Nor Jnl Ling 32.1, 137–164 © Nordic Association of Linguists 2009 doi:10.1017/S0332586509002066
Printed in the United Kingdom

Paggio, Patrizia. 2009. The information structure of Danish grammar constructions. *Nordic Journal of Linguistics* **32.1**, 137–164.

# The information structure of Danish grammar constructions

Patrizia Paggio

This paper addresses the issue of how information structure can be accounted for in a formal grammar of Danish. Three information structure features – topic, focus and background – are discussed, and it is shown how they are instantiated in a number of different grammatical constructions from a corpus of spoken Danish. Prosodic, syntactic and information structure constraints characterising the various constructions are represented as typed feature structures following Head-driven Phrase Structure Grammar (HPSG), and the constructions themselves are ordered in a type hierarchy. The proposed approach modifies and extends earlier HPSG-based accounts by integrating information structure as a dimension of phrasal and clausal grammar constructions.

**Keywords** Danish grammar, focus, information structure, topic

*Patrizia Paggio, Københavns Universitet, Center for Sprogteknologi, Njalsgade 140–142, bygning 25, 2300 Københavns S. Denmark. paggio@hum.ku.dk*

## 1. INTRODUCTION

The purpose of this study is to develop an account of Danish information structure which combines insights coming from the functional approach to information structure represented mainly by Lambrecht (1994) with a formal framework inspired by work on grammatical constructions in Head-driven Phrase Structure Grammar (HPSG) (Ginzburg & Sag 2000). The approach I propose builds on previous accounts of information structure within HPSG to be found in Vallduví & Engdahl (1996), Engdahl & Vallduví (1996) and Wilcock (2005).

Information structure is treated as an integral part of grammar, and formalised as a hierarchy of feature structure types which interact with grammatical constructions. This interaction is constrained by prosodic and syntactic features depending on the construction type. It is cast in the HPSG framework, which lends itself well to the task since it allows the conflation of syntactic, semantic and pragmatic features into a single complex feature structure.

The view of information structure on which the proposal is based is introduced in section 2, which also gives examples of information structure types in Danish. Topic-comment and topicless sentences are then discussed in sections 3 and 4, where I address the issue of how various grammatical sentence types relate to the two kinds

of information structure. Section 5 deals with focus, in particular the relation between accenting and focusing. The main claim of this section is that focus projection in standard Danish does not rely on the prominence of one sentence accent, but must be phrased in terms of projection from the rightmost accent. In section 6 a formal account of how information structure can be included in an HPSG grammar of Danish is presented. This is done by defining the interaction of information structure with the semantic interpretation of the sentence on the one hand, and with phrasal as well as clausal syntactic projections on the other. In addition to accounting for the empirical data discussed in the paper, it is argued that the present approach overcomes a problem inherent to earlier proposals, which failed to deal correctly with the difference between all-focus sentences and topic-comment sentences with wide focus. Section 7 sums up and offers suggestions for future work.

## 2. CATEGORIES OF INFORMATION STRUCTURE

Information structure concerns the organisation of the information conveyed by the sentence, in other words, the pragmatic function of the various sentence elements in building up discourse information. The literature on information structure is vast and the terminology used in different linguistic traditions is not uniform. However, it is fairly uncontroversial to posit two basic pragmatic dimensions, one of which, topicality, deals with the identifiability of discourse referents, while the other, focality, concerns the fact that one or more elements of an utterance are emphasised to indicate a modification of the pragmatic information built by the overall discourse.

Based on focality, the sentence can be divided into a FOCUS (abbreviated *F* in the following examples) and a BACKGROUND (abbreviated *BG*). Focus is often described as the NEW information. However, as shown by Lambrecht (1994:257–263) among others, this is too simplistic a view: what is new is not necessarily the information itself (for instance the entity that a focused NP refers to), but the relation established between this information and the rest of the sentence. For instance in example (1), *din søster* 'your sister' cannot be considered new information in itself: what is new, or rather non-presupposed, is its participating as an object in the meeting event. At least in one of the possible readings of the sentence, the subject and the verb are both background information.

(1) Mødte du nogen i går? Jeg mødte [$_F$ din søster].
*met you anybody yesterday I met your sister*
'Did you meet anybody yesterday? I met your sister.'

Different languages have different indicators of focality. One that is often mentioned in the literature is sentence accent; others may be a certain position, for instance final position, or morphological marking. As shown in example (1), identifiable sentence topics may occur in the focus domain.

It is often claimed that one sentence topic in the background part of the sentence has a privileged status in terms of information structure. This primary topic – or TOPIC for short (abbreviated *T*) – can be characterised as the entity which the sentence expresses pertinent information about, thereby increasing the hearer's knowledge of it. For instance in (2), which could be a continuation of (1), the subject pronoun refers to the topic (coreferential with 'your sister' in the preceding sentence).

(2) [$_T$ Hun] [$_F$ havde travlt].
*she had hurry*
'She was in a hurry.'

Let us now consider the way in which different information structures are expressed in a number of constructions in Danish. Consider the sentences in (3).

(3) a. (Hvad lavede børnene?) [$_T$ De] [$_F$ spiste is].
(*what did children.DEF*) *they ate icecream*
'What did the children do? They ate icecream.'

b. (Hvad spiste børnene?) [$_{BG}$ [$_T$ De] spiste] [$_F$ is].
(*what ate children.DEF*) *they ate icecream*
'What did the children eat? They ate icecream.'

c. (Hvem har spist isen?)[$_{BG}$ Det har] [$_F$ børnene].
(*who has eaten icecream.DEF*) *that have children.DEF*
'Who has eaten the icecream? The children did.'

d. (Hvad skete der?)[$_F$ Børnene spiste is].
(*what happened there*) *children.DEF ate icecream*
'What happened? The children ate icecream.'

All four answers denote the same event – the fact that the children ate icecream. However, their information structures differ. The answer in (3a) is a topic-comment sentence, i.e. a sentence in which something is predicated of the topic (typically the subject of the sentence): the topical status of the subject is indicated by its being an unaccented pronoun. There is no other element in the background than the topic itself, and the focus spans over the entire predicate. This construction has also been called a predicate-focus structure. Another term used in the generative tradition is wide focus. However, sentences with wide focus do not necessarily have a topic, therefore I will call this construction a *topic-focus* construction. Note that the topic, as will be seen in relation to topicalisation, is not necessarily the subject, and the focused part of the sentence does not always coincide with the verb and all its complements.

In (3b), the subject is still the topic, but the focus domain is restricted to one of the complements of the verb. The construction is still a topic-comment structure, but it has narrow focus. I will call this construction *topic-focus-bg*. In practice it is difficult to distinguish between wide and narrow focus out of context, and often even the context is too vague to clearly single out one of the two interpretations.

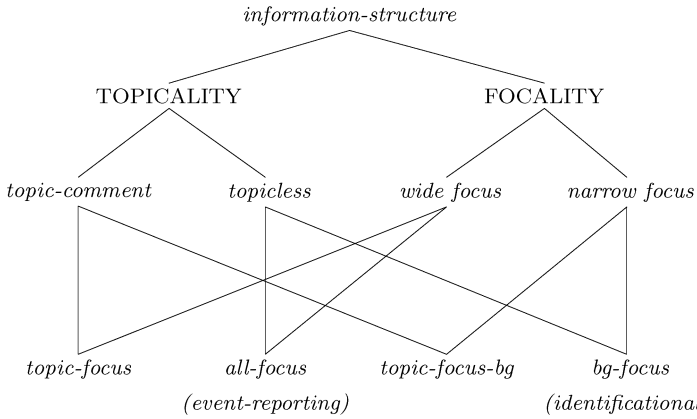


**Figure 1. A hierarchy of information structure types.**

In (3c), the information structure is inverted in that the content expressed by the topicalised object and the verb is presupposed, while the subject in post-verbal position is the focus. There is no topic. Note in fact that the subject pronoun refers not to the icecream, but to the open proposition 'eat icecream', and an equally good answer to the same question would be *Børnene* 'The children', where only the focus is mentioned. Lambrecht calls this construction identificational, i.e. a construction that 'serves to identify a referent as the missing argument of an open proposition' (Lambrecht 1994:122). He argues convincingly that the open proposition cannot be considered a topic since it is semantically incomplete and cannot therefore be said to have a referent. I will call this construction *bg-focus*.

Finally, the sentence type in (3d) has received several names in the literature, e.g. thetic or event-reporting sentence. The whole event denoted by the sentence is in focus, so I will call it an *all-focus* construction (although as I have already noted, the semantic arguments participating in the 'new' event, in this case the children, may be recoverable information). As in (3a) and (3b), the information structure is not marked formally by syntactic or prosodic means, but emerges from the context. I will return to syntactic and prosodic marking of information structure in the following sections.

The four information structure types illustrated in (3) can be organised in a hierarchy as shown in Figure 1. We will see shortly how the information structure types interact with grammar constructions.

## 3. TOPIC-COMMENT SENTENCES

In his work on the grammar of Danish, Togeby (2003) identifies the topic with the constituent in the foundation field, which is the position preceding the finite verb. The foundation is occupied by the subject in a non-inverted sentence or by the extracted

constituent in an inverted one (Diderichsen 1946). In other words, the topic is defined by its position like the theme in Halliday (1967:212). Hansen & Heltoft (2000) are more cautious, in that they claim that the foundation is not necessarily identical with the topic, although a subject in the foundation is the prototypical topic, followed by the comment in the predicate.

On the basis of empirical data from spoken and written Danish corpora, Thomsen (1998) argues that the initial constituent is very likely to be the topic expression. He notes in fact that out of 205 occurrences, 90.78% are pronominals. Additional evidence of the fact that the foundation field is reserved for topical information is provided in this study by the fact that when the initial constituent is long (defined in Thomsen's study in terms of number of words), it tends to be left-dislocated and reduplicated by a pronoun in the foundation field, as in (4).

(4) den urtepotte den vil vi placere i midten af stuevinduet
*that flowerpot that will we put in middle.DEF of living-room.window.DEF*
'That flowerpot we will put in the middle of the living room window.'

This example and a number of others in this paper are taken from the Danish corpus DanPASS (Grønnum 2005), and in particular from a series of monologues in which speakers instruct interlocutors in how to assemble the drawing of a house from pre-existing pieces.

The average length of the dislocated constituent in the examples examined by Thomsen is 4.2 words compared to 1.068 words for the foundation. Following Lambrecht (1991:188), the left-dislocated constituent can be considered a topic-announcing expression, which serves the purpose of bringing a discourse entity to the foreground so that it can be referred to as the topic immediately afterwards.

The fact that sentence position is a strong indication of topicality in Danish is confirmed by the data presented in Paggio, Diderichsen & Elming (2004). In this study, a fragment of the Danish BySoc corpus (Henrichsen 1998) was annotated by two independent subjects: the task was to tag all nominal sentence topics based on a subjective ' aboutness' criterion. It was then found that subject NPs in sentence initial position had the highest probability of being perceived as topics in both annotated data sets. Similarly in Diderichsen & Elming (2004), it is found on the basis of annotated Danish dialogue data that there is a strong tendency for definite pronouns occurring in pre-verbal position in main clauses to be perceived as topic expressions (around 80% precision and 60% recall).

From a syntactic point of view, a sentence displaying a topic-comment information structure can be constructed in Danish either (i) as a declarative main clause with subject-verb word order, where the topic coincides with the subject, or (ii) as a topicalisation construction characterised by subject-verb inversion, where the topic corresponds to the topicalised complement. Both possibilities are illustrated in the examples in (5), from the DanPASS corpus. In both examples, the preceding

context shows that the entity referred to in the foundation of the second sentence is the current discourse topic.

(5) a. så er der et vindue til ... [$_T$ det] [$_F$ skal sidde oppe i trekanten]
*then is there a window more ... it must sit up in triangle.DEF*
'Then there is one more window ... It must be placed up inside the triangle.'

b. dette vindue ... skulle ... også have en blomst[$_T$ den] [$_F$ kan jeg ikke finde].
*this window ... should ... also have a plant that can I not find*
'This window ... should also have a plant. I cannot find it.'

In (5a) the topic in the second sentence is the subject. In (5b) the topic in the second sentence is still the initial constituent, but this time it coincides with the topicalised object. Note that in Danish a topicalised constituent may or may not be contrastive, and topicalisation is therefore used more liberally in Danish than e.g. English, where it is subject to different semantic constraints (Prince 1999). In the example under consideration, for example, the topic expression simply refers to the plant just mentioned in the preceding sentence. By far the majority of the examples of topicalisation in the DanPASS corpus are of the non-contrastive type. An example in which we can see a contrast between two discourse referents, one of which is referred to by means of topicalisation, is the following:

(6) der er derimod to kapper ... den blå den skal du placere i det venstre
*there are however two pelmets ... the blue that must you put in the left*
vindue
*window*
'There are, however, two pelmets ... The blue one you must put in the left window.'

Prepositional complements can also be topicalised, as in the following example, where after having talked about the upper windows, the speaker contrasts them with the lowest one:

(7) og [$_T$ i det nederste vindue] [$_F$ tager man og sætter urtepotten]
*and in the lowest window takes one and puts flowerpot.DEF*
'And in the lowest window you put the flowerpot.'

Subject-verb inversion is also present in Danish when the initial constituent is a verbal or sentence adjunct, as shown in (8), from the same corpus:

(8) [$_{BG}$ så tager du] [$_F$ den store trekant]
*then take you the big triangle*
'Then you take the big triangle.'

Following Lambrecht (1994:117–131), I have defined topic as an entity about which the sentence predicates pertinent information. Therefore, I do not consider the adjunct *så* 'then' a topic. Firstly, it is not clearly referential, and secondly, it is not an integral part of the predicate-argument structure of the sentence. In general, adjuncts

(whether phrasal or clausal) can be seen as scene-setting expressions against which the main predication of the sentence is to be interpreted. They are treated here as background elements within a topic-comment or an identificational structure. The sentence in (8) is an example of identificational (background-focus) construction, with the final object as the only element in focus, and is thus similar to (3c), repeated below for convenience:

(3c) (Hvem har spist isen?)[$_{BG}$ Det har] [$_{F}$ børnene].
(*who has eaten icecream.*DEF) *that have children.*DEF
'Who has eaten the icecream? The children did.'

In other words, topicalisation constructions with verb-subject word order can be construed either as topic-comment constructions, where the topic is constituted by the extracted complement, or as topicless constructions. The latter can be the case if the topicalised constituent is an adjunct, as in (8), or an abstract anaphor, as in (3c).

## 4. TOPICLESS SENTENCES

Topicless sentences can be divided into two types based on their information structure. One is the identificational construction, of which we saw an example in the preceding section; the other is the all-focus, or thetic construction.

As we have seen, there is a strong tendency in Danish for the constituent preceding the verb to be construed as the topic of the sentence. This implies that focused subjects are quite rare, as clearly indicated by the data from Thomsen (1998) reported in the preceding section. Non-topical subjects are, however, not ungrammatical. An example are generic subjects, as in the opening sentence from the website of the Copenhagen zoo:

(9) En mæt løve er en tilfreds løve.
*a full lion is a satisfied lion*

The focus here includes not only the subject, but the whole sentence, which is construed as an all-focus structure. Another context where the subject is part of the focus is provided by event-reporting sentences, which I will return to.

An example of narrow, contrastive focus on the subject is given in the next section, where we will see that it is associated with deaccentuation in the rest of the sentence. Narrow (identificational) focus on the constituent in the foundation is possible, however, in connection with predicate topicalisation, as noted i.a. by Line Mikkelsen in connection with the following examples (Mikkelsen 2004:21):

(10) Solen var på himlen men skyfrit var det ikke.
*sun.*DEF *was on sky.*DEF *but cloudless was it not*
'The sun was out, but there were clouds.'

A construction in which the logical subject (and not the surface subject in the foundation) is interpreted as a narrow focus, is the cleft. In cleft sentences, in fact, the

identificational focus structure is grammaticalised, in that the cleft head (also called clefted constituent) corresponds to the missing argument in the presupposed open proposition expressed by the cleft clause.[1]

A corpus example is given in (11) below. Here the cleft head in the second part of the excerpt singles out one of the two curtains mentioned at the beginning. It is the missing argument of the open proposition expressed by the rest of the sentence (the cleft clause), which, although not necessarily known to the listener, must be accommodated as presupposed.

(11) og det er sådan nogle røvballegardiner som du skal vende
*and that is such some bum.shaped.curtains which you must turn*
'And that's some bum-shaped curtains which you must turn round.'
ja [$_{BG}$ det er] [$_{F}$ de største] [$_{BG}$ der skal være øverst]
*yes it is the biggest that must be highest*
'Yes, it's the biggest that must go at the top.'

The cleft conveys the presupposition that there are curtains and it asserts that the biggest curtain must be on top. The fact that the proposition expressed by the cleft clause (the part of the cleft following the cleft head) is a presupposition and not just an entailment can be shown by embedding the cleft in a presupposition preserving context, where ordinary entailments do not survive:

(12) Det er ikke/muligvis de største der skal være øverst.
'It is not/possibly the biggest that must go at the top.'

In both contexts, it is still true that a curtain 'must go at the top'. In contrast, there is no such presupposition in the non-cleft counterpart of the same sentences:

(13) De største skal ikke/muligvis være øverst.
'The biggest must not/possibly go at the top.'

The fact that the cleft clause expresses presupposed information, however, does not mean that it does not contain stress. This fact has brought several authors working with English data to consider clefts as cases of double focus, see e.g. Delin (1989) and Quirk, Greenbaum, Leech & Svartvik (1995). As regards Danish, both Nølke (1984) and Togeby (1992) seem to assume that the head is always in focus and do not discuss the relation between stress and focus in the cleft clause. As will be explained in section 5, full NPs in Danish are accented in the unmarked case, and the rightmost sentence accent usually corresponds to the righthand focus boundary. In the case of clefts, however, there is no such correspondence, as shown by the infelicity in terms of information structure of the following exchange:

(14) ?Hvor skal hvad være? Det er de største der skal være øverst.
'Where should what be? It's the biggest that must go at the top.'

Let us now turn to all-focus constructions. Below are two examples of event-reporting sentences, where the whole event expressed by the proposition is in focus.

Both sentences can be uttered out of the blue, and although they presuppose the existence of objects such as telephones and snow, their contribution in terms of information structure is not to tell us something new about the telephone or the snow, but to report an event.

(15) a. [$_F$ Telefonen ringer].
*phone.DEF rings*
'The phone is ringing.'
b. [$_F$ Det sner udenfor].
*it snows outside*
'It is snowing outside.'

While the all-focus structure is grammaticalised in the impersonal construction (15b), example (15a) does not differ in either syntax or prosody from an unmarked topic-comment sentence. Let us compare it with similar sentences in English and Italian:

(16) a. [$_F$ Tele′fonen ′ringer].
b. [$_F$ The PHONE is ringing].
c. [$_F$ Sta squillando il TELEFONO].
*is ringing the phone*

Event-reporting sentences of this kind are the only case in English where sentence accent falls on the subject, and projects to the entire sentence (focus projection will be addressed in the next section). Thus, as observed *inter alia* by Lambrecht (1994:143), in English the information structure of this sentence type is grammaticalised. In Italian the right focus articulation is achieved by placing the subject in post-verbal position. In Danish, on the other hand, the word order is the unmarked subject-verb one, and both subject and predicate are accented as they would be in the unmarked case. In other words, we understand the sentence as a thetic judgement only from the semantico-pragmatic context in which it is uttered.

As a rule, the opening sentence of a text is an all-focus construction. Nearly all the house-building monologues of the DanPASS corpus begin by mentioning the first piece that the speaker is expected to put in place:

(17) a. [$_F$ du begynder med at lægge den hvide firkant]
*you start with to lay-down the white square*
'You start by placing the white square.'
b. [$_F$ du har en stor firkant]
*you have a large square*

Another example of initial all-focus sentence from the DanPASS monologues is the *der* 'there' construction below:

(18) [$_F$ der er en stor hvid firkant ]
*there is a large white square*

As in English *there* constructions, the NP following the copula in Danish *der* sentences, although the logical subject, is not construed as a topic. The following examples all conform to this pattern.

(19) a. og så er der [$_F$ en lille hvid rektangel]
*and then is there a small white rectagle*
'And then there is a small white rectagle.'
b. og der er [$_F$ ikke nogen andre gule]
*and there is not any other yellow*
'And there are no other yellow ones.'

Both examples comply with the observation originally made by Milsark (1974) that the NP following the copula in what he calls existential sentences has to be weakly-quantified. Milsark in fact uses existential sentences as a context to distinguish between weak and strong quantifiers, a distinction rephrased by Lumsden (1988) in terms of presuppositional vs. non-presuppositional. This view is consistent with the claim made by Dahl (1974) that the NP following the copula cannot carry existential presupposition, and with his definition of existential sentences as detopicalising constructions. Several authors, however, have noted that presuppositional quantification is indeed possible if the NP following the copula is a member of a presupposed set of referents (Lumsden (1988) provides evidence for English and Heltoft (1986) for Danish). The following is an example of such an enumerative context from the DanPASS corpus:

(20) og til ′venstre for ′døren [$_T$ der] ′er der [$_F$ det ′største ′vindue]
*and to left of door.DEF there is there the largest window*
'And to the left of the door there is the largest window.'

This example is particularly interesting because it combines left-dislocation with the *there* construction. The initial locative adverbial is left-dislocated and resumed by the locative pronoun *der*. In section 3 it was argued, following Lambrecht (1991), that left dislocation serves the purpose of bringing to the foreground a referent so that this can function as the topic of the sentence predication. This motivates the topic interpretation of the adverb in the foundation field.

Thus, although the non-inverted *there* sentence is a topicless construction, it can be combined with left dislocation to yield a topic-comment information structure.

## 5. FOCALITY

In the examples discussed in the previous sections, we have seen that the focus sometimes is a narrow one, as in cleft sentences and identificational constructions in general, sometimes a wide one, extending over the entire predicate in topic-comment

sentences and the whole sentence in thetic judgements. In this section, I look more closely at what factors account for the extension of the focus domain in a sentence.

As I already mentioned, a strong indicator of focality in many languages is sentence accent. But Grønnum (1992:154–155) claims that standard Danish does not have a compulsory sentence accent. At least in what she calls 'pragmatically and emotionally neutral speech', there is no stressed syllable that can be observed to be more prominent than the others. According to this author, this is a distinguishing characteristic of standard Danish compared to regional variants of Danish as well as other Scandinavian languages, for example Swedish. Let us consider the following Danish sentence, also discussed by Grønnum:

(21) Køben′havn er ′hovedstaden i ′Danmark.
*Copenhagen is the capital in Denmark*
'Copenhagen is the capital of Denmark.'

None of the accents is more prominent than the others, and the focus right-hand boundary is indicated instead by the rightmost accent. Both narrow and wide analyses are available, so that the sentence can answer all of the following questions, each carving out a wider and wider focus domain:

(22) a. Hvilket land er Københavnhovedstaden i?
*which country is Copenhagen the capital in*
'Which country is Copenhagen the capital of?'
Køben′havn er ′hovedstaden [$_F$ i ′Danmark].
b. Hvad ved du om Københavnn?
*what know you about Copenhagen*
'What do you know about Copenhagen?'
Køben′havn [$_F$ er ′hovedstaden i ′Danmark].
c. Hvad har du lært i geografi i dag?
*what have you learnt in geography today*
What did you learn in geography class today?
[$_F$ Køben′havn er ′hovedstaden i ′Danmark].

Narrow focus on the subject, on the other hand, is only possible if the post-verbal constituents are deaccented as shown below (deaccenting is indicated by ,):

(23) Hvilken by er hovedstaden i Danmark?
*which city is the capital in Denmark*
'Which city is the capital of Denmark?'
[$_F$ Køben′havn] er ,hovedstaden i ,Danmark.

Danish is, in other words, subject to the principle of focus projection, whereby focus projects from a peripheral non-head daughter to an entire phrase. Focus projection has been much discussed in the literature (Bresnan 1971, Chomsky 1971, Selkirk 1984, Rochemont 1989), and there isn't total agreement as to whether the

principle obeys word order constraints, whether it works differently for complement and adjuncts, or whether it is subject to configurational restrictions.

The examples under consideration in (22) above show, in addition to the lack of a most prominent sentence accent, two other general facts characterising Danish prosody in pragmatically neutral contexts: (i) full NPs preceding the focus domain or included in it are accented; and (ii) the finite verb may be unaccented if it occurs in certain syntactic environments. The copula-subject predicate construction in the example is one of these environments; another is the construction verb + indefinite object, as in *spise′is* 'eat icecream'.[2] Therefore, while focus projection is often described, at least for English, as a projection from the most prominent non-head daughter, the correct interpretation of the phenomenon for Danish is rather that focus can project from the RIGHMOST ACCENTED non-head daughter, and that any other preceding constituent, including the predicate, may or may not be accented mostly depending on syntactic factors.[3] The fact that the rightmost daughter in our example is an adjunct shows that focus projection does not only concern complements, as has been claimed for English (Rochemont 1989).

Although full NPs are accented in the unmarked case, we have seen that deaccenting can be used to convey a narrow focus. Hansen & Lund (1983) claim that a constituent can be emphasised by reducing the stress of other constituents in the same sentence:

(24) a. Anker ′Jørgensen rejser på ′ferie i au′gust.
*Anker Jørgensen travels on holiday in August*
'Anker Jørgensen is going on holiday in AUGUST.'
b. Nej, Kjeld ′Olesen rejser på ˌferie i auˌgust.
*no Kjeld Olesen travels on holiday in August*
'No, Kjeld OLESEN is going on holiday in August.'

(Hansen & Lund 1983:13)

The same authors also note that a constituent which would not be stressed according to lexical or syntactic rules, e.g. a preposition, can be stressed in contrastive expressions:

(25) Avisen ligger ′på bordet, ikke nedenunder.
*newspaper.DEF lies on table.DEF not under*
'The newspaper is ON the table, not under it.'

(Hansen & Lund 1983:14)

All these cases are in fact subsumed by the general rule according to which focus projection is only possible if the rightmost non-head daughter is accented. In (24a), the focus can project from the final adverbial to the whole sentence, so that both narrow and wide focus readings are possible. In (24b) on the other hand, the rightmost accent falls on the subject, which is to the left of the verb, and focus projection is therefore ruled out. We already saw in our discussion of thetic sentences

that focus in Danish cannot project from subject to verb, so that both subject and verb bear an accent in examples such as *Tele'fonen'ringer* 'The phone is ringing'. Finally, in (25), the accented word is the head of the phrase, so projection is again ruled out.

We have also seen, however, that there are contexts in which accented lexical material in the verbal predicate does not fall within the focus domain. One such context is provided by clefts, discussed earlier; another by subordinate clauses, which contribute as a whole to the information structure of the matrix clause, but for which I do not postulate an internal information structure.

## 6. REPRESENTING INFORMATION STRUCTURE IN HPSG

Typed feature structure formalisms lend themselves well to the task of representing the information structure of a sentence as they allow us to conflate phonological, syntactic, semantic and pragmatic analyses into a single complex feature structure. Thus, the contribution made by the various words to the information structure of a sentence can be represented by means of features belonging to the pragmatic description of the sign, while the necessary links are established between these pragmatic features and features pertaining to the other linguistic dimensions. In the remainder of the paper, a proposal is made for how this can be done in the HPSG framework (Pollard & Sag 1994). First it is shown how the representation of information structure interacts with that of the semantics of the sentence, then the inheritance of information structure values in the syntactic structure is formalised, and finally it is explained how the defined feature types account for the information structure of some of the sentence constructions discussed in the preceding sections.

### *6.1 Information structure and MRS*

The basic unit of linguistic representation in HPSG is the sign, which is modelled in terms of a typed feature structure with features accounting for phonological, syntactic and contextual aspects. To represent information structure, the attribute INFOSTR is introduced as part of the context (CTXT) of the linguistic sign. The value of INFOSTR is a feature structure with the three attributes TOPIC, FOCUS and BG. This is quite similar to what is proposed in Vallduví & Engdahl (1996) and Engdahl & Vallduví (1996). When it comes to the value of the three features, however, the present proposal differs in that each feature takes as value a list of semantic indices rather than content feature types as proposed by those authors. The reason for this difference will become clear shortly.

(26)
$$\begin{bmatrix} \text{INFOSTR} & \begin{bmatrix} \text{TOPIC (list of indices)} \\ \text{FOCUS (list of indices)} \\ \text{BG (list of indices)} \end{bmatrix} \end{bmatrix}$$

By means of the semantic indices the information structure features can refer to the elements that make up the semantic representation of the sentence, constructed here by means of Minimal Recursion Semantics (Copestake, Flickinger, Pollard & Sag 1997, Flickinger & Bender 2003). In Minimal Recursion Semantics (MRS), the two main components of the semantics of a sentence are (i) a BAG of relations, where each relation can be referred to by a label and can in turn refer to a number of argument variables (or indices), and (ii) a number of scope assignments over pairs of relations. In addition to the relation bag and the scope assignments, an MRS representation also contains an LTOP HANDLE referring to the relation with the highest scope, and an instance or event INDEX introduced by the semantic head, analogous to lambda. For example, let us consider a topic-comment sentence from the DanPASS corpus:

(27) [$_T$ døren] [$_F$ er en firkant]
*door.DEF is a square*
'The door is a square.'

The content of the sentence will be represented in MRS as follows:

(28)
$$\begin{bmatrix} \textit{mrs} \\ \text{LTOP h1} \\ \text{INDEX e1} \\ \text{RELS}\left\langle \begin{bmatrix} \textit{være\_v\_rel} \\ \text{LBL h2} \\ \text{ARG0 e1} \\ \text{ARG1 x1} \\ \text{ARG2 x2} \end{bmatrix}, \begin{bmatrix} \textit{udef} \\ \text{LBL h3} \\ \text{ARG0 x2} \\ \text{RESTR h4} \\ \text{BODY h10} \end{bmatrix}, \begin{bmatrix} \textit{firkant\_n\_rel} \\ \text{LBL h5} \\ \text{ARG0 x2} \end{bmatrix}, \right. \\ \left. \begin{bmatrix} \textit{def} \\ \text{LBL h11} \\ \text{ARG0 x1} \\ \text{RESTR h6} \\ \text{BODY h9} \end{bmatrix}, \begin{bmatrix} \textit{dør\_n\_rel} \\ \text{LBL h7} \\ \text{ARG0 x1} \end{bmatrix}, \begin{bmatrix} \textit{message\_m\_rel} \\ \text{LBL h1} \\ \text{MARG h8} \end{bmatrix} \right\rangle \\ \text{HCONS}\left\langle \begin{bmatrix} \textit{qeq} \\ \text{HARG h8} \\ \text{LARG h2} \end{bmatrix}, \begin{bmatrix} \textit{qeq} \\ \text{HARG h4} \\ \text{LARG h5} \end{bmatrix}, \begin{bmatrix} \textit{qeq} \\ \text{HARG h6} \\ \text{LARG h7} \end{bmatrix}, \begin{bmatrix} \textit{qeq} \\ \text{HARG h9} \\ \text{LARG h3} \end{bmatrix} \right\rangle \end{bmatrix}$$

The LTOP handle, which allows the clause to be embedded in a longer sentence, is structure-shared with the label (LBL) handle of the *message* relation at the end of the relation list (RELS), whilst the argument of the message (MARG) is equalled to the event relation introduced by the main verb via a *qeq* ('equality modulo quantifiers') constraint in the HCONS list. The verbal relation *være* 'be' has three arguments

referring to an event index and to the instances associated with the relations *firkant* 'square' and *dør* 'door', respectively. One of the strengths of MRS is the fact that the semantic representations may be underspecified in case of scope ambiguity, for example when several quantifiers are present. In the example under consideration, however, the interpretation is enforced where the quantifier introducing the topic outscopes the quantifier of the NP belonging to the focus, so that the value of the BODY of the *def* quantifier preceding the subject topic, is equalled to the handle of the *udef* one preceding the object. Note that in these feature structures variable sharing is indicated by coincident variable values, where a variable value is a digit preceded by either 'h' for *handle* variable, 'x' for *individual index* variable or 'e' for *event* variable.

If a specification of the information structure of the sentence is now added, the representation will look as shown in (29). The values of the information structure features coincide with indices in the MRS representation. Both TOPIC and BG coincide with the index associated with the subject, whilst the FOCUS is a list of two indices corresponding to the event introduced by the main verb and the instance corresponding to the object.

(29)
$$\left[\begin{array}{l}
\textit{mrs} \\
\text{LTOP h1} \\
\text{INDEX e1} \\
\text{CONT}\left[\begin{array}{l}
\text{RELS}\left\langle \left[\begin{array}{l}\textit{være_v_rel}\\ \text{LBL h2}\\ \text{ARG0 e1}\\ \text{ARG1 x1}\\ \text{ARG2 x2}\end{array}\right], \left[\begin{array}{l}\textit{udef}\\ \text{LBL h3}\\ \text{ARG0 x2}\\ \text{RESTR h4}\\ \text{BODY h10}\end{array}\right], \left[\begin{array}{l}\textit{firkant_n_rel}\\ \text{LBL h5}\\ \text{ARG0 x2}\end{array}\right], \right. \\
\qquad \left. \left[\begin{array}{l}\textit{def}\\ \text{LBL h11}\\ \text{ARG0 x1}\\ \text{RESTR h6}\\ \text{BODY h9}\end{array}\right], \left[\begin{array}{l}\textit{dør_n_rel}\\ \text{LBL h7}\\ \text{ARG0 x1}\end{array}\right], \left[\begin{array}{l}\textit{message_m_rel}\\ \text{LBL h1}\\ \text{MARG h8}\end{array}\right] \right\rangle \\
\text{HCONS}\left\langle \left[\begin{array}{l}\textit{qeq}\\ \text{HARG h8}\\ \text{LARG h2}\end{array}\right], \left[\begin{array}{l}\textit{qeq}\\ \text{HARG h4}\\ \text{LARG h5}\end{array}\right], \left[\begin{array}{l}\textit{qeq}\\ \text{HARG h6}\\ \text{LARG h7}\end{array}\right], \left[\begin{array}{l}\textit{qeq}\\ \text{HARG h9}\\ \text{LARG h3}\end{array}\right] \right\rangle
\end{array}\right] \\
\text{CTXT | INFOSTR}\left[\begin{array}{l}\text{TOPIC}\langle \text{x1}\rangle \\ \text{FOCUS}\langle \text{x2, e1}\rangle \\ \text{BG}\langle \text{x1}\rangle\end{array}\right]
\end{array}\right]$$

It may be objected that this style of representation does not do justice to the notion of wide focus as is usually described in the literature, i.e. as a projection of a focus value from a daughter to either the verb phrase or the whole sentence. On this view, it would seem that the feature structure representation corresponding to the analysis of example (27) should express token-identity between the focus and the content of the VP as a whole, rather than a list of the individual contents of the verb and its complements. In HPSG, however, the content of a headed phrase is token-identical to that of the head. In other words, the content of a VP cannot be distinguished from that of the head verb or the whole sentence. Therefore, the approach proposed in Engdahl & Vallduví (1996), where the focus is claimed to be token-identical with the content of the VP node, cannot distinguish between sentence focus, VP focus and narrow focus on the verb. The problem is noted by Cooper (2004), who in fact uses it as an argument to propose a more elaborate linguistic representation (based on record types rather than feature types) where the content of a VP is a function of the content of the arguments.

It may be argued, in fact, that Engdahl & Vallduví's proposal is too syntax-based, an issue raised by Wilcock (2005). He argues that a semantics-oriented representation of information structure is more adequate for topic tracking and the tracking of old and new information e.g. in dialogue systems. The interaction between information structure and quantification, which we will return to below, is an additional argument in favour of a semantic treatment of information structure. Wilcock represents information structure values as variables over handles in the MRS representation. Yoshimoto, Kobayashi, Nakamura & Mori (2005) presents an alternative solution where information structure values are unified with whole MRS relations rather than just indices. Both these proposals are similar to mine. In particular Wilcock's seems based on a very similar intuition, since it also uses variables to map information structure values onto parts of the semantic representation. Using a linear notation, he proposes the following MRS representation for a wide focus interpretation of the sentence *every dog chased some cat* (Wilcock 2005:275):

(30) 1:every(x,3,4),3:dog(x),7:cat(y),5:some(y,7,1),4:chase(e,x,y)
TOP-HANDLE:1, LINK:{1}, FOCUS:{4,5}

The top handle corresponds to my LTOP, while the link is roughly the same as my TOPIC. The value of the link is a set whose unique element is the handle of the quantifier that has scope over the subject *dog*. The value of the focus, on the other hand, is a set that includes two elements: the handle of the *chase* event, and the handle of the quantifier that has scope over the object *cat*.

As noted by Ericsson (2005:193–196), however, using handle variables has undesired effects. For example, in (30), by referring to the feature structure corresponding to *chase*, which includes subject and object as arguments of the event relation, the focus seems to span not only over verb and object but over the whole

sentence. Similarly the link, by unifying with the quantifier preceding the subject, seems to refer not only to the subject but to the whole event, which is also included in the scope of the quantifier.

In (27), if the TOPIC is unified with the handle that points to the definite article, it ends up referring not only to the subject *døren*, that shares the semantic argument with the article, but also to the subject predicate *en firkant* via the BODY feature in the *def* feature structure. By using semantic indices, this effect is avoided. However, an undesired consequence of this choice noted by one of the anonymous reviewers of this article, is the fact that information structure is made to refer to bare concepts rather than quantified expressions. Although MRS provides a flexible means of representing quantifier scope, it does not seem possible to refer to a semantic element corresponding to just a quantifier and its restrictor. Therefore, here I do not include quantifiers in the information structure values.

As already mentioned, one of the strengths of MRS is its ability to cope with quantifier scope ambiguity by allowing scope underspecification. In the analysis of (27) proposed here, the quantifier associated with the topic has widest scope. But the question is whether there is always a one-to-one correspondence between quantifier scoping and information structure. Wilcock claims, in his discussion of the sentence *every dog chased some cat*,[4] that if *some cat* is interpreted as a narrow focus, it outscopes *every dog*. On the contrary, if the focus also includes the verb, the opposite scoping would hold. This claim, however, is disputed by Ericsson (2005), who maintains that regardless of whether the focus is narrow or wide, both a reading in which the dogs chase a unique cat and one in which several cats are chased are possible. Ericsson concludes that quantifier scope and information structure appear to be independent, at least in this particular example.

In the sentence under discussion here, however, it seems impossible to provide an interpretation in which the quantifier preceding the subject predicate should outscope the quantifier preceding the topic. I believe the approach taken in McNay (2005) provides an initial answer to the issue. Looking at German data, McNay notes in fact that the relation between quantifier scope and information structure is a complex one. Scope preferences seem to depend on a mixture of information structure features that are lexically assigned to the quantifier themselves (e.g. *alle* 'all' and *jeder* 'every' are lexically marked as links) and word order constraints. I will not pursue the issue any further here.

To conclude this section, Figure 2 illustrates the syntactic tree corresponding to the analysis of example (27), with an indication of the information structure values. The variables 1, 2 and 3 (in boxes in the figure) show structure sharing between the features in the tree. The ultimate values of these variables will correspond to individual instances or events as shown in (29) depending on whether the lexical item that introduces the index is a nominal or a verbal object. Note that in this and subsequent trees, feature paths are abbreviated for reasons of space. The complete

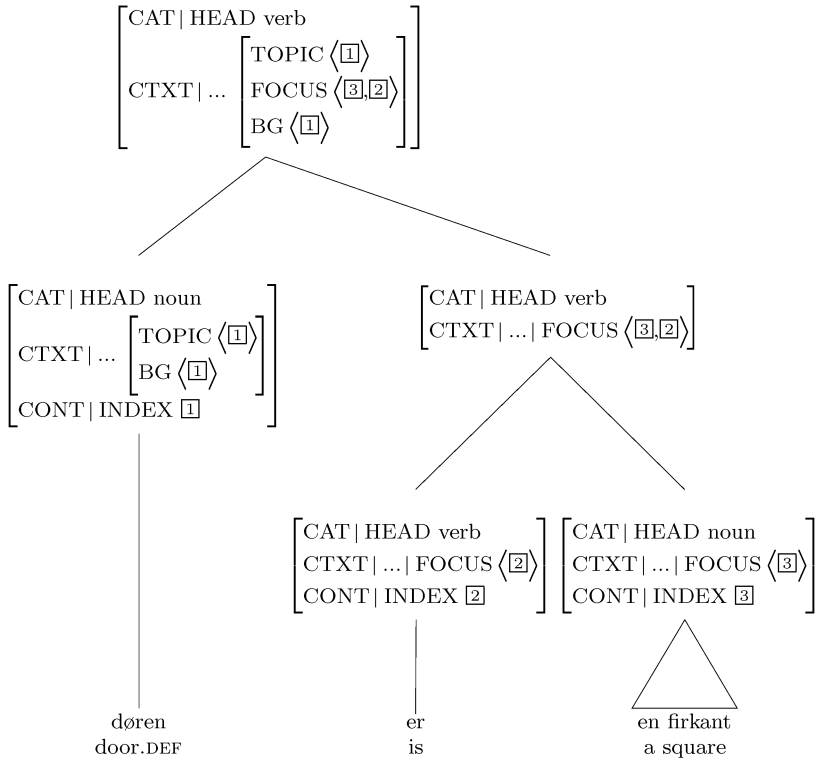


**Figure 2. Collecting indices in wide VP focus.**

path to the information structure features is SYNSEM|LOC|CTXT|INFOSTR, and the path to the content features is SYNSEM| LOG | CONT.

### *6.2 Inheritance of information structure values*

In the preceding section I proposed a representation of information structure in terms of semantic indices. This section and the next deal with how this semantics-oriented representation can be constructed based on the syntactic derivation of a sentence.

Inheritance of information structure values refers to the mechanisms that instantiate topic, background and focus values of mother and daughter nodes in a syntactic structure. In Engdahl & Vallduví (1996), it is assumed that the instantiation of the information structure of the main sentence is achieved by the combination of two general principles called INHERITANCE and PROJECTION: inheritance applies to narrow focus and what these authors call links and tails (roughly corresponding to

our topic and to the part of the background which does not include the topic), whilst projection applies to wide focus.

In the present proposal, on the contrary, the instantiation of focus and background values is obtained by positing two different feature types accounting for focus and background inheritance. Subtypes of the basic HPSG phrasal types (*hd-comp*, *hd-spec*, *hd-adj*, etc.) are then defined that unify with either focus or background inheritance. These subtypes give rise to binary projections in which the information structure of the mother is constructed by adding the relevant value of the non-head daughter (i.e. either the FOCUS or the BG value) to the information structure of the head. We will see how this works for a concrete example, but let us first discuss information structure inheritance in more detail. I will start from *focus-inheritance*, which accounts for a binary projection in which the non-head daughter is a focused constituent and follows the head daughter.

(31)
$$\left[\begin{array}{l} \textit{focus-inheritance} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \left[\begin{array}{l} \text{FOCUS} \left\langle \boxed{2}, \boxed{1} \right\rangle \\ \text{BG} \ \boxed{3} \end{array}\right] \\ \text{HD} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \left[\begin{array}{ll} \text{FOCUS} & \boxed{1} \\ \text{BG} & \boxed{3} \end{array}\right] \\ \text{NON-HD} \left[\begin{array}{l} \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots . \mid \text{FOCUS} \left\langle \boxed{2} \right\rangle \\ \text{ACCENT } \textit{true} \end{array}\right] \end{array}\right]$$

In the corresponding feature type in (31) the FOCUS value of the mother daughter is a list consisting of the focus values of the non-head and that of the head (which may or may not contain additional focused elements).[5] The BG value of the mother node is structure-shared (unified) with that of the head. This ensures the accumulation of focus values on the top node of the syntactic tree as well as inheritance of the background values. Finally, the ACCENT feature constraint on the focused non-head makes sure that only accented constituents are assigned to the focus domain.

This will prevent wrong focus analyses of unaccented constituents which should rather belong to the background, as the unaccented pronominal subject *du* 'you' in the following example:

(32) (Hvad skal jeg med trekanterne?)
(*what shall I with triangles.DEF*)
'(What shall I do with the triangles?)'

a. [$_T$ den o′range] [$_F$ ′anbringer][$_{BG}$ du] [$_F$ for′oven].

b. ? [$_T$ den o′range] [$_F$ ′anbringer du for′oven].
*the orange place you above*
'The orange one you place above.'

Background inheritance is very similar: it ensures that the background value of a non-head daughter is unified into the background value of the mother node. In addition, background inheritance obeys different constraints depending on whether the background constituent occurs to the right or to the left of the focus domain.

A constituent to the right of the focus is in fact only given a background interpretation if it is not accented. This constraint, which is formalised in the feature type in (33), applies to final unaccented (or deaccented) constituents. We saw an example in (24b).

(33)
$$\left[\begin{array}{l} \textit{right-bg-inheritance} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \left[\begin{array}{l} \text{FOCUS } \boxed{2} \\ \text{BG} \left\langle \boxed{4}, \boxed{3} \right\rangle \end{array}\right] \\ \text{HD} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \left[\begin{array}{l} \text{FOCUS } \boxed{2}\ \textit{non-empty} \\ \text{BG } \boxed{3} \end{array}\right] \\ \text{NON-HD} \left[\begin{array}{l} \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \mid \text{BG} \left\langle \boxed{4} \right\rangle \\ \text{ACCENT } \textit{false} \end{array}\right] \end{array}\right]$$

If a constituent occurs to the left of the focus, on the other hand, it may be analysed as belonging to the background whether or not it is accented. Therefore, the feature type that accounts for background constituents to the left of the focus in (34) has no accent specification on the background daughter. The constraint that the focus list of the head daughter must be empty (as indicated by the symbol '⟨⟩') ensures, on the other hand, that the feature type only applies to background constituents interposed between the verbal head and the focus.

(34)
$$\left[\begin{array}{l} \textit{left-bg-inheritance} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \left[\begin{array}{l} \text{FOCUS } \langle\rangle \\ \text{BG} \left\langle \boxed{3}, \boxed{2} \right\rangle \end{array}\right] \\ \text{HD} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \left[\begin{array}{l} \text{FOCUS } \langle\rangle \\ \text{BG } \boxed{2} \end{array}\right] \\ \text{NON-HD} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CTXT} \mid \ldots \mid \text{BG} \left\langle \boxed{3} \right\rangle \end{array}\right]$$

For instance, the type will provide a background analysis of the subject in example (8), repeated below, so that the final costituent – the direct object – in turn receives a narrow focus interpretation by the application of *focus-inheritance*.

(8) [$_{BG}$ så ′tager du] [$_{F}$ den ′store ′trekant]
*then take you the big triangle*
'Then you take the big triangle.'

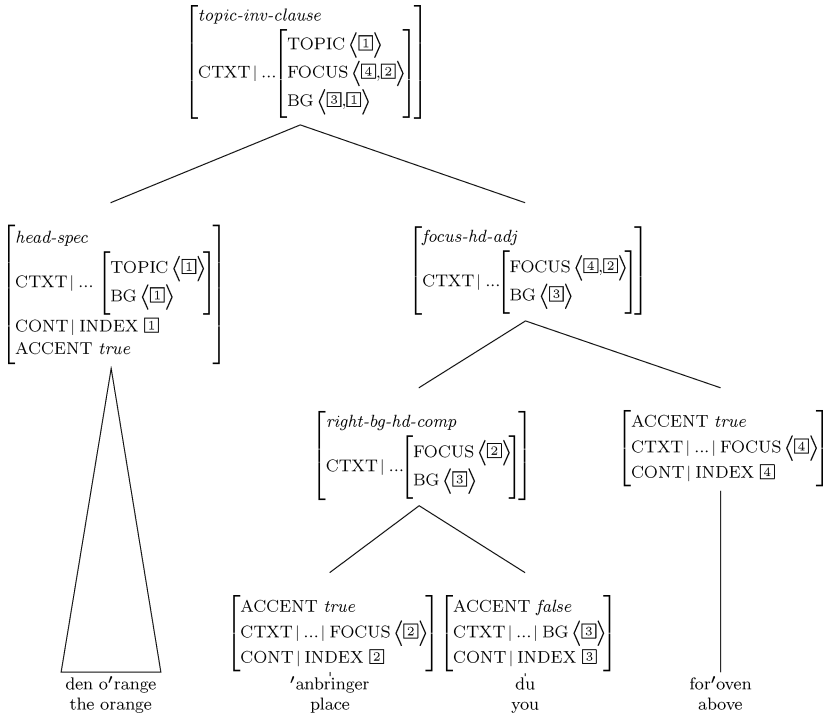


**Figure 3. Inheritance of information structure values.**

Both inheritance principles take care of the unification of information structure features between mother and daughters in a binary projection. As for the verbal head, it was mentioned earlier that finite verbs in Danish are generally accented, but certain syntactic environments allow deaccentuation. Thus, verbs may or may not be in the focus domain regardless of whether they are accented. In (8) the verb is accented but belongs to the background since it is repeated many times in the preceding discourse (consisting in fact of a list of similar events: 'you take. . . you take. . .'). This seems to indicate that lexical entries for verbs must be underspecified with respect to whether they contribute focus or background values, or that they should be given a default assignment that could be overridden based on discourse factors.

Both focus and background information structure inheritance types can be seen at work in Figure 3, where they are applied to the example in (32a). Each binary projection carries the name of the type responsible for the projection. Some of these types combine an inheritance principle with a phrasal type. For instance, the type *focus-hd-adj* combines focus-inheritance with *hd-adj*, the phrasal type that accounts for a head-adjunct structure. The type *right-bg-hd-comp*, on the other hand, combines background inheritance for constituents to the right of the focus with a

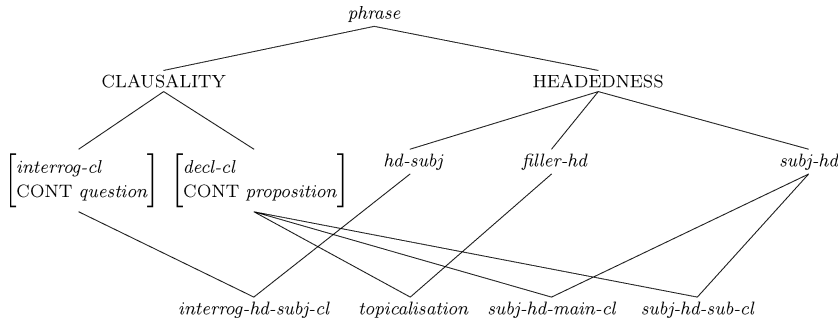


**Figure 4. Clausal constructions.**

head-complement structure. The topic of the sentence is made to unify with the TOPIC feature of the extracted object by the application of the *topic-inv-clause*, which I will return to. The verb is given a default focus interpretation.

### 6.3 Information structure in clausal constructions

Within HPSG, constructions are modelled in terms of phrasal types that are cross-classified along several dimensions, the two most important ones being CLAUSALITY and HEADEDNESS (Ginzburg & Sag 2000, Sag 2001). Multiple inheritance ensures that the features of the more general types are also shared by the more specific ones. Such an approach makes it possible to define clausal as well as phrasal constructions, both of which interact with the information structure principles and types presented earlier. In the preceding section we saw that phrasal projection types combine with different types of information structure inheritance. Here I look at information structure and clausal constructions.

Following the HPSG view of grammatical constructions, Neville & Paggio (2001) define a typology of Danish clause types which accounts for the different word orders and extraction patterns allowed in the various types. The most general (leaving out in particular relative clauses) are shown in Figure 4.

The type *interrog-hd-subj-cl* accounts for questions, *topicalisation* for topicalisation constructions, *subj-hd-main-cl* for main clauses with subject-verb word order and *subj-hd-sub-cl* for subordinate clauses.

It was proposed earlier that declarative sentences should be constrained to have a non-empty focus domain.[6] This constraint can be added to the specification of the *decl-cl* type, which also carries the syntactic constraint [IC *true*], which characterises independent main clauses.

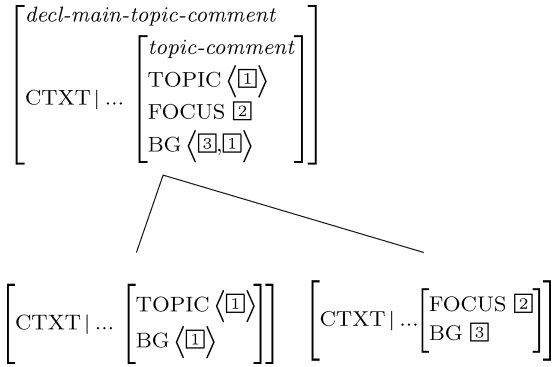


**Figure 5. Declarative topic comment construction.**

(35) $\begin{bmatrix} \textit{decl-cl} \\ \text{SYNSEM} \mid \text{LOC} \quad \begin{bmatrix} \text{CAT} \mid \text{HEAD} \mid \text{IC } \textit{true} \\ \text{CTXT} \mid \text{FOCUS } \textit{non-empty} \end{bmatrix} \end{bmatrix}$

Furthermore, the clausal types also account for the information structure values at the top level of the syntactic projection: in other words, the projection that directly dominates the constituent in the foundation and the node that spans over the rest of the sentence. Here I will be concerned with only two types of declarative clauses, namely non-inverted declarative main clauses and topicalisation constructions. They are syntactically different, and at the top level of the syntactic derivation they are dealt with by different rules.

To start with declarative main clauses with subject-verb word order, marked by the feature [BCO *sv*] (Basic Constituent Order subject-verb), we saw that the obvious interpretation of this clause type is a topic-comment sentence with either wide or narrow focus, but that all-focus interpretations are also possible. The first type is shown in Figure 5. It applies e.g. to example (27). Note that the feature structure type *topic-comment* allows for wide as well as narrow focus interpretations of the predicate. The all-focus interpretation is provided by the projection type in Figure 6, where the information structure is set to be *all-focus*. This means that both BG and TOPIC are empty lists. The type applies e.g. to example (17b).

Let us now turn to topicalisation constructions. They are characterised by verb-subject constituent order, i.e. [BCO *vs*]. Their interaction with the information structure types gives again rise to two different projections, depending on whether the extracted constituent is interpreted as the topic or as part of the background. The two projections are shown in Figure 7 and Figure 8, respectively. They apply e.g. to examples (7) and (8), both discussed earlier.

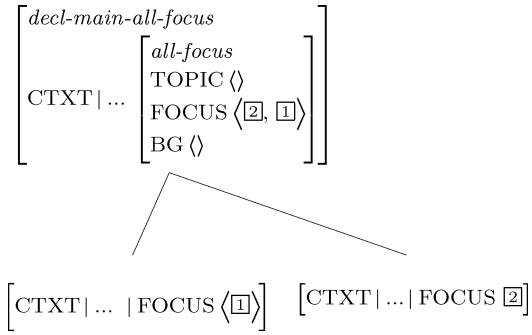


**Figure 6. Declarative all-focus construction.**

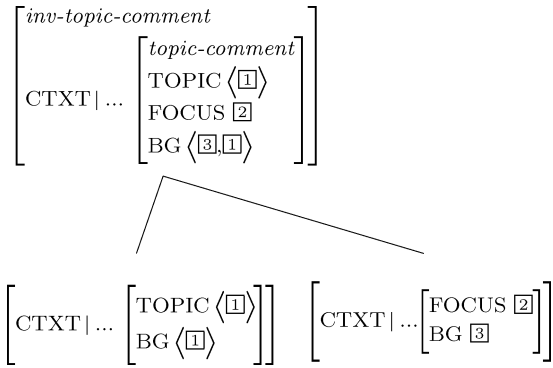


**Figure 7. Topicalisation construction: extracted topic.**

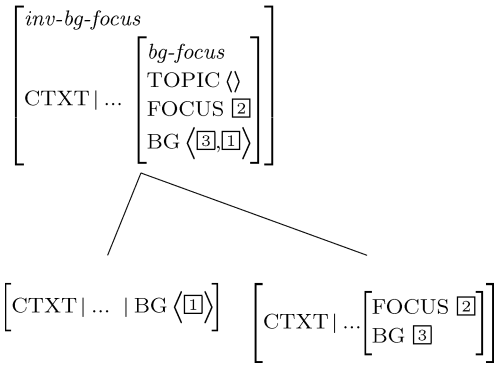


**Figure 8. Topicalisation construction: extracted background.**

Other sentence types that were discussed earlier are clefts and *der* sentences. The placement of both sentence types in the clausal hierarchy shown in (32) is not trivial in that clefts can syntactically be constructed in different ways, and *der* sentences can interact with topicalisation. An attempt to add them to the hierarchy will not be made here. Concerning their interaction with the information structure dimension, in the case of clefts we saw that the cleft head constitutes the narrow focus. *Der* sentences, on the other hand, although generally described as topicless, can be construed as topic-comment sentences if the location is topicalised, as in example (20).

## 7. CONCLUSION

Several aspects of the relation between information structure and the grammar of Danish have been addressed in this paper, and a proposal has been made for how the information structure of Danish sentences can be expressed by way of feature types in a typed feature structure formalism. In particular, focus, topic and background are represented in terms of lists of semantic indices corresponding to nominal referents and events in the semantics of the sentence, expressed in the MRS formalism. It was argued that, contrary to earlier approaches, this proposal has the advantage of allowing a clear distinction between sentence focus, VP focus and narrow focus.

It was described how information in Danish sentences can be organised in different ways with respect to topicality and focus, in particular how information structure is expressed in declarative main clauses, topicalisation constructions, clefts and *der* sentences. The relation between accentuation and focus was then discussed, and it was argued that focus in Danish projects not from a nucleus bearing the most prominent sentence accent, but from the constituent carrying the rightmost accent.

The interaction between information structure and clausal types was then formalised in a hierarchy where different phrasal and clausal types are enriched with information structure constraints. It was shown that the feature types defined could account for prosodic and syntactic phenomena, for instance the different stress patterns displayed by background constituents preceding or following the focus, or the different interpretations of topicalised complements vs. adjuncts.

The present proposal can be extended in several directions. One is the formal treatment of syntactically marked topicless sentences such as clefts, and detopicalising constructions such as *der* sentences. Both contruction types, although discussed in the paper, are not yet included in the hierarchy. They constitute an interesting task because of their syntactic complexity as well as their information structural properties.

Another important issue which certainly merits further investigation is the complex interaction between information structure and quantifier scope. It was observed that an adequate treatment of this phenomenon must take into account

lexical and constructional aspects, and that the MRS representation chosen here to represent information structure offers an apparently flexible formal means of accounting for the relevant scope interactions. At the same time, however, it was pointed out that, because of the way in which quantifier scope is represented in MRS, including quantifiers in the information structure representation has the undesired effect of making this representation less precise. Therefore, in the present proposal information structure features refer to semantic indices and leave out quantification, and the issue is left open for future research.

## ACKNOWLEDGEMENTS

I would like to thank my colleagues at CST, in particular Bolette Sandford Pedersen and Costanza Navarretta, as well as the three anonymous reviewers, for their knowledgeable comments and suggestions.

This work has been funded by the Carlsberg Foundation.

## NOTES

1. Here I am using argument in the logical sense of the term. The point made applies to grammatical arguments and adjuncts alike, cf. *Det var i går jeg ringede til ham* 'It was yesterday I called him'.
2. For a detailed account, see Grønnum (2001).
3. The fact that full NPs but not pronouns are accented in the normal case can be explained as a consequence of pragmatic activation.
4. This sentence is also discussed in Copestake et al. (1997).
5. Throughout the paper, the precedence order of the nodes in the feature structures is assumed to correspond to surface order. To account for focus-inheritance in a case in which the head follows the non-head, therefore, an additional feature structure must be posited. An example is focus -inheritance from the past participle to the preceding adverb in a sentence like *Peter har igen spist is* 'Peter has again eaten icecream'.
6. The present proposal leaves out the treatment of focus in interrogative clauses, in particular *wh*-interrogatives, where the focus is the unaccented interrogative pronoun (Lambrecht & Michaelis 1998).

## REFERENCES

Bresnan, Joan. 1971. Sentence stress and syntactic transformations. *Language* 47, 257–281.

Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In Danny Steinberg & Leon A. Jakobovitz (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, 183–216. Cambridge: Cambridge University Press.

Cooper, Robert. 2004. Using type theory with records for HPSG. Presented at Fest Colloquium for Uwe Mönnich, Freudenstadt.

Copestake, Anne, Dan Flickinger, Carl Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3, 281–332. New York: Springer. http://lingo.stanford.edu/sag/papers/copestake.pdf (accessed 3 April 2009).

Dahl, Östen. 1974. Topic–comment structure revisited. In Östen Dahl (ed.), *Topic and Comment, Contextual Boundness and Focus* (Papiere Zur Textlinguistik 6), 1–24. Hamburg: Helmut Buske.

Delin, Judy. 1989. *Cleft Constructions in Discourse*. Ph.D. thesis, University of Edinburgh.

Diderichsen, Paul. 1946. *Elementær dansk grammatik*. Copenhagen: Gyldendal. [3rd edn. 1962.]

Diderichsen, Philip & Jakob Elming. 2004. Topik i spontan dansk dialog. Master's thesis, Copenhagen Business School, Institute of Computational Linguistics.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover & Enric Vallduví (eds.), *Studies in HPSG* (Edinburgh Working Papers in Cognitive Science 12), 1–32. Edinburgh: Centre for Cognitive Science, University of Edinburgh.

Ericsson, Stina. 2005. *Information Enriched Constituents in Dialogue*. Ph.D. thesis, University of Gothemburg.

Flickinger, Dan & Emily M. Bender. 2003. Compositional semantics in a multilingual grammar resource. In *Workshop on Ideas and Strategies for Multilingual Grammar Development*, 33–42. Vienna: European Summer School in Logic, Language and Information (ESSLLI).

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative Investigations: Form, Meaning and Use of English Interrogatives*. Stanford, CA: CSLI Publications.

Grønnum, Nina. 1992. *The Groundworks of Danish Intonation*. Copenhagen: Museum Tusculanum Press, University of Copenhagen.

Grønnum, Nina. 2001. *Fonetik og Fonologi: Almen og Dansk*. Copenhagen: Akademisk Forlag.

Grønnum, Nina. 2005. May. DanPASS – Danish phonetically annotated spontaneous speech. Presented at FONETIK 2005, Gothenburg.

Halliday, M. A. K. 1967. Notes on transitivity and theme in English (part 2). *Journal of Linguistics* 3, 199–243.

Hansen, Erik & Lars Heltoft. 2000. Grammatik over det danske sprog. Preprint.

Hansen, Erik & Jørn Lund. 1983. *Sæt Tryk På: Syntaktisk tryk i dansk*. Copenhagen: Lærerforeningernes Materialeudvalg.

Heltoft, Lars. 1986. The pragmatic syntax of Danish *der*-constructions. *ROLIG-papir* 40, 1–23.

Henrichsen, Peter J. 1998. Peeking into the Danish living room: Internet access to a large speech corpus. In *NODALIDA 1998*, 119. Copenhagen: University of Copenhagen.

Lambrecht, Knud. 1991. Dislocation. In Martin Haspelmath, Ekkehard König, Wulf Osterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals: An International Handbook* (vol. 2), 1050–1078. Berlin & New York: Walter de Gruyter.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

Lambrecht, Knud & Laura A. Michaelis. 1998. Sentence accent in information questions: Default and projection. *Linguistics and Philosophy* 21(5), 447–544.

Lumsden, Michael. 1988. *Existential Sentences*. Routledge: London.

McNay, Anna. 2005. Information structure, word order, and quantifier scope in German (Technical Report 11). *Durham and Newcastle Working Papers in Linguistics* 11. http://www.clp.ox.ac.uk/pages/ANNAMCNAY/durhamarticle.pdf (accessed 3 April 2009).

Mikkelsen, Line. 2004. *Specifying Who: On the Structure, Meaning, and Use of Specificational Copular Clauses*. Ph.D. thesis, University of California, Santa Cruz.

Milsark, Gary. 1974. *Existential Sentences in English*. Cambridge, MA: MIT Press.

Neville, Anne & Patrizia Paggio. 2001. Developing a Danish grammar in the GRASP project: A construction-based approach to topology and extraction in Danish. *Electronic Notes in Theoretical Computer Science* 53, 1–13.

Nølke, Henning. 1984. Clefting in Danish? *NyS* 14, 72–111.

Paggio, Patrizia, Philip Diderichsen & Jakob Elming. 2004. The information structure of Danish constructions. *Abstracts of the Third International Conference on Constructional Grammar*, 81. University of Provence.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Prince, Ellen F. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. In *Texas Linguistics Forum*. Austin, TX: University of Texas.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1995. *A Comprehensive Grammar of the English Language*. London: Longman.

Rochemont, Michael S. 1989. *Implementing Focus in Machine Translation* (Eurotra-D Working Papers 9). Saarbrücken: IAI/Eurotra-D.

Sag, Ivan A. 2001. Aspects of a theory of grammatical constructions. Presented at the First International Construction Grammar Conference, Berkeley, CA.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Thomsen, Ole N. 1998. Syntactic processing and word order in Danish. In Elisabeth Engberg-Pedersen & Peter Harder (eds.), *Acta Linguistica Hafniensia* 30, 129–166.

Togeby, Ole. 1992. *PRAXT: Pragmatisk tekstteori*. Århus: Aarhus Universitetsforlag.

Togeby, Ole. 2003. *Fungerer denne sætning? Funktionel dansk sproglære*. Copenhagen: Gads Forlag.

Vallduví, Enric & Elisabet Engdahl. 1996. The linguistic realisation of information packaging. *Linguistics* 34, 459–519.

Wilcock, Graham. 2005. Information structure and minimal recursion semantics. In Antti Arppe, Lauri Carlson, Krister Lindén, Jussi Piitulainen, Mickael Suominen, Martti Vainio, Hanna Westerlund & Anssi Yli-Jyrä (eds.), *Inquiries into Words, Constraints and Contexts: Festschrift for Kimmo Koskenniemi on his 60th Birthday* (CSLI Studies in Computational Linguistics Online), 268–277. Stanford, CA: CSLI Publications.

Yoshimoto, Kei, Masahiro Kobayashi, Hiroaki Nakamura & Yoshiki Mori. 2005. Processing of information structure and floating quantifiers in Japanese. In Katsuhiko Yabushita (ed.), *2nd International Workshop of Logic and Engineering of Natural Language Semantics (LENLS 2005), Kitakyushu, Japan*, 37–49.