M.   Modals and functional categories

The English modals (as we have already established - see above) form a clearly defined class with distinctive syntactic, morphological and semantic properties. Given that in earlier stages in the evolution of English, many of these special properties were absent, it is clear that the modals have undergone what is known as a **diachronic reanalysis** (i.e. they have changed their syntactic status over time), the result of which has been to 'grammaticalise' them, transforming them from (relatively) normal lexical verbs into 'functional heads'. As elements that realise the head of the IP functional projection (or of *one* of the functional projections of the verb, for those who posit more than one), they are assumed to be directly generated in this position (i.e. in I^0) rather than 'raising' to it from a lexical projection (i.e. from a VP). In this respect they differ syntactically from the corresponding verbs in Romance and other Germanic varieties (which, of course, can appear in VP, with the appropriate infinitive or participial morphology). Their status as grammaticalised elements realising I^0 accounts for their deficient morphological properties (lack of infinitival and participial morphology) - these fall out as a natural consequence if we assume that these morphological elements are located in positions lower than I^0, positions that the English modals never occur in.

Given their status as functional heads, it is intuitively clear that the English modals function in some way as 'substitutes' for the mood inflections which we find in other languages whose verbal morphology is richer. As proof of this we can cite the evidence of the English version of 'subjunctive clauses' (subordinate clauses selected by intensional predicates): in these I^0 is filled not by an inflectional affix (as it would be in Italian or French) but by an overt modal verb (*should*) or else by an abstract modal operator - something which we have hypothesised is a 'null' (i.e. 'not pronounced') equivalent of a modal - and crucially also a word-level element, like the modals. In addition to this we may cite the case of unreal conditional sentences: in the apodosis clause a modal must be present ([*If Tom had applied*,] he **would** *probably have got the job* - the normal choice is *would* but certain other modals are possible) - this is exactly the syntactic environment where, in languages with morphology-based verbal systems, we would expect to find a conditional mood inflection (...*avrebbe ottenuto il posto*). Stated in the terms of section G above, the English modals are undoubtedly along the lines of what we might expect to find in a language that works on the principle of 'lexical insertion' (i.e. inserting word-level elements in functional categories) rather than using morphological suffixes.

From the semantic point of view, the English modals have a number of important properties that are consistent with this functional/grammaticalised status: (i) they are almost without exception 'transparent' in the sense of having no argument structure (i.e. they project no θ roles); (ii) they are associated (though not exclusively) with 'speaker-evaluation' (epistemic and deontic judgements) & with the accessing of unreal worlds and thus counterfactuality (inherently evaluative uses, which clearly overlap with epistemic and deontic modality); (iii) they have a problematic relationship with past time reference: what are traditionally considered their 'past tense' forms (*should, would, might, could*), in fact rarely (and in certain cases never) appear with past time reference. The past tense morpheme, when present (as it transparently is in these cases), is interpreted as realising a modal value [+ remote], rather than a temporal one (see sections K & L above); (iv)in pronunciation they may be phonetically reduced (/kən/ instead of the full form /kæn/).

It will be clear, then, that one major problem in analysing the English modals consists in bringing a (unique) series of syntactic and morphological properties (which can roughly be summarised as 'reanalysis as functional heads') into relation with a distinctive set of semantic

properties[1]. This means trying to understand to what extent the semantic properties can be held responsible for the change in the syntactic status of the verbs in question. Assuming that modals have undergone a 'diachronic reanalysis' (as functional heads), would we be right in thinking that this reanalysis was 'driven' by semantic factors? We will return to this question shortly.

Functional categories and their realisations

In this section we briefly explore the implications of the fact (which we observed above in our brief discussion of 'subjunctive clauses') that languages appear to differ in the sort of realisation they give to the heads of functional projections. Some languages (French, Italian, German) have a system based on **affixal heads**, i.e. elements that are not 'word-level' elements (free morphemes or X^0 elements) but rather affixes (bound morphemes or X^{-1} elements) and thus need to attach to a lexical root. In this type of system, the affixes in question are assumed to be generated in the head position of the functional projection (IP) and to 'attract' the lexical verb (or aspectual auxiliary), which systematically moves to this position (where it unites with - or 'incorporates into' - the affixes). This movement is what we have referred to as 'verb raising'. We will expect to find this system for realising functional categories in languages which have a rich morphological endowment (since it is the morphology - the affixes - that are responsible for triggering the raising). Other languages (the prime example being English) have a system that we have characterised as **lexical insertion**, meaning that values in the mood/aspect system will be expressed by 'free morphemes' (i.e. full lexical items) rather than affixes. In such a system, as we have seen, the free morpheme representing the mood value is inserted directly ('generated') in the head position of the functional projection (I^0). The result of this is that the position is filled (by a word-level element), and thus no verb raising will be needed or indeed possible[2]. This is the

---

[1] Of course, part of the work in bringing these two sets of properties into relation is already done thanks to the positing of 'functional projections', which are seen as the syntactic locus of important semantic choices such as Tense and Mood selection. The recognition that in English auxiliary verbs (when finite) and lexical verbs never occupy the same syntactic position, and the consequent need to recognise a higher position (originally called INFL but later rebaptised IP, "Inflection Phrase") for auxiliaries (verbal elements whose functional character is apparent) are the first steps in this direction. The positing of functional projections in any case provides an answer to the long-standing question of how to treat 'closed class' elements in general - besides auxiliary verbs, other closed class categories are determiners (*the, il, le* etc) and complementisers (*that, whether, if, que, che, se, dass, ob, oti, pou* etc). In the original formulations of X-bar Theory these were not recognised as projecting to full double-bar status, and thus basically remained unaccounted for. The positing of functional categories (IP for verbal elements, DP for determiner elements & CP for complementiser elements) removes this anomaly and provides a clear account of how these functional categories relate to the principal lexical categories.

[2] It is reasonable to expect that the lexical items which become specialised in this function (i.e. 'grammaticalised' as heads of functional categories and as expressions of a value in the mood system) will display a series of particular properties - especially if it is right to think of them as lexically incorporating an operator of mood. Notice that it has long been assumed that modal verbs are grammaticalised forms: they constitute a closed class (new modal verbs

system typically used by languages whose verbal systems are morphologically poor, having weakened or reduced inflectional paradigms. A further refinement of this system, which is suggested by our discussion of English 'subjunctive clauses', is that the inserted element may (in certain cases) have a null phonetic realisation. Thus we may obtain a structure where I^0 appears superficially to be empty but where a value in the functional system is nevertheless expressed. Summarising, we may say that languages have 3 different ways of dealing with their functional categories:

(i) realise them with affixal elements (X^{-1}) and move lexical items into the functional head position to unite with the affixes;

(ii) realise them directly with dedicated word-level (X^0) elements;

(iii) realise them with abstract elements (i.e. elements with no phonetic matrix).

Comparing IP with another functional domain, that of the determiner (DP), we may observe the same range of possible realisations:

(i) D^0 realised with affixal elements (triggering movement of N) - **Rumanian**
   [D^0 **-ul**] cesta frumos baiat
   [D^0 baiat-**ul**] cesta frumos
   (boy-the this nice - 'this nice boy')

(ii) D^0 realised with dedicated word-level elements - **English**
   [D^0 **the**] Roman people

(iii) D^0 realised with an abstract element (or perhaps not realised at all) - **Latin**
   [D^0 ∅] populus Romanus
   ('the Roman people')

The point about this hypothesis regarding (three-way) variation in the realisation of functional categories is that it brings together two important areas where languages are known to differ: (i) in the order of elements; (ii) in their morphological endowments.

Movement - i.e. change in the basic order of elements, with elements moving to positions in functional projections - is assumed to be triggered by morphological elements (often referred to as 'cues' for movement); absence of adequate morphological cues means that movement is not triggered and thus the basic order of elements in not changed. Some languages are endowed with sufficiently rich morphology (agreement morphology, tense aspect and mood morphology) to trigger movement consistently. These languages will consequently display a quite different order of elements from those languages whose

---

cannot be invented in the way that new lexical verbs can) characterised by well-defined shared morphosyntactic behaviour (obligatory occurrence in INFL, followed by infinitive of lexical verb without *to*, no agreement). They appear to have undergone 'semantic bleaching' (losing their original lexical meanings and the argument structures that accompanied this - as a result they no longer select NP or CP complements). At a certain point in the evolution of English they begin to appear systematically in environments where inflected Subjunctive forms had previously appeared. Intuitively, then, it is clear that modal verbs develop logico-semantic content similar to that associated with mood inflections and finally come to take the place of those inflections in the system.

morphological endowment is poor and insufficient to trigger movement.

This correlation between morphological richness and movement (or between deficient morphology and lack of movement) offers considerable explanatory potential. A strong hypothesis would be that *all* variation in the order of elements (from one language to another) is connected with the different realisations given to functional categories in those languages, and thus with their differing (morphology-based) capacities to trigger movement. In terms of a theory that distinguishes between fixed principles of linguistic structure common to all languages (i.e. the principles of UG) and areas where variation (albeit highly constrained variation) is possible (i.e. the 'Principles and Parameters Theory'), this amounts to saying that *parametric variation will be variation in the realisation of functional categories*. And, since, in a theory of this type, the task of the child learner consists principally in determining the values for the parameters in his/her language, it follows that his/her main task will in effect be that of determining the realisation of functional categories (which categories are instantiated in his/her language and by what system).

An important presupposition is that languages have the same basic range of functional categories, i.e that however many separate functional projections one ends up positing in the various functional domains C, I, D), these will be present - at least potentially - in *all* languages (which means that they will be included in Universal Grammar and thus in our genetic endowment). A degree of support for this idea comes from certain recent studies that have explored the I domain (i.e. the functional categories of V): it is well known that crosslinguistically this domain hosts (i) a variety of auxiliary verbs and morphological inflections corresponding to various semantic categories (modal, aspectual, frequency), and (ii) a variety of adverbs belonging to a limited number of semantic categories (modal, aspectual, frequency, evidential, evaluative etc). As will be clear from the lists just given in brackets, on a semantic level there is much in common between auxiliary verbs/inflectional systems and the categories of adverb that occur in the I domain: both auxiliaries/inflectional systems and I-domain adverbs seem to be chiefly concerned with expressing the same type of meanings:(roughly speaking) modal and aspectual. Given this, it seems not unreasonable to posit an interdependence between the two types of elements (auxiliaries/inflectional systems on the one hand and I-domain adverbs on the other). What is particularly interesting is that both types of element display relatively fixed orders cosslinguistically, with modal/mood elements (whether adverbs or auxiliaries/inflectional elements) preceding aspectual elements (whether adverbs or auxiliaries/inflectional elements). We illustrate this on the basis of English, where evidential adverbs (*evidently*) precede modal adverbs (*probably*), which in turn precede aspectual adverbs (*already*). Similarly modal auxiliaries precede aspectual auxiliaries:

> Tom may evidently quite probably already have left
> *Tom may quite probably evidently already have left
> *Tom may already evidently quite probably have left
> *Tom has may left

The high degree of replicability of such orders across languages strongly suggests that there is a relatively fixed endowment of functional categories (as a part of UG).

If - as we prospected above - one examines a variety of languages and assembles (on the basis of overt evidence in the form of inflectional systems, auxiliaries and adverb types) a crosslinguistic inventory functional categories - which presumably constitutes the range of

functional categories that UG makes available (for any given functional domain) - it is clear that in any given language only a subset of the functional categories listed in the UG inventory will have any sort of overt realisation for their heads, whether in the inflectional system (i.e. affixal realisation) or the system of auxiliary verbs (lexical-level realisation - see above). In other words, only a few of the possible functional categories will have corresponding elements (exponents of their heads) in the listed lexical-level (i.e. X^0) or sublexical (i.e. X^{-1}) material available in that language[3]. Seen from this point of view, language acquisition (syntactically a question of parameter setting) begins to look like a process of trying to establish correspondences between listed material (closed class lexical elements & sublexical material (affix systems)) on the one hand and functional categories (starting from the range of possibilities provided by UG) on the other. In other words, parameter fixing is a task that sees the learner wrestling with the listed material he has discovered (particularly those elements whose behaviour is visibly different from that of elements belonging to the major lexical classes) and trying to accommodate it in the functional categories that he 'knows' - on the basis of his UG genetic endowment- are theoretically available for realisation in his (or any) language. His aim is to establish which of these functional categories can be correlated with which inflectional paradigms or closed class lexical elements (including abstract ones!). In short, learning the syntax (which functional categories are realised and by what) and learning the lexicon (the listed material and how it is to be dealt with) go hand in hand and cannot be prised apart. What this means is that, since parameter setting is limited to functional categories and functional categories are dependent on the material that realises them, *there is no purely syntactic acquisition.* Acquisition of syntax is - to some extent - *driven by the lexicon* (the list of material, lexical and sublexical).

Language change and diachronic reanalysis
　　　　The basic fact about language change is that a given generation of speakers end up with a slightly different internalised grammar ('I- grammar') from that of the previous generation. Given that the raw linguistic data that the younger generation is exposed to (during their acquisition phase) is nothing other than the speech of the older generation(s), it is not at first sight easy to understand why they do not come up with exactly the same internalised grammar as their parents (the grammar that produced this speech, as it were). This is especially puzzling given the fact that language acquisition, driven as it is by exposure to raw language data, must necessarily be a highly deterministic process. In other words, assuming the learner is exposed to naturally occurring data in sufficient quantities, the outcome - in the sense of acquisition of an I-grammar capable of generating structures similar to those encountered - is to some degree guaranteed. Indeed, it is difficult to imagine how acquisition could function as efficiently as it does if it were not highly deterministic in this sense. But the other side of this coin is that if acquisition is 100% efficient, exposure to the raw data of speech of the older generation should - it seems - result in the learner acquiring *exactly the same I-grammar as speakers of the adult generation.* But this would mean that significant language change would be excluded, which is clearly contrary to fact.
　　　　Since syntactic change *does* occur, we may formulate the problem (the 'Logical

---

[3]We use the term 'listed' to suggest that both these series of elements are basically form-meaning pairs (like any lexical item) and as such should be thought of as being listed in the repository of such items, the lexicon.

problem of language change') as follows (using the formulation of Clark & Roberts (1993, 1994)):

> If the trigger experience[4] of one generation, say g_1, permits members of g_1 to set parameter p_k to value v_i, why is the trigger experience produced by g_1 insufficient to cause the next generation, g_2, to set p_k to value v_i?

What we have to have clear, however, is that the internalised grammars constructed during the process of language acquisition involve establishing mental representations of syntactic constructions. The learner has to (unconsciously) analyse and categorise the raw data to which he/she is exposed. Crucially, however, a given surface string may be compatible with more than one analysis (i.e. more than one mental representation in structural terms). In extreme cases, this may happen with very common syntactic structures. Thus a given syntactic string of elements may be common in the speech of a given linguistic community (for instance strings involving combinations of modal and lexical verbs in English) and yet be *analysed differently* by speakers of different generations. In other words, for speakers of the older generation a string such as *must work* is a token of one construction (for instance they give it a mental representation in terms of the abstract structure 'lexical verb + lexical verb'), while for the immediately following generation it is the token of another construction (they give it a mental representation in terms of the abstract structure 'functional head + lexical verb'). This shift in the abstract mental representation associated with a given type of syntactic string (which as a surface form remains constant across generations) is the basis of language change and evolution.

Of course, *something concrete* must have changed in the data to prompt the younger generation to construct a different mental representation from the one that the older generation gave to the same structural string. But it may not need to be very much. Indeed, even minor changes (from one generation of speakers to another) in the perceptibility of certain crucial grammatical elements - for instance inflectional morphology, which may weaken as a result of phonological change - may be enough to make it impossible for the younger generation to adopt the same mental representation as the older generation. In other words, a small change in the perceptibility of morphology in speech may make the adult grammar unlearnable (obscuring crucial information) and consequently *force* the younger generation of learners to make a different set of assumptions and thus fall back on a different mental representation.

There is a further refinement to this: since in syntax everything is interdependent, the way one element is categorised in our mental representation may depend (in the last resort) on the elements that surround it. Thus, while the problem we are trying to solve (viz. understanding why modal verbs ceased to be analysed as lexical verbs and began to be analysed as functional heads) appears to be centred on the modal verbs themselves, the crucial change that triggers this reanalysis may in fact be located elsewhere. Imagine a structure 'verb + verb', where the first verb is something like *must, should, can* etc and the second verb any member of the V class. What we want to know is what *stops* the younger generation from adopting the analysis of such a string as 'lexical verb + lexical verb' and

---

[4] By 'trigger experience' we mean naturally occurring examples - in the crude linguistic data produced by adult generations to which the child learner is exposed - of structures which constitute 'evidence' for the setting of a parameter to one value or another.

forces them to analyse it as 'functional head + lexical verb'. The higher verb (one of the series *must, should, can*..) may be a very atypical member of the V class (and indeed these verbs all had strange properties long before their reanalysis as functional heads) but this may not be enough - on its own - to force learners (when constructing their mental representation of such structures) to reject the possibility that the element in question is a lexical verb. It may be that this analysis only becomes untenable when something changes in the lower verb (a morphological change - see the following section) that rules out an analysis in which the higher verb has the status of a full lexical verb. What this amounts to is arguing that the English modals - with all their strange properties - owe their diachronic reanalysis not principally to those strange properties but to changes elsewhere in the system.

Functional categories and syntactic change

        Given the two contrasting ways of realising functional categories described above (affixal and lexical insertion), it is reasonable to suppose that (in a historical perspective of language change) a given language may - in the course of time - evolve from using one of them to using the other. In other words, a system based on affixal heads may cede its place to one based on lexical insertion. This sort of diachronic development is exactly what we find in English. At a certain point in its development, English (which had already lost all inflectional realisations of mood) undergoes a change, as a result of which the central modal verbs acquire the morphosyntactic properties that they display today. They become heads of functional projections, i.e. they are directly inserted in I^0 (or some other functional projection), rather than being raised to this position from some lower VP.

        The crucial point that needs to be addressed in the account we have just sketched out is the question of what *triggers* the change from an IP that functions through affixation to one that functions through lexical insertion. Here the semantic properties of the English modals, though undoubtedly significant, are arguably not in themselves sufficient to account for the change. After all, in languages that work on the inflectional affixation system (for instance Italian and French) we find 'modal verbs', in the loose sense of verbs whose lexico-semantic content is 'necessity' or 'possibility' and which are associated with epistemic and deontic judgements, as well as with unreal worlds and counterfactuality etc. We also find that modals are generally 'transparent' in the sense of not normally having argument structures (and where they do appear to allow structures that suggest some residual argument structure, such as *Non **ne** posso più* in Italian, it generally turns out that these structures are marginal and not fully productive). Roughly speaking, the modal-like verbs of inflection-based languages such as Italian and French have many (though perhaps not all) of the same semantic characteristics as the 'functional head' modals of English. And yet in these languages they have *not* been reanalysed as heads of functional projections as they have in English. They have pointedly *not* been singled out for this special syntactic status (with the syntactic and morphological properties that arguably follow from it). Even more significantly, in earlier varieties of English (when the modals had not yet undergone their diachronic reanalysis as functional heads) many of these semantic properties were already present but the modal verbs of that period behaved syntactically as lexical verbs (see data above). Conversely, even in contemporary English, certain modals retain what are arguably 'descriptive' readings (*When he was young Tom <u>could</u> swim very well* - compare: *Tom <u>was able to</u> swim very well*); the subjective/evaluative readings they are otherwise associated with do not have the complete monopoly one might expect them to have, although it remains the case that with those verbs which display a clear 'lexical split' (i.e. verbs like *need* that appear sometimes with the syntax

of modals and sometimes with that of lexical verbs) counterfactual readings[5] are only systematically available when the verb is treated syntactically as a modal. Compare *Tom needn't have redone all that work*, which is clearly counterfactual and thus necessarily evaluative rather than descriptive, and *Tom didn't need to redo all that work*, which would normally be interpreted descriptively (as simply describing a situation in which there was no need for Tom to redo the work - and as implying that he did not in fact redo it).

The above facts make it essential to take the analysis a stage further and provide some account of what brought this change about.

A proposal of this type regarding the English modals:
Roberts (1993/1999) accepts that the 2 important semantic factors in the diachronic reanalysis of the modals as functional heads were: (i) their association with epistemic modality & (ii) their precarious association with past time reference (as we have seen, there is every reason to believe that these two factors are closely connected). But he claims that *neither of these factors was sufficient to trigger the change that occurred at the time that it occurred* (= around 1500). Had these been the crucial factors, Roberts argues, then we would expect the change to have occurred much earlier than it did. This reflection allows Roberts to turn the problem on its head: thus - to use the words of Roberts & Roussou (1999) - 'we do not need to ask what caused the change; what we need to see is what prevented it from taking place sooner'. Rephrased in the more precise terms of Clark & Roberts (1993,1994 - see quotation above), this becomes:

> What was present in the trigger experience of acquirers until 1500 that provided 'robust evidence' for treating modals as lexical verbs (rather than as inflectional heads in I^0)?

It is important to understand that 'treating modals as lexical verbs' means analysing them as elements which are generated in a V head and subsequently raised to I^0, like any verb in contemporary Italian, French, German etc. This in turn means that a string consisting of a modal followed by a lexical verb - for instance *The kynge mote speken* ('The king must speak') - must be analysed as 'biclausal', i.e. as containing *two* full IPs, each with its respective VP. In other words:

> The kynge mote speken
> [$_{IP1}$ The kynge [$_I$ mote [$_{VP1}$ t$_{mote}$ [$_{IP2}$ [$_I$ spek-*en* [$_{VP2}$ t$_{spek}$-]]]]]

This representation shows *two* instances of verb raising: raising of the modal *mote* from VP1 (where it is generated) to the IP1 projection, and movement of the lexical verb root *spek-* from VP2 to the IP2 projection, where it is united with the infinitive suffix *-en* (the base positions of the two moved elements are indicated by t$_{mote}$ & t$_{spek}$- respectively).

What Roberts is looking for is some factor that *forced* speakers of the earlier

---

[5]The example that follows in the text (*Tom needn't have redone all that work*) would normally be understood as implicating that Tom did in fact redo all the work. The explicit content of the sentence is the opposite of this: 'Tom not redo the work'. The term 'counterfactual' is intended to capture this tension between the state of affairs as evaluated ('not necessary - Tom redo the work') and the real situation as implicated ('Tom redid the work').

generation to retain this representation and stopped them abandoning it in favour of a monoclausal analysis (involving one IP projection and one VP), in which the modal is generated directly in the functional head I^0 and the lexical verb simply remains in VP (which is how this sort of string is represented in contemporary English):

> The king must speak
> [IP The king [I must [VP speak]]]

Roberts & Roussou note that second representation (the contemporary English one) is much simpler than the pre-1500 one: the latter involves two instances of movement (of lexical material to a functional head), while the former involves no movement whatever (being a case of lexical insertion directly in I^0). Roberts and Roussou assume that movement always produces more complex - and less transparent - structures than lexical insertion, for the simple reason that what it leaves us with are 'hybrid' categories. The result of movement is an element that is a combination of lexical material (the lexical root) and functional material (the inflectional suffix)[6]. Lexical insertion, by contrast, is more transparent: the functional head is realised transparently by a functional element and the lexical material remains where it belongs (out of the way of any functional material). No hybrid category is created. Given, then, that the monoclausal representation of contemporary English is simpler[7] than that of pre-1500 English, child learners, who are assumed to be 'conservative' in the sense of being predisposed to choose the simpler representation, would - all other things being equal - have a strong reason for preferring it. So, *à plus forte raison* Roberts & Roussou see the problem as one of understanding what stops them (until 1500 circa) adopting the conceptually simpler representation and simply classifying modals as functional heads.

In Roberts' opinion the crucial factor *impeding* the change to the simpler representation was morphological, the retention of the explicit infinitive ending (*spek -en*). As long as infinitives occurred with this *-en* ending, there was clear evidence for the existence of the lower IP (IP2) and thus for a biclausal structure with *mote* in a separate IP (IP1). Clear evidence because the infinitive affix *-en* is itself a functional element, so any lexical verb that has combined with it must have already executed a movement from VP to IP, thus exhausting its functional potential. It follows that any further verbal material higher up (in our example the modal *mote*) cannot possibly be functional material relating to this same lexical verb; it must be something else (i.e. the lexical head of a whole new VP+ IP combination). Thus the presence of an overt infinitive affix on the lower verb forces the child learner (whose task is to construct a mental representation of all this) to assume that the higher verb (*mote*) is another lexical verb, not simply a functional head. It obliges him/her to adopt a biclausal representation (involving two instances of verb movement), despite its greater complexity and despite the fact that from the semantic point of view everything would favour the new analysis.

---

[6]Another instance of a hybrid category being created is the 'N into D' structure in Rumanian (mentioned earlier in the text): *baiat-ul cesta frumos* (boy-the this nice - 'this nice boy'). Here lexical material *baiat* ('boy') is incorporated into functional material *-ul* having typical functional content ('definiteness').

[7]Roberts & Roussou formalise this concept, speaking of a 'simplicity metric'.

But, once the infinitive ending weakens (first through loss of the final /n/ and then through loss of the vowel), everything changes: the crucial evidence that might have sustained the biclausal analysis for the current generation of learners is no longer there. The way is open for treating the modal as functional material relating directly to the lexical verb *speak*, i.e. the way is open for adopting the simpler (movement-free) analysis. In this simpler analysis we have only one IP, and no verb movement: the modal is directly inserted as head of the IP functional projection, and the lexical verb heads the VP that is the syntactic complement of this IP.

The new analysis involves the creation of new functional material: the modal verbs are 'grammaticalised', receiving an new syntactic status as functional heads.

A further merit of Roberts' account is that it alllows two descriptive facts to be correlated:

(i)   English is the only Germanic language *with* a syntactically defined class of modals;

(ii)   English is the only Germanic language *without* an infinitive ending.

What happens in ordinary finite declarative clauses in English?

Our analysis of subjunctive clauses (see above) has shown that English has a sort of abstract mood operator (i.e. an element with the same interpretational characteristics as modal *should* but no phonetic matrix) that can fill I^0 and give the same value as a [+ Subjunctive] inflectional morpheme in other languages. This abstract mood operator - presumably a word-level or X^0 element - is in complementary distribution with an overt modal realisation (using *should*), which is perhaps the preferred choice in British English. Crucially - as we saw - in Subjunctive clauses there is no verb raising (and this even extends to aspect/voice auxiliaries such as *be/have*).

Now, in not having verb raising Subjunctive clauses are basically no different from ordinary independent clauses in English (of the finite, declarative type). And there is a further resemblance: both environments allow explicit realisation as an alternative to an non-phonetically realised one. In the case of Subjunctive clauses (as we have seen) this involves using *should*, while in independent clauses the explicit element that appears is *do* (presumably a spell-out of the value [- mood], i.e. 'indicative'). This parallelism at the level of overt realisations - both environments have an overt element capable of realising I^0 and expressing a value in the mood system - suggests that we might posit a similar parallelism at the level of non-overt (i.e. null) elements. Thus, if Subjunctive clauses have an abstract [+ mood] operator that fills I^0 (and blocks movement of any verb to this position), maybe independent clauses, instead of just appearing with an empty I^0 position, also use an abstract operator, which fills I^0 in similar fashion to what happens in Subjunctive clauses and realises the value [- mood], blocking the I^0 position (as regards movement of other verbal elements).

This would give us a situation as follows:

Subjunctive clauses:

3a.   We insist [$_{CP(SUBJ)}$ that the boys [$_{INFL}$ $\varnothing$] [$_{VP}$ come home at three in the afternoon]

3b.   We insist [$_{CP(SUBJ)}$ that the boys [$_{INFL}$ *should*] [$_{VP}$ come home at three in the afternoon]

Independent clauses:

4a.   The boys [$_{INFL}$ $\varnothing$] [$_{VP}$ prefer swimming in the sea]

4b.	The boys [INFL *do*] [VP prefer swimming in the sea]

Summarising, the mood system is as follows:

Mood system (binary choice [+/- mood])
              null realisation   **overt realisation**
[+ mood] = insertion of null operator   *should*
[- mood] = insertion of null operator   *do*