1. Introduction

In this paper I want to discuss the analysis of auxiliaries and modals proposed by Butt, Niño and Segond (1996) and its implications for the theoretical status of f-structures in LFG. The discussion will be based partly on some Norwegian data and partly on more general considerations.

2. Butt, Niño and Segond’s analysis of auxiliaries

Butt, Niño and Segond’s analysis of auxiliaries and modals is presented first in Butt & al. (1996), and later in a slightly modified form in A Grammar Writer’s Cookbook from 1999 by themselves and Tracy King. They address the old question whether auxiliaries should be analysed as a special subset of main verbs or as special AUX categories of a more grammatical or functional nature – for instance in examples like the following:

(1) The driver will have turned the lever
    Der Fahrer wird den Hebel gedreht haben
    Le conducteur aura tourné le levier

As they point out, the traditional analyses within HPSG and LFG treat the auxiliaries as elements that are similar to main verbs, with their own PREDs and XCOMPs. Thus, the English sentence traditionally (i.e., since Falk 1984) gets a c-structure representation roughly like (2) and an f-structure representation roughly like (3):
The French example, however, expresses future tense morphologically, and therefore gets representations with one level less:

The authors’ proposal is to discard these traditional analyses and rather give the English, French and German examples the same, flat, f-structure analysis, in which the auxiliaries do not introduce their own PREDs or take their own XCOMPS, but rather are analysed as functional categories just contributing tense and aspect features to the f-structure. Accordingly the f-structure in (5) will do for all three languages. (The feature TENSE with an atomic value has been changed to a feature TNS-ASP with a complex value in accordance with the revision in the Cookbook.)

The authors’ arguments why we don’t need the traditional analysis are:

(i) VP-ellipsis and VP-topicalisation, which presuppose a syntactical hierarchy of the traditional kind, can still be handled in the c-structure, where the hierarchical information remains.

(ii) The restrictions on the form of the complement verbs (‘have’ takes past participle and ‘will’ base form in their respective complements, for example), hitherto handled straightforwardly in the f-structure by restrictions on XCOMPs, can be handled by a new projection called Morphological Structure orm-structure (see (7) below).

The reasons given by Butt & al. for opting for the flat f-structure analysis are (partly in my paraphrases):

(iii) Crosslinguistic evidence indicates that elements bearing only tense/aspect, mood or voice should belong to a distinct syntactic category.

(iv) The structural complexity of the traditional analysis is unmotivated and falsely indicates that there is a deep difference in predicational structure of auxiliaries like will and have on the one hand and the French aura on the other.

(v) The difference between the analyses of translationally corresponding structures is not helpful for MT.

(vi) Relegating the constraints on morphological forms to a separate t-structure is an advantage since the information does not really belong in the f-structure, being mostly unrelated to the grammatical relations and function-argument structure which are such stuff as f-structures are made of.

In m-structure the hierarchical information is preserved, and the morphological form of words and their dependents are represented:
The lexical entries for auxiliaries now specify the required form of their dependents in the m-structure rather than adding the same kind of restrictions to the f-structure.

In the *Cookbook* it is stated that the “flat” analysis only applies to the future and perfect auxiliaries, while modal verbs like German *müssen* and *können* are still given the traditional hierarchical f-structure analysis. Thus the analysis involves treating auxiliaries and modal verbs as fundamentally different kinds of categories.

### 3. Auxiliaries and modals in Norwegian

Before commenting further on Butt & al.’s analysis I would like to give a sketch of the corresponding grammatical phenomena in Norwegian.

Norwegian has a set of modal verbs which includes *ville*, *kunne*, *måtte*, *skulle*. We will consider some examples.

(8) a. Han *vil* dreie håndtaket
   *he* will/wants to Pres *turn.Inf* the-lever

b. Han *kan* dreie håndtaket
   *he* may/can/is able to Pres *turn.Inf* the-lever

c. Han *må* dreie håndtaket
   *he* must/is obliged to Pres *turn.Inf* the-lever

d. Han *skal* dreie håndtaket
   *he* is said to/has a duty to Pres *turn.Inf* the-lever

The semantic range of the modals is to some extent paralleled by the corresponding modals in English, French and German, but we may note the systematicity of the alternatives in Norwegian: Every modal can be interpreted either as a one-place epistemic modal or as a two-place root modal. Under the epistemic interpretations the subject referent is not an argument of the modal, which only takes the entire proposition as an argument: “It is going to be the case that/may be the case that/must be the case that he turns the lever.” Under the root interpretation the subject referent is an argument of the modal: “He wants to/is allowed to/is able to/is obliged to/has a duty to turn the lever.”

The epistemic meaning of *ville* comes close to ‘future tense’, but considering the systematic relationship between *ville* and the other modals, Norwegian grammar seems to classify this meaning as the epistemic counterpart of volition, i.e., as a modal rather than as atemporal kind of meaning.

Under the epistemic interpretation the modals meet a universal criterion of an auxiliary, namely, that it should not impose any semantic restrictions on the subject, as pointed out by Helge Lødrup (1996). Thus, the modals can occur with formal subjects, but then only with the epistemic interpretation:

(9) a. Det *vil* komme noen
   *it* will come *someone* ‘Someone will come’

b. Det *kan* komme noen
   *it* may come *someone* ‘Possibly, someone comes’
c. Det må komme noen
   it must come someone
   ‘It must be the case that someone comes’/‘May someone come!’

d. Det skal komme noen
   it is said to come someone
   ‘Allegedly, someone is coming’/‘I promise you that someone will come’

The modals and the perfect auxiliary can be embedded under each other in complex phrases. The perfect auxiliary in Norwegian is ha ‘have’ (and være ‘be’ with “ergative” verbs in the Bokmål variety). When a modal takes the perfect auxiliary as a complement, the reading of the modal is always epistemic:

(10) a. Han vil ha dreiet håndtaket
       he will have turned the-lever
b. Han kan ha dreiet håndtaket
   he may have turned the-lever
c. Han må ha dreiet håndtaket
   he must have turned the-lever
d. Han skal ha dreiet håndtaket
   he is said to have turned the-lever

However, the modals have infinitive and participle forms and can also be complements of the perfect auxiliary and of each other, and then usually with the two-place root meanings that take the subject referent as an argument. (11) shows modals as complements of the perfect auxiliary:

(11) a. Han har villt dreiet håndtaket
       he has wanted to turn the-lever
b. Han har kunnet dreiet håndtaket
   he has been able to turn the-lever
c. Han har måttet dreiet håndtaket
   he has been obliged to turn the-lever
d. Han skal ha dreiet håndtaket
   he has had a duty to turn the-lever

(12) exemplifies more complex cases:

(12) a. Han må ha villt dreiet håndtaket
       he must have turned the-lever
b. Han vil ha måttet dreiet håndtaket
   he will have had a duty to turn the-lever
c. Han må ha kunnet ville dreiet håndtaket
   he must have been able to want to turn the-lever

In the previous examples epistemic modals are never complements. Examples were they are seem possible, but then only as a complement of another epistemic modal, and most clearly before the perfect auxiliary ha:

(13) a. Han vil kunne ha reist imorgen
       he will may have travelled tomorrow
       ‘Tomorrow it will be the case that he may have gone away’
b. Han vil kunne reise imorgen
   he will be able to/may.Inf travel tomorrow
   ‘Tomorrow he will be able to go away’/’Tomorrow it will be the
case that he may go away’

c. Han vil ha kunnet reise imorgen
   he will have been able to travel tomorrow
   ‘Tomorrow he will have been able to go away’

From these syntactic facts it follows that epistemic modals only occur in finite forms
(present and past tense) and the infinitive, while the past participle is reserved for the root
modals.

How should these phenomena be analysed? Among the questions to be answered, are:

(14) (i) Should the epistemic and the root varieties be considered distinct lexemes or
alternative readings of single lexemes?

(ii) If the epistemic varieties are considered distinct lexemes, should they then be
classified as a subclass of verbs or as belonging to a functional ‘auxiliary’ category without
predicational content of their own?

Helge Lødrup (1996) discusses question (14.ii) and adduces several arguments why
the epistemic modals should be considered as a subclass of raising verbs and the root modals
as a subclass of control verbs. Among the things he observes is the fact that root modals can
appear not only with verbal complements, but also with NP objects:

(15) Jeg vil/kan/må/skal dette
   I want/am able to do/am obliged to do/have a duty to do this

He also observes that even the epistemic modals can have pronominalised complements – a fact which (as he points out himself) poses a slight problem for his analysis of them as raising verbs, but which on the other hand supports the assumption that the epistemic modals, too, are verbs:

(16) (Vil det regne?) Det vil det
   (Will it rain?) It will that

As for question (14.i), Lødrup seems to opt for homonymy rather than polysemy and
presuppose that the epistemic and the root varieties are distinct lexemes (although he is not
quite explicit on this point). Under such an analysis the epistemic modals would be slightly
‘defective’ verbs without past participle forms.

But the analysis of epistemic and root modals as distinct lexemes would give rise to a
puzzlingly systematic homonymy linking pairs of epistemic and root modals in Norwegian, a
systematicity which would then be unaccounted for. The formal identity of all morphosyntactic
forms which they both have, combined with their obvious semantic relatedness, would appear
accidental. In fact, the modal meanings are not simply ‘related’ – one might claim that there is
a semantic continuum linking the epistemic and the root meanings. Usages may frequently be
vague and difficult to classify along such a scale. Thus, kunne can be said to have ‘possibility’
as its central meaning, with ‘epistemic’, ‘deontic’ and ‘individual property’ as possible
further specifications. Deontic possibility would equal ‘permission’, and possibility as an
individual property would equal ‘ability’:

(17) a. Han kan være syk [epistemic]
he may be ill

b. Han kan komme inn [deontic: permission]
he may come in

c. Han kan svømme [individual property: habilitative]
he can swim

Contextual factors determine the possibilities – thus, permission and habilitative are possible only with agentive verbs. Furthermore, permission and habilitative associate ‘possibility’ specifically with the subject referent, with the result that these meanings correspond to two-place relations among the grammaticalised participants in the situation: the subject referent and a state of affairs. The epistemic meaning, on the other hand, only takes a state of affairs as argument among the entities referred to in the sentence. One might say that epistemic possibility is conceived as ability abstracting away from the able participant.

The situation with vīle and the other modals is quite parallel and provides no reason to give epistemic vīle a separate treatment as a special future tense auxiliary. To put it alliter impressionistically: one of the ways in which ‘future’ is expressed in Norwegian is as volition abstracting away from the willing participant. Grammatically this is not a tense category at all.

The indicated solution, therefore, is to bring the epistemic and the root meanings together by deriving the epistemic varieties from the root varieties by lexical rules operating on semantic forms and XCOMP constraints. The constraints on complements cannot be entirely relegated to a separate m-structure, since in Norwegian these constraints are not limited to morphological form. We also need to state, for instance, that the complements of root modals and the perfect auxiliary ha can only be root modals or main verbs and not epistemic modals or the auxiliary ha itself, while epistemic modals can take all kinds of complements.

The unorthodox aspect of this analysis will be the derivation of the epistemic relation from the root relation within the semantic forms – i.e., having lexical rules operate on the semantic relations themselves, which would hence have to be decomposed along the lines already sketched. Without going into this problem, let us assume that the perfect auxiliary, the epistemic modals and the root modals carry the features PERF, MOD1 and MOD2, respectively. We would then have lexical entries like the following. First, the perfect auxiliary ha would contain the specifications in(18):

\[
(18) \quad \text{ha} \quad \text{V} \quad (↑ \text{PRED}) = \text{‘PERF<XCOMP>SUBJ’} \\
(↑ \text{PERF}) = + \\
(↑ \text{SUBJ}) = (↑ \text{XCOMPSUBJ}) \\
(↑ \text{XCOMP FORM}) = \text{PASTPTC} \\
(↑ \text{XCOMP MOD1}) = - \\
(↑ \text{XCOMP PERF}) = -
\]

Each modal verb would have pairs of entries like (19a-b), in which the epistemic b-entry is assumed to be derivable from the roota-entry:

\[
(19) \\
\text{a. <modvrb> V} \quad (↑ \text{PRED}) = \text{‘ROOTREL<SUBJ,XCOMP>’}
\]

---

1 For simplicity the PRED introduced by ha is called “PERF” in (18), although this glosses over a semantic analysis not relevant to the present discussion. The Norwegian perfect is semantically very close to the English perfect, and less close to the French and German perfects, which can be used to refer to specific past times (“Ich habe ihn gestern gesehen” ‘I saw him yesterday’). The meaning of Norwegian (and English) perfect is neither deictic past tense nor perfective aspect, but rather non-referential relative past – the category existentially quantifies over times preceding the time indicated by the tense of the finite verb: “Jeg har sett ham” ‘I have seen him’ – ‘There exists a time in the past such that I saw him then’.
b. 

\[
\text{\textless modvrb\textgreater } \ V \quad (\uparrow \text{PRED}) = \text{\textquoteleft EPISTREL\textless XCOMP\textgreater SUBJ}\text{\textquoteright} \\
(\uparrow \text{MOD1}) = + \\
(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMPSUBJ}) \\
(\uparrow \text{XCOMP FORM}) = \text{INF}
\]

For a sentence like *Han vil ha villet dreie håndtaket* ‘He will have wanted to turn the lever’ we would then get ac-structure along the lines of (20)\(^2\) and an f-structure like (21):

(20)

---

\(^2\) In (20) functional categories are used in accordance with Bresnan (1996). I assume some language specific variation in the interpretation of the category I, head of IP: in English, I comprises the finite forms of the special class of AUX items, whereas in Norwegian I comprises the finite forms of all verbs, i.e., I = V: (\(\uparrow\text{FORM}\))=FIN.
4. f-structures, semantic representations and universality

The claim made here, then, is that this traditional type of analysis captures the linguistic facts of Norwegian better than the flat analysis suggested for at least English, French and German by Butt & al.: in Norwegian the perfect auxiliary and the epistemic modals have the properties of complement-taking verbs, and future time is grammaticalised by a modal verb as the epistemic counterpart of volition, and not as grammatical tense.

Still, it would no doubt be technically possible to provide an alternative analysis of the Norwegian constructions along the lines suggested by Butt & al., with a flatter f-structure in which epistemic ville is treated as a future tense auxiliary only contributing a tense feature, and ha is treated similarly as a perfect auxiliary only contributing an aspectual feature. Such an analysis certainly captures a certain relationship between the Norwegian constructions and the corresponding constructions in English, French, and German, not captured by the analysis proposed here. However, the alternative would carry different implications with respect to the theoretical status attributed to f-structures in LFG. Ultimately the chosen analysis reflects a certain view of what f-structures are meant to be.

So what are f-structures meant to be? In Bresnan (1996) Joan Bresnan discusses the principles of variability and universality across languages, and relates the principle of universality especially to f-structures:

(22)
The internal structure of a language represents the meaningful grammatical relations of sentences (how their syntactic functions are associated with semantic predicate argument relations); this structure is determined by generalizations about case government, pronominal binding, and agreement relations among the predicates and arguments of a sentence. The principle of universality states that internal structures are largely invariant across languages. The formal model of internal structure in LFG is the f-structure, ‘functional structure’. (1996:34 f.)

As we have already seen in Ch. 3, the principles of completeness and coherence require full representation of grammatical relations in f-structure. Full representation might be thought of as a universal iconicity requirement between syntax and semantics at f-structure. (1996:83)

The basic question to be asked here is whether the assumption that f-structure captures what is universally invariant is to be treated as a definition, i.e., as a stipulation dictating how f-structures should be constructed, or as an empirical hypothesis to be tested against f-structures constructed at least partly on independent grounds. I believe there are reasons to opt for some version of the second alternative.

Certain properties of f-structures are basic and probably beyond dispute:

(23) (i) F-structures abstract away from constituent order and to some extent from the hierarchical embedding relations of c-structures.

(ii) F-structures represent the predicates that are lexicalised and grammaticalised in the language and their complete set of linguistically expressed arguments, as well as the syntactical relations contracted by the constituents expressing such arguments.

Universality does not follow from (i) and (ii). Hence it emerges as an empirical question whether f-structures with these properties will also be universal in some interesting sense. We may here disregard the weak sense of ‘universal’ by which it simply means that f-structures are constructed within a universal format, i.e., using terms and formal properties that are not tied to particular languages but defined language-independently. Such ‘universality’ is true of c-structures as well and a precondition for even raising the question of
universality in a stronger sense. One such stronger sense which sometimes seems to be presupposed is the following:

(23) (iii) F-structures are universal in the sense that translationally corresponding expressions across languages are assigned the same (or closely similar) f-structures.

Universality – property (iii) – as an empirical hypothesis could then be the hypothesis that properties (i) and (ii) generally lead to property (iii) – something which does not follow logically and which would be an interesting discovery if true.

Universality as a stipulation, on the other hand, would mean that (iii) would be taken not as a hypothesis, but as one of the criteria to be met when grammars are written and f-structures constructed. One possible consequence would clearly be that it might sometimes be impossible to meet all three criteria at once. If we then let (iii) win over the other two criteria in such cases, we arrive at the situation which motivates my claim that taking universality as a stipulation is a bad idea. For then f-structures would be pure semantic representations, and their universality would be trivialised. Having identical f-structures for expressions in different languages would then just amount to stating the rather boring fact that the same things can be said in different languages; there would be no implied claim that the same things are also said in the same way on some level of abstraction.

F-structures are generally taken to be syntactic representations. A syntactic representation represents some of the properties of a linguistic expression that one has to refer to in order to justify that the expression is a well-formed expression of the language in question. Hence a syntactic representation cannot be universal by definition (in sense (iii) of ‘universal’).

A semantic representation, on the other hand, is exactly that: universal, or at least cross-linguistic, by definition (possibly restricted to a limited set of languages). If we assume that there is a discoverable relation of ‘literal translation’ among expressions of different languages, one could approach a comparatively theory-neutral characterisation of semantic representations based on such a translational relation. That is, rather than saying that a semantic representation denotes entities in a model, or cognitive structures, or some other highly theory-dependent objects, one could say that it denotes a set of linguistic expressions that is held together by a relation of literal translation. The language of semantic representations is then conceived as a kind of theoretical interlingua. Such a characterisation accords well with the way we normally treat semantic representations in a multilingual context, for instance in the context of machine translation. Thus, if we are only dealing with a set of closely-related languages, such as, say, Norwegian and Swedish, then our formal language of semantic representations need not draw very fine-grained distinctions of tense and aspect. Since the grammatical categories of the languages are in a very close correspondence with each other semantically, the semantic terms can be almost isomorphic with the grammatical ones and need not be much more fine-grained than they are. Include a significantly different language, however – such as Russian – and the semantic representations of Norwegian and Swedish expressions immediately need to be more fine-grained in order to capture the new set of translational relations. This common experience in the field of machine translation gets a principled basis if we assume that the task of semantic representations simply is to keep sets of translationally corresponding expressions apart – in other words, if we assume that the semantic analysis reflected in a semantic representation will always be implicitly or explicitly relative to a presupposed set of possible languages.

Hence, in the semantic representation of an expression e in a language L a given distinction means that such a distinction is drawn by lexical or grammatical means in some relevant language, but not necessarily in L itself, which may be more coarse-grained. Including more languages in the set of relevant languages may therefore lead to new distinctions being introduced in old semantic representations. In a syntactic representation of e,
on the other hand, a given distinction means that such a distinction is drawn by lexical or grammatical means in L itself. Hence the syntactic representation of e does not change with the introduction of new languages in the field of vision. I am therefore skeptical to the argument advanced for a certain f-structure analysis in the Cookbook by Butt & al., where they write:

(24)
This treatment of tense/aspect information was found to be inadequate as it was difficult to devise a standardized system that properly reflected the interplay between tense and aspect in all three languages. It was therefore decided to separate the dimensions of tense and aspect. (1999:74)

This kind of argument is perfect if f-structures are taken to be semantic rather than syntactic representations – but then, as we have seen, their universality has no empirical content, and furthermore there would be no representation showing how a particular language structures the temporal and modal content.

Another argument adduced by Butt & al. in favour of their flat f-structure analysis is that it facilitates machine translation (1996:2, 5). As far as I can see, the engineering advantage of flat, common structures could just as well be attained by deriving semantic representations alongside the f-structures. I would be a little wary of using the MT argument at the f-structure level, because I believe that such an argument runs the risk of undermining the basic motivation behind linguistic approaches to language engineering.

The assumption behind linguistic approaches to language engineering such as the PARGRAM project (as opposed to purely statistical approaches, for instance) is that in the long run linguistically motivated language descriptions will turn out to yield the most generalisable, robust and sophisticated practical solutions to a range of language engineering problems. If we consider this belief a hypothesis, the question arises what it takes to give it empirical content. At least one thing seems clear: if the hypothesis is not to be tautologically true and hence empirically empty, then the concept ‘motivated by linguistic considerations’ must somehow be distinct from the concept ‘motivated by engineering considerations’. Granted, considerations of language processing have been valuable sources of motivation for linguistic theories in the past couple of decades, and linguistic theories must obviously be allowed to be motivated also by some processing insights and still remain linguistic theories. Still, for the reasons I have discussed I suggest that the MT argument for the flat f-structure may be a case of favouring efficient processing of a limited set of cases by disregarding linguistic insights.

One might perhaps question my assumption that property (22.iii) – universality – does not follow from property (22.ii), which states that f-structures represent predicates and arguments – what has been called the predicational structure of an expression. Aren’t predicational structures of sentences universal, in a translational sense, so that translationally corresponding sentences are assigned the same predicational structure?

Not necessarily, if we take into account the way this concept is often understood in the context of f-structures. Bresnan speaks about “a universal iconicity requirement between syntax and semantics at f-structure” (1996:83). I take this to mean, intuitively speaking, that f-structure represents the particular way a given language carves up denoted reality. The format in which to represent this common reality in a language-independent way is the format of the semantic representations – in practice (I claim) only graspable denoting a set of translational relations among languages. The f-structure predicates must hence be analysable as complexes of the more basic predicates of the semantic representations, predicates which have been

---

3 It is a different matter that one’s syntactic meta-theory, and as a consequence of that one’s representations, may change with such a widened field of vision. That is the way research progresses and insight grows – it is a ‘once-and-for-all’ change which does not imply that a given syntactic representation continues to be relative to a given set of languages.
factored out by applying the ‘prisms’ of other languages to the linguistically encoded predicates of the f-structures. But the predicates of the f-structures themselves need not correspond one-to-one to each other in translationally corresponding sentences.

5. References


