Predicate Complements

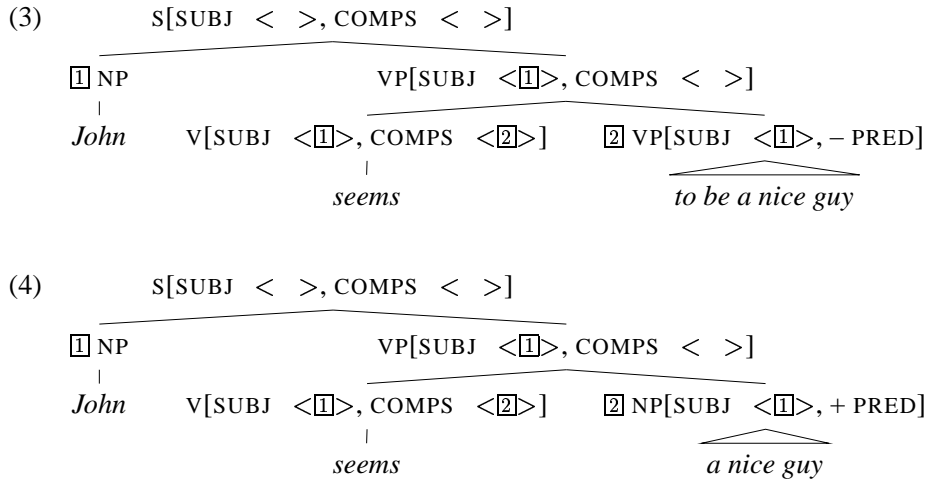## Frank Van Eynde

University of Leuven

Abstract

This paper briefly presents the canonical HPSG treatment of predicate complements (Section 1) and points out a number of problems with it (Section 2). Then it presents an alternative (Section 3) and shows how it avoids or solves the problems with the canonical treatment (Section 4).

1 The raising treatment of the predicate complements

In mainstream HPSG predicate complements are treated along the same lines as complements of raising verbs. Prima facie evidence is provided by the minimal pairs in (1) and (2).

(1) a. John seems to be a nice guy.

　　b. John seems a nice guy.

(2) a. Bob considers his brother to be a genius.

　　b. Bob considers his brother a genius.

After analogy with the analysis of (1a), in which the unexpressed subject of the VP complement is identified with the subject of the raising verb, as illustrated in (3), the NP complement in (1b) is assumed to have an unexpressed subject, which is identified with the subject of the predicate selecting verb, as illustrated in (4).[1]

(3)　　　S[SUBJ < >, COMPS < >]

　　[1] NP　　　　　　　　VP[SUBJ <[1]>, COMPS < >]

　　John　　V[SUBJ <[1]>, COMPS <[2]>]　　[2] VP[SUBJ <[1]>, − PRED]

　　　　　　　　　seems　　　　　　*to be a nice guy*

(4)　　　S[SUBJ < >, COMPS < >]

　　[1] NP　　　　　　　　VP[SUBJ <[1]>, COMPS < >]

　　John　　V[SUBJ <[1]>, COMPS <[2]>]　　[2] NP[SUBJ <[1]>, + PRED]

　　　　　　　　　seems　　　　　　*a nice guy*

[1]This treatment is advocated in chapter 3 of Pollard and Sag (1994). For an application to German, see (Müller, 2002, 103-9). It can be extended straightforwardly to the sentences in (2), where the unexpressed subject of the most oblique complement is identified with the direct object of *considers*.

What differentiates (3) from (4) is captured by the PRED feature. Its value is positive when the complement is used in predicate position, and negative otherwise. The feature is included in the HEAD value of the nouns, the verbs, the adjectives and the prepositions. Words which must be (the head daughter of) a predicate complement are lexically marked as [+ PRED], words which cannot be (the head of) a predicate complement are marked as [– PRED], and words which may but need not be (the head of) a predicate complement have the underspecified value *boolean*. Most words belong to the third class, but there are some which are inherently [+ PRED], such as the adjectives *ready, asleep* and *ablaze*, and some which are inherently [– PRED], such as the adjectives *wooden, mere, utter* and *former*. The negative value is also assigned to the infinitives, the gerunds, the perfect participles and the finite verbs; only the present and the passive participles can be [+ PRED], see (Gazdar et al., 1985, 111), (Pollard and Sag, 1987, 64) and (Ginzburg and Sag, 2000, 25).

2 Problems

There are several problems with the canonical HPSG treatment of the predicate complements. I will group them in four classes.

2.1 Nominal predicates

Since nouns are not the kind of words that are commonly thought of as taking a subject, special measures are needed to provide them with a non-empty SUBJ list when they are used as predicates. To this end, the lexical rules which map nominal lexemes onto words are enriched with information about the value of PRED and SUBJ, as in the Singular Predicative Noun Lexical Rule (Ginzburg and Sag, 2000, 409).

$$(5) \quad \begin{bmatrix} lexeme \\ \text{ARG-ST} & \left\langle \boxed{1} \right\rangle \oplus \boxed{A} \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \quad noun \end{bmatrix} \Longrightarrow_{LR} \begin{bmatrix} word \\ \text{ARG-ST} & \left\langle \boxed{2}, \boxed{1} \right\rangle \oplus \boxed{A} \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{AGR} \mid \text{NUM} \quad sing \\ \text{PRED} \quad + \end{bmatrix} \\ \text{SPR} \quad \left\langle \boxed{1} \right\rangle \\ \text{SUBJ} \quad \left\langle \boxed{2} \right\rangle \end{bmatrix} \end{bmatrix}$$

In words, the noun is marked as singular and [+ PRED] and its ARG-ST value gets an extra member that is realized as a subject.

This rule does the job it is tailored to do, but it has the undesirable consequence of introducing a systematic ambiguity. Since every noun which can be used in predicative position can also be used in non-predicative positions (and vice versa), we need two entries for every noun: a [+ PRED] one with a singleton SUBJ list and

a [– PRED] one with an empty SUBJ list. The latter are derived by the Singular Attributive Noun Lexical Rule (Ginzburg and Sag, 2000, 190). Besides, the plural nouns show the same systematic ambiguity. Admittedly, the postulation of lexical ambiguities is unavoidable in any grammar with nontrivial coverage, but this one is particularly unfortunate since it affects a very large and constantly growing class of words.

Another problem concerns the combinations of a possessive determiner with a deverbal noun, as in (6).

(6)   This hotel was her first choice.

Since the possessive denotes the one who makes the choice, it would seem logical to treat it as the subject of the noun, but instead it is the subject of the copula, *this hotel*, which has to be treated as such. To model combinations of this kind Stowell (1983) has proposed to allow deverbal predicative nouns to have two subjects, but this is not compatible with the HPSG claim that the SUBJ list contains at most one *synsem* object.

A third problem for (5) is that it only applies to common nouns. Proper nouns and pronouns are not subsumed, since they have an empty ARG-ST list. Yet they do occur in predicate position.

(7)   a.   The winner is Daniel D. Lewis.
       b.   That must be her.

That there is a problem with predicative proper nouns is also acknowledged in (Pollard and Sag, 1987, 66). The usual escape route involves the postulation of an ambiguity between the predicating use of the copula, as in *be a nice guy*, and the identifying use, as in (7). What this distinction is about and how the ambiguity can be resolved in context is usually left implicit, but even if these questions can be answered satisfactorily, the problem remains, since the predicating *be* is compatible with pronominal complements, as in (8a–8b), and with proper nouns, as in (8c–8d).

(8)   a.   He is somebody with an impeccable track record.
       b.   No matter what happens, she will remain herself.
       c.   This is typically Microsoft.
       d.   This movie is SO Woody Allen.

Besides, the introduction of an ambiguity between predicating and identifying *be* raises thorny issues for the treatment of noun phrases with a definite determiner. The one in (9a), for instance, could be argued to be a complement of the identifying *be*, given the synonymy with (7a), but for those in (9b–9c) it seems more appropriate to treat them as complements of the predicating *be*.

(9)   a.   Daniel D. Lewis is the winner.
       b.   These plumbers are the best we could find in this town.

c. Every lawyer is the star of his own universe.

In sum, the postulation of an ambiguity between identifying and predicating *be* raises more problems than it solves.

2.2 Gerundial, infinitival and clausal predicates

The raising treatment is also problematic for gerundial and infinitival predicates. The unexpressed subjects of the gerund in (10a) and of the infinitive in (10b), for instance, are not identical to the subject of the copula: it is not the pleasure that is eating and drinking and it is not the worry that gets rid of detractors.[2]

(10) a. The greatest pleasure on earth is eating oysters and drinking champagne.
      b. His main worry now is to get rid of his detractors.

Instead, the unexpressed subjects have arbitrary reference. Another problem with these complements concerns their PRED feature. Gerunds are explicitly claimed to be [– PRED] in (Ginzburg and Sag, 2000, 203) and infinitives are [– PRED] too, but (10a) and (10b) are undeniably grammatical.

Also the *that* clause in (11) does not fit the raising mould for the obvious reason that it already contains a subject.

(11) The main problem is that we have run out of cash.

Complements like those in (10) and (11) are also discussed in Dalrymple et al. (2004) where they are called 'closed' predicates.

2.3 Lack of generality

The proposal to treat the predicate complements in the same way as the complements of raising verbs (modulo the PRED value) is inspired by the fact that there are verbs, such as *seem* and *consider*, which are used in both ways, as shown in (1) and (2). The force of this argument is weakened, though, by the fact that there are also many predicate selectors that cannot be used as raising verbs. *Stay* and *become*, for instance, combine with predicate complements, but not with nonfinite VPs.

(12) a. Let us stay (*to be) calm.
      b. He became (*to be) a lawyer.

In other words, the fact that some of the predicate selecting verbs are also used as raising verbs does not by itself demonstrate that predicate selecting verbs ARE raising verbs. For comparison, the fact that some of the predicate selecting verbs

---

[2]This is explicitly confirmed in (Pollard and Sag, 1994, 115): "We assume, following Postal (2974,1977) that the subjects of gerunds never undergo raising."

are also used as ordinary transitive verbs, as illustrated for *grow* and *make* in (13–14), does not by itself demonstrate that they ARE ordinary transitive verbs.

(13) a. He grew angry.

   b. He grows potatoes.

(14) a. He made us crazy.

   b. He made us a cake.

2.4 The status of the PRED feature

The presence of the boolean PRED feature in the HEAD value suggests that it has the same status as the other information that is canonically found in HEAD values, such as part of speech, verb form and case. A closer look, though, reveals more differences than similarities. Part of speech, for one, concerns an inherent property of a lexical item which it furthermore shares with its phrasal projection. *Teacher*, for instance, is a noun and *a teacher* is a noun phrase, no matter whether it is used as a subject, a complement or a conjunct. By the same token, *has* is a finite verb and *she has a dog* is a finite clause, no matter whether it is used as a main clause, a complement clause or a conjunct. Being predicative, by contrast, does not concern an inherent property of a lexical item. *Teacher* is predicative in *be a teacher* and nonpredicative in *tease a teacher*. There is, hence, nothing in the word *teacher* itself, which makes it predicative or nonpredicative. This suggests that the [+/– PRED] dichotomy is a functional (or a relational) distinction rather than a categorial one, and that it, therefore, belongs in the features which model selection (ARG-ST and the valence features) rather than in the HEAD feature.[3]

Something along these lines is in fact also pondered in (Pollard and Sag, 1987, 66): "why posit a new feature PRED to distinguish those words and phrases that can appear in postcopular position? Why not simply say that the copula just happens to subcategorize for a disjunction of catgeories, including NPs, PPs, certain APs, passive-participial VPs and present-participial VPs?" The answer given there is that precisely the same collection of categories has the ability to occur in a number of other syntactic functions or positions, such as the controlled adjunct function in (15) and the most oblique complement position of the *there* constructions in (16).

(15) a. Kim came back from Texas a Republican.

   b. Kim came back from Texas in a boxcar.

   c. Kim came back from Texas ready for anything.

   d. Kim came back from Texas driving a Bentley.

   e. Kim came back from Texas pursued by lawmen from four states.

---

[3]For comparison, to distinguish subjects from other dependents, one does not include a boolean SUBJ feature in the HEAD values all nouns, which is then resolved to positive in case they are used as a subject, and to negative otherwise. Instead, the distinction is canonically made in terms of valence features.

(16)   a.   There is a donkey in the garden.

          b.   There is a donkey asleep in the garden.

          c.   There is a donkey sleeping in the garden.

          d.   There is a donkey stuffed with kapok.

This argument is weakened, though, by the fact that many of the words and phrases which can be used in predicate position cannot be used in these positions. Controlled adjuncts, for instance, cannot take the form of a definite or pronominal NP, nor of an adjective which denotes an individual level predicate, as shown in (17).

(17)   a.   * Kim came back from Texas my neighbor.

          b.   * Kim came back from Texas the man we need.

          c.   * Kim came back from Texas somebody with good taste.
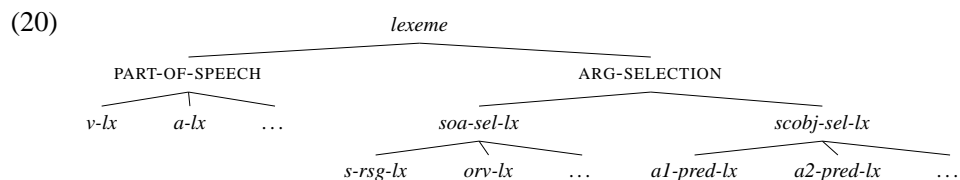
          d.   * Kim came back from Texas male.

Conversely, there are words and phrases which can be used as controlled adjuncts but not as predicate complements.

(18)   a.   Kim came back from Texas owning a ranch and an oil well.

          b.   Kim came back from Texas in order to participate in the election.

          c.   Kim came back from Texas after having bought a home there.

(19)   a.   * Kim is owning a ranch and an oil well.

          b.   * Kim is in order to participate in the election.

          c.   * Kim is after having bought a home there.

Moreover, the most oblique complement of the *there* construction cannot be nominal, no matter whether it is definite or indefinite.

3   Analysis

I will now propose a treatment of the predicate complements which avoids the problems which were presented in Section 2. The starting point is the assumption that raising verbs and predicate selecting verbs belong to mutually distinct classes of lexemes. In terms of the bidimensional hierarchy of lexemes, proposed in (Ginzburg and Sag, 2000, 20), they belong to separate subtypes of the ARG-SELECTION partition.

(20)

                                  lexeme

       PART-OF-SPEECH   ARG-SELECTION

   v-lx   *a-lx*   …   *soa-sel-lx*   *scobj-sel-lx*

                      s-rsg-lx   *orv-lx*   …   *a1-pred-lx*   *a2-pred-lx*   …

While the raising verbs belong to the class of lexemes which select a state of affairs as their most oblique argument, the predicate selecting verbs belong to the class of lexemes which select a scope-object as their most oblique argument.

(21) a. *soa-sel-lx* $\implies$ $\left[\text{ARG-ST} \quad nelist \;\oplus\; \left\langle \left[\text{CONTENT} \quad soa \right] \right\rangle \right]$

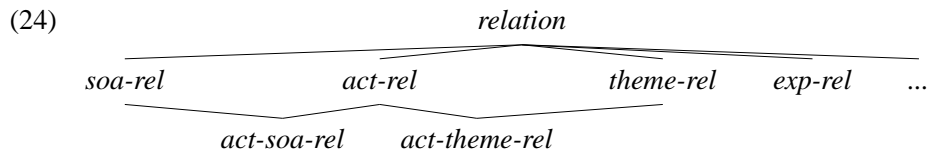 b. *scobj-sel-lx* $\implies$ $\left[\text{ARG-ST} \quad nelist \;\oplus\; \left\langle \left[\text{CONTENT} \quad scope\text{-}obj \right] \right\rangle \right]$

To spell out what these terms mean I start from the hierarchy of CONTENT values in (Ginzburg and Sag, 2000, 386).

(22)
$$sem\text{-}obj$$
$$soa \qquad relation \qquad scope\text{-}obj \qquad index \qquad message$$
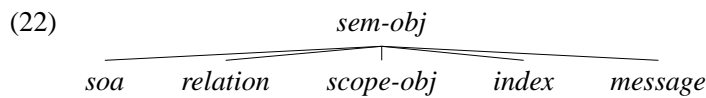
Objects of type *s(tate-)o(f-)a(ffairs)* consist of a nucleus and a list of quantifiers.

(23) $\begin{bmatrix} soa \\ \text{NUCLEUS} \quad relation \\ \text{QUANTS} \quad list\left(quant\text{-}rel\right) \end{bmatrix}$

The NUCLEUS value is an object of type *relation* and contains the semantic roles which a lexeme assigns. The inventory of those roles can be very large, as in Pollard and Sag (1994) and Ginzburg and Sag (2000), which employ such roles as VISITOR and VISITED, but there are also proposals which employ a rather limited inventory, such as those of Wechsler (1995), Davis (2001) and Müller (2002). It is the latter approach that I will adopt. To spell out which roles are relevant for which classes of lexemes I use the following hierarchy of relations.

(24)
$$relation$$
$$soa\text{-}rel \qquad act\text{-}rel \qquad theme\text{-}rel \qquad exp\text{-}rel \qquad ...$$
$$act\text{-}soa\text{-}rel \qquad act\text{-}theme\text{-}rel$$

Act-rel and *soa-rel* are taken from (Davis, 2001, 92). They stand, respectively, for relations which involve an actor and relations which involve a state of affairs.

(25) $\begin{bmatrix} act\text{-}rel \\ \text{ACTOR} \quad index \end{bmatrix}$ $\begin{bmatrix} soa\text{-}rel \\ \text{SOA-ARG} \quad soa \end{bmatrix}$

Theme-rel replaces Davis' *undergoer-rel*. The reason for this change is that the notion of 'undergoer' is defined in such a way that only action verbs can assign this role, whereas the verbs of interest in this paper include many stative verbs. I also add *exp-rel* for relations which involve an experiencer, as in Müller (2002).

(26)
$$
\begin{bmatrix} \textit{theme-rel} \\ \text{THEME} \quad \textit{index} \end{bmatrix}
\quad
\begin{bmatrix} \textit{exp-rel} \\ \text{EXPERIENCER} \quad \textit{index} \end{bmatrix}
$$

Lexemes which assign more than one role are subsumed by types which inherit from the respective simple types, such as *act-soa-rel* and *act-theme-rel*.

The CONTENT values of type *scope-object* consist of an index and set of restrictions, as illustrated by the one of *sailor* in (27).

(27)
$$
\begin{bmatrix}
\textit{scope-obj} \\
\text{INDEX} \quad \boxed{1}\ \textit{index} \\
\text{RESTR} \left\{
\begin{bmatrix}
\textit{fact} \\
\text{PROP}
\begin{bmatrix}
\textit{proposition} \\
\text{SIT} \quad s \\
\text{SOA}
\begin{bmatrix}
\textit{soa} \\
\text{QUANTS} \quad \langle\ \rangle \\
\text{NUCLEUS}
\begin{bmatrix}
\textit{sailor-rel} \\
\text{INSTANCE} \quad \boxed{1}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\right\}
\end{bmatrix}
$$

The index stands for an entity in the universe of discourse, and the RESTR(ICTION) value adds constraints on its denotation. In this case it consists of the fact that the entity is a sailor. The indices are comparable to predicate logic variables and the scope-objects to formulae like 'x | *sailor*(x)'.

As pointed out in (Ginzburg and Sag, 2000, 136), the restrictions are usually abbreviated as *sailor*($\boxed{1}$), but it is crucial for a proper understanding of the notion *scope-object* to realize that its restriction is a set of facts.[4] Since facts invariably contain the NUCLEUS attribute and since the latter's value is of type *relation*, it follows that also nouns denote a relation, even if they do not take any arguments to assign a semantic role to. In that case they have one attribute, INSTANCE, whose value is identified with their index, as in (27). To model this I assume that all objects of type *relation* have this attribute.

(28)
$$
\begin{bmatrix} \textit{relation} \\ \text{INSTANCE} \quad \textit{index} \end{bmatrix}
$$

For the verbs, this implies that the relations which they denote contain both the names of the semantic roles which they assign and the INSTANCE attribute. The NUCLEUS value of the verb *bite*, for instance, looks as follows:

---

[4]In the type hierarchy facts and propositions are subtypes of *message* Ginzburg and Sag (2000).

(29)
$$\begin{bmatrix} \textit{act-theme-rel} \\ \text{INSTANCE} \quad \textit{index} \\ \text{ACTOR} \quad \textit{index} \\ \text{THEME} \quad \textit{index} \end{bmatrix}$$

The values of the ACTOR role and the THEME role are supplied by respectively the first and the second argument of the verb and the value of the INSTANCE attribute stands for the situation which the verb denotes. It is, hence, comparable to a Davidsonian event variable.

Having introduced the basic notions, I now focus on the raising lexemes (Section 3.1) and the predicate selectors (Section 3.2).

3.1 Raising lexemes

The English raising lexemes are either subject-to-subject or subject-to-object raisers. The specific properties of the former are spelled out in the following constraint (Ginzburg and Sag, 2000, 22).[5]

(30) $\textit{s-rsg-lx} \quad \Rightarrow \quad \begin{bmatrix} \text{ARG-ST} \quad \left\langle \begin{bmatrix} \text{LOCAL} \quad \boxed{1} \end{bmatrix}, \begin{bmatrix} \text{SUBJ} \quad \left\langle \begin{bmatrix} \text{LOCAL} \quad \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \right\rangle \end{bmatrix}$

This type subsumes among others the use of *seem* in (1a). To spell out the CONTENT value I make a finer-grained distinction between two subtypes. The first subsumes the lexemes which denote a relation of type *soa-rel* and which, hence, assign the SOA-ARG role to their complement.[6]

(31) $\textit{s-s-rsg-lx} \quad \Rightarrow \quad \begin{bmatrix} \text{ARG-ST} \quad \textit{nelist} \quad \oplus \quad \left\langle \begin{bmatrix} \text{CONTENT} \quad \boxed{1} \, \textit{soa} \end{bmatrix} \right\rangle \\ \\ \text{SS} \,|\, \text{LOCAL} \,|\, \text{CONTENT} \begin{bmatrix} \textit{soa} \\ \\ \text{NUCLEUS} \begin{bmatrix} \textit{soa-rel} \\ \text{SOA-ARG} \quad \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$

This type subsumes the modal and the aspectual subject raisers, including the modal *be* in (32).

(32) a. The best is yet to come.
      b. They are to be informed on a daily basis.

---

[5]Nothing in my treatment depends on whether the shared information is limited to the LOCAL values, as in (Ginzburg and Sag, 2000, 21), or extended to the SYNSEM value as a whole, as in older versions of HPSG.

[6]Notice that *soa* is not the same as *soa-rel*. While the former stands for a state of affairs, the latter stands for a relation in which one of the arguments is a state of affairs.

The second subtype subsumes the lexemes which are semantically vacuous, such as the auxiliary *do* and the progressive and the passive *be*. They simply inherit the CONTENT value of their complement.

(33)   *v-s-rsg-lx* $\Rightarrow$ $\begin{bmatrix} \text{ARG-ST} & nelist & \oplus & \left\langle \begin{bmatrix} \text{CONTENT} & \boxed{1}\,soa \end{bmatrix} \right\rangle \\ \text{SS} \mid \text{LOCAL} \mid \text{CONTENT} & \boxed{1} \end{bmatrix}$

The subject-to-object-raisers belong to the type *orv-lx* and have the properties which are spelled out in the corresponding constraint (Ginzburg and Sag, 2000, 22). Since they all denote a state of affairs, I include the CONTENT value in the definition of the type.

(34)   *orv-lx* $\Rightarrow$ $\begin{bmatrix} \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}} , \begin{bmatrix} \text{LOC} & \boxed{3} \end{bmatrix}, \begin{bmatrix} \text{SUBJ} & \left\langle \begin{bmatrix} \text{LOC} & \boxed{3} \end{bmatrix} \right\rangle \\ \text{CONT} & \boxed{2}\,soa \end{bmatrix} \right\rangle \\ \text{SS} \mid \text{LOCAL} \mid \text{CONT} \mid \text{NUCLEUS} & \begin{bmatrix} exp\text{-}soa\text{-}rel \\ \text{EXPERIENCER} & \boxed{1} \\ \text{SOA-ARG} & \boxed{2} \end{bmatrix} \end{bmatrix}$

This subsumes among others the use of *considers* in (2a). The first argument supplies the EXPERIENCER role.

Since the ARG-SELECTION partition and the PART-OF-SPEECH partition are mutually independent, it is possible to define more specific types by multiple inheritance. The non-vacuous subject raising verbs, for instance, belong to a type that inherits from *v-lx*, on the one hand, and *s-s-rsg-lx* and its supertypes, on the other hand. In the same way, the subject raising adjectives, such as *likely*, belong to a type that inherits from *a-lx* and *s-s-rsg-lx*. The vacuous subject raisers and the subject-to-object raisers are all verbs, at least in English.

3.2   Predicate selectors

I assume that the predicate selectors belong to the lexemes which select a *scope-object* as their most oblique argument. This implies that the complement of the copula in *I am a sailor* is of the same semantic type as the complememt of the transitive verb in *I met a sailor*. This is not undesirable, since there is nothing in the noun phrase itself which singles it out as either a predicate or a direct object, see Section 2.4. The difference in interpretation is entirely due to the selecting verb, and it is hence in the CONTENT value of the verb that the distinction between predicate and object complements has to be captured.

For verbs which select an object complement, such as *met*, I adopt the canonical treatment, as in:

(35)
$$\begin{bmatrix} \text{ARG-ST} & \left\langle \text{X}_{\boxed{1}}, \text{Y}_{\boxed{2}} \right\rangle \\ \\ \text{SS} \mid \text{LOCAL} \mid \text{CONTENT} \mid \text{NUCLEUS} & \begin{bmatrix} \textit{act-theme-rel} \\ \text{INSTANCE} & \textit{index} \\ \text{ACTOR} & \boxed{1} \ \textit{index} \\ \text{THEME} & \boxed{2} \ \textit{index} \end{bmatrix} \end{bmatrix}$$

For verbs which select a predicate complement the representation is more complex, since there is not only the relation between the verb and its arguments which has to be captured, but also the relation between the predicate and its target. To accommodate this I make a distinction between the selectors of a predicate whose target is the first argument, such as the use of *seems* in (1b), and the selectors of a predicate whose target is the second argument, such as the use of *considers* in (2b).

The former denote a relation of type *exp-soa-rel*. The value of the EXPERIENCER attribute is the index of an optional PP complement, such as the one in *he seems an excellent choice to me*, and the value of SOA-ARG is a state of affairs. The nucleus of this state of affairs is the requirement that the indices of the predicate and its target (the subject) be coreferential.

(36) *a1-pred-lx* $\Rightarrow$
$$\begin{bmatrix} \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}} \left(, \text{PP}_{\boxed{2}}\right), \text{Z}_{\boxed{3}} \right\rangle \\ \\ \text{SS} \mid \text{LOC} \mid \text{CONT} \mid \text{NUCL} & \begin{bmatrix} \textit{exp-soa-rel} \\ \text{EXPERIENCER} & \boxed{2} \ \textit{index} \\ \\ \text{SOA-ARG} \mid \text{NUCL} & \begin{bmatrix} \textit{coref-rel} \\ \text{INST} & \boxed{3} \ \textit{index} \\ \text{THEME} & \boxed{1} \ \textit{index} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Notice that the constraint requires coreferentiality of the indices, rather than identity. Identity would be too restrictive, since it implies agreement for person, number and gender, whereas the predicate does not need to share the person, number and gender values of the target, as illustrated by (37).[7]

(37) a. I am a linguist.
      b. We are a team.

This style of analysis in which the predicate and its target have different indices which are required to be coreferential is similar to the one that Richard Montague proposed in Montague (1970): "the 'is' of such formulas as 'v_0 is a horse' may be identified with the 'is' of identity, and the indefinite singular term 'a horse' treated, as usual, existentially." (Allegranza, 2006, 78), from which this quote is taken, comments: "the example 'Mary is a woman' will allow the FOL-style translation (38a), which is logically equivalent to (38b).

---

[7]The proposal to include features for person, number and gender in the indices is defended at length in chapter 2 of Pollard and Sag (1994) and has been taken on board in most of the HPSG literature.

(38) a. *for at least one x*: woman(x) *and* Mary $= x$
  b. woman(Mary)

Given the simplicity of (38b), it is not surprising that this was often adopted as direct logical translation of the English sentence, skipping the step (38a). But theoretically, the more elaborate analysis proves rewarding insofar as it contributes to a uniform semantic treatment of NPs and their determiners, which can be worked out in a compositional fashion."

The selectors of predicates whose target is identified with the second argument are treated along the same lines.

$$
(39) \quad \textit{a2-pred-lx} \quad \Rightarrow \quad
\begin{bmatrix}
\text{ARG-ST} \quad \langle \text{NP}, \text{NP}_{\boxed{2}}, \text{Z}_{\boxed{3}} \rangle \\[2ex]
\text{SS} \mid \text{LOC} \mid \text{CONT} \mid \text{NUCL}
\begin{bmatrix}
\textit{soa-rel} \\[1ex]
\text{SOA-ARG} \mid \text{NUCL}
\begin{bmatrix}
\textit{coref-rel} \\
\text{INST} \quad \boxed{3} \; \textit{index} \\
\text{THEME} \quad \boxed{2} \; \textit{index}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$

This class of lexemes can further be partitioned depending on whether their first argument is assigned the ACTOR role, as in *he drives me crazy*, or the EXPERIENCER role, as in *she considers him an idiot*.

In contrast to the raising lexemes, which contain a few semantically vacuous verbs, the predicate selectors all denote a relation of type *soa-rel*, also the copula. Since this deviates from the canonical GPSG/HPSG treatment, in which the copula is treated as semantically vacuous, I add some arguments.

First, the copula belongs to a paradigm which also includes verbs like *remain, become* and *seem*, and its meaning contrasts with those of the other members. *He was ill*, for instance, has another meaning than *he became ill*, *he remained ill* and *he seemed ill*, and this difference must be due to the fact that the verbs have different meanings, since the subject, the predicate and the tense are all identical. Second, in many languages the copula can be combined with a complement that denotes the experiencer, as in the Dutch *dat is me een raadsel* 'that is a riddle to me' and *het is me om het even* 'it is all the same to me', and the German *es ist mir zu kalt* 'it is too cold for me'. It is not clear what other word could assign this role. Third, the fact that the copula can be omitted in certain contexts in certain languages, such as Russian and Hungarian, does not by itself prove that it is semantically vacuous. Otherwise, one might as well conclude from the absence of the articles in Latin that the articles are semantically vacuous in the languages that have them.

Since the ARG-SELECTION partition is orthogonal to the PART-OF-SPEECH partition, we can define more specific types by multiple inheritance. The selectors of subject-oriented predicates, for instance, belong to a type that inherits from *v-lx* and *a1-pred-lx*. There is also a type which inherits from *p-lx* and *a1-pred-lx*. It subsumes the so-called absolutive *with*, as used in *with John ill we have to find a*

substitute.[8] The selectors of predicates whose target is the second argument must have at least three arguments and are all verbs: They, hence, inherit from *v-lx* and *a2-pred-lx*.

4 Solving the problems

Let us now return to the problems that were mentioned in section 2 and check whether they are avoided or solved by the new treatment.

4.1 Nominal predicates

The three problems with the nominal predicates are no longer there. First, the nouns are not treated as systematically ambiguous, since the distinction between their predicative uses and their other uses is not made in the AVMs of the nouns, but in the CONTENT value of the lexemes which select them. Second, the problem with pronouns and proper nouns disappears, since their CONTENT value is of the required type, i.e. *scope-object*. Third, there are no complications with the interpretation of the notion 'subject' in deverbal nouns, since predicative nouns are not required to select a subject.

4.2 Gerundial, infinitival and clausal predicates

Solving the problems with the closed predicate complements is less obvious. I'll discuss the three types one by one, starting with the gerunds.

Gerunds typically occur in positions which are canonically taken by NPs, such as the subject position, the direct object position, the prepositional object position and the predicate position.

(40)   a. Swimming in the open sea is dangerous.

      b. He recommended taking a hot bath.

      c. She is not afraid of walking in the dark.

      d. The greatest pleasure on earth is eating oysters and drinking champagne.

To capture this commonality I assume that they have the same type of CONTENT values as the nouns. More specifically, I assume that the lexical rule which derives gerunds from verbal lexemes changes the latter's CONTENT value from *soa* to *scope-object*.[9]

---

[8]To model the fact that the first argument must be realized as a complement I assume that the preposition is assigned the empty SUBJ list in the lexicon.

[9]The addition of the suffix is entirely regular from a phonological point of view, but orthographically, adjustments are needed in certain cases, such as consonant gemination (*swimming*), *-e* deletion (*having*) and *-ie/-y* substitution (*dying*). It is immaterial at this point whether the gerunds are treated as nouns, as verbs or as members of a separate part of speech. All that matters at this point is their semantic type.

(41)
$$
\begin{bmatrix}
lexeme \\
\text{PHON} \quad \boxed{A} \\
\text{SS} \,|\, \text{LOC} \begin{bmatrix} \text{CAT} \,|\, \text{HEAD} \quad verb \\ \text{CONTENT} \quad \boxed{1}\, soa \end{bmatrix}
\end{bmatrix}
\implies_{LR}
$$

$$
\begin{bmatrix}
\text{PHON} \quad \boxed{A} \ \oplus\ \left\langle \text{ing} \right\rangle \\[2ex]
\text{SS} \,|\, \text{LOC} \,|\, \text{CONT}
\begin{bmatrix}
scope\text{-}obj \\
\text{INDEX} \quad \boxed{2}
\begin{bmatrix} index \\ \text{PERSON} \quad 3 \\ \text{NUMBER} \quad sing \end{bmatrix} \\
\text{RESTR} \quad \left\{ \begin{bmatrix} fact \\ \text{PROP} \,|\, \text{SOA} \quad \boxed{1} \begin{bmatrix} \text{NUCL} \,|\, \text{INST} \quad \boxed{2} \end{bmatrix} \end{bmatrix} \right\}
\end{bmatrix}
\end{bmatrix}
$$

The CONTENT value of the verbal lexeme, which is of type *soa*, is integrated in the RESTR attribute of the gerund. This is straightforward since the facts which make up the RESTR value invariably contain an *soa* object. Besides, the gerund is assigned an index. It stands for instances of whatever the verb denotes, such as swimming, walking or eating. Just like in the case of the nouns, the index recurs in the object of type *relation*, as the value of the INSTANCE attribute. Assuming that the indices of gerunds are third person and singular we account for the fact that they only combine with third person singular verbs when they are used in subject position, as in (40a).

Having a CONTENT value of type *scope-object*, the gerund can be used as the complement of a predicate selector, but not as the complement of a raising verb.

Infinitives have CONTENT values of type *soa* and are, hence, compatible with raising lexemes but not with predicate selectors. However, in the same way as the *ing* forms are differentiated between a gerundial and a participial use, we can distinguish between two uses of the infinitives. In one use they denote a state of affairs and occur as complements of raising lexemes, but in another use they denote a scope object and occur as predicate complements or as subjects, as in (42).

(42) a. His main worry now is to get rid of his detractors.
      b. To eat smoked salmon at this time of the year is dangerous.

Since the presence of *to* is obligatory in these positions, we can treat it as the word that triggers the change from the typically verbal CONTENT value (*soa*) to the typically nominal value (*scope-object*). Its properties are captured in (43).

(43)
$$\begin{bmatrix} \text{PHON} & \langle \text{to} \rangle \\[2pt] \text{ARG-ST} & \left\langle \begin{bmatrix} \text{VFORM} & \textit{bare-inf} \\ \text{SUBJ} & \langle X_{ref} \rangle \\ \text{CONTENT} & \boxed{1}\ \textit{soa} \end{bmatrix} \right\rangle \\[2pt] \text{SS | LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{prep} \\ \text{SUBJ} & \langle\ \rangle \end{bmatrix} \\[2pt] \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2}\ \begin{bmatrix} \textit{index} \\ \text{PERSON} & \textit{3} \\ \text{NUMBER} & \textit{sing} \end{bmatrix} \\[2pt] \text{RESTR} & \left\{ \begin{bmatrix} \text{PROP | SOA} & \boxed{1} \end{bmatrix} \begin{bmatrix} \text{NUCL | INST} & \boxed{2} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

In words, *to* selects a bare infinitive which denotes a state of affairs and integrates the latter in its RESTRICTION value. Its third person singular index accounts for the agreement with the finite verb in (42b).

The unexpressed subject of the selected infinitive must be referential (Ginzburg and Sag, 2000, 50–57), but its reference is arbitrary. Crucially, it is not identified with the subject of the verb that selects the *to* infinitive. The SUBJ list of *to* itself is declared empty. Given the Argument Realization Principle this implies that the selected bare infinitive is realized as a complement.

The treatment of this use of *to* as a preposition may at first look outlandish, given the wide-spread acceptance in GPSG/HPSG of Geoff Pullum's proposal to treat it as a nonfinite auxiliary verb Pullum (1982). Nonetheless, there are several arguments which support it. First, *to* is a preposition in all of its other uses, including the combination with gerunds, as in *object to being treated badly*. Second, it is morphologically invariant, just like the other English prepositions (but unlike the verbs). Third, it shows some of the typical properties of the argument marking prepositions, such as an empty SUBJ list and a CONTENT value of type *scope-object*. Fourth, the positions in which these *to*-infinitives occur (predicate and subject) can also be taken by PPs.

(44) a. They are not in the mood.

  b. After dinner is a good time for a walk.

The arguments which Pullum (1982) adduces against the prepositional status are numerous (ten), but their relevance for the analysis of this particular use of *to* is limited, since they all pertain to its other use, i.e. the one in which it yields a VP which denotes a state of affairs.[10]

---

[10]Also for that use, it is not necessary to assume that *to* is an auxiliary verb. If one treats the bare infinitival VP as the head and *to* as its functor, the combination is an SOA denoting VP, no matter what the part of speech of *to* is, see Van Eynde (2004).
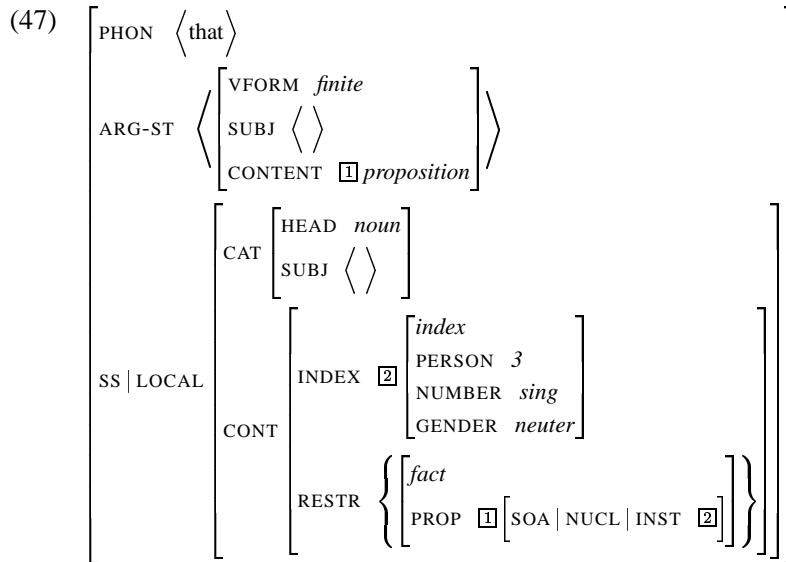
Clauses have CONTENT values of type *message* and, hence, do not match the requirements which the predicate selectors impose on their most oblique argument. This accounts for the ill-formedness of (45).

(45)   * The main problem is we have run out of cash.

The sentence becomes well-formed, though, if the clausal predicate is introduced by the complementizer *that*. The same holds for clauses in subject position.

(46)   *(That) she did that all on her own still puzzles me.

This suggests that the addition of the complementizer can be treated as the factor that triggers a change of the CONTENT value. More specifically, the complementizer takes a finite clause which denotes a proposition as its argument and integrates the proposition in its RESTR value.

(47)

$$
\begin{bmatrix}
\text{PHON} \quad \langle \text{that} \rangle \\[2ex]
\text{ARG-ST} \quad \left\langle
\begin{bmatrix}
\text{VFORM } \textit{finite} \\
\text{SUBJ } \langle \rangle \\
\text{CONTENT } \boxed{1}\ \textit{proposition}
\end{bmatrix}
\right\rangle \\[6ex]
\text{SS} \mid \text{LOCAL}
\begin{bmatrix}
\text{CAT}
\begin{bmatrix}
\text{HEAD } \textit{noun} \\
\text{SUBJ } \langle \rangle
\end{bmatrix} \\[4ex]
\text{CONT}
\begin{bmatrix}
\text{INDEX } \boxed{2}
\begin{bmatrix}
\textit{index} \\
\text{PERSON } \textit{3} \\
\text{NUMBER } \textit{sing} \\
\text{GENDER } \textit{neuter}
\end{bmatrix} \\[6ex]
\text{RESTR } \left\{
\begin{bmatrix}
\textit{fact} \\
\text{PROP } \boxed{1}\left[\text{SOA} \mid \text{NUCL} \mid \text{INST } \boxed{2}\right]
\end{bmatrix}
\right\}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$

The constraint that the index is third person singular accounts for the agreement with the finite verb in (46), and the assignment of an empty SUB list to the complementizer accounts for the fact that its argument must be realized as a complement.

The assumption that the complementizer is a pronoun flies in the face of a long-standing practice to treat the complementizers as members of a separate part of speech, but it is not unmotivated. First, it is homophonous to a demonstrative pronoun, not only in English, but also in Dutch (*dat*) and German (*dass*).[11] Second, the index of the complementizer has the same agreement values as the homophonous demonstrative pronoun. Third, the complementizer shares the typically pronominal properties of having an empty SUBJ list and a CONTENT value of

---

[11]Also in the Romance languages the corresponding complementizers are homophonous to pronouns: the French *que* and the Italian *che*, for instance, are also used as interrogative pronouns.

type *scope-obj*. Fourth, it entails that the *that* clauses in subject and predicate position are NPs, so that it is not necessary to express generalizations about subjects in disjunctive terms (CP or NP).

The possible objection that pronouns are not the kind of words that normally take a complement is not so strong, since it is not uncommon for a pronoun to take a postnominal dependent, as illustrated by the AP in *something soft and slimy*, the NP in *who the hell*, the PPs in *those of you*, *everybody in this room* and *what on earth* and the clause in *those who are about to leave*.

Just like *to*, the complementizer has another use in which it yields a clause which denotes a proposition, rather than an NP which denotes a scope-object. That use is not relevant, though, for the topic of this paper.

4.3 Generality

The raising treatment of the predicate complements has been argued to lack generality, in that the co-existence of raising and predicate selecting uses is limited to a small subset of the predicate selectors. Whether the alternative treatment fares better in terms of generality depends on whether it is applicable to all constructions which are canonically thought of as predicative. To check this let us first restate the criterial properties of the predicate complements in the present treatment: first, they denote a scope-object and, second, their index is co-referential with the index of the target of the predicate. So far, we have seen that these properties subsume the nominal, gerundial, infinitival and clausal predicates. What is left then are the adjectival and prepositional predicates.

For the adjectival predicates it is easy to demonstrate that they are subsumed by the criterial properties, since the adjectives are canonically assigned the same type of CONTENT values as the nouns. The one of *clever*, for instance, is of type *scope-object* and is, hence, comparable to the formula 'x | *clever*(x)'. Semantically speaking, its contribution is of the same kind as that of a common noun.

Prepositional predicates fit the mould as well if the predicative PPs have a CONTENT value of type *scope-object*. To ensure this their head daughter must have the properties spelled out in (48).

(48)
$$
\begin{bmatrix}
\text{ARG-ST} & \left\langle \text{NP}_{\boxed{2}} \right\rangle \\[2ex]
\text{SS} \mid \text{LOCAL} &
\begin{bmatrix}
\text{CAT} &
\begin{bmatrix}
\text{HEAD} & \textit{prep} \\
\text{SUBJ} & \left\langle \ \right\rangle
\end{bmatrix} \\[3ex]
\text{CONT} &
\begin{bmatrix}
\text{INDEX} & \boxed{1}\ \textit{index} \\[2ex]
\text{RESTR} & \left\{
\begin{bmatrix}
\text{PROP} \mid \text{SOA} \mid \text{NUCL} &
\begin{bmatrix}
\textit{prep-rel} \\
\text{INST} & \boxed{1} \\
\text{THEME} & \boxed{2}
\end{bmatrix}
\end{bmatrix}
\right\}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$

In words, the preposition takes an NP argument which must be realized as a complement since its SUBJ list is declared empty, and it denotes the set of entities which

stand in the relevant relation to the index of the complement. The CONTENT value of *in Paris*, for instance, is comparable to the formula 'x | *in*(x, paris)'. When combined with a predicate selector, index $\boxed{1}$ is required to be co-referential with the index of the target of the predicate, just like in the case of the nominal and adjectival predicates.

Independent evidence for this treatment is provided by the fact that the argument marking prepositions are standardly assigned an empty SUBJ list and a CONTENT value of type *scope-object*. Since predicates are arguments, it is no surpise that they have these properties as well. Notice also that they behave in the same way with respect to binding.

(49) a. They$_i$ are talking to themselves$_i$/*them$_i$.

     b. They$_i$ do not agree with one another$_i$/*them$_i$.

(50) a. They$_i$ stay among themselves$_i$/*them$_i$.

     b. They$_i$ are after one another$_i$/*them$_i$.

In sum, there is no need for any stipulations to make sure that the predicative APs and PPs are assigned a CONTENT value of type *scope-object*, since that is the value which they are canonically assigned anyway.

4.4 The omission of the PRED feature

The problems with the status of the PRED feature are no longer an issue, since the analysis of Section 3 does not employ this feature. What is an issue, though, is the question whether the functions of the PRED feature are also fulfilled in the new treatment.

First, its role of distinguishing predicate complements from other types of complements is taken over by the CONTENT value of the selecting verbs. Whether *a sailor*, *clever* or *in the box* is a predicate complement or not is made explicit in the semantic properties of the lexemes that select them.

Second, the role of singling out the words that must or cannot be used as predicates by the assignment of a positive, c.q. negative, value for PRED in their lexical entry can be taken over by the independently motivated SELECT feature, as used in Van Eynde (2004). It is part of the HEAD value of all signs and models the requirements which an adjunct, specifier or marker imposes on its head sister. Its value is either an object of type *synsem* which then has to unify with the SYNSEM value of the selected head sister or *none*.[12]

(51) $\begin{bmatrix} head \\ \text{SELECT} \quad synsem \quad \vee \quad none \end{bmatrix}$

---

[12] In terms of the type hierachy of Pollard and Sag (1994) SELECT replaces the features MOD and SPEC.

The value *none* is assigned when the nonhead daughter does not select its head sister. Subjects and complements, for instance, do not select their head sister. Now, since predicates are complements, we can single out the words that must be used as (heads of) predicates by assigning them a SELECT value of type *none*. Likewise, we can single out the words that cannot be used as (heads of) predicates by assigning them a SELECT value of type *synsem*. The former include *ready, ablaze* and *asleep*; the latter include *wooden, mere, utter* and *former*.

5   Conclusion

In the canonical HPSG treatment the criterial properties of predicate complements are defined in syntactic terms: their HEAD value contains the [+ PRED] feature and their selectors are treated as raising verbs (Section 1). This treatment is infelicitous for nominal predicates and inappropriate for gerundial, infinitival and clausal predicates. Besides, the assumption that predicate selecting uses are a kind of raising uses lacks generality and the PRED feature lacks proper motivation (Section 2).

The alternative treatment defines the criterial properties of predicate complements in semantic terms: they denote a scope-object and their index is co-referential with that of the target of the predicate (Section 3). The resulting treatment solves or avoids the problems that were raised in Section 2, and is sufficiently general to apply to all constructions which are canonically treated as predicative (Section 4).

References

Allegranza, V.(2006), *The Signs of Determination. Constraint-based Modelling across Languages*, Peter Lang.

Dalrymple, M., Dyvik, H. and Holloway King, T.(2004), Copular complements: closed or open?, *in* M. Butt and T. Holloway King (eds), *Proceedings of the LFG04 Conference*, CSLI Publications, Stanford, pp. 188–198.

Davis, A.(2001), *Linking by types in the hierarchical lexicon*, CSLI Publications, Dordrecht.

Gazdar, G., Klein, E., Pullum, G. and Sag, I.(1985), *Generalized Phrase Structure Grammar*, Basil Blackwell, Oxford.

Ginzburg, J. and Sag, I.(2000), *Interrogative Investigations*, CSLI, Stanford.

Montague, R.(1970), English as a formal language, *in* B. Visentini (ed.), *Linguaggi nella socìteà e nella tecnica'*, Edizioni di Comunità, Milano.

Müller, S.(2002), *Complex predicates*, CSLI, Stanford.

Pollard, C. and Sag, I.(1987), *Information-based Syntax and Semantics*, CSLI Publications, Stanford.

Pollard, C. and Sag, I.(1994), *Head-driven Phrase Structure Grammar*, CSLI Publications and University of Chicago Press, Stanford/Chicago.

Pullum, G.(1982), Syncategorematicity and English infinitival *to*, *Glossa* **16**, 181–215.

Stowell, T.(1983), Subjects across categories, *Linguistic Review* **2**, 285–312.

Van Eynde, F.(2004), Minor Adpositions in Dutch, *Journal of Comparative Germanic Linguistics* **7**, 1–58.

Wechsler, S.(1995), *The semantic basis of argument structure*, CSLI Publications, Stanford.