A Functional Typology of Copular "Be": Towards an HPSG Formalisation

Tania Avgustinova

DFKI and Saarland University

Abstract

A functional typology of copular "be" in Russian allows us to systematically relate variants of predication with and without copula. The analysis sketched in this article does not need empty categories; neither does it have to stipulate categories, category changes or constituents that are not morphologically signalled. With regard to HPSG formalization, the presented approach independently motivates the use of features and mechanisms that are already available in this framework.

1.   Introduction

The wide range of morphosyntactic variation in verbless clauses cross-linguistically reveals that they are not a single structural type at all. In the Slavic language family, Russian offers the broadest spectrum of potentially copula-less constructions, comprising not only lexically predicative categories (1a), but also ascriptive (1b) and identificational (1c) predication, as well as locative (1d), existential (1e) and possessive (1f) constructions.

(1)   a. *On   gord   rezul'tatami.*
           he.NOM.SG.M   proud.PRD-ADJ.SG.M   results.INST.PL
           He is proud of the results.
       b. *On   durak | tolstyj | vysokogo rosta.*
           he.NOM.SG.M   fool.NOM.SG.M | fat.NOM.SG.M | high height.GEN
           He is a fool | fat | of a high height (i.e. tall).
       c. *On   –   brat   Maksima.*
           he.NOM.SG.M   brother.NOM.SG.M Maksim.GEN
           He is Maksim's brother.
       d. *Boris   na   sobranii.*
           Boris.NOM   at   meeting.LOC
           Boris is at a meeting.
       e. *Za   uglom   (est') magazin*
           behind   corner.SG.M.INST   (is)   store.NOM.SG.M
           There is a store around the corner.
       f. *U Kati   (est')   samovar.*
           at Katia.GEN   (is)   samovar.NOM.SG.M
           Katia has a samovar.

Distributional and periphrastic tests suggest that these distinctions are plausible cross-linguistically, as they systematically correspond to truth-conditional semantic differences. In all these constructions there will be an overt copular 'be' as soon as the tense and mood information is different from the present-indicative default. What this data demonstrates is that the

possibility of the copula being absent, and therefore of non-verbal syntactic predication, is not limited to one particular semantic type of copula construction, but is widely available as a syntactic strategy.

The mainstream linguistic research would often downplay the challenge by approaching it piecemeal and assuming that the respective constructions were headed in the unmarked case by a phonologically empty category. In this contribution we step back to reconsider fundamental aspects of linguistic classification in order to formulate a comprehensive alternative to such ad hoc analyses. We will show how a slightly different perspective on the way classification is performed leads to a straightforward HPSG formalisation of the desired degree of granularity, and allows us not only capture functional similarities but also predict what distinctions should be possible cross-linguistically.

2.   Proposal

Following the approach in (Avgustinova and Uszkoreit 2003), where different types of constructions containing non-verbal predicates are classified on the basis of the relational ontology of (Avgustinova and Uszkoreit 2000), we present a typology of copula for Russian and show how the corresponding semantics can be encoded in the HPSG framework. As the analysed constructs differ in their syntactic (e.g., case marking of arguments) and semantic properties, these differences can now be made explicit and linked to the proposed classification.

The lowest (most informative) types, i.e. the leaves of the hierarchy in (Figure 1), can be straightforwardly motivated, as they correspond to empirical distinctions. The intermediate types factorise the information common to the subclasses of a class, and constraints associated with the specific sub-types provide the appropriate linguistic generalisations.

At the highest level of abstraction, linguistic objects of type *copula* are partitioned according to their function as *inflectional-cop(ula)*, which occurs with lexically/morphologically predicative categories (e.g., Russian short-form adjectives), or as *assembling-operator*, which puts together two non-verbal and lexically non-predicative categories. Overt forms of 'be' in the former case tend to function as mere inflectional tense-mood markers. Recall that according to (Pollard and Sag 1994, p. 44-45), "… a marker is a word that is 'functional' or 'grammatical' as opposed to substantive, in the sense that its semantic content is purely logical in nature (perhaps even vacuous)". In turn, the copula as assembling operator is further partitioned into *copular-functor* and *copular-predicator*. Such a key distinction would find strong

cross-linguistic motivation within the Slavic language family. As a matter of fact, Bulgarian instances of copular functor would correspond to forms of "to be", while those of copular predicator to forms of "to have"
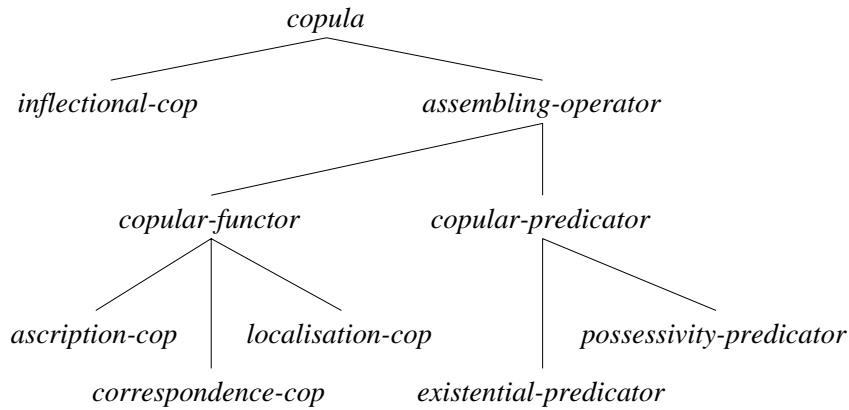

Figure 1: A hierarchy of copula types

More specifically, the copular functor can be of type *ascription-cop(ula)*, *correspondence-cop(ula)* or *localisation-cop(ula)*; and the copular predicator – of type *existentail-predicator* and *possessivity-predicator*. The resulting feature structures are sketched below. Following (Copstake, et al. 1999), the CONTENT value encodes the central predication of a phrase as its KEY, the semantic INDEX of a phrase, and a list of relevant semantic relations RELS.

Semantically, the assembling operator in ascriptive predication (Figure 2) identifies (the INDEX value in) its content with that of the non-verbal (predicative) complement.

Figure 2: Ascriptive predication

The semantic contribution of the assembling operator in identificational predication (Figure 3) is to introduce a key relation of correspondence [5]

(supplying an event variable $\boxed{4}$) whose first argument is identified with the index of the subject $\boxed{3}$ and its second argument with the index of the non-verbal (predicative) complement $\boxed{2}$.

$$
\begin{bmatrix}
correspondence-cop \\[4pt]
\text{CAT | VAL}
\begin{bmatrix}
\text{SUBJ} \left\langle \begin{bmatrix} \text{INDEX } \boxed{3} \end{bmatrix} \right\rangle \\[6pt]
\text{COMPS} \left\langle \begin{bmatrix} non-verbal \\ \text{INDEX } \boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix} \\[20pt]
\text{CONT}
\begin{bmatrix}
\text{INDEX } \boxed{4} \\
\text{KEY } \boxed{5} \\[6pt]
\text{RELS} \left\langle \boxed{5} \begin{bmatrix} correspond-rel \\ \text{EVENT } \boxed{4} \\ \text{ARG1 } \boxed{3} \\ \text{ARG2 } \boxed{2} \end{bmatrix} \right\rangle
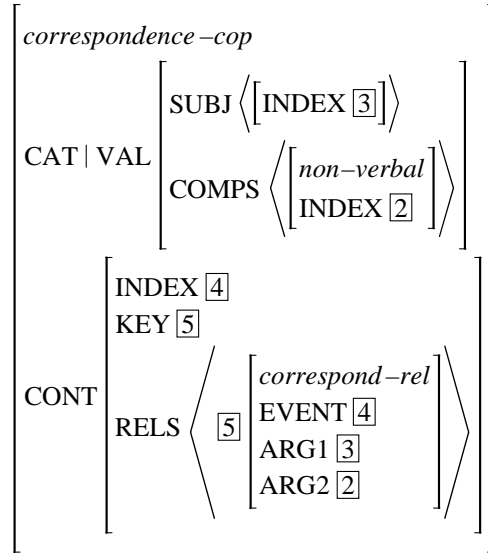\end{bmatrix}
\end{bmatrix}
$$

Figure 3: Identificational predication

The semantic a contribution of the assembling operator in localisational predication (Figure 4) is a key relation of localisation $\boxed{5}$ (supplying an event variable $\boxed{4}$) whose first argument is identified with the index of the subject $\boxed{3}$ and its second argument with the index of the non-verbal (predicative) complement $\boxed{2}$.

Semantically, the assembling operator in existential predication (Figure 5) introduces a key relation of existence $\boxed{5}$ (supplying an event variable $\boxed{4}$) with only one argument the existence of which is actually predicated. This argument is identified with the index of the subject $\boxed{3}$. The semantic contribution $\boxed{6}$ of the non-verbal (predicative) complement – i.e. of the locative adverbial $\boxed{2}$ – is integrated in (the RELS list of) the content.

The semantic contribution of the assembling operator in possessive predication (Figure 6) introduces a key relation of possession $\boxed{5}$ (supplying an event variable $\boxed{4}$) whose first argument is identified with the index of the non-verbal (predicative) complement $\boxed{2}$ – the possessor – and its second argument with the index of the subject $\boxed{3}$ – the possessed entity.
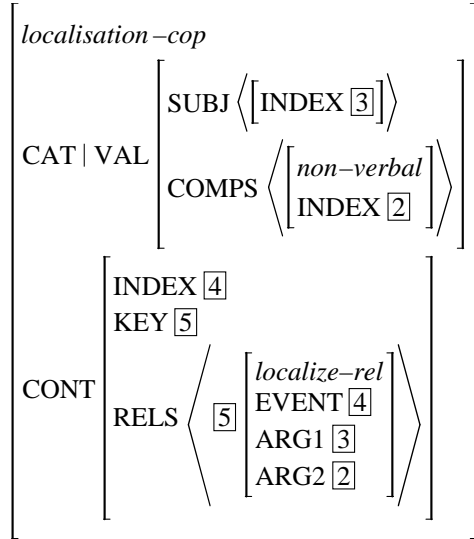
$$
\begin{bmatrix}
localisation-cop \\[4pt]
\text{CAT} \mid \text{VAL}
\begin{bmatrix}
\text{SUBJ} \left\langle \begin{bmatrix} \text{INDEX} \;\boxed{3} \end{bmatrix} \right\rangle \\[6pt]
\text{COMPS} \left\langle \begin{bmatrix} non-verbal \\ \text{INDEX} \;\boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix} \\[20pt]
\text{CONT}
\begin{bmatrix}
\text{INDEX} \;\boxed{4} \\
\text{KEY} \;\boxed{5} \\[6pt]
\text{RELS} \left\langle \boxed{5} \begin{bmatrix} localize-rel \\ \text{EVENT} \;\boxed{4} \\ \text{ARG1} \;\boxed{3} \\ \text{ARG2} \;\boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

Figure 4: Localisational predication

$$
\begin{bmatrix}
existential-predicator \\[4pt]
\text{CAT} \mid \text{VAL}
\begin{bmatrix}
\text{SUBJ} \left\langle \begin{bmatrix} \text{INDEX} \;\boxed{3} \end{bmatrix} \right\rangle \\[12pt]
\text{COMPS} \left\langle \begin{bmatrix} non-verbal \\ \text{CONT} \begin{bmatrix} \text{INDEX} \;\boxed{2} \\[4pt] \text{RELS} \;\boxed{6} \left\langle \begin{bmatrix} location-rel \\ \text{ARG1} \;\boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix} \\[24pt]
\text{CONT}
\begin{bmatrix}
\text{INDEX} \;\boxed{4} \\
\text{KEY} \;\boxed{5} \\[6pt]
\text{RELS} \left\langle \boxed{5} \begin{bmatrix} exist \\ \text{EVENT} \;\boxed{4} \\ \text{ARG1} \;\boxed{3} \end{bmatrix} \right\rangle \oplus \boxed{6}
\end{bmatrix}
\end{bmatrix}
$$
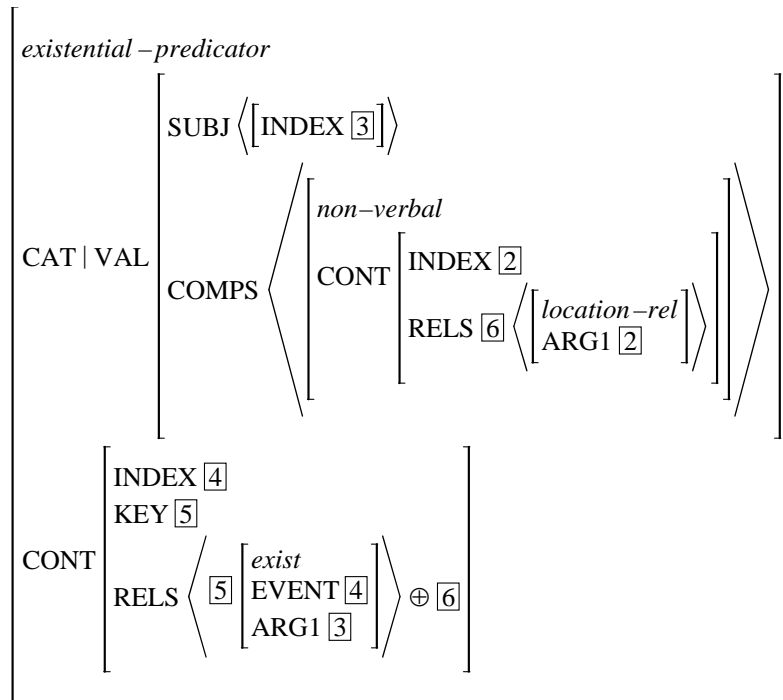
Figure 5: Existential predication

$$
\begin{bmatrix}
possessivity-predicator \\[4pt]
\text{CAT} \mid \text{VAL}
\begin{bmatrix}
\text{SUBJ} \left\langle \begin{bmatrix} \text{INDEX}\ \boxed{3} \end{bmatrix} \right\rangle \\[6pt]
\text{COMPS} \left\langle \begin{bmatrix} non-verbal \\ \text{INDEX}\ \boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix} \\[20pt]
\text{CONT}
\begin{bmatrix}
\text{INDEX}\ \boxed{4} \\
\text{KEY}\ \boxed{5} \\[4pt]
\text{RELS} \left\langle \boxed{5} \begin{bmatrix} possess-rel \\ \text{EVENT}\ \boxed{4} \\ \text{ARG1}\ \boxed{2} \\ \text{ARG2}\ \boxed{3} \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$
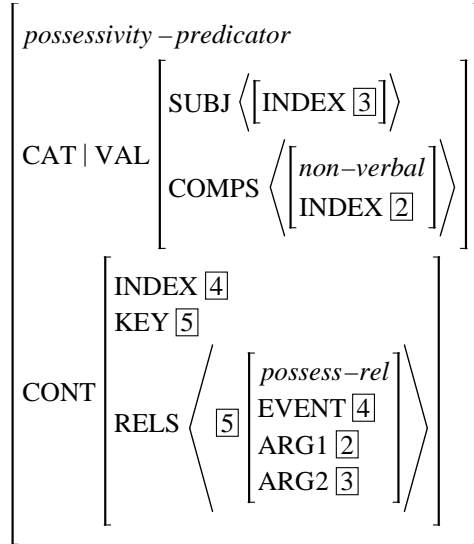
Figure 6: Possessive predication

As a result, two principally different instances of non-verbal predication can be distinguished. Morphologically signalled predicative categories are heads selecting the contingent copula as a specifier (cf. Section 3). Otherwise, the copula is the head (cf. Section 4) – when it is overt, this trivially results in a headed phrase; if there is no overt copula the result is a special type of non-headed phrase.

3.   Copular "be" as inflectional marker

In a reasonably large number of languages it is in fact the case that the absence of an overt copula stands in a paradigmatic opposition to the presence of non-present tense copula forms within a particular construction.

So, in Russian, the present tense copula is ungrammatical in combination with the predicative short adjectives (2a), but is required to encode tense in past and future tense constructions (2b). While verbs are inherent predicators with non-verbal categories this is a *derived* property. Russian short adjectives are exclusively used as predicates. As the contrast in (2c) illustrates, their attributive use is ungrammatical.

(2) a. *Otec* (**est'*) *gord* *rezul'tatami.*
      father.NOM proud.PRD-ADJ.SG.M results.INST.PL
      Father is proud of the results.

b. *Otec   byl | budet   gord   rezul'tatami.*
   father.NOM   was | will-be   proud.PRD-ADJ.SG.M   results.INST.PL
   Father was | will be proud of the results.

c. *gordyj   otec | *gord   otec*
   proud.NOM SG.M   father.NOM | proud.PRD-ADJ.SG.M   father.NOM
   a proud father

The two clauses in (2a-b) are apparently functionally equivalent – differing only in temporal features, it seems correct to propose an analysis under which the predication relations will be the same across both clauses. Given that the role of the copula here is solely functional, we take it that these cases are suggestive of a lexical approach to such tense-related paradigmatic alternation. Being morphologically signalled, the combinatorial potential of Russian short adjectives is derived lexically as a *diathesis alternation* in the sense of (Avgustinova 2001a, b), which is illustrated in (Figure 7).

$$
\begin{bmatrix}
prd-drv \\
\text{RESULT} \begin{bmatrix} prd-adjective \\ \text{DEPS} \left\langle \boxed{1}, \boxed{3}infl-cop \mid \boxed{2} \right\rangle \end{bmatrix} \\
\text{SOURCE} \begin{bmatrix} adjective \\ \text{MOD } \boxed{1} \\ \text{DEPS } \boxed{2} \end{bmatrix}
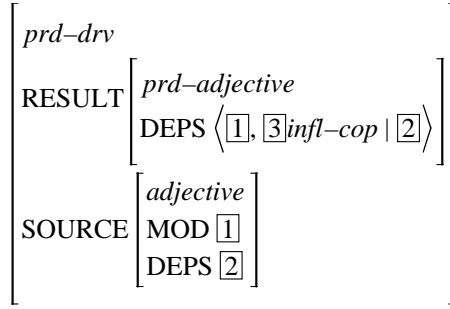\end{bmatrix}
$$

Figure 7: Russian predicative adjective derivation

The initial element $\boxed{1}$ on the DEPS list of the resulting predicative adjective is identified with the MOD value of the source adjective. This encodes the linguistic generalisation that the subject of a *predicatively* used adjective corresponds to the nominal category modified by this adjective when it is used *attributively.* The observed systematicity justified the assumption in (Avgustinova and Uszkoreit 2003) that the predicative short adjective itself is heading the construction and its VALENCE includes, in addition to SUBJ(ECT) and COM(PLEMENT)S, the attribute SP(ECIFIE)R of the type *infl(ectional)-cop(ula).* The latter is introduced as a new dependent $\boxed{3}$ of the predicative adjective. Finally, the dependents list $\boxed{2}$ of the source adjective is appended to the DEPS value of the predicative adjective. Note that the value of the ARG-ST feature is not mentioned in the constraint because nothing changes on this level. In accord with the Argument Realisation constraint of (Bouma, et al. 2001), the valence of a predicative adjective is then organised as in (Figure 8).
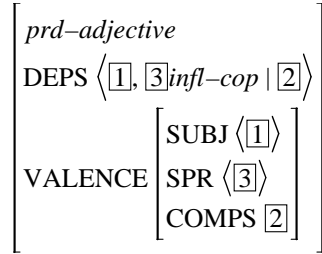
$$
\begin{bmatrix}
\textit{prd–adjective} \\
\text{DEPS } \left\langle \boxed{1}, \boxed{3}\textit{infl–cop} \mid \boxed{2} \right\rangle \\
\text{VALENCE }
\begin{bmatrix}
\text{SUBJ } \left\langle \boxed{1} \right\rangle \\
\text{SPR } \left\langle \boxed{3} \right\rangle \\
\text{COMPS } \boxed{2}
\end{bmatrix}
\end{bmatrix}
$$

Figure 8: Combinatorial potential of Russian predicative adjectives

In HPSG terms, Russian constructions with an overt *inflectional copula* are headed phrases which can be built as instances of the type *head-all-valence-phrase* (Figure 9). The head daughter is of type *prd-adjective*, as derived lexically in (Figure 7). So, the copula is taken as an optional specifier (i.e. dependent) of the adjectival predicate.

$$
\begin{bmatrix}
\textit{head–all–val–ph} \\
\text{TENSE } \boxed{5} \\
\text{MOOD } \boxed{6} \\
\text{VALENCE }
\begin{bmatrix}
\text{SUBJ } \langle\,\rangle \\
\text{SPR } \langle\,\rangle \\
\text{COMPS } \langle\,\rangle
\end{bmatrix} \\
\text{HD–DTR } \boxed{4}
\begin{bmatrix}
\textit{prd–adjective} \\
\text{VAL}
\begin{bmatrix}
\text{SUBJ } \left\langle \boxed{1} \right\rangle \\
\text{SPR } \left\langle \boxed{2}
\begin{bmatrix}
\text{SPEC } \boxed{4} \\
\text{TENSE } \boxed{5} \\
\text{MOOD } \boxed{6}
\end{bmatrix} \right\rangle \\
\text{COMPS } \boxed{3}
\end{bmatrix}
\end{bmatrix} \\
\text{NH–DTRS } \left\langle \boxed{1}, \boxed{2} \mid \boxed{3} \right\rangle
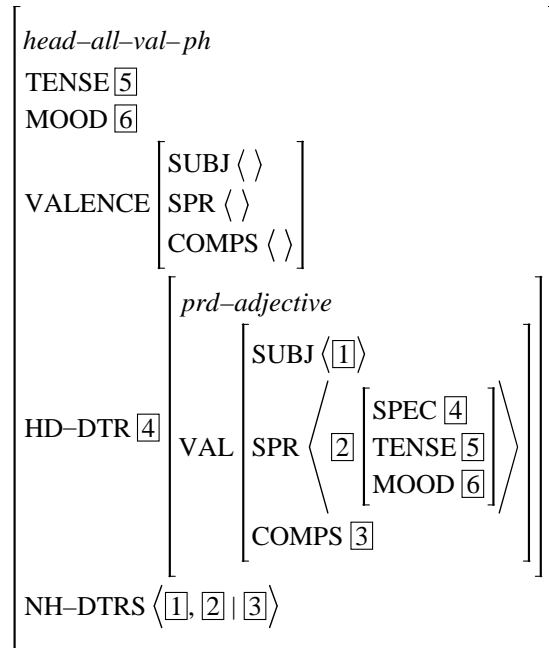\end{bmatrix}
$$

Figure 9: Construction headed by the predicative adjective

Alternatively, for a language like Russian, a language-specific constraint on type *clause* has to ensure a *default* present-tense indicative-mood interpretation in the copula-less variant whenever the specifier valence is not discharged, i.e. the VAL|SPR value is a non-empty list (Figure 10).

$$clause \Rightarrow \begin{bmatrix} \text{TENSE} & present \\ \text{MOOD} & indicative \\ \text{VAL|SPR} & \langle[2]\rangle \end{bmatrix}$$

Figure 10: 'Copula-less' constraint

4.   The syntactic structure: silent vs. overt assembling operator

With prototypical adjectives, nominals or adverbials in predicative use no morphological signalling of the predicative status is available. A constructional analysis inspired by the silent-copula-phrase approach of (Sag and Wasow 1999) is more adequate than yet another lexical derivation with no observable formal manifestation. A construction with a silent assembling operator is obtained as headless construction in (Figure 11).

$$\begin{bmatrix} silent-copula-ph \\ \text{CAT} \begin{bmatrix} \text{TENSE } present \\ \text{MOOD } indicative \\ \text{HEAD} \begin{bmatrix} assembling-operator \\ \text{FORM } fin \end{bmatrix} \\ \text{VAL} \begin{bmatrix} \text{SUBJ}\langle\,\rangle \\ \text{SPR }\langle\,\rangle \\ \text{COMPS }\langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{NON-HD-DTRS} \left\langle [A], [B]\begin{bmatrix} non-verbal \\ \text{EXT-ARG } [A] \end{bmatrix}\right\rangle \end{bmatrix}$$

Figure 11: Headless construction

The corresponding construction headed by an overt assembling operator is illustrated in (Figure 12). Intuitively, as soon as a given non-predicative category occurs in the predicate, it acquires the property of subcategorising for a subject (broadly understood as the topic of the predication). Introducing an *external argument* for non-verbal categories to be identified with the subject (Figure 13a) models the intuition of opening a slot when these categories are used predicatively.

$$\begin{bmatrix} head\text{--}all\text{--}val\text{--}ph \\[2pt] \text{CAT} \begin{bmatrix} \text{TENSE}\;\boxed{1} \\ \text{MOOD}\;\boxed{2} \\ \text{VAL} \begin{bmatrix} \text{SUBJ}\;\langle\,\rangle \\ \text{COMPS}\;\langle\,\rangle \end{bmatrix} \end{bmatrix} \\[6pt] \text{HD--DTR}\;\boxed{C} \begin{bmatrix} assembling\text{--}operator \\ \text{TENSE}\;\boxed{1} \\ \text{MOOD}\;\boxed{2} \\ \text{VAL} \begin{bmatrix} \text{SUBJ}\;\langle\boxed{A}\rangle \\ \text{COMPS}\;\langle\boxed{B}\rangle \end{bmatrix} \end{bmatrix} \\[6pt] \text{NON--HD--DTRS}\;\Big\langle\;\boxed{A},\boxed{B}\begin{bmatrix} non\text{--}verbal \\ \text{EXT--ARG}\;\boxed{A} \end{bmatrix}\Big\rangle \end{bmatrix}$$

Figure 12: Headed construction

(a) $\begin{bmatrix} \text{HEAD}\;non\text{--}verbal \\ \text{EXT--ARG}\;\boxed{1} \\ \text{SUBJ}\;\langle\boxed{1}\rangle \end{bmatrix}$ (b) $\begin{bmatrix} \text{HEAD}\,|\,\text{MOD}\;\boxed{1} \\ \text{EXT--ARG}\;\boxed{1} \end{bmatrix}$ (c) $\begin{bmatrix} \text{HEAD}\;nominal \\ \text{EXT--ARG}\;\boxed{1} \end{bmatrix}$

Figure 13: Generalised external argument

With adjectival and adverbial categories, which are specified for the head feature MOD, the external argument is the modified category (Figure 13b). With nominal categories, however, the external argument has to be explicitly introduced (Figure 13c).

5.   Conclusions and outlook

A well-known challenge to any grammatical description is posed by predicative constructions in which there is no overt copular verb interpretable as a syntactic head. Empty categories used to be designed for one or several types of copula. The HPSG formalisation sketched in this contribution allows for encoding the significant distinctions as well as for capturing the linguistic generalisations without postulating any empty categories.

The lexical derivation of Russian predicative adjectives systematically differs from the constructional treatment of non-verbal predicates with no morphological signalling of predicative status. In the latter case, the contingent copular item not only marks verbal inflection but functions as an assembling operator putting together two categories that are prototypically non-verbal. Intuitively, as soon as a given non-predicative category occurs in the predicate, it acquires the property of subcategorising for a subject (broadly understood as the topic of the predication).

Related future research has to concentrate on drawing more connections to other Slavic languages, inasmuch as the approach presented here allows linguistically adequate modelling of minimal differences between related languages. From a more general perspective, it is crucial to consider other languages with non-verbal predicative constructions, e.g., Hebrew. And finally, further development of the "generalised external argument" approach within the theoretical model of HPSG is called for.

References

Avgustinova, T. (2001a). Arguments, grammatical relations, and diathetic paradigm. 7th International Conference on HPSG, UC Berkeley, CSLI Online Publications: 23-42.

Avgustinova, T. (2001b). Distinguishing argument structure, syntactic dependents and valence in HPSG: relevance for Slavic. In: G. Zybatow et al. (eds.) Current Issues in Formal Slavic Linguistics. Peter Lang GmbH: Linguistik International 5: 554-567.

Avgustinova, T. and H. Uszkoreit (2000). An ontology of systematic relations for a shared grammar of Slavic. 18th International Conference on Computational Linguistics COLING'2000, Saarbrücken. 1: 28-34.

Avgustinova, T. and H. Uszkoreit (2003). Reconsidering the relations in constructions with non-verbal predicates. In: P. Kosta et al. (eds.) Investigations into Formal Slavic Linguistics. Contributions of the Fourth European Conference on Formal Description of Slavic Languages. Peter Lang GmbH: Linguistik International 10, Part II: 483-498.

Bouma, G., R. Malouf and I. A. Sag (2001). Satisfying constraints on extraction and adjunction. Natural Language and Linguistic Theory 19(1): 1-65.

Copestake, A., D. Flickinger, I. Sag and C. Pollard (1999). Minimal recursion semantics: An introduction. Unpublished manuscript, Stanford University and OSU. http://lingo.stanford.edu/pubs.html.

Pollard, C. and I. Sag (1994). Head-Driven Phrase Structure Grammar. Chicago, University of Chicago Press.

Sag, I. and T. Wasow (1999). Syntactic theory: a formal introduction. CSLI Publications. Stanford. CSLI.