epistemic)
	*'It is possible that Jean has not left.'	*can > not
	'It is not the case that it is possible for Jean to have left.'	not > can

Cormack and Smith's analysis has two important consequences. First, since both necessity and possibility modals precede negation at the surface, it must be the case that possibility modals raise across negation at PF. Second, since possibility modals scoping under negation can have an epistemic reading whereas necessity modals scoping over negation can have a root interpretation, there cannot be a syntactic explanation for the cross-linguistic impossibility for root modals to take scope over epistemic modals. Indeed, Cormack and Smith (2002 : 155) claim that *root > epistemic is a conceptual constraint. A point in favor of this claim is that modal main verbs in English cannot have root > epistemic scope either.

A culmination of distinctions between modal interpretations and corresponding syntactic positions can be found in Cinque (1999). Cinque proposes the syntactic hierarchy in (26).

(26) $Mod_{Epistemic}$ . . . > $Mod_{Necessity}$ > $Mod_{Possibility}$ > $Mod_{Volition}$ > $Mod_{Obligation}$ > $Mod_{Ability/Permission}$

It is clear that Cormack and Smith and Cinque agree that necessity modals are higher than possibility modals and that obligation modals are higher than ability modals. Three positions that Cinque adds involve volition modals and a distinction between epistemic and alethic modality. Evidence for the syntactic position of volition modals comes from Italian sentences such as (27a, b) (Cinque 1999 : 80), where the only possible scope relations are possibility > volition and volition > ability/permission, respectively.

(27)

a.	Gianni potrebbe voler uscire.	Italian
	'John could want to go out.'	
b.	Gianni vorrebbe poter cantare.	
	'John would want to be allowed to sing.'	

As for the second distinction, alethic modality semantically involves purely logical necessity or possibility, whereas epistemic modality involves speaker's deductions or opinions. Syntactically, epistemic modals are in a position higher than Tense, whereas alethic modals are lower than Tense. As evidence for the first point, Cinque provides, *inter alia*, the Una example in (28a).[14] According to Baker's (1985a) Mirror Principle, if X is a functional head generated higher in the structure than the functional head Y, X will occur in a position peripheral to Y when a complex word is derived by successive cyclic movement of the verb to the respective functional heads. Since *darib* in (28a) is to the right of the future-tense morpheme, it must originate in a higher position. Evidence that alethic modals are lower than Tense comes from the Hawick Scots sentence in (28b) (from Brown 1992). In this sentence *can* is in the scope of the future tense marker *will* and can be interpreted either in the ability or in the possibility sense.

(28)

a.	Er bin-kwan-de-darib.	Una
	she go-Fut-3Sg-PROBAB	
	'She might go.'	
b.	He will can do it.	Hawick Scots

In Cinque's theory, every head in (26) provides a Spec position for a corresponding modal adverb. This predicts that different types of modal adverb observe the same

hierarchy as the heads in (26). For evidence, see Cinque (1999 : chapter 4).

3.4 Different positions at PF and LF

As we have seen in the preceding section, the position in which a modal is pronounced should be distinguished from the position in which it is interpreted. This can be done in various ways, depending on the set of assumptions adopted. Cormack and Smith (2002) argue that a modal is inserted in the position where it is interpreted, while the position of pronunciation may be different as a result of PF movement. A second possibility is that modals are not generated in the position where they are interpreted, but that they reach their interpretive position by LF-movement.

It is common to assume that modal auxiliaries are quantifiers. Necessity/obligation modals can be shown to involve universal quantification over possible worlds, whereas possibility/ability/permission modals are existential quantifiers over possible worlds. If modals are quantifiers, they should be able to raise at LF. McDowell (1987) argues that this is the case for epistemic modals. Epistemic modals would differ from root modals in that they raise to a clause initial position at LF where they take scope over the entire proposition. According to McDowell, the incompatibility of epistemic modality with yes/no and *wh*-questions supports this analysis, as the epistemic operator and the *wh* or question operator compete for the same position within COMP. Whereas McDowell's claim may be correct for the incompatibility of epistemic necessity with yes/no questions (cf. 29b), epistemic possibility and yes/no questions are clearly not incompatible; see (29c, d). Moreover, Brennan (1993) shows that epistemic modals occur in *wh*-questions quite freely (see 29e, f).

(29)

a.	Jan	moet	wel	een	moedertaalspreker	van	het	Fins	zijn.
	John	must	AFF	a	native speaker	of	the	Finnish	be
	‘(John speaks Finnish fluently, so) John must be a native speaker of Finnish.’								

b.	*Moet	Jan	wel	een	moedertaalspreker	van	het	Fins	zijn?
	must	John	AFF	a	native speaker	of	the	Finnish	be
	Intended meaning: ‘Is it necessarily true that John is a native speaker of Finnish?’								

c.	Kan	Jan	een	moedertaalspreker	van	het	Fins	zijn?

	can	John	a	native speaker	of	the	Finnish	be
	'Is it possible that he is a native speaker of Finnish?'							

(29)

d.	Could the keys be on John's table? (Brennan 1993)
e.	What may he have done?
f.	Where must the murder weapon be hidden, in your view?

Butler (2003) is another attempt at capturing the scope of modals in terms of LF-raising. He argues that there are two scope positions for modals, one at the periphery of vP giving rise to root interpretations and one at the periphery of CP giving rise to epistemic interpretations. Both vP and CP are phases, hence propositional, and Butler's proposal thus captures Kratzer's (1977) insight that modals are propositional operators.

As Brennan (1993) shows, the interpretation of symmetric predicates embedded under a modal supports the claim that epistemic modals are in a higher position at LF than root modals. Brennan's example of a symmetric predicate is given in (30). The predicate is symmetric because (30a) implies (30b) and vice versa. When embedded under an epistemic modal, symmetry is retained (30c). When embedded under a root modal, symmetry is destroyed. This follows if only root modals are low enough to establish a relation between the subject and the embedded predicate.

(30)

a.	The governor shook hands with all the prisoners.	
b.	All the prisoners shook hands with the governor.	
c.	The governor may shake hands with all the prisoners.	Epistemic
	'It is possible that the governor will shake hands with all the prisoners.'	
	'It is possible that all the prisoners will shake hands with the governor.'	
d.	The governor may shake hands with all the prisoners.	Root
	'The governor is allowed to shake hands with all the prisoners.'	
	≠'All the prisoners are allowed to shake hands with the governor.'	

The scopal interaction between modals and other quantifiers could be taken to provide further evidence for the higher position of epistemic modals at LF (von Fintel and Iatridou 2003). Whereas quantifiers show scope ambiguities with root modals (31a), they do not show such ambiguities with epistemic modals (31b, c).

(31)

a.	Most of our students must get outside funding –		
	i.	for the department budget to work out.	must > most
	ii.	the others have already been given university fellowships.	most > must
b.	#Every student is the tallest person in the department.		
c.	#Every student may be the tallest person in the department.		
	i.	every student x (may x be the tallest)	
	ii.	may (every student be the tallest)	

The first interpretation of (31c) would make sense but is not available; the second interpretation does not make sense but is available. This shows that a quantified subject cannot raise over an epistemic modal whereas it can raise over a root modal. These facts would follow if the position of epistemic modals at LF is higher than the position to which the quantifier raises. However, von Fintel and Iatridou argue that this is not the right analysis, since even low possibility modals do not allow a quantifier to take wide scope. They conclude that a quantifier cannot bind its trace across an epistemic modal at LF.

4 The complement of modals

1 Introduction 2 Argument structure of modals 3 The structural position of modals 5 Conclusion NOTES REFERENCES

The syntactic representation of the epistemic–root distinction can also be studied from the perspective of the modal's complement. Relatively little attention has been paid in the literature to the influence of the modal's complement on modal interpretation. This section provides an overview.

4.1 Non-disambiguating complements

At first sight it might seem plausible to expect that modals with a root interpretation require verbs assigning an agent role as their complements. An agent may seem to be necessary to carry the permission, obligation, ability, or will to perform the action denoted by the embedded verb. Passive, unaccusative, stative complements and complements with a perfect might then be expected to force epistemic interpretations. In passives, the agent role is demoted, in unaccusatives there is no agent role, stative verbs denote a state of affairs, not an action that can be performed by an agent, and in perfectives the event denoted by the verb has already been completed. However, these types of complement all allow root interpretations:

(32)

a.	This room must be cleaned immediately.				Passive
	'You must clean this room immediately'				
b.	Je	mag	niet	weglopen.	Unaccusative
	you	may	not	away-walk	
	'You're not allowed to walk away.'				
c.	You must know the answer before five o'clock.				Stative
	'You're required to know the answer before five o'clock.'				
d.	You must have cleaned your room before five o'clock.				Perfective
	'You're obliged to have cleaned your room before five o'clock.'				

4.2 Complements that force an epistemic interpretation

Complements denoting a state or event that cannot change (anymore) force an epistemic interpretation (Barbiers 1995, 2002). There are two types of complement for which this is the case:

(33)

(33)

a.	provided that the subject has fixed reference.[15]
b.	Perfect complements, but only if the completion stage of the event has taken place in the past.

An example of both cases is given in (34):

(34)

a.	John must be a native speaker of Finnish.
	#'John has the obligation to be a native speaker of Finnish.'
b.	They must have cleaned this room yesterday.
	#'They must have had the obligation to clean this room yesterday.'

These two observations do not seem to have any consequences for the syntactic analysis of modal ambiguity, as disambiguation is the result of semantic properties of the complements. Nevertheless, the two types of complement in (34) are useful diagnostics to establish whether an epistemic interpretation is possible in a certain construction. An anonymous reviewer correctly points out that this test is not watertight, however. If an *if*-clause is added to sentences like (34), root interpretations are available (35).

(35)

a.	John must be a native speaker of Finnish if he wants to apply for the position.
b.	They must have cleaned this room yesterday if they want to go to the zoo today.

4.3 Complements that force root interpretations

For an epistemic interpretation to be available, it is necessary that the complement contain an infinitival verb. This can be seen in languages in which modals allow non-verbal complements, such as Dutch. Such constructions can have only root interpretations, as the examples in (36) show (from Barbiers 1995 : 153).

(36)

a.	Jan	mag	(dan)	weggaan	(hij zal ooit terugkeren).

(36)

	John	may	(then)	away-go	(he will once return)
	i. ‘John has permission to leave.				
	ii. ‘It may be true that John is leaving (but some day he will return).’				
b.	Jan	mag	(dan)	weg	(hij zal ooit terugkeren).
	John	may	(then)	away	(he will once return)
	i. ‘John has permission to leave.’				
	ii. *‘It may be true that John is leaving (but some day he will return).’				

As was noted in , the claim that such constructions involve non-verbal complements is controversial. Vanden Wyngaerd (1994) proposes that they involve PF deletion of an infinitival verb. Van Riemsdijk (2002) argues that the lexicons of languages that allow ‘non-verbal complements’ contain empty verbs that are licensed by modals. While there is good evidence that this claim is correct for Swiss German, the impossibility of an epistemic interpretation in cases like (36b) is a strong argument against an empty verb analysis for Dutch. For more arguments, cf. Barbiers (1995, 2002).

5 Conclusion

1 Introduction 2 Argument structure of modals 3 The structural position of modals 4 The complement of modals NOTES REFERENCES

In the past forty years, generative syntactic research on modal auxiliaries has focussed on the syntactic correlate of the semantic epistemic–root distinction. The original idea that epistemic modals are monadic predicates while root modals are dyadic has proved to be too simple, as monadic instances of root modals have been shown to exist as well. The influential analysis that takes epistemic modals to be raising predicates and root modals to be control predicates has turned out to be problematic too, given that monadic root modals behave like raising predicates on a number of tests. In current analyses, the leading idea is that epistemic modals are interpreted in a position higher in the clause than root modals. So far, it appears that there is not any conclusive evidence to decide whether this higher position is the result of base generation or LF-movement. Recent research also suggests that finer positional distinctions are necessary, such as a distinction between necessity and possibility modals, obligation and ability/permission modals, alethic and epistemic modals. Issues for future research include cross-linguistic

parameterization, such as the difference between main verb and auxiliary modals, and the question why the epistemic–root ambiguity occurs only with modal auxiliaries, not with modal adverbs.[16]

NOTES

1 Introduction 2 Argument structure of modals 3 The structural position of modals 4 The complement of modals 5 Conclusion REFERENCES

I thank Werner Abraham, Frits Beukema, Wim van der Wurff, and an anonymous reviewer for useful comments on previous versions of this chapter. All remaining errors are mine.

1 Caps indicate English modals and their equivalents in other languages.

2 I use an underlying object different from the one in (4d) to avoid complications caused by the fact that BE is also the perfect auxiliary of the passive in Dutch.

3 Note that this is true at the observational level. The theory-internal hypothesis that PRO cannot occur in governed positions is irrelevant for the problem under discussion.

4 See Barbiers (1995, 2002) for arguments that this construction does not involve ellipsis; see van Riemsdijk (2002) for the opposite view; see also .

5 The verb in the complement can be passivized if it has an external argument. The point at issue here is that the modal verb itself cannot be passivized.

5 The construction can also be found in Afrikaans, German, Norwegian, Yiddish, and West-Frisian.

6 An indication that modal auxiliaries are stative is the fact that they cannot occur in the progressive (i), like other stative verbs (ii), but unlike dynamic verbs (iii).

(i) *Jan is aan het kunnen schaatsen.

John is to the can-INF skating

#‘John is being able to skate.’

(ii) *Jan is het antwoord aan het kennen

John is the answer to the knowing

(iii) Jan is aan het schaatsen.

John is to the skating

'John is skating.'

7 In Danish, root modals do not seem to allow non-arguments such as expletives, weather IT and idiom chunks as their subjects. Nevertheless, Thráinsson and Vikner (1995 : 63–66) claim that in Danish both root and epistemic modals involve raising structures. They account for the impossiblity of non-arguments as subjects of root modals by assuming that root but not epistemic modals assign an 'additional' theta-role to the surface subject. This 'additional' theta role is subject to an 'additional' theta criterion that precludes the assignment of more than one additional role to an argument. This explains why embedded predicates that assign an additional theta role as well, such as BECOME and GET block the root interpretations. According to Thráinsson and Vikner, in Icelandic the epistemic–root distinction does straightforwardly correspond to raising and control structures. A less attractive aspect of this analysis is that it assumes different mechanisms for theta role assignment, and cross-linguistic parameterization of the availability of additional theta-role assignment.

8 Examples in (11) from Wurmbrand and Bobaljik (1999).

9 The construction can also be found in Afrikaans, German, Norwegian, Yiddish, and West-Frisian.

10 Example from Wurmbrand and Bobaljik. Pragmatic context is provided to favor the wide-scope reading of the modal.

11 I owe this observation to Gertjan Postma.

12 Barbiers (1995) follows Picallo's assumption that epistemic modals are generated in a higher position than root modals, but develops an analysis on the basis of a split-VP-hypothesis, in which what is usually taken to be little v is identified as a D-head. According to this analysis, root modals are generated below D and epistemic modals, on top of DP. The restrictions on double modals and the interaction between perfective auxiliaries and modal interpretation follow straightforwardly from this analysis. It is further shown that there are a number of parallels between modals with a nominal DP-complement and modals with a verbal DP-complement.

13 Importantly, Cormack and Smith distinguish three types of negation: echoic, polarity, and adverbial (i.e., VP) negation. The observations in (25) all involve scope relative to polarity negation.

14 Una is an Irian Jaya language; the example is from Louwerse (1988).

15 Names are DPs with fixed reference. In the unmarked case, the referent of a name is constant across time and across different possible worlds (but see section 4.4 for irrealis contexts).

16 I owe the latter question to an anonymous reviewer.

REFERENCES

1 Introduction 2 Argument structure of modals 3 The structural position of modals 4 The complement of modals 5 Conclusion NOTES

- Abraham, Werner (2000). The Aspect–Case Typology Correlation: Perfectivity and Burzio's Generalization. In: *Arguments and Case: Explaining Burzio's Generalization*. Eric Reuland (Ed.), 131–193. Amsterdam: John Benjamins.
- Baker, Mark (1985a). The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry* **16**: 373–416.
- Barbiers, Sjef (1995). The Syntax of Interpretation. PhD dissertation, Leiden University.
- Barbiers, Sjef (2002). Modality and Polarity. In: *Modality and its Interaction with the Verbal System*. Sjef Barbiers, Frits Beukema, and Wim van der Wurff (eds.), 51–74. Amsterdam/Philadelphia: John Benjamins.
- Bennis, Hans (1986). *Gaps and Dummies.* Dordrecht: Foris.
- Bennis, Hans (1990). TI: A Note on Modal Passives. In: *Grammar in Progress: GLOW Essays for Henk van Riemsdijk*. Juan Mascaró and Marina Nespor (eds.), 33–40. Dordrecht: Foris.
- Brennan, Virginia (1993). Root and Epistemic Modal Auxiliary Verbs. PhD dissertation, University of Massachusetts.
- Brennan, Virginia (1997). Quantificational Modals. *Linguistic Inquiry* **28**: 165–169.

- Brown, Keith (1992). Double Modals in Hawick Scots. In: *Dialects of English: Studies in Grammatical Variation*. Peter Trudgill and Jack K. Chambers (eds.), 74–103. London: Longman.
- Burzio, Luigi (1986). *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.
- Butler, Johnny (2003). A Minimalist Treatment of Modality. *Lingua* **113**: 967–996.
- Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York: Oxford University Press.
- Cormack, Annabel and Neil Smith (2002). Modals and Negation in English. In: *Modality and its Interaction with the Verbal System*. Sjef Barbiers, Frits Beukema, and Wim van der Wurff (eds.), 133–164. Amsterdam/Philadelphia: John Benjamins.
- Diewald, Gabriele (1993). Zur Grammatikalisierung der Modalverben im Deutschen. *Zeitschrift für Sprachwissenschaft* **12**: 218–234.
- Ernst, Thomas (1992). The Phrase Structure of English Negation. *Linguistic Review* **9**: 109–144.
- Feldman, Fred (1986). *Doing the Best We Can.* Dordrecht: Reidel.
- Fintel, Kai von and Sabine Iatridou (2003). Epistemic Containment. *Linguistic Inquiry* **34/2**: 173–198.
- Hoekstra, Teun (1984b). *Transitivity.* Dordrecht: Foris.
- Hofmann, T. Ronald (1966). Past Tense Replacement and the Modal System. In: *Notes from the Linguistic Underground*. James McCawley (Ed.), 85–100. New York: Academic Press.
- Klooster, Wim (1986). Problemen met Complementen. (Problems with Complements.) *Tabu* **16**: 112–132.
- Kratzer, Angelika (1977). What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy* **1**: 337–355.
- Louwerse, John (1988). The Morphosyntax of *Una* in Relation to Discourse Structure: A Descriptive Analysis. *Pacific Linguistics* Series B, 100. Canberra: Australian National University.

- Lyons, John (1977). *Semantics.* Cambridge: Cambridge University Press.
- May, Robert (1985). *Logical Form: Its Structure and Derivation*. Cambridge: MIT Press.
- McDowell, Joyce (1987). Assertion and Modality. PhD dissertation, University of Southern California.
- Palmer, Frank (1986). *Mood and Modality.* Cambridge: Cambridge University Press.
- Perlmutter, David (1970b). The Two Verbs *Begin*. In: *Readings in English Transformational Grammar*. Roderick Jacobs and Peter Rosenbaum (eds.). Boston: Ginn.
- Picallo, Carme (1990). Modal Verbs in Catalan. *Natural Language and Linguistic Theory* **8**: 285–312.
- Riemsdijk, Henk C. van (2002). The unbearable lightness of GOing. *The Journal of Comparative Germanic Linguistics* **5**: 143–196.
- Roberts, Ian (1985). Agreement Parameters and the Development of English Modal Auxiliaries. *Natural Language and Linguistic Theory* **3**: 21–58.
- Ross, John Robert (1969a). Auxiliaries as Main Verbs. In: *Studies in Philosophical Linguistics 1*. William Todd (Ed.), 77–102. Evanston: Great Expectations Press.
- Roussou, Anna (1999). Modals and the Subjunctive. In: *Studies in Greek Syntax*. Artemis Alexiadou, Geoffrey Horrocks, and Melita Stavrou (eds.), 169–183. Dordrecht: Kluwer.
- Stowell, Tim (1981). Origins of Phrase Structure. PhD dissertation, MIT.
- Thráinsson, Höskuldur and Sten Vikner (1995). Modals and Double Modals in the Scandinavian Languages. *Working Papers in Scandinavian Syntax* **55**: 51–88.
- Tomić, Olga Mišeska (2002). Modality and Mood in Macedonian. In: *Modality and its Interaction with the Verbal System*. Sjef Barbiers, Frits Beukema, and Wim van der Wurff (eds.), 261–278. Amsterdam/Philadelphia: John Benjamins.
- Warner, Anthony (1993). *English Auxiliaries.* Cambridge: Cambridge University Press.

- Wurmbrand, Susi and Jonathan Bobaljik (1999). Modals, Raising and A-Reconstruction. Technical report, Leiden University.
- Wyngaerd, Guido Vanden (1994). *Pro-Legomena, Distribution and Reference of Infinitival Subjects.* Berlin/New York: Mouton.