predications indicates that *every* in (37) quantifies over dogs; $\boxed{3}$ is the variable bound by the quantifier. And similarly, the value of ARG1 of *sleep_rel* is lexically constrained to corresponds to the index of the subject, itself constrained to be identical to the value of ARG0 for the *dog_rel* predication (i.e, $\boxed{3}$). Second, *prbly_rel* is required to outscope *sleep_rel* (a *qeq* constraint either identifies its HARG or LARG or it constrains its HARG to outscope its ARG). Similarly, the restriction of *every_rel* is constrained to outscope *dog_rel* as $\boxed{4} =_{qeq} \boxed{6}$. Finally, the global top (the value of GTOP) is constrained to outscope the local top (the value of LTOP). (To simplify, the local top is the handle of the elementary predication that is not a quantifier with the widest scope.) The semantic representation that the grammar of English motivates remains underspecified, as it does not specify what the value of the BODY of *every_rel* is, in particular whether it is the handle of the *prbly_rel* or *sleep_rel* elementary predications. Resolving this scope ambiguity amounts to adding an HCONS that identifies the value of BODY with either handle, i.e., $\boxed{7}$ or $\boxed{9}$.

Examples that include multiple quantifiers work in a similar way. Take the sentence in (39) and the elementary predications for *every*, *chases*, and *some* (we only include relevant elementary predications and attributes for simplicity). We know that the body of *every_rel* and *some_rel* each outscope *chase_rel* (so $\boxed{1} =_{qeq} \boxed{3}$ and $\boxed{1} =_{qeq} \boxed{3}$, where the left hand side of the equality corresponds to the HARG and the right hand side to the LARG). But, we do not know if *every_rel* outscopes *some_rel* or the reverse; adding either HCONS $\boxed{1} =_{qeq} \boxed{2}$ or $\boxed{2} =_{qeq} \boxed{1}$ specifies which is the case. (This example illustrates that $=_{qeq}$ is not commutative, as it is meant to encode greater or equal scope.)

(39) Every dog chases some cat.

(40) $\begin{bmatrix} \textit{every\_rel} & \\ \text{RESTR} & \textit{handle} \\ \text{BODY} & \boxed{1} \end{bmatrix} \quad \begin{bmatrix} \textit{some\_rel} & \\ \text{RESTR} & \textit{handle} \\ \text{BODY} & \boxed{2} \end{bmatrix} \quad \begin{bmatrix} \textit{chase\_rel} & \\ \text{LBL} & \boxed{3} \end{bmatrix}$

Semantic composition within MRS is relatively simple and is stated in (41) (Copestake et al. 2005: p. 313–314); the second clause of this semantic composition rule amounts to a case-based definition, as is true of all Semantics Principles since Pollard & Sag (1987), as different constructions determine differently the HOOK of the head-daughter (Copestake et al. 2005 only discuss intersective and scopal constructions in their paper). (A slot in (41) is defined as “a semantic argument position in a word or phrase A that is associated with syntactic constraints on the word or phrase B whose semantics will supply that argument when the relevant grammar rule combines A and B” (op.cit., p.313).)

(41)
1. The RELS value of a phrase is the concatenation (append) of the RELS values of the daughters.
2. The HCONS of a phrase is the concatenation (append) of the HCONS values of the daughters.
3. The HOOK of a phrase is the HOOK of the semantic head daughter, which is determined uniquely for each construction type.
4. One slot of the semantic head daughter of a phrase is identified with the HOOK in the other daughter.

This quite brief description of MRS illustrates what is attractive about it from an engineering point of view. Semantic composition is particularly simple, concatenation of lists (lists of elementary predications and constraints), percolation of the HOOK from the semantic head, and some general constraint on connectedness between the head daughter and the non-head daughter. Furthermore, resolving scope means adding $=_{qeq}$ constraints to a list of $=_{qeq}$, thus avoiding traversing the semantic tree to check on scope relations. Furthermore, a flat representation makes translation easier, as argued in Copestake et al. (1995), and has several other advantages from an engineering perspective as detailed in Copestake (2009). The ease flat representations provide comes at a cost, though, namely semantic representations are cluttered with uninterpretable symbols (handles) and, more generally, do not correspond to sub-pieces of a well-formed formula. For example, we would expect the value of a quantifier restriction and nuclear scope to be, say, formulas denoting sets (as per Barwise & Cooper 1981), not pointers to or labels of predications. This is not to say that a compositional, "standard" interpretation of MRS structures is not possible (see, for example, Copestake et al. 2001); it is rather that the model-theoretic interpretation of MRS requires adding to the model hooks and holes, abstract objects of dubious semantic import. While it is true, as Copestake et al. point out, that abstract objects have been included in the models of other semantic approaches, DRT in particular (Zeevat 1989), abstract objects in compositional specifications of DRT and other such dynamic semantic approaches are composed of semantically interpretable objects. In the case of DRT, the set of variables (discourse referents) that form the other component of semantic representations (aside from predicative conditions) are anchored to individuals in the "traditional" model theoretic sense. Holes and hooks, on the other hand, are not necessarily so anchored, as labels (handles) do not have any interpretation in the universe of discourse.

(42) $\begin{bmatrix} \textit{every\_rel} & \\ \textsc{restr} & \textit{handle} \\ \textsc{body} & \textit{handle} \end{bmatrix}$

An example of the model-theoretic opacity of handles is provided by the compositional semantics of intersective attributive adjectives. The RELS value of *white horse*, for example, is as shown in (43) (after identification of handles due to the meaning composition performed by the *intersective_phrase* rule that (intersective) adjectival modification is a subsort of).

(43) $$\left[\text{RELS}\ \left\langle \begin{bmatrix} \textit{white\_rel} & \\ \text{LBL} & \boxed{1} \\ \text{ARG0} & \boxed{2} \end{bmatrix}, \begin{bmatrix} \textit{horse\_rel} & \\ \text{LBL} & \boxed{1} \\ \text{ARG0} & \boxed{2} \end{bmatrix} \right\rangle \right]$$

The fact that the value of ARG0 is the same for both elementary predications ($\boxed{2}$) is model-theoretically motivated: both properties are predicated of the same individual. The fact that the value of LBL is identical ($\boxed{1}$) is also motivated if labels are used to help determine the scope of quantifiers; in a quantifier like *every white horse*, the content of *white* and *horse* have the same scope, they conjunctively serve as the restriction of *every_rel*. But, the identity of the two elementary predications' labels is not *directly* model-theoretically motivated. It is a consequence of the semantic representation language that is used to model the meaning of sentences, not a consequence of the sentences truth-conditions.

6.2 Lexical Resource Semantics

Whereas MRS emphasizes underspecification in semantic representations and expresses the syntax of underspecified representations in HPSG as typed feature structures, LRS focuses primarily on fine-grained linguistic analyses with explicit higher-order logics for meaning representation and utilizes underspecification prominently in the architecture of the syntax-semantics interface. Instead of encoding underspecified representations as denotations of grammar principles, it uses the feature logic itself as a tool for underspecifying fully specific logical representations in the symbolic languages of the literature on formal semantics. This means that a grammar with LRS semantics denotes sets of syntactic structures that comprise fully explicit meaning representations in a standard logical language, but it does so with means of underspecification in the grammar principles, enlisting the techniques which HPSG developed for writing general descriptions in grammar principles in the definition of the relationship between syntactic structure and semantic representation. Grammar principles may admit a large number of structures, which in this case can be multiple semantic representations compatible with one and the same syntactic structure. An LRS analysis may then represent the readings of (33) with two generalized quantifiers as value of a semantic feature as shown in (44).

(44) a. $\forall\left(\lambda x.dog_w(x), \lambda x.\exists\left(\lambda y.cat_w(y), \lambda y.chase_w(x,y)\right)\right)$
b. $\exists\left(\lambda y.cat_w(y), \lambda y.\forall\left(\lambda x.dog_w(x), \lambda x.chase_w(x,y)\right)\right)$

The syntactic format of semantic representations is flexible and can be adapted to the purposes of the linguistic analysis at hand. While (44) chooses predicates with an argument for possible worlds, lambda abstraction over the unary predicates which translate the nominal arguments, and categorematic quantifiers of type $\langle\langle et\rangle\,\langle\langle et\rangle\, t\rangle\rangle$, in many contexts less elaborate representations will suffice, and the two readings would be rendered in a notational variant of first order languages. Other phenomena might necessitate more semantic structure. The LRS framework makes a selection of choices available to linguists to decide what is most adequate to spell out a semantic analysis.

6.2.1 Basic architecture

Lexical items contribute semantic resources to utterances; every semantic representation of an utterance must use up all and only the semantic resources provided by the lexical items in the utterance in all their legitimate combinations.[2] What is legitimate is determined by semantic principles which restrict at each phrase how the semantic resources of its daughters may be combined. What these restrictions do not rule out is permitted. Scope ambiguities between co-arguments of a verb can be seen as arising from the lack of a principled restriction to the effect that one outscope the other. In the absence of restrictions, LRS expects ambiguity. As a special property setting LRS apart from other semantic underspecification frameworks, LRS semantics exploits HPSG's notion of structure sharing in its semantic representations by permitting that semantic contributions of different lexemes may in fact be identical. For example, if two words in a clause contribute negation in their meaning, the two negations may in fact turn out to be the same negation, in which case we observe a negative concord reading. The implementation of this idea is based on the fundamental structure-sharing mechanism of HPSG, which is available throughout all levels of grammatical description.

The combinatorial semantics of phrases is encoded with structures of sort *lrs*:

(45) $\left[\textsc{sem}\ \begin{bmatrix} \textit{lrs} & \\ \textsc{excont} & \textit{me} \\ \textsc{incont} & \textit{me} \\ \textsc{parts} & \textit{list(me)} \end{bmatrix}\right]$

[2]Lexical items may be phrasal.

Signs have an attribute SEMantics with value *lrs*. External content (EXCONT) and internal content (INCONT) designate two prominent aspects of the semantics of signs. Both of these attributes have values of sort *meaningful_expression*, short *me*. The attribute EXCONT contains a term that represents the meaning of the maximal syntactic projection of the sign and is built from semantic material contributed within the projection. The INCONT is that part of a lexical sign's representation which is outscoped by any scope-taking operator that it combines with within its syntactic projection. The PARTS list records all semantic resources contributed by a given sign. The LRS Projection Principle governs the percolation of these attribute values along the syntactic head path of phrases, whereas the EXCONT and INCONT principles determine the relationship of the respective attribute values to other semantic attribute values within local syntactic trees. The most important relationships are those of term identity and of subtermhood of one term relative to another or to some designated part of another term. Subterm restrictions are in essence similar to the *qeq* constraints of MRS.

(46) a. LRS Projection Principle
In each phrase,
1. the EXCONT values of syntactic head and mother are identical,
2. the INCONT values of syntactic head and mother are identical,
3. the list in the PARTS value contains all and only the elements of the PARTS values of the daughters.

b. INCONT Principle
In each *lrs*, the INCONT value is an element of the PARTS list and a component of the EXCONT value.

c. EXCONT Principle
First, in every phrase, the EXCONT value of the non-head daughter is an element of the non-head daughter's PARTS list. Second, in every utterance, every subexpression of the EXCONT value of the utterance is an element of its PARTS list, and every element of the utterance's PARTS list is a subexpression of the EXCONT value.

The Projection Principle guarantees the percolation of EXCONT and INCONT values along the head path of syntactic phrases, and it records the semantic resources available at each phrase based on the semantic contributions of their daughters (46a). The INCONT Principle and the EXCONT Principle manage the properties of the respective attribute values. The term with minimal scope of each lexeme must be contributed by the lexeme itself and must be semantically realized within the representation of the maximal syntactic head projection (46b).

The maximal semantic meaning contribution of a maximal syntactic projection must originate from within that maximal projection, and an utterance (as a distinguished maximal projection) consists of all and only those pieces of semantic representation which are contributed by some lexeme in the utterance (46c). The meaning of an utterance is given by the semantic representation which is its EXCONT value. An ambiguous utterance receives structural analyses that are potentially only distinguished by different EXCONT values of their root node.

The constraints in (46) take care of the integrity of the semantic combinatorics. The task of the clauses of the Semantics Principle is to regulate the semantic restrictions on specific syntactic constructions (as in all previously discussed versions of semantics in HPSG). A quantificational determiner, represented as a generalized quantifier, which syntactically combines as non-head daughter with a nominal projection, integrates the INCONT of the nominal projection as a subterm into its restrictor and requires that its own INCONT (containing the quantificational expression) be identical with the EXCONT of the nominal projection. This clause makes the quantifier take wide scope in the noun phrase and forces the semantics of the nominal head into the restrictor. In (44) we observe the effect of this clause by the placement of the predicate dog in the restrictor of the universal and the predicate cat in the restrictor of the existential quantifier.

Another clause of the Semantics Principle governs the combination of quantificational NP arguments with verbal projections. If the non-head of a verbal projection is a quantificational NP, the INCONT of the verbal head must be a subexpression of the scope of the quantifier. Since this clause does not require immediate scope, other quantificational NPs which combine in the same verbal projection may take scope in between, as we can again see with the two possible scopings of the two quantifiers in (44), in particular in (44b), where the subject quantifier intervenes between the verb and the object quantifier.

The local semantics of signs is split from the combinatorial *lrs* structures in parallel to the separation of local syntactic structure from the syntactic tree structure. The local semantics remains under the traditional CONTENT attribute, where it is available for lexical selection by the valence attributes. The LOCAL value of the noun *dog* illustrates the relevant structure:

(47) $$\begin{bmatrix} \textit{local} & \\ \textsc{cat}|\textsc{val} & \begin{bmatrix} \textsc{spr} & \left\langle \textrm{Det}_{\boxed{1}} \right\rangle \end{bmatrix} \\ \textsc{content} & \begin{bmatrix} \textsc{index}|\textsc{dr} & \boxed{1}\ \mathsf{x} \\ \textsc{main} & \mathsf{dog} \end{bmatrix} \end{bmatrix}$$

The attribute DISCOURSE-REFERENT (DR) contains the variable that will be the argument of the unary predicate dog, which is the MAIN semantic contribution of the lexeme. The variable, x, does not come from the noun but is available to the noun by selection of the determiner by the valence attribute SPR. The subscripted tag $\boxed{1}$ on the SPR list indicates the identity of DR values of the determiner and the nominal head *dog*. A principle of local semantics says that MAIN values and DR values are inherited along the syntactic head path.

The semantics of phrases follows from the interaction of the (lexical) selection of local semantic structures and the semantic combinatorics that results from the principles in (46) and the clauses of the Semantics Principle:

(48)

$$\left[\begin{array}{ll} \text{PHON} & \langle \text{EVERY, DOG}\rangle \\ \text{... CONT} & \left[\begin{array}{ll}\text{INDEX}|\text{DR} & \boxed{2a}\ \mathsf{x}\\ \text{MAIN} & \boxed{3a}\ \mathsf{dog}\end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll}\text{EXC} & \boxed{2}\ \forall \mathsf{x}\,(\phi,\psi)\\ \text{INC} & \boxed{3}\ \mathsf{dog(x)} \\ \text{PTS} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{3}, \boxed{3a}\rangle\end{array}\right]\end{array}\right] \\ \& \boxed{3} \triangleleft \phi$$

$$\left[\begin{array}{ll} \text{PHON} & \langle \text{EVERY}\rangle \\ \text{... CONT} & \left[\begin{array}{ll}\text{INDEX}|\text{DR} & \boxed{2a}\ \mathsf{x}\\ \text{MAIN} & \boxed{2b}\ \forall\end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll}\text{EXC} & \boxed{2}\\ \text{INC} & \boxed{2}\ \forall \mathsf{x}\,(\phi,\psi) \\ \text{PTS} & \langle \forall, \mathsf{x}, \forall \mathsf{x}\,(\phi,\psi)\rangle\end{array}\right]\end{array}\right] \quad \left[\begin{array}{ll} \text{PHON} & \langle \text{DOG}\rangle \\ \text{SS}|\text{L} & \left[\begin{array}{ll}\text{CAT}|\text{VAL} & \left[\text{SPR}\ \langle \text{DET}_{\boxed{2a}}\rangle\right]\\ \text{CONT} & \left[\begin{array}{ll}\text{INDEX}|\text{DR} & \boxed{2a}\\ \text{MAIN} & \boxed{3a}\ \mathsf{dog}\end{array}\right]\end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll}\text{EXC} & \boxed{2}\\ \text{INC} & \boxed{3}\ \mathsf{dog(x)} \\ \text{PTS} & \langle \mathsf{dog}, \mathsf{dog}(\boxed{2a})\rangle\end{array}\right]\end{array}\right]$$

For ease of readability, the notation above omits the lambda abstractions from the generalized quantifier and chooses a notation from first order logic and not all structure sharings between pieces of the logical representation are made explicit in (48). The head noun *dog* contributes (on PARTS, PTS), the predicate dog and the application of the predicate to a lexically unknown argument, $\boxed{2a}$, identical with the DR value of *dog*. As shown in (47), the DR value of the noun is shared with the DR value of the selected determiner, which is the item contributing the variable x to the representation. In addition, *every* contributes the quantifier and the application of the quantifier to its arguments. The clause of the Semantics Principle which restricts the combination of quantificational determiners with nominal projections identifies the INC of *every* with the EXC of *dog*, and requires that the INC of dog ($\boxed{3}$) be a subterm of the restrictor of the quantifier, ϕ (notated as '$\boxed{3} \triangleleft \phi$', conjoined to the AVM describing the phrase). The identification of the EXC and INC of *every* follows from (46b-c). According to this analysis, the semantic representation of the phrase *every dog* is a universal quantification with $\mathsf{dog(x)}$ in the restrictor and unknown scope (ψ). The scope will be determined when the noun phrase combines with a verb phrase. For example, such a verb phrase could be *barks*. If its semantics is represented as a unary predicate bark, the predicate and its application to a single argument are contributed by the verb phrase, and local syntactic selection of the subject *every dog* by the verb *barks* identifies this argument as variable x, parallel to the selection of the quantifier's variable by *dog* above. The relevant clause of the Semantics Principle requires that $\mathsf{bark(x)}$ be a subterm of ψ, and the EXC, $\boxed{2}$, of the complete sentence receives the value $\forall\mathsf{x}\ (\mathsf{dog(x)}, \mathsf{bark(x)})$ as the only available reading in accordance with the EXCONT Principle:

(49)

$$\begin{bmatrix} \text{PHON} & \langle \text{EVERY, DOG, BARKS} \rangle \\ \text{... CONT} & \begin{bmatrix} \text{MAIN} & \boxed{4a}\ \mathsf{bark} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{EXC} & \boxed{2}\ \forall\mathsf{x}\ (\phi,\psi) \\ \text{INC} & \boxed{4}\ \mathsf{bark(x)} \\ \text{PTS} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{3}, \boxed{3a}, \boxed{4}, \boxed{4a} \rangle \end{bmatrix} \end{bmatrix}$$
$$\& \ \boxed{4} \triangleleft \psi$$

$$\begin{bmatrix} \text{PHON} & \langle \text{EVERY, DOG} \rangle \\ \text{... CONT} & \begin{bmatrix} \text{INDEX}|\text{DR} & \boxed{2a}\ \mathsf{x} \\ \text{MAIN} & \boxed{3a}\ \mathsf{dog} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{EXC} & \boxed{2}\ \forall\mathsf{x}\ (\phi,\psi) \\ \text{INC} & \boxed{3}\ \mathsf{dog(x)} \\ \text{PTS} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{3}, \boxed{3a} \rangle \end{bmatrix} \end{bmatrix}$$
$$\& \ \boxed{3} \triangleleft \phi$$

$$\begin{bmatrix} \text{PHON} & \langle \text{BARKS} \rangle \\ \text{SS}|\text{L} & \begin{bmatrix} \text{CAT}|\text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \text{NP}_{\boxed{2a}} \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{MAIN} & \boxed{4a}\ \mathsf{bark} \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{EXC} & \boxed{2} \\ \text{INC} & \boxed{4}\ \mathsf{bark(x)} \\ \text{PTS} & \langle \mathsf{bark}, \mathsf{bark}(\boxed{2a}) \rangle \end{bmatrix} \end{bmatrix}$$

The identity of the restrictor ϕ of the universal quantifier with $\boxed{3}$ $\mathsf{dog(x)}$ and of its scope ψ with $\boxed{4}$ $\mathsf{bark(x)}$ are determined at utterance level by the lack of other material that could be added to the two arguments of the quantifier. For example, an extraposed relative clause which belongs to *every dog* could consistently contribute its meaning representation to the restrictor, and only the absence of such additional semantic material leads to the inferred identity of $\boxed{3}$ with ϕ.

Underspecification of the structure of meaning representations in the clauses of the Semantics Principle and in lexical entries interacts with the possibility of structure sharing. If two pieces of meaning representation have the same shape and obey compatible structural conditions (as determined by relevant subterm contraints) they can be identical. Even stronger, in certain grammatical constellations, principles of grammar may even strictly require their identity. Lexical underspecification of meaning contributions moreover permits the shared construction of functors such as the construction of a polyadic quantifier from several lexical items in a sentence. These two applications of LRS lead to new possibilities of semantic composition compared to standard compositional semantics in generative grammar, because functors can be composed in (logical) syntax which cannot be semantically decomposed or cannot be decomposed within the structural limits of a surface-oriented syntax, i.e. a syntactic structure which only reflects syntactic but not semantic composition.

$$\left[\begin{array}{ll} \text{PHON} & \langle \text{NIKT, NIE PRZYSZEDŁ} \rangle \\ \ldots\ \text{CONT} & \left[\text{MAIN}\ \boxed{4a}\ \text{come}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{EXC} & \boxed{1}\ \neg\exists x\,(\phi,\psi) \\ \text{INC} & \boxed{4}\ \text{come}(x) \\ \text{PTS} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{3}, \boxed{3a}, \boxed{4}, \boxed{4a} \rangle \end{array}\right] \end{array}\right]$$
$$\& \ \boxed{4} \triangleleft \psi$$

(50)

$$\left[\begin{array}{ll} \text{PHON} & \langle \text{NIKT} \rangle \\ \ldots\ \text{CONT} & \left[\begin{array}{ll} \text{INDEX}|\text{DR} & \boxed{2a}\ \text{x} \\ \text{MAIN} & \boxed{3a}\ \text{person} \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{EXC} & \boxed{2}\ \exists x\,(\phi,\psi) \\ \text{INC} & \boxed{3}\ \text{person}(x) \\ \text{PTS} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{2c}\neg\beta, \boxed{3}, \boxed{3a} \rangle \end{array}\right] \end{array}\right]$$
$$\& \ \boxed{3} \triangleleft \phi \ \& \ \boxed{2} \triangleleft \beta$$

$$\left[\begin{array}{ll} \text{PHON} & \langle \text{NIE PRZYSZEDŁ} \rangle \\ \text{SS}|\text{L} & \left[\begin{array}{ll} \text{CAT}|\text{VAL} & \left[\text{SUBJ}\ \langle \text{NP}_{\boxed{2a}} \rangle\right] \\ \text{CONT} & \left[\text{MAIN}\ \boxed{4a}\ \text{come}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{EXC} & \boxed{1} \\ \text{INC} & \boxed{4}\ \text{come}(x) \\ \text{PTS} & \langle \text{come}, \text{come}(\boxed{2a}), \boxed{4b}\neg\alpha \rangle \end{array}\right] \end{array}\right]$$
$$\& \ \boxed{4b} \triangleleft \boxed{1} \ \& \ \boxed{4} \triangleleft \alpha$$

Consider the semantic representation of the Polish sentence *nikt nie przyszed* 'nobody came' in (50). Negated finite verbs in Polish contribute a negation that must be realized within the verbs EXCONT ($\boxed{4b} \triangleleft \boxed{1}$) and outscopes the INCONT of the verb ($\boxed{4} \triangleleft \alpha$). Similarly, the existential quantifier of the n-word *nikt* 'nobody' is outscoped by negation ($\boxed{2} \triangleleft \beta$). However, in addition to the familiar restriction when the quantificational subject combines with the finite verb, Polish as a strict negative concord language requires that a negated finite verb be in the scope of at most one negation in its EXCONT, entailing identity of the two negations, $\boxed{2c} = \boxed{4b}$, and the single negation reading *nobody came* as the only admissible reading of the sentence shown in (50). To capture obligatory negation marking at finite words in Polish, a second principle of negative concord rules that if a finite verb is in the scope of negation in its EXCONT, it must itself be a contributor of negation (Richter & Sailer 2001).

The idea of identifying contributions form different constituents in an utterance is even more pronounced in cases of unreducible polyadic quantification. The reading of (51a) in which each unicorn from a collection of unicorns has a set of favorite meadows that is not the same as the set of favorite meadows of any other unicorn is known to be expressible by a polyadic quantifier taking two sets and a binary relation as arguments (51d) but it cannot be expressed by two independent monadic quantifiers (Keenan 1992).

(51) a. Every unicorn prefers different meadows.
b. different meadows: $(\gamma', \Delta)(\sigma_1, \lambda y.meadow(y), \lambda v_1 \lambda y.\rho')$
c. every unicorn: $(\forall, \gamma)(\lambda x.unicorn(x), \sigma_2, \lambda x \lambda v_2.\rho)$
d. $(\forall, \Delta)(\lambda x.unicorn(x), \lambda y.meadow(y), \lambda x \lambda y.prefer(x, y))$

(51) sketches the LRS solution to this puzzle in Richter (2016). The adjective *different* contributes an incomplete polyadic quantifier of appropriate type which integrates the representation of the nominal head of its NP into the second restrictor but leaves open a slot in the representation of its functor for another quantifier it must still combine with (51b). The determiner *every* underspecifies the realization of its quantifier in such a way that one of the possible representations yields (51c) for *every unicorn*, which is exactly of the right shape to be identified with the representation of *different meadows*, leading to the expression in (51d) for (51a). Lahm (2016) presents an alternative account of such readings with *different* using Skolem functions which also hinges on LRS-specific techniques. Iordăchioaia & Richter (2015) study Romanian negative concord constructions and represent their readings using polyadic negative quantifiers; Lahm (2018) develops a lexicalized theory of plural semantics.

6.2.2 Representation languages and notational conventions

Any LRS grammar relies on an encoding of the syntax of an appropriate semantic representation language in the feature logic. In principle, any finitary logical language can be encoded in RSRL, which covers every language that has been proposed for meaning representations in linguistics. Work in LRS has so far been couched mostly in variants of Two-sorted Type Theory (Ty2, Gallin 1975) as one of the standard languages of formal semantics, or in Montague's Intensional Logic. The type system of these logical languages are useful for underspecified descriptions in semantic principles, since relevant groups of expressions can be generalized over by their type without reference to their internal structure. For example, a clause of the Semantics Principle can use the type of generalized quantifiers to distinguish quantificational complement daughters of verbal projections and state the necessary restrictions on how they are integrated with the semantics of the verbal head daughter, while other types of complement daughters are treated differently and may even not at all be restricted by a clause in the Semantics Principle in how they integrate with the verbal semantics. The latter is often the case with proper names and definite descriptions, which can be directly integrated with the semantics of the verb by lexical argument selection.

Encodings of semantic representations in feature logic are usually assumed as given by the background LRS theory. Examples of encodings can be found in Sailer (2000) and Richter (2004). Sailer (2000) offers a correspondence proof of the encoded structures with a standard syntax of languages of Ty2. As descriptions of logical terms in literal feature logic are very cumbersome to read and write, and offer no practical advantage or theoretical insight, all publications use notational shortcuts and employ logical expressions with metavariables for their descriptions instead. As nothing depends on feature logical notation, the gain in readability outweighs any concerns about notational precision.

References

Asudeh, Ash & Richard Crouch. 2002. Glue Semantics for HPSG. In Beerman Dorothee van Eynde Frank Hellan Lars (ed.), *Proceedings of the 8th International HPSG Conference*, 1–19. Stanford: CSLI Publications.

Bach, Emmon. 1976. An extension of classical transformational grammar. In *Problems of linguistic metatheory (proceedings of the 1976 conference)*, 183–224. Michigan State University.

Baker, C. L. 1970. *Indirect questions in english*. University of Illinois at Urbana-Champaign dissertation.

Barwise, John & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4. 159–219.

Cooper, Robin. 1975. *Montague's semantic theory and transformational syntax*. University of Massachusetts, Amherst dissertation.

Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel.

Copestake, Ann. 2007. Semantic composition with (robust) minimal recursion semantics. In *DeepLP '07 proceedings of the workshop on deep linguistic processing*, 73–80. Prague.

Copestake, Ann. 2009. Slacker semantics: why superficiality, dependency and avoidance of commitment can be the right way to go. In *Proceedings of the 12th conference of the european chapter of the ACL*, 1–9. Athens: Association for Computational Linguistics.

Copestake, Ann, Dan Flickinger, Rob Malouf, Susanne Riehemann & Ivan Sag. 1995. Translation using minimal recursion semantics. In *Proceedings of the sixth international conference on theoretical and methodological issues in machine translation (TMI95)*. Leuven, Belgium.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3. 281–332. DOI:10.1007/s11168-006-6327-9

Copestake, Ann, Alex Lascarides & Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th ACL meeting*. Toulouse, France.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics and Information-based Morphology. *Journal of Linguistics* 52(2). 311–374.

Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxfor: Oxford University Press.

Egg, Markus. 2004. Flexible semantic construction: the case of reinterpretation. Saarbrücken.

Fillmore, Charles. 1999. Inversion and constructional inheritance. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 113–128. Stanford: CSLI Publications.

Gallin, Daniel. 1975. *Intensional and higher-order modal logic*. Amsterdam: North Holland.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Groenendijk, Jeroen & Martin Stokhof. 1997. Questions. In J. Van Benthem & A Ter Meulen (eds.), *Handbook of logic and language*, 1055–1124. Cambridge, MA: MIT Press.

Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Wiley-Blackwell.

Hornstein, Norbert. 1995. *Logical form, from GB to minimalism*. Oxford: Blackwell.

Iordăchioaia, Gianina & Frank Richter. 2015. Negative concord with polyadic quantifiers. The case of Romanian. *Natural Language and Linguistic Theory* 33(2). 607–658.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer.

Kasper, Robert. 1997. The semantics of recursive modification. Ohio State University.

Kay, Martin. 1979. Functional grammar. In Christine et al. Chiarello (ed.), *Proceedings of the fifth meeting of the berkeley linguistics society*, 142–158. Berkeley: Berkeley Linguistics Society.

Kay, Martin, Mark Gawron & Peter Norvig. 1992. *Verbmobil: a translation system for face-to-face dialog*. Stanford: CSLI Publications.

Keenan, Edward. 1992. Anaphora, language variation, and language universals. In J. Guenter, B. Kaiser & C. Zoll (eds.), *Proceedings of the 19th annual meeting of the berkeley linguistics society*, 503–516. Berkeley: Berkeley Linguistics Society.

Keenan, Edward & Leonard Faltz. 1985. *Boolean semantics for natural language*. Dordrect: Reidel.

Klein, Ewan & Ivan Sag. 1985. Type-driven translation. *Linguistics and Philosophy* 8. 163–201.

Koenig, Jean-Pierre & Anthony Davis. 2003. Semantically transparent linking in HPSG. In Stefan Müller (ed.), *Proceedings of the HPSG 2003 conference*, 222–235. Stanford: CSLI Publications.

Ladusaw, William. 1988. A proposed distinction between *Levels* and *Strata*. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm 2*. Seoul: Hanshin Publishing Co.

Lahm, David. 2016. ‘‘different” as a restriction on Skolem functions. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burdorf (eds.), *Proceedings of the 26th semantics and linguistic theory conference (SALT 26)*, 546–565. LSA. DOI:10.3765/salt.v26i0.3793

Lahm, David. 2018. Plural in Lexical Resource Semantics. In Stefan Müller & Frank Richter (eds.), *Proceedings of the 25th International Conference on Head-Driven Phrase Structure Grammar, University of Tokyo*, 89–109. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2018/hpsg2018-lahm.pdf.

May, Robert. 1985. *Logical form: its structure and derivation*. Cambridge, MA: MIT Press.

Montague, Richard. 1974. *Formal philosophy*. New Haven: Yale University Press.

Partee, Barbara. 1984. Compositionality. In Fred Landman & Frank Veltman (eds.), *Varieties of formal semantics*, 281–311. Dordrecht: Foris.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Pollard, Carl & Ivan Sag. 1987. *Information-based syntax and semantics*. Stanford: CSLI Publications.

Pollard, Carl & Eun Jung Yoo. 1998. A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics* 34. 415–445.

Przepiórkowski, Adam. 1998. 'A Unified Theory of Scope' revisited: Quantifier retrieval without spurious ambiguities. In Gosse Bouma, Geert-Jan Kruijff & Richard T. Oehrle (eds.), *Procedings of the FHCG98. Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar and Categorial Grammar*, 185–195. Saarbrücken. http://nlp.ipipan.waw.pl/~adamp/Papers/1998-fhcg/.

Richter, Frank. 2004. *A mathematical formalism for linguistic theories with an application in Head-Driven Phrase Structure Grammar*. Universität Tübingen Phil. Dissertation (2000). https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46230, accessed 2018-2-25.

Richter, Frank. 2016. Categorematic unreducible polyadic quantifiers in Lexical Resource Semantics. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 599–619. Stanford, CA: CSLI Publications.

Richter, Frank & Manfred Sailer. 2001. Polish negation and lexical resource semantics. In Geert-Jan Kruijff, Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings FG-MOL 2001* (Electronic Notes in Theoretical Computer Science 53). Elsevier Science.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In *ESSLLI 2003 course material*, vol. 5 (Collegium Logicum). Kurt Gödel Society Wien.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sailer, Manfred. 2000. *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar*. Eberhard-Karls-Universität Tübingen Dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46191, accessed 2018-2-25.

Searle, John R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge, UK: Cambridge University Press.

Zeevat, Henk. 1989. A compositional approach to discourse representation theory. *Linguistics and Philosophy*. 1st ser. 12. 95–131.

Chapter 24

Information structure

Kordula De Kuthy

Universität Tübingen

Information structure as the hinge between sentence and discourse has been at the center of interest for linguists working in different areas such as semantics, syntax or prosody for several decades. A constraint-based grammar formalism such as HPSG that encodes multiple levels of linguistic representation within the architecture of signs opens up the possibility to elegantly integrate such information about discourse properties. In this chapter, I discuss a number of approaches that have explored how to best integrate information structure as a separate level into the representation of signs. I discuss which lexical and phrasal principles have been implemented in these approaches and how they constrain the distribution of the various information structural features. Finally, I discuss how the various approaches are used to formulate theories about the interaction of syntax, prosody and information structure. In particular, we will see several cases where (word order) principles that used to be stipulated in syntax can now be formulated as an interaction of syntax and discourse properties.

1 Introduction

The *information structure* of a sentence captures how the meaning expressed by the sentence is integrated into the discourse. The *information structure* thus encodes which part of an utterance is informative in which way, in relation to a particular context. A wide range of approaches exists with respect to the question of what should be regarded as the primitives of the information structure.

It is now commonly assumed that there are three basic dimensions of information structure[1] that are encoded in natural languages and that have been assumed as the basic primitives: (i) a distinction between what is new information

[1]For a comprehensive overview of the different research strands with respect to the information structural dimension, see Kruijff-Korbayová & Steedman (2003).

Kordula De Kuthy. N.d. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*. prepublished version. Berlin: Language Science Press. [Preliminary page numbering]

advancing the discourse (*focus*) and what is known, i.e., anchoring the sentence in existing (or presupposed) knowledge or discourse (*background*), (ii) a distinction between what the utterance is about (*topic*, *theme*) and what the speaker has to say about it (*comment*, *rheme*), and (iii) a dimension referred to as *information status* where entities that have already been mentioned in the discourse (*given*) are distinguished from those that have not been mentioned (*new*). For all three ways of partitioning the information structure, we find approaches within the HPSG framework. Example (1) illustrates how one utterance in the context of a question can be structured according to different partitionings of information structure.

(1) Q: What does Sarah drink?

<table>
<tr><td></td><td colspan="2">background</td><td>focus</td></tr>
<tr><td>A:</td><td>Sarah</td><td>drinks</td><td>TEA.</td></tr>
<tr><td></td><td>topic</td><td colspan="2">comment</td></tr>
</table>

The focus/background division with focus as the part of an utterance that is informative with respect to the discourse is one of the most commonly adopted partitionings when studying information structure, and thus many approaches within the HPSG architecture as well assume a division into focus and background, such as the ones that will be discussed in this article: Engdahl & Vallduví (1996), De Kuthy (2002), Webelhuth (2007), Paggio (2009), Bildhauer (2008), Song & Bender (2012) and Song (2017). Less common within the HPSG framework are approaches that take topic, i.e. the material that an utterance is about, as the central notion and assume topic and comment (or theme and rheme) as the primitives of the information structure. Most approaches discussed here assume that the background has one designated (mostly referential) element functioning as the topic (or link), among them Engdahl & Vallduví (1996), De Kuthy (2002), Paggio (2009) and Song (2017).

With respect to information status (including primitives such as new and given mentioned above), the discourse status of referential elements is of interest, i.e. whether they can be linked to previously mentioned items, i.e. whether they are (discourse) old or given, or whether they haven't been mentioned before and are thus (discourse) new. The representation of information status has received comparatively little interest within the HPSG community; the approach by De Kuthy & Meurers (2011) is one of the few that explicitly integrate this dimension into their information structural architecture.

The need to represent discourse properties in a grammar architecture of signs results from the insight that in many, if not all, languages, the way utterances are realized via their syntactic structure, morphological patterns and prosody very often interacts with discourse requirements of these utterances. In other words, approaches dealing with constraints on word order in a particular construction need to encode that this particular word order is only grammatical given a particular context, or a particular accent pattern has to be connected to a particular discourse status of the accented elements.[2] Most of the approaches discussed here deal with such interface questions, and I therefore discuss the particular word order and phonetic theories that have been implemented in Sections 6 and 7 in detail. As a starting point, however, I will first discuss the various architectural designs that have been implemented in order to be able to formulate the specific theories integrating discourse constraints.

2 Information structure in the architecture of signs

Several ways of representing information structure within the architecture of signs have been pursued as part of the HPSG framework: one of the earliest approaches, which is similar to the idea of F-marking as pursued in many syntax-based approaches to information structure in Generative Grammar (such as Jackendoff 1972; Selkirk 1982), has been proposed by Manandhar (1994b). He assumes that all signs have an additional appropriate feature INFO-STRUC which takes as its value objects of the sort *info-type*. A sign can then have one of the subtypes of *info-type* shown in (2) as its informational marking.

(2) Type hierarchy under *info-type* of Manandhar (1994b):

info-type

focus *background*

link *tail*

The distribution of the INFO-STRUC values in a sign is determined by the *Focus Inheritance Principle*, which enforces that in every phrase, the INFO-STRUC value of the mother subsumes the values of the INFO-STRUC of all of its daughters. The consequence of this principle is that if one daughter in a phrase is in the focus

[2] For some examples in the literature where this has been explored for word order phenomena, see for example Ambridge & Goldberg (2008), De Kuthy & Konietzko (2019) and Culicover & Winkler (2019).

and the other one in the background, then the mother's INFO-STRUC value is the smallest common supertype of both, namely *info-type*.

There are two problematic aspects of such an architecture. Firstly, it leads to a proliferation of syntactic markup of non-syntactic properties, in particular once one considers the full range of information structural notions, such as focus and focus projection, multiple foci and the marking of other discourse functions such as topic. And secondly, the perspective of information structure as resulting from an independent interpretation process of syntactic markup does not support a view of syntax, information structure and intonation as directly interacting modules, a view that can be nicely implemented in a multi-layer framework such as HPSG. More common are thus approaches that encode the information structure as a separate layer, i.e., a feature with its own structural representation.

In the original setup of signs introduced in Pollard & Sag (1994), the feature CONTEXT is introduced as part of *local* objects as a place to encode information relating to the pragmatic context (and other pragmatic properties) of utterances. In Engdahl & Vallduví (1996) it is argued that it would be most natural to also represent information structural information as part of this CONTEXT feature. Engdahl & Vallduví (1996) thus introduce the feature INFO-STRUC as part of the CONTEXT and since they couch their approach into Vallduví's (1992) information packaging account, INFO-STRUC is further divided into FOCUS and GROUND. All INFO-STRUC features take entire signs as their values. The complete specification is shown in (3).

(3) Information structure in Engdahl & Vallduví (1996: 56)

$$\begin{bmatrix} \textit{sign} & \\ \text{SYNSEM}|\text{LOCAL}|\text{CONTEXT}|\text{INFO-STRUC} & \begin{bmatrix} \text{FOCUS} & \textit{sign} \\ \text{GROUND} & \begin{bmatrix} \text{LINK} & \textit{sign} \\ \text{TAIL} & \textit{sign} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Another approach locating the representation of information structure within the CONTEXT feature is the one by Paggio (2009) as part of a grammar of Danish. The INFOSTR features TOPIC, FOCUS and BG take as their values lists of indices. Since Paggio (2009) includes Minimal Recursion Semantics (MRS, Copestake et al. 2005) as the semantic representation,[3] these indices can be structure-shared with the argument indices of the semantic relations collected on the RELS list of the content of a sign. The basic setup is illustrated in (4).

(4) Information structure in Paggio (2009: 149):

[3] A detailed discussion of the properties and principles of MRS as implemented in HPSG can be found in Koenig & Richter (2019), Chapter 23 of this volume.

$$\begin{bmatrix} \textit{sign} & \\ \textsc{synsem}|\textsc{local}|\textsc{context}|\textsc{infostr} & \begin{bmatrix} \textsc{focus} & \textit{list-of-indices} \\ \textsc{topic} & \textit{list-of-indices} \\ \textsc{bg} & \textit{list-of-indices} \end{bmatrix} \end{bmatrix}$$

Several approaches encode information structure as part of the CONTENT, such as Song (2017) and Song & Bender (2012). Since they also use MRS as the semantic representation language, they enrich the architecture of *mrs* structures. The information structure itself is encoded via a feature ICONS (individual constraints) that is introduced parallel to HCONS (handle constraints) as part of the CONTENT, as shown in (5). Song (2017) and Song & Bender (2012) use *diff-list* as values for the features RELS, HCONS and ICONS (expressed by the "!" at the beginning and the end of the list). This type of list includes an explicit pointer to the last element of the list.

(5) Information structure in Song & Bender (2012) and Song (2017: 116):

$$\begin{bmatrix} \textit{sign} & \\ \textsc{synsem}|\textsc{local}|\textsc{content} & \begin{bmatrix} \textit{mrs} & \\ \textsc{hook} & \begin{bmatrix} \textit{hook} & \\ \textsc{icons-key} & \textit{info-str} \\ \textsc{clause-key} & \textit{event} \end{bmatrix} \\ \textsc{rels} & \textit{diff-list} \\ \textsc{hcons} & \textit{diff-list} \\ \textsc{icons} & \langle !\ \ldots, \begin{bmatrix} \textit{info-str} & \\ \textsc{clause} & \textit{individual} \\ \textsc{target} & \textit{individual} \end{bmatrix}, \ldots\ ! \rangle \end{bmatrix} \end{bmatrix}$$

The type *info-str* used as the value for elements on the ICONS list is divided into an elaborate hierarchy with several subtypes, such as *semantic-focus, constrast-focus, focus-or-topic, non-focus*, etc. (cf. Song 2017: 114). The elements of type *info-str* on the ICONS list have two appropriate features CLAUSE and TARGET. TARGET is always structure-shared with the respective sign's ARG0 value, and the value of CLAUSE is always structure-shared with the INDEX value of the predicate that is the semantic head of the clause.

As pointed out by De Kuthy (2002), assuming that the information structure is part of *local* objects (which it is if it is part of the CONTEXT in HPSG as proposed by Engdahl & Vallduví 1996 or part of the CONTENT) is problematic in connection with a trace-based account of unbounded dependencies. Traces should not contribute anything to the information structure of a sentence. If one wants to develop an information structure approach which is independent of the decision of which kind of UDC theory one assumes, the only options for placing the information structure attribute are under *synsem* objects or at the top level.

Information structure as part of *synsem* objects would suggest that it plays a role in syntactic selection. This possibility is assumed in Bildhauer & Cook (2011), and they thus represent INFO-STRUC as a feature appropriate for *synsem* objects (their account will be discussed in more detail in Section 6). A third possibility is argued for in De Kuthy (2002) and Bildhauer (2008), namely that information structure should not be part of *synsem* objects. As a result, they encode information structure again as an additional feature of signs (similar to the approach by Manandhar 1994b discussed above), but it is argued that the appropriate values should be semantic representations. Using indices as the value of information structure-related features (as in the approaches by Paggio 2009, Song & Bender 2012 and Song 2017) is again problematic whenever two constituents share their index value, but only one of them is assigned a particular information structural function. For example, under the assumption that in a head-adjunct phrase the index is structure-shared between an intersective adjective and the nominal head (as in *red car*), there is no way to relate a particular information structure function (e. g., contrast) to the adjective alone (as in *RED car*).

In De Kuthy (2002), a tripartite partition of information structure into focus, topic and background is introduced. As to the question of what kinds of objects should be defined as the values of these features, De Kuthy proposes the values of the INFO-STRUC features to be chunks of semantic information. It is argued that the semantic representation proposed in Pollard & Sag (1994) is not appropriate for her purpose, because the semantic composition is not done in parallel with the syntactic build-up of a phrase. Instead, the Montague-style (cf. Dowty et al. 1981) semantic representation for HPSG proposed in Sailer (2000) is adopted, in which CONTENT values are regarded as representations of a symbolic language with a model-theoretic interpretation. As the semantic object language under CONTENT the language Ty2 (cf. Gallin 1975) of two-sorted type theory is chosen. The resulting feature architecture is shown in (6).

(6) The structure of INFO-STRUC in De Kuthy (2002: 165):

$$\begin{bmatrix} \textit{sign} & \\ \text{PHON} & \textit{list} \\ \text{SYNSEM} & \textit{synsem} \\ \text{INFO-STRUC} & \begin{bmatrix} \textit{info-struc} & \\ \text{FOCUS} & \textit{list-of-mes} \\ \text{TOPIC} & \textit{list-of-mes} \end{bmatrix} \end{bmatrix}$$

The information structure is encoded in the attribute INFO-STRUC that is appropriate for signs and has the appropriate features FOCUS and TOPIC, with lists of so-called meaningful expressions (semantic terms, cf. Sailer 2000) as values. These meaningful expressions (that are also used as the representation of logical forms as the CONT values) are lambda terms formulated in a predicate logic language as discussed in more detail in Section 3.2.2 in (13).

3 Information structure principles

The approaches sketched above all assume that signs contain some kind of structural representation of information structure, with the consequence that they need to introduce principles that constrain the values of the information structural features. Most approaches thus formulate two types of principles as part of their grammar fragment: one set of principles on the lexical level tying information structure to word level properties such as accents, and another set of principles on the phrasal level determining the distribution of information structure values between mother and daughters in a phrase.

3.1 Instantiating information structure on the word level

In the approach of Engdahl & Vallduví (1996), prosodic properties of English, in particular accent placement, are tied to specific information structural properties of words and phrases. On the word level, they thus introduce two principles that instantiate the information structure FOCUS and LINK when the word has a particular accent. The two principles are shown in (7).

(7) Information structure of words (Engdahl & Vallduví 1996: 56):

$$\textit{word} \rightarrow \boxed{1}\begin{bmatrix}\textsc{phon}|\textsc{accent} & A\\ \textsc{info-struc}|\textsc{focus} & \boxed{1}\end{bmatrix}$$

$$\textit{word} \rightarrow \boxed{1}\begin{bmatrix}\textsc{phon}|\textsc{accent} & B\\ \textsc{info-struc}|\textsc{ground}|\textsc{link} & \boxed{1}\end{bmatrix}$$

Words with an A accent always contribute focal information, i.e. the entire sign is structure-shared with the INFO-STRUC|FOCUS value; words carrying a B accent contribute link information, i.e. the entire sign is structure-shared with the INFO-STRUC|GROUND|LINK value.[4]

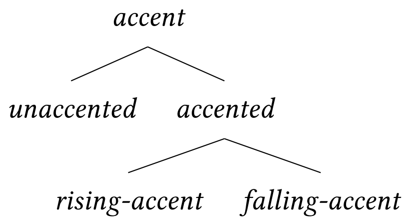
A similar set of word level principles is introduced in the approach of De Kuthy (2002), where the information structure of utterances in German is also tied to

[4] The usage of the terms “A accent” and “B accent” goes back to Jackendoff (1972).

words carrying particular accent patterns. The phonology of signs is altered as shown in Figure 8 to include an ACCENT attribute to encode whether a word receives an accent or not, and whether it is a rising or falling accent, should it receive one.

(8) Representing pitch accents and accent type hierarchy according to De Kuthy (2002: 166):

$$\begin{bmatrix} \textit{sign} & \\ \textsc{phon} & \begin{bmatrix} \textsc{phon-string} & \textit{list} \\ \textsc{accent} & \textit{accent} \end{bmatrix} \end{bmatrix}$$



The information structure of words is defined through the principle shown in Figure 9 which assigns the semantic contribution of the word to the focus or topic specification in the information structure representation of that word, depending on the type of accent the word receives.

(9) Principle assigning information structure to words (De Kuthy 2002: 167):

$$\textit{word} \rightarrow \begin{bmatrix} \textsc{phon}|\textsc{accent} & \textit{falling-accent} \\ \textsc{ss}|\textsc{loc}|\textsc{cont}|\textsc{lf} & \boxed{1} \\ \textsc{info-struc} & \begin{bmatrix} \textsc{focus} & \langle \boxed{1} \rangle \\ \textsc{topic} & \langle \rangle \end{bmatrix} \end{bmatrix} \vee$$

$$\begin{bmatrix} \textsc{phon}|\textsc{accent} & \textit{unaccented} \\ \textsc{info-struc} & \begin{bmatrix} \textsc{focus} & \langle \rangle \\ \textsc{topic} & \langle \rangle \end{bmatrix} \end{bmatrix} \quad \vee \ldots$$

Here only two cases are spelled out, one for *falling-accent* signalling focus, and one for unaccented words not contributing anything to the information structure. Other possible cases could for example be a specific accent (like a fall-rise) signalling topic, i.e. a non-empty TOPIC list.

In the approach of Song (2017), lexical items are subtypes of four different *icons-lex-item* types, which specify whether lexical items can contribute any information structural information to the *icons* list, and if yes, how many items can do this. These four lexical subtypes are shown in (10).

(10) Lexical types specifying ICONS values (Song 2017: 137)

a. $\begin{bmatrix} \textit{no-icons-lex-item} \\ \textsc{mkg} & \begin{bmatrix} \textsc{fc} & \textit{na} \\ \textsc{tp} & \textit{na} \end{bmatrix} \\ \textsc{icons} & \langle\,!\;!\,\rangle \end{bmatrix}$

b. $\begin{bmatrix} \textit{basic-icons-lex-item} \\ \textsc{icons}\ \langle\,!\;!\,\rangle \end{bmatrix}$

c. $\begin{bmatrix} \textit{one-icons-lex-item} \\ \textsc{icons}\ \langle\,![\,]!\,\rangle \end{bmatrix}$

d. $\begin{bmatrix} \textit{two-icons-lex-item} \\ \textsc{icons}\ \langle\,![\,],[\,]!\,\rangle \end{bmatrix}$

Lexical entries for elements that cannot be marked with respect to information structure are of type *no-icons-lex-items*, such as relative pronouns or expletives in English. Nominal items, such as common nouns, proper nouns and pronouns, have lexical entries of type *basic-icons-lex-item.* These types of words can have an information structural marking, but don't have to. The two other lexical subtypes are used for verbs with one clausal argument (*one-icons-lex-item*) or two clausal arguments (*two-icons-lex-items*). The information structural contribution of these clausal arguments then has to be part of the verb's ICONS list. All other verbs are not required to have any elements on their ICONS list and can thus also be of type *basic-icons-lex-item.*

To capture further constraints on the information structure properties on the word level, such as accent patterns triggering focus or topic, lexical rules are formulated in Song (2017) that derive lexical entries with the respective specifications. One such set of lexical rules for A and B accents in English is discussed in Section 7.

3.2 Information structure principles on the phrasal level

3.2.1 Information packaging (Engdahl & Vallduví 1996)

One of the first approaches integrating an explicit representation of information structure into the HPSG architecture, Engdahl & Vallduví (1996) encode the information structure as part of the CONTEXT of signs with the help of an additional feature INFO-STRUC. As discussed above, on the lexical level the instantiation of these features can be triggered by phonetic properties, such as certain accents, for intonation languages like English. Phrasal signs must then satisfy the INFO-STRUC instantiation constraints in (11).[5]

[5]Engdahl and Vallduví's formulation of the principle is incompatible with the model theoretic view of HPSG in Pollard & Sag (1994). Feature structures are complete models of objects, thus

(11) INFO-STRUC *instantiation principles for English:*

Either (i) if a DAUGHTER's INFO-STRUC is instantiated, then the mother inherits this instantiation (for narrow foci, links and tails),

or (ii) if the most oblique DAUGHTER's FOCUS is instantiated, then the FOCUS of the mother is the sign itself (wide focus).

An example including a wide VP focus licensed by the principle in (11) with the relevant INFO-STRUC values is shown in Figure 1.

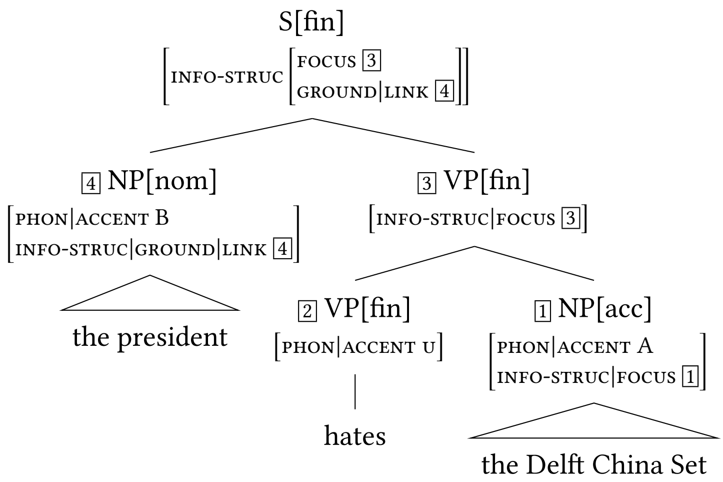


Figure 1: An example for VP focus in Engdahl & Vallduví (1996: 59)

In this example, the rightmost NP daughter *the Delft China Set* carries an A accent. According to the principle in (7) shown earlier, the entire sign is thus structure-shared with the focus value (or, in Engdahl & Valluvi's terms, the focus value "is instantiated"). As a consequence, the second clause of the principle in (11) applies and the focus value of the VP mother is the sign itself, which is then inherited by the sentence. Several aspects of the licensing of the structure in Figure 1 are not properly spelled out in Engdahl & Vallduvi's approach. For example, the analysis seems to presuppose a set of additional principles for focus inheritance in nominal phrases which do not straightforwardly follow from the principles formulated in (11).

there is no way in which a value can not be instantiated in a feature structure. Only descriptions of feature structures can be underspecified, but not the feature structures themselves.

3.2.2 Information structure as structured meanings (De Kuthy 2002)

The so-called structured meaning approach (von Stechow 1981; Jacobs 1983; Krifka 1992) to information structure provides a compositional semantic mechanism based on separate representations of the semantic contribution of the focus and that of the background. De Kuthy (2002), De Kuthy & Meurers (2003) and Webelhuth (2007) worked out how such a structured meaning approach can be integrated into the HPSG architecture.

As discussed above, in De Kuthy (2002), the information structure is encoded in the attribute INFO-STRUC that is appropriate for signs and has the appropriate features FOCUS and TOPIC, with lists of so-called meaningful expressions as values. The background of a sentence in De Kuthy's approach is then defined to be that part of the logical form of the sentence which is neither in focus nor in topic. This characterization of background closely resembles the definition of background employed by the *structured meaning* approach to focus (cf. Krifka 1992). The INFO-STRUC value of a simple sentence with the focus as indicated in (12) is thus structured as shown in (13).

(12) Peter [[liest ein BUCH]]$_F$.
Peter reads a book

(13) A sign representation including information structure (De Kuthy 2002: 163):

$$\begin{bmatrix} \text{S}|\text{LOC}|\text{CONT}|\text{LF} & \exists x[book'(x) \wedge read'(p,x)] \\ \text{INFO-STRUC} & \begin{bmatrix} \text{FOCUS} & \langle \lambda y \exists x[book'(x) \wedge read'(y,x)] \rangle \\ \text{TOPIC} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

The information-structure values of phrases are constrained by principles such as the one in (14). The original principle formulated in De Kuthy (2002: 169) only contains the first two disjuncts shown in (14). The third disjunct is added in De Kuthy & Meurers (2003). Sentences where the focus or the topic does not project represent the most basic case: only those words bearing an accent are in the topic or in the focus of an utterance.

(14) Principle 1: Extended focus projection principle (De Kuthy & Meurers 2003):

$$phrase \rightarrow \begin{bmatrix} \text{INFO-STR}|\text{FOCUS } \boxed{1} \oplus \textit{collect-focus}\left(\boxed{2}\right) \\ \text{HEAD-DTR}|\text{INFO-STR}|\text{FOCUS } \boxed{1} \\ \text{NON-HEAD-DTRS } \boxed{2} \end{bmatrix} \vee$$

$$\begin{bmatrix} \text{PHON}|\text{PHON-STR } \textit{list} \oplus \boxed{2} \\ \text{SS}|\text{LOC} \begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{noun} \vee \textit{prep} \\ \text{CONT}|\text{LF} & \boxed{3} \end{bmatrix} \\ \text{INFO-STR}|\text{FOCUS } \langle\boxed{3}\rangle \\ \textit{any-dtr}\left(\begin{bmatrix} \text{PHON}|\text{PHON-STR} & \boxed{2} \\ \text{SS}|\text{L}|\text{CONT}|\text{LF} & \boxed{4} \\ \text{INFO-STR}|\text{FOCUS} & \langle\boxed{4}\rangle \end{bmatrix}\right) \end{bmatrix} \vee$$

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC} \begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{verb} \\ \text{CONT}|\text{LF} & \boxed{3} \end{bmatrix} \\ \text{INFO-STR}|\text{FOCUS } \langle\boxed{3}\rangle \\ \text{NON-HEAD-DTRS } \left\langle \ldots, \begin{bmatrix} \text{SYNSEM} \begin{bmatrix} \text{FPP } \textit{plus} \\ \text{LOC}|\text{CONT}|\text{LF } \boxed{4} \end{bmatrix} \\ \text{INFO-STR}|\text{FOCUS } \langle\boxed{4}\rangle \end{bmatrix}, \ldots \right\rangle \end{bmatrix} \vee \ldots$$

In this case, the mother of a phrase just collects the focus values of all her daughters as ensured by the principle in (14).[6] The relation *collect-focus* ensures that from the list of non-head daughters, the FOCUS value of every non-head daughter is added to the list of FOCUS values of the entire phrase. A similar principle is needed to determine the TOPIC value of phrases.

For cases of so-called focus projection[7] in NPs and PPs, it is assumed in De Kuthy (2002: 169) that it is sufficient to express that the entire NP (or PP) can be focused if the rightmost constituent in that NP (or PP) is focused, as expressed by the second disjunct of the principle in (14). If focus projection is possible in a certain configuration then this is always optional, therefore the focus projection principle for nouns and prepositions is formulated as a disjunct. The second

[6] The presentation differs from that in De Kuthy (2002); it is the one from De Kuthy & Meurers (2003). Definitions of the auxiliary relations:

$$\textit{any-dtr}\left(\boxed{1}\right) := \begin{bmatrix}\text{HEAD-DTR } \boxed{1}\end{bmatrix}.$$
$$\textit{any-dtr}\left(\boxed{1}\right) := \begin{bmatrix}\text{NON-HEAD-DTRS element}\left(\boxed{1}\right)\end{bmatrix}.$$

$$\textit{collect-focus}\left(\langle\rangle\right) := \langle\rangle.$$
$$\textit{collect-focus}\left(\left\langle\begin{bmatrix}\text{INFO-STRUC}|\text{FOCUS } \langle\boxed{1}\rangle\end{bmatrix} \mid \boxed{2}\right\rangle\right) := \left\langle\boxed{1} \mid \textit{collect-focus}\left(\boxed{2}\right)\right\rangle.$$

[7] Focus projection is a term commonly used to describe the fact that in an utterance with prosodic marking of focus on a word, this marking can lead to ambiguity, in that different constituents containing the word can be interpreted as focused (cf. Gussenhoven 1983; Selkirk 1995).

disjunct of the principle in (14) ensures that a phrase headed by a noun or a preposition can only be in the focus (i.e., its entire logical form is token identical to its FOCUS value) if the daughter that contributes the rightmost part of the phonology of the phrase is entirely focused itself. The relation *any-dtr* is a description of a sign with a head daughter or a list of non-head daughters and thereby ensures that it can be either the head (i.e. head daughter) of the phrase itself, or any non-head daughter that meets the condition of being focused. Again, a similar principle needs to be provided for the TOPIC value of nominal and prepositional phrases.

For the verbal domain, the regularities are known to be influenced by a variety of factors, such as the word order and lexical properties of the verbal head (cf., e.g., von Stechow & Uhmann 1986). Since verbs need to be able to lexically mark which of their arguments can project focus when they are accented, De Kuthy & Meurers (2003) introduce the boolean-valued feature FOCUS-PROJECTION-POTENTIAL (FPP) for objects of type *synsem*. Figure 15 shows the relevant part of the lexical entry of the verb *lieben* 'love' which allows projection from the object but not from the subject:

(15) The focus projection potential of *lieben* (De Kuthy & Meurers 2003):

$$\begin{bmatrix} \textsc{phon}|\textsc{phon-str}\ \langle \textit{lieben} \rangle \\ \textsc{arg-s}\ \left\langle \begin{bmatrix} \textsc{loc}|\textsc{cat}|\textsc{head} \begin{bmatrix} \textit{noun} \\ \textsc{case}\ \ \textit{nom} \end{bmatrix} \\ \textsc{fpp}\ \textit{minus} \end{bmatrix}, \begin{bmatrix} \textsc{loc}|\textsc{cat}|\textsc{head} \begin{bmatrix} \textit{noun} \\ \textsc{case}\ \ \textit{acc} \end{bmatrix} \\ \textsc{fpp}\ \textit{plus} \end{bmatrix} \right\rangle \end{bmatrix}$$

The third disjunct of the principle in (14) then specifies under which circumstances focus can project in the verbal domain: a phrase headed by a verb can only be in the focus (i.e., its entire logical form is token identical to an element of its FOCUS value) if the daughter that has the focus projection potential (FPP *plus*) is entirely focused itself.

3.2.3 Information structure principles in MRS

As introduced above, in the MRS based approach of Paggio (2009), the information structure is part of the CONTEXT, consisting of FOCUS, TOPIC and BACKGROUND features which are structure-shared with the respective INDEX values of the semantic representation of a phrase. Paggio (2009) connects the distribution of information structure values to particular clausal types and introduces new phrasal subtypes which constrain the distribution of information structure in the respective phrase. One such new phrasal subtype is the type *focus-inheritance* as defined in (16), which then has to be cross-classified with every basic phrasal subtype (such as *hd-comp, hd-spec, hd-adj*, etc.) in order to constrain the distribution of focus values across all phrasal subtypes.

(16) Principle for focus-inheritance (Paggio 2009: 155):

$$
\textit{focus-inheritance} \longrightarrow \begin{bmatrix}
\textsc{synsem}|\textsc{loc}|\textsc{context} & \begin{bmatrix} \textsc{focus} & \langle \boxed{2}, \boxed{1} \rangle \\ \textsc{bg} & \boxed{3} \end{bmatrix} \\
\textsc{hd}|\textsc{synsem}|\textsc{loc}|\textsc{context} & \begin{bmatrix} \textsc{focus} & \boxed{1} \\ \textsc{bg} & \boxed{3} \end{bmatrix} \\
\textsc{non-hd} & \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{context}|\textsc{focus} & \langle \boxed{2} \rangle \\ \textsc{accent} & \textit{true} \end{bmatrix}
\end{bmatrix}
$$

The principle in (16) ensures that for signs of type *focus-inheritance*, the list of focus values of the mother is the list of focus values of the head daughter[8] plus the focus value of the non-head daughter, in case it is accented. Similar principles are defined for the inheritance of background values, also depending on the accent status of the non-head daughter. Paggio further assumes that each phrasal subtype has further subtypes connecting it to one of the information structure inheritance phrasal types. For example, she assumes that there is a phrasal subtype *focus-hd-adj* that is a subtype both of *hd-adj* and of *focus-inheritance*. Finally, clausal types are introduced that account for the information structure values at the top level of a clause. For example, the specification for *decl-main-all-focus* as shown in (17) is a clause in which both the background and the topic values are empty and the mother collects the focus values from the head and the non-head daughters.[9]

(17) Declarative all-focus construction (Paggio 2009: 160):

$$
\begin{bmatrix}
\textit{decl-main-all-focus} & \\
\textsc{ctxt}|\ldots & \begin{bmatrix} \textit{all-focus} & \\ \textsc{topic} & \langle\rangle \\ \textsc{focus} & \langle \boxed{2}, \boxed{1} \rangle \\ \textsc{bg} & \langle\rangle \end{bmatrix}
\end{bmatrix}
$$

$$
\begin{bmatrix} \textsc{ctxt}|\ldots|\textsc{focus} & \langle \boxed{1} \rangle \end{bmatrix} \qquad \begin{bmatrix} \textsc{ctxt}|\ldots|\textsc{focus} & \boxed{2} \end{bmatrix}
$$

[8] This is not correctly specified in the original principle as formulated by Paggio (2009). If the head daughter can have a list with more than one element as its FOCUS value, then this entire list would have to be added to the list of FOCUS values of the mother, and not just be one element of that list.

[9] Again, the list specifications as formulated by Paggio (2009) are not entirely correct: if the head daughter's FOCUS value $\boxed{2}$ is a list with more than one element, this list has to be added to the list of FOCUS values of the mother.

Different from Paggio's approach, Song & Bender (2012) and Song (2017) locate the representation of information structure within the MRS-based CONTENT value of signs. The list elements of information structural values that are built up for a phrase consist of focus, background or topic elements co-indexed with the semantic INDEX values of the daughters of that phrase. The main point of their approach is that they want to be able to represent underspecified information structural values, since very often a phrase, for example with a certain accent pattern, is ambiguous with respect to the context in which it can occur and thus is ambiguous with respect to its information structure values. An example they discuss is the one in (18), where the first sentence could be an answer to the question *What barks?* and thus signal narrow focus, whereas the second utterance could be an answer to the question *What happened?* and signal broad focus.

(18) a. [[The DOG]]$_F$ barks.
 b. [[The DOG barks]]$_F$.

The approach pursued in Song & Bender (2012) thus assumes that the two possible readings in (18) are further specializations of one MRS which is associated with one syntactic structure and includes underspecified values, in particular the type of the ICONS element for the constituent *barks*, leaving it open whether that is part of the focus or not.

In Song (2017), this approach is further spelled out and lexical rules allowing transitive and ditransitive verbs to be a possible source for focus projection. In an example such as (19), Song (2017) assumes that focus can only project if the last argument is accented as in (19b) (here accent is shown on the noun *book* in small caps), but not if some other argument is accented, as in (19a), where the proper noun *Lee* is accented.

(19) a. Kim sent LEE the book.
 b. Kim sent Lee the BOOK.

Accordingly, there are two lexical entries for the verb *send*, which are derived by the lexical rules shown in (20).

(20) Focus projection lexical rules (Song 2017: 227):
 a. *no-focus-projection-rule* →

$$\begin{bmatrix} \textsc{index} & \boxed{1} \\ \textsc{icons-key} & \boxed{2} \\ \textsc{val} & \begin{bmatrix} \textsc{subj} & \left\langle \begin{bmatrix} \textsc{icons-key non-focus} \end{bmatrix} \right\rangle \\ \textsc{comps} & \left\langle \begin{bmatrix} \textsc{mkg}|\textsc{fc} + \end{bmatrix} \begin{bmatrix} \textsc{mkg}|\textsc{fc} - \\ \textsc{icons} \ \langle\ !\ !\rangle \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{c-cont}|\textsc{icons} & \langle\ !\ \boxed{2} \begin{bmatrix} \textit{non-focus} \\ \textsc{target}\ \boxed{1} \end{bmatrix} !\rangle \\ \textsc{dtr}\ \textit{lex-rule-infl-affixed} & \end{bmatrix}$$

b. *focus-projection-rule* $\longrightarrow$

$$\begin{bmatrix} \textsc{clause-key} & \boxed{1} \\ \textsc{val}|\textsc{comps} & \left\langle \begin{bmatrix} \textsc{mkg}|\textsc{fc} - & \\ \textsc{index} & \boxed{2} \end{bmatrix} \begin{bmatrix} \textsc{mkg}|\textsc{fc} + \\ \textsc{icons} \ \langle\ !\ \begin{bmatrix} \textit{semantic-focus} \end{bmatrix} !\rangle \end{bmatrix} \right\rangle \\ \textsc{c-cont}|\textsc{icons} & \left\langle\ ! \begin{bmatrix} \textit{non-focus} \\ \textsc{target}\ \boxed{2} \\ \textsc{clause}\ \boxed{1} \end{bmatrix} ! \right\rangle \\ \textsc{dtr}\ \textit{lex-rule-infl-affixed} & \end{bmatrix}$$

The lexical rule *no-focus-projection-rule* requires lexical entries to have a non-focus-marked element as the last element on the COMPS list, and in addition the word itself has an ICONS-KEY of type *non-focus* preventing the word itself from being focused. The lexical rule *focus-projection-rule* has a focus-marked element as the last element in the COMPS list. It is not further specified whether only that focussed complement or also the word itself contributes anything to the ICONS value. In the example (19b), if the verb *sent* is licensed by the rule *focus-projection-rule*, either only *the book*, or the entire VP *sent Lee the book*, or even the entire sentence *Kim sent Lee the book* could be focused.

Since the approach of Song (2017) is part of a larger grammar fragment (the LinGO Grammar Matrix; Bender et al. 2010) with the aim of parsing and generating sentences from a large number of different languages, it contains a multitude of lexical and phrasal types and principles. Some of these specifications are introduced to capture very language-specific information structure properties (such as morphological markings, word order constraints, etc.), while others are necessary for the specific way in which grammar fragments in the LinGO Grammar Matrix are implemented and processed. It would be far beyond the scope of this article to discuss all these principles and specifications in detail and I therefore only included the most essential aspects of Song's approach in my discussion here.

4 Topics

Most HPSG approaches are based on a focus/background division of the information structure. To capture aspects of a topic vs. comment distinction, or to be able to specify topics as a special element in the background, they include an additional feature or substructure for topics. Engdahl & Vallduví (1996), for example, divide the GROUND into LINK and TAIL, where the link is a special element of the background linking it to the previous discourse, just like topics. In the approaches of De Kuthy (2002) and Paggio (2009), an additional feature TOPIC is introduced, parallel to FOCUS and BACKGROUND, in order to distinguish discourse referents as topics from the rest of the background.

Most approaches don't introduce separate mechanisms for the distribution of TOPIC values, but rather assume that similar principles as the ones introduced for focus can constrain topic values, as mentioned above for the approach of De Kuthy (2002). A more specific example can be found in Paggio (2009), where a constraint on topicalization constructions including a topic-comment partitioning is formulated, as illustrated in Figure 2. This *inv-topic-comment* phrasal type

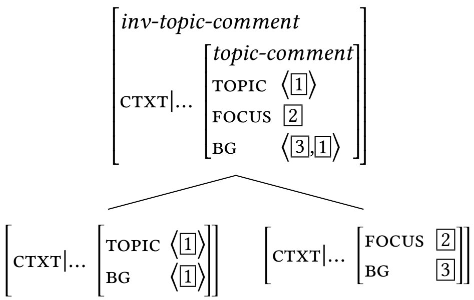


Figure 2: Topicalization construction with extracted topic

constrains the information structure values of topicalization constructions in Danish characterized by subject verb inversion,[10] where the topic corresponds to the topicalized complement, as illustrated by the example in (21) from Paggio (2009).

[10] Although Danish is generally considered to be a V2 language, where any kind of constituent (not only the subject) can occur in the position before the finite verb, Paggio (2009) seems to assume that clauses in which a dependent different from the subject, i.e. an object or some adjunct phrase, occurs before the finite verb have a different structure than those where the subject occurs in the sentence-initial position.

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

In Song (2017), a number of lexical and phrasal principles are provided with the purpose of licensing topic-comment structures. The principles and lexical entry in (22) are spelled out in order to license *wa* constructions in Japanese with a left dislocated topic phrase.

(22) Licensing topic-comment structures in Song (2017):

a. $$\textit{topic-comment} \longrightarrow \begin{bmatrix} \textsc{l-periph} & + \\ \textsc{mkg} & \textit{tp} \\ \textsc{hd}|\textsc{mkg}|\textsc{tp} & - \\ \textsc{nhd} & \begin{bmatrix} \textsc{mkg} & \textit{tp} \\ \textsc{l-periph} & + \end{bmatrix} \end{bmatrix}$$

b. $$\textit{top-scr-comp-head} \longrightarrow \begin{bmatrix} \textsc{ha}|\textsc{val}|\textsc{comps}\ \langle\rangle \\ \textsc{nhd}|\textsc{icons-key}\ \textit{constrast-topic} \end{bmatrix}$$

c. $$\textit{wa-marker} \longrightarrow \begin{bmatrix} \textsc{stem} & \langle \textit{wa} \rangle \\ \textsc{incons-key} & \boxed{2} \\ \textsc{mkg} & \textit{tp} \\ \textsc{comps} & \langle \begin{bmatrix} \textsc{index}\ \boxed{1} \end{bmatrix} \rangle \\ \textsc{incons} & \langle !\ \boxed{2} \begin{bmatrix} \textit{contrast-or-topic} \\ \textsc{target}\ \boxed{1} \end{bmatrix} ! \rangle \end{bmatrix}$$

The constraint in (22a) on the phrasal subtype *topic-comment* ensures that only the non-head daughter is marked as a topic, whereas the head daughter functions as the comment (and presumably contains some focused material). The specification [L-PERIPH +] indicates that a constituent with this feature value cannot be combined with another constituent leftward.

A Japanese topic-comment structure, such as the one in (23) (Song 2017: 198), is licensed by the phrasal subtype *top-scr-comp-head*, i.e. it is assumed that the fronted complement, the *wa*-marked NP *sono hon wa* 'the book' is scrambled to the left peripheral position and is interpreted as a contrastive topic phrase.

(23) sono hon wa Kim ga yomu.
Det book wa Kim nom read
'This book, Kim read.'

The topic marker *wa* in Japanese is treated as an adposition with the lexical specifications shown in (22c). The entire sentence is thus licensed as a head complement structure, where the object NP is scrambled to the sentence initial position and functions as a contrastive topic. The *tp* marking of the entire *topic-comment* phrase ensures that this phrase cannot be embedded as the comment in another *topic-comment* phrase.

5 Givenness

In De Kuthy & Meurers (2011), it is shown how the HPSG approach to information structure of De Kuthy (2002) and colleagues can be extended to capture givenness and to make the right predictions for so-called *deaccenting*, which has been shown to be widespread (Büring 2006). In contrast to Schwarzschild (1999), who spells out his approach in the framework of alternative semantics (Rooth 1992), they show how the notion of givenness can be couched in a standard structured meaning approach – thereby preserving the explicit, compositional representations of focus.

The example in (24) illustrates the necessity to include information about givenness into the information structural setup.

(24) The conference participants are renting all kind of vehicles. Yesterday, Bill came to the conference driving a red convertible and today he's arrived with a blue one.

   a. What did John rent?

   b. He (only) rented [[a GREEN convertible]]$_F$.

The context in (24) introduces some conference participants, Bill, the rental of vehicles and red and blue convertibles into the discourse. Based on this context, when considering the question (24a) asking for the object that John is renting as the focus, one can answer this question with sentence (24b), where *a green convertible* is the focus: out of all the things John could have rented, he picked a green convertible. In this focus, only *green* is new to the discourse, whereas convertibles were already given in the context, and still the entire NP is in the focus.

To capture such cases of focus projection, an additional feature GIVEN is introduced as part of the setup of De Kuthy (2002), as already discussed in Section 3.2.2. The relation between pitch accents and the information structure of words is still defined by the principle shown in (25), depending on the type of accent the word receives.

(25) Relating intonation and information structure for words (De Kuthy & Meurers 2011):

$$\textit{word} \;\longrightarrow\; \begin{bmatrix} \textsc{phon|accent} & \textit{accented} \\ \textsc{ss|loc|cont|lf} & \boxed{1} \\ \textsc{struc-meaning} & \begin{bmatrix} \textsc{focus} & \langle \boxed{1} \rangle \\ \textsc{given} & \langle\rangle \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \textsc{phon|accent} & \textit{unaccented} \\ \textsc{struc-meaning} & \begin{bmatrix} \textsc{focus} & \langle\rangle \\ \textsc{given} & \langle\rangle \end{bmatrix} \end{bmatrix} \vee \ldots$$

In addition, the Focus Projection Principle originally introduced in De Kuthy (2002) and then extended in De Kuthy & Meurers (2003) is extended with a disjunct capturing focus projection in the presence of givenness (De Kuthy & Meurers 2011). Figure 26 shows the resulting principle.

(26) Extended Focus Projection Principle including Givenness (De Kuthy & Meurers 2011):

$$\textit{phrase} \longrightarrow \begin{bmatrix} \textsc{struc-meaning|focus}\ \boxed{1} \oplus \textit{collect-focus}\big(\boxed{2}\big) \\ \textsc{head-dtr|info-str|focus}\ \boxed{1} \\ \textsc{non-head-dtrs}\ \boxed{2} \end{bmatrix} \vee$$

$$\begin{bmatrix} \textsc{phon|phon-str}\ \textit{list} \oplus \boxed{2} \\ \textsc{ss|loc} \begin{bmatrix} \textsc{cat|head} & \textit{noun} \vee \textit{prep} \\ \textsc{cont|lf} & \boxed{3} \end{bmatrix} \\ \textsc{struc-meaning|focus}\ \langle \boxed{3} \rangle \\ \textit{any-dtr}\left(\begin{bmatrix} \textsc{phon|phon-str} & \boxed{2} \\ \textsc{ss|l|cont|lf} & \boxed{4} \\ \textsc{struc-meaning|focus} & \langle \boxed{4} \rangle \end{bmatrix} \right) \end{bmatrix} \vee$$

$$\begin{bmatrix} \textsc{synsem|loc} \begin{bmatrix} \textsc{cat|head} & \textit{verb} \\ \textsc{cont|lf} & \boxed{3} \end{bmatrix} \\ \textsc{struc-meaning|focus}\ \langle \boxed{3} \rangle \\ \textsc{non-head-dtrs}\ \Big\langle \ldots, \begin{bmatrix} \textsc{synsem} \begin{bmatrix} \textsc{fpp}\ \textit{plus} \\ \textsc{loc|cont|lf}\ \boxed{4} \end{bmatrix} \\ \textsc{struc-meaning|focus}\ \langle \boxed{4} \rangle \end{bmatrix}, \ldots \Big\rangle \end{bmatrix} \vee$$

$$\begin{bmatrix} \textsc{ss|loc|cont|lf}\ \boxed{3} \\ \textsc{struc-meaning|focus}\ \langle \boxed{3} \rangle \\ \textsc{dtrs-list}\left(\textit{given-sign-list} \bigcirc \left\langle \begin{bmatrix} \textsc{ss|l|cont|lf} & \boxed{4} \\ \textsc{struc-meaning|focus} & \langle \boxed{4} \rangle \end{bmatrix} \right\rangle \right) \end{bmatrix} \vee \ldots$$

The new fourth disjunct of the Extended Focus Projection Principle[11] captures the cases previously unaccounted for where given material in a focused phrase is deaccented. Focus in those examples can project from a focused daughter in a position which normally does not allow focus projection. This only is an option if all other daughters in that focused phrase are *given*. Spelling this out, the fourth

[11] The auxiliary relations are defined as:

disjunct of the principle in (26) specifies that the mother of a phrase can be in the focus (i.e., the entire LF value of the mother's CONTENT is token identical to an element on the mother's FOCUS list) if it is the case that the list of all daughters (provided by *dtrs-list*, a relational description of a list containing signs that are given) consists of *given* signs into which a single *focused* sign is shuffled (◯).[12] [13] As before, a sign is focused if its LF value is token identical to an element of its FOCUS value; and a sign is given if its LF value is token identical to an element of its GIVEN value.

The pitch accent in this example is on the adjective *green* so that the principle in (9) on p. viii licenses structure sharing of the adjective's content with its FOCUS value. In the context of the question (24a), the entire NP *a green convertible* from example (24b) is in the focus. In the phrase *green convertible*, the clause licensing focus projection in NPs does not apply, since the adjective *green*, from which the focus has to project in this case, is not the rightmost element of the phrase. What does apply is the fourth disjunct of the principle licensing focus projection in connection with givenness. Since the noun *convertible* is given, the adjective *green* is the only daughter in the phrase that is not given and focus is allowed to project to the mother of the phrase. In the phrase *a green convertible*, focus projection is again licensed via the clause for focus projection in noun phrases, since the focused phrase *green convertible* is the rightmost daughter in that noun phrase.

6 Information structure and word order

The explicit representation of information structure as part of signs in HPSG opens up the possibility of providing explanations for constraints previously stipulated in syntax, such as word order constraints, by deriving the constraints from the nature of the integration of a sentence into the discourse. Many of the ap-

$$\textit{dtrs-list}\left(\left\langle \boxed{1} | \boxed{2} \right\rangle\right) := \begin{bmatrix} \text{HEAD-DTR} & \boxed{1} \\ \text{NON-HD-DTRS} & \boxed{2} \end{bmatrix}$$

$$\textit{given-sign-list} := \langle\rangle.$$

$$\textit{given-sign-list} := \left\langle \begin{bmatrix} \text{SS}|\text{L}|\text{CONT}|\text{LF} & \boxed{1} \\ \text{STRUC-MEANING} & \begin{bmatrix} \text{GIVEN} & \left\langle \boxed{1} \right\rangle \end{bmatrix} \end{bmatrix} \middle| \ \textit{given-sign-list} \right\rangle.$$

[12] The relation "shuffle" ◯ is used as originally introduced in Reape (1994): the result is a list that contains all element from the two input lists and the order of elements from the original lists is preserved.

[13] If only binary structures are assumed, as in the examples in this chapter, the principle can be simplified. Here, I kept the general version with recursive relations following De Kuthy & Meurers (2003), which also supports flatter structures.

proaches discussed in the previous section employ the information structural architecture exactly in this way and formulate principles linking word order to discourse properties.

One first such approach is presented in Engdahl & Vallduví (1996), where word order constraints for Catalan are couched into the information structure setup discussed in Section 3.2. The basic observation is that in Catalan, the word order within the sentential core is and that every constituent within this sentential core is interpreted as focal. If an argument of the main verb of a sentence is to be interpreted as non-focal, it must be clitic-dislocated. The example in (27) from Engdahl & Vallduví (1996) illustrates the two possible cases: the argument *a Barcelona* 'to Barcelona' can be topicalized as in (27b) or positioned at the end of the sentence as in (27c) in order to be interpreted as non-focal.

(27) a. Ahir [[va tornar a Barcelona el PRESIDENT]]$_F$.
yesterday 3s-past-return to Barcelona the president

b. A Barcelona$_1$ [[hi$_1$ va tornar el PRESIDENT]]$_F$.
to Barcelona there returned the president

c. [[Hi$_1$ va tornar el PRESIDENT]]$_F$ a Barcelona$_1$.
there returned the president to Barcelona
'Yesterday, the president returned to Barcelona.'

With respect to modeling this within the HPSG account, they assume that phrases associated with a LINK interpretation should be constrained to be left dislocated, whereas phrases associated with a TAIL interpretation should be right attached. They thus introduce the following ID schema for Catalan:

(28) *Head-Dislocation Schema for Catalan:*
The DTRS value is an object of sort *head-disloc-struc* whose HEAD-DTR|SYNSEM|LOCAL|CATEGORY value satisfies the description
$\left[\text{HEAD } \textit{verb} \left[\text{VFORM } \textit{finite}\right], \text{SUBCAT } \langle\rangle\right]$, and whose DISLOC-DTRS|CONTEXT|INFO-STRUC value is instantiated and for each DISLOC-DTR, the HEAD-DTR|SYNSEM|LOCAL|CONTENT value contains an element which stands in a *binding* relation to that DISLOC-DTR.

The principle requires that the information structure value of dislocated daughters of a finite sentence has to be GROUND. An additional LP statement then captures the relation between the directionality of the dislocation and a further restriction of the GROUND value, as illustrated in (29).

(29) LP constraint on information structure in Catalan (Engdahl & Vallduví 1996):

LINK > FOCUS > TAIL

This LP statement is meant to ensure that link material must precede focus material and focus material must precede tails. By this, Engdahl & Vallduví (1996) want to ensure that left-detached constituents are always interpreted as links and right-detached constituents as tails.

The insights from Engdahl & Vallduvi's approach are the basis for an approach accounting for clitic left dislocation in Greek presented by Alexopoulou & Kolliakou (2002). The representation of information structure with the features FOCUS and GROUND (further divided into LINK and TAIL) is taken over as well as the phonological constraints on words and the information structure instantiation principle. In order to account for clitic left dislocation, as illustrated in (30), an additional feature CLITIC is introduced as appropriate for *nonlocal* objects.

(30) a. Pii simetehoun s' afti tin paragogi?
who take part in that the production
'Who contributed to this production?'

b. Tin parastasi *ti* skinothetise o Karolos KOUN …
the performance FEM.3SG.ACC directed the Karolos Koun
'Karolos Koun directed the performance …'

The Linkhood Constraint shown in (31) ensures that links (i.e. elements whose INFO-STRUC|LINK value is instantiated) can only be fillers that are "duplicated" in the morphology by a pronominal affix, i.e. it is required that there is an element $\boxed{1}$ on the CLITIC list of the head daughter that is structure-shared with the filler's HEAD value. The use of the disjoint union relation $\uplus$[14] ensures that the singleton element $\boxed{1}$ representing the doubled clitic is the only element on the phrase's clitic list with these specifications. In addition, it is required that the filler-daughter $\boxed{2}$ is structure-shared with the LINK attribute in the information structure of the mother.

(31) The Linkhood Constraint for clitic left dislocation phrases (Alexopoulou & Kolliakou 2002: 238):

$$\begin{bmatrix} \textit{clitic-left-disloc-phrase} & \\ \text{INFO-STRUC}|\text{LINK} & \{\boxed{2}\} \\ \text{CLITIC} & \boxed{\Sigma_2} \end{bmatrix} \rightarrow \boxed{2}\begin{bmatrix} \text{PHON}|\text{ACCENT} & u \\ \text{HEAD} & \boxed{1} \end{bmatrix}, \mathbf{H}\begin{bmatrix} \textit{phrase} & \\ \text{HEAD} & \textit{verb} \\ \text{CLITIC} & \{\boxed{1}\} \uplus \boxed{\Sigma_2} \end{bmatrix}$$

[14] Alexopoulou & Kolliakou (2002) provide no exact definition for the use of the symbol $\uplus$ (disjoint union), but a definition that is often used within HPSG approaches can be found in Manandhar (1994a).