Czech and Polish Raising/Control
with or without Structure Sharing*

Adam Przepiórkowski and Alexandr Rosen

13th April 2004

1   Introduction

Modern syntactic theories, including Principles and Parameters (P&P, i.e., Government and Binding, GB, Chomsky 1981, 1986, and the Minimalist Program, MP, Chomsky 1995) and Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag 1987, 1994), acknowledge a clear empirical distinction between raising to subject (SR) and so-called Exceptional Case Marking (ECM; here called raising to object, or OR) constructions, as in (1), on the one hand, and subject control (SC) and object control (OC) constructions, as in (2), on the other hand.

(1)   a.   John seems to support Mary.   (SR; E.)

   b.   Honza   přestal podporovat Marii.   (SR; Cz.)
        Honza-NOM stopped support-INF Marie-ACC
        'Honza stopped supporting Marie.'

   c.   Jan   przestał wspierać   Marię.   (SR; P.)
        Jan-NOM stopped support-INF Maria-ACC
        'Jan stopped supporting Maria.'

   d.   I expect John to support Mary.   (OR; E.)

   e.   Viděl jsem   Honzu   podporovat Marii.   (OR; Cz.)
        seen   AUX-1.SG Honza-ACC support-INF Marie-ACC
        'I saw Honza supporting Marie.'

   f.   *no ECM in Polish*

(2)   a.   John tries to support Mary.   (SC; E.)

   b.   Honza   zkouší podporovat Marii.   (SC; Cz.)
        Honza-NOM tries   support-INF Marie-ACC
        'Honza tries to support Marie.'

   c.   Jan   próbuje wspierać   Marię.   (SC; P.)
        Jan-NOM tries   support-INF Maria-ACC
        'Jan tries to support Maria.'

   d.   I ordered John to support Mary.   (OC; E.)

   e.   Nařídil jsem   Honzovi   podporovat Marii.   (OC; Cz.)
        ordered AUX-1.SG Honza-DAT support-INF Marie-ACC

'I ordered Honza to support Marie'

f. Kazałem   Janowi wspierać   Marię.   (OC; P.)
ordered-1.SG Jan-DAT support-INF Maria-ACC
'I ordered Jan to support Maria'

Two of the most robust cross-linguistic tests distinguishing raising and control[1] involve passivisation (e.g., Pollard and Sag 1994 and, for Czech, Skoumalová 2002) and idiom chunks (e.g., Postal 1974 and, for Polish, Zabrocki 1981): i) when the lower verb is in the passive, the meaning of the sentence is the same as in the active voice in case of raising constructions, but not in case of control constructions, e.g., (3)–(4), and ii) chunks of sentential idioms can be raised arguments, but they cannot be controllers, e.g., (5)–(6).

(3) a. Mary seems to be supported by John.   (SR; E.; $\approx$(1a))

     b. I expect Mary to be supported by John.   (OR; E.; $\approx$(1d))

(4) a. Mary tries to be supported by John.   (SC; E.; $\napprox$(2a))

     b. I ordered Mary to be supported by John.   (OC; E.; $\napprox$(2d))

(5) a. The cat seems to be out of the bag.   (SR; E.)

     b. I expect the cat to be out of the bag.   (OR; E.)

(6) a. *The cat tries to be out of the bag.   (SC; E.)

     b. *I ordered the cat to be out of the bag.   (OC; E.)

In some modern syntactic theories, including most versions of P&P and HPSG, two systematic theoretical differences between raising and control are postulated: i) semantically, raising verbs have one argument fewer than the corresponding control verbs, e.g., *seem* is a (semantically) 1-argument verb, while *try* is a (semantically) 2-argument verb; ii) structurally, the raised argument and the subject of the infinitival verb are the same element (so-called structure sharing; henceforth SS), while the controller and the subject of the infinitival verb are two different elements (the latter realised as PRO in P&P). The strong correlation between i) and ii) is ensured by, in one version of P&P, the interaction of Theta criterion (which implies the case filter), Move α and the properties of PRO, and in HPSG, by appropriate lexical entries of raising and control verbs, by the control theory and the Raising Principle.

Other syntactic accounts, including, e.g., Dziwirek 1994, 1998 within Relational Grammar and Hornstein 1999 within MP, uniformly treat obligatory control via SS (NP-movement in the latter case), i.e., assign essentially identical syntactic structures to raising and control constructions.

The destructive aim of this paper is to show, on the basis of data from Czech (Cz.) and Polish (P.), that both approaches to control are fatally flawed, i.e., on the one hand, that the purported correlation between raising and structure sharing blatantly does not hold and, thus, that the Theta criterion of P&P is empirically wrong,[2] and, on the other hand, that control is not movement or its cognate in non-transformational theories.[3]

On the constructive side, we propose an explicit and precise account of case transmission in control and raising constructions which builds on the standard HPSG approach to control and raising and on an earlier independently motivated analysis of syntactic case assignment. It turns out that only minor additions to these two grammar modules are required to successfully account for the curious case transmission facts in Czech and Polish. The analysis proposed here follows the insights of Hudson (1998, 2003), who — on the basis of Icelandic, Ancient Greek, and, in the latter article, Russian — claims that whether control involves structure-sharing or not "is ultimately an empirical matter" (Hudson 1998, p. 151).

---

[1]In this paper we focus on obligatory control within infinitival environments.

[2]Theta criterion is also abolished in Hornstein 1999, essentially on theory-internal grounds.

[3]This is the main thesis of Culicover and Jackendoff 2001, which contains a comprehensive critique of Hornstein 1999, but does not adduce the type of argument proposed here.

The following section, §2, presents the relevant case transmission data from Polish and Czech, and makes some initial empirical generalisations. The main theoretical section of the present paper, §3, contains a contrastive HPSG analysis of the data in terms of dissociating structure sharing from raising, and briefly discusses possible revisions of that analysis in view of apparent lack of SS in some Czech raising constructions. Finally, §4 summarises the main conclusions of the current article.

2 Empirical Generalisations

2.1 Polish

Consider the case transmission examples (7)–(8).

(7)   a.   Piotr   wydawał się być   niespokojny.   (SR; P.)
                Piotr-NOM seemed   be-INF uneasy-NOM
                'Piotr seemed to be uneasy.'

           b.   Pięć   kobiet   wydawało się być   niespokojnych / niespokojne.   (SR; P.)
                five-ACC women-GEN seemed   be-INF uneasy-GEN   uneasy-ACC
                'Five women seemed to be uneasy.'

(8)   a.   Piotr   bał się przyjść   niespokojny.   (SC; P.)
                Piotr-NOM feared come-INF uneasy-NOM
                'Piotr was afraid to come uneasy.'

           b.   Pięć   kobiet   bało się być   niespokojnych / niespokojne.   (SC; P.)
                five-ACC women-GEN feared   be-INF uneasy-GEN   uneasy-ACC
                'Five women were afraid to be uneasy.'

Both in SR examples, (7), and in SC examples, (8), the predicative adjective complement of infinitival copula agrees in case with the matrix subject; the fact that it is agreement and not assignment of the nominative case is confirmed by the quirky (genitive or accusative) agreement with numeral subjects in examples (7b) and (8b).[4] Examples (7) are expected: the subject of the copula is structure-shared with (raised to) the matrix subject, so they have common case value, and the case on the adjective is the result of the ordinary local case agreement between the subject of the copula and its predicative adjective complement. On the other hand, examples in (8) are not expected on the standard assumptions of P&P or HPSG: in control constructions, only the content of the controller and the subject of the infinitival is shared, not case value, so there is no appropriately cased NP/DP (henceforth, NP) with which the predicative adjective could locally agree.

Such case transmission data were previously discussed for Polish by Franks 1995 and Przepiórkowski 1999a, 2004b. In order to analyse them, Franks 1995 argues that PRO must bear case, contrary to P&P assumptions, and sketches a few possible solutions, noting their various drawbacks. Przepiórkowski 1999a, 2004b attempts to eliminate those drawbacks and proposes HPSG principles which, in the spirit of Hudson 1998, amount to ensuring that, in Polish, SC involves SS, just as SR does, and — hence — case transmission occurs obligatorily.

The problem of case transmission in SC does not arise in SS analyses of control, such as Dziwirek 1994, 1998 and Hornstein 1999, where the controller and the controllee are the same objects (Dziwirek 1994, 1998), or one is a trace of the other (Hornstein 1999). What is unexpected for such analyses, though, is the behaviour of object control, illustrated with (9) below: contrary to predictions made by

---

[4]This quirky agreement with (a class of) numeral subjects occurs also in simple predicative constructions, e.g., (i) below, and thus requires an independent explanation (Przepiórkowski 2001, 2000).

(i) Pięć   kobiet   było miłe   / miłych.   (P.)
      five-ACC women-GEN were nice-ACC   nice-GEN
        'Five women were nice.'

For the arguments that Polish numeral subjects are in fact accusative, see, e.g., Przepiórkowski 1999a, 2004a.

such accounts, there is no case transmission in Polish OC and the predicative adjective occurs in the 'default' instrumental case, so-called 'instrumental of predication'.[5]

(9)   Kazałem   Piotrowi być miłym / *miłemu.   (OC; P.)
       ordered-1.SG.MASC Piotr-DAT be-INF nice-INS   nice-DAT
       'I ordered Peter to be nice.'

Such parochial language-specific behaviour of SS or no SS in control constructions seems to be the rule cross-linguistically, cf., e.g., the fact that in Lithuanian accusative controllers do not allow SS while their genitive of negation counterparts do (Timberlake 1988); cf. also Hudson 1998, 2003 and the work cited therein on Icelandic and Ancient Greek. In the next section, we will see that also Czech exhibits interesting SS behaviour in control constructions, considerably more complex than in case of Polish.

2.2   Czech

2.2.1   Subject Control/Raising

As far as SR and SC are concerned, Czech behaves in a similar way as Polish: the Czech SR (10) and SC (11) examples parallel their Polish counterparts presented as (7) and (8) above — there is case transmission between matrix subject and predicative complement, a phenomenon unexpected in the SC examples (11).[6],[7]

(10)   a.   Petr   se zdál být   nespokojený.   (SR; Cz.)
            Petr-NOM seemed be-INF dissatisfied-NOM
            'Petr seemed to be dissatisfied'

          b.   Pět   poslanců se zdálo být   nespokojených.   (SR; Cz.)
            Five-NOM MPs-GEN seemed be-INF dissatisfied-GEN
            'Five MPs seemed to be dissatisfied'

(11)   a.   Petr   se bál přijít   neohlášený.   (SC; Cz.)
            Petr-NOM feared come-INF unannounced-NOM
            'Petr was afraid of arriving impromptu.'

          b.   Pět   poslanců se bálo být   upřímných.   (SC; Cz.)
            five-NOM MPs-GEN feared   be-INF frank-GEN
            'Five MPs were afraid to be frank.'

The analysis proposed in §3 correctly accounts for this similarity between Polish and Czech.

2.2.2   Object Control/Raising

What sets Czech apart from Polish are OC constructions, where Czech turns out to provide even more striking evidence for the lack of correlation between raising/control and SS/no SS. In Czech, the non-agreeing case of predicative adjectives is the nominative, with instrumental, the non-agreeing case in Polish, being a much more restricted option, acceptable to some extent only with copula, as in (12).

---

[5]We ignore here the instrumental of predication, whose distribution in Slavic is subject to complex semantic conditions (e.g., Pisarkowa 1965, Nichols 1981, Filip 2001, Ionin and Matushansky 2002).

[6]Again, as in Polish (cf. fn. 4), the quirky agreement with numeral phrases is a matter independent of subject/raising. Moreover, there is a difference between Polish and Czech: in Czech predicative complements agreeing with a numeral phrase, only the genitive case is possible (compare (10b) and (11b) to (7b) and (8b) above).

[7]In the absence of clear arguments for a different solution, subject numerals in Czech examples are glossed as nominative according to usual assumptions, e.g., as in Daneš *et al.* 1987, p. 43.

(12) a. Být opilý znamená být hloupý. (Cz.)
be-INF drunk-NOM means be-INF stupid-NOM
'Being drunk means being stupid.'

b. ??Být opilým znamená být hloupým. (Cz.)
be-INF drunk-INS means be-INF stupid-INS

Apart from the morphological realisation of the non-agreeing predicative case, Czech OC with dative objects is similar to Polish OC (cf. (9) and (13)) — there is no case transmission.[8]

(13) a. Marie nařídila Honzovi přijít střízlivý / *střízlivému. (OC; Cz.)
Marie-NOM ordered Honza-DAT come-INF sober-NOM sober-DAT
'Marie ordered Honza to come sober.'

b. Poručil pěti pacientům přijít svlečení / *svlečeným.
ordered-3.SG five-DAT patients-DAT come-INF undressed-NOM undressed-DAT
(OC; Cz.)
'He ordered five patients to come undressed.'

The situation is more complicated in case of OC with accusative objects. When the object is a plain (non-numeral) NP, as in (14), the case of the predicative complement can be either nominative (no case transmission, as in the dative object case) or accusative (case transmission).[9]

(14) a. Marie naučila Honzu chodit domů střízlivý / střízlivého. (OC; Cz.)
Marie taught Honza-ACC go-INF home sober-NOM sober-ACC
'Marie taught Honza to come home sober.'

b. Donutil jsem ho přijít samotného. (OC; Cz.; Franks 1998 after Toman 1991)
forced AUX-1.SG he-ACC come-INF alone-ACC
'I forced him to come alone.'

However, when the object is a numeral NP, as in (15), the preferred case of the predicative complement seems to be genitive, i.e., case transmission takes place.

(15) a. Marie přiměla pět pacientů přijít svlečených / ??svlečení /
Marie made five-ACC patients-GEN come-INF undressed-GEN undressed-NOM
??svlečené. (OC; Cz.)
undressed-ACC
'Marie made five patients come undressed.'

b. Odnaučila mnoho žáků přicházet do školy ??nevyspalí /
untaught-3.SG.FEM many-ACC pupils-GEN come-INF to school sleepy-NOM
??nevyspalé / nevyspalých. (OC; Cz.)
sleepy-ACC sleepy-GEN
'She taught many pupils to stop coming to school without having proper sleep.'

On the other hand, OR examples including accusative objects, both plain NPs and numeral phrases,[10] seem to behave as expected, i.e., case transmission takes place. In (16a) the predicative complement is accusative, in (16b) it is genitive.

(16) a. Marie viděla Honzu přijít *střízlivý / střízlivého. (OR; Cz.)
Marie saw Honza-ACC come-INF sober-NOM sober-ACC

---

[8]Examples (13) with predicative complements in dative are acceptable when predicative complements are understood as complementing the higher verb.

[9]Similarly as examples (13) above, examples (14) in the version with the predicative complement in accusative can have the meaning 'While he was sober, she taught him to come home'. Nevertheless, the glossed reading is the preferred option.

[10]Only OR verbs with accusative objects have been found.

'Marie saw Honza come sober.'

b. Marie viděla pět   pacientů   ležet   neošetřených / *neošetření /
    Marie saw five-ACC patients-GEN lie-INF unattended-GEN   unattended-NOM
    ??neošetřené.   (OC; Cz.)
       unattended-ACC
    'Marie saw five patients lying unattended.'

As the tentative acceptability ratings of the examples above show, an adequate description would profit from more representative and reliable data. This is especially true about sentences with accusative object control verbs. After corpus search, introspective research and judgments of a few speakers had failed to produce conclusive results, we resorted to a web-based survey.[11] Visitors of the site were asked to rate 26 Czech sentences. The sentences exemplified the presence or absence of structure sharing by means of the presence or absence of case transmission from matrix object to predicative complement of embedded infinitive. Table 1 gives the number of sentences in the survey listed according to the type of matrix verb (control or raising), the case of the *anchor* (the controller or the raised element)[12] and the absence or presence of structure sharing (as shown by the case of the predicative complement).

	anchor	no SS	SS
object control	*dative*	1	1
	plain accusative	4	3
	accusative numeral	2	3
object raising	*plain accusative*	2	2
	accusative numeral	2	2

Table 1: Number of sentences in the survey

In object control, two kinds of accusative controllers were tested: plain and numeral NPs, in order to verify the hypothesis that the latter support structure sharing more readily. Dative numeral NPs have not been included — their agreement pattern does not differ from that of plain dative NP.

Respondents, whose number reached 699, had to choose one of four options on an acceptability scale (fine, acceptable, strange, impossible). Some of them provided comments, mostly pointing out the difficulty of judgment due to stylistic preference for alternative syntactic structures (finite clauses instead of infinitival clauses, adverbs instead of predicative adjectives). The judgments may have been influenced by factors other than the control/raising contrast, the form of the matrix object, and the presence or absence of structure sharing. Such factors may include:

- linear distance between the matrix object and the predicative complement: the acceptability of the non-sharing option increases with longer distance;[13]

- lexical setting (e.g., infinitival copula tends to make non-sharing option more acceptable);

- idiomatic nature of the embedded predicative;

- verbal aspect of either of the verbs.

Table 2 shows results for sentences with accusative anchors. The third column gives the mean rating value on the four-point scale, while the last column gives the share of respondents for whom the examples were fine or acceptable (i.e., 1 or 2). The intervals reflect the range of acceptability judgments

example type	structure sharing	mean	fine or acceptable for
accusative OC	yes	1.8	30–74%
	no	2.8	1–50%
numeral acc. OC	yes	2.4	28–74%
	no	3.4	10–11%
accusative OR	yes	1.6	84–88%
	no	3.6	3–5%
numeral acc. OR	yes	2.2	54–71%
	no	3.2	12–20%

Table 2: The survey results for accusative OC/OR

for multiple examples of the same type.

The results of the survey support the following conclusions:

1. In OC with accusative objects, SS is slightly preferred when the object is an ordinary NP, as in (14).

2. However, when the object in an OC construction is a numeral phrase, SS is strongly preferred, cf. (15).

3. In OR, SS is the strongly preferred option, cf. (16).

2.3 Summary of the Data

Hudson 2003 provides a summary of sharing (SS) and non-sharing (PRO) options for Russian, Icelandic and Ancient Greek. In Table 3, the summary is extended to Polish and Czech.[14] Czech and Polish data are presented in more detail in Table 4.

Anchor		Structure (Structure Sharing or PRO)				
θ-marked	fun	Russian	Icelandic	A. Greek	Polish	Czech
no	subj	SS	SS	SS	SS	SS
yes	subj	SS	SS (??PRO)	SS	SS	SS
no	obj	–	SS	SS,PRO	–	SS
yes	obj	(SS),PRO	SS,PRO	SS,PRO	PRO	SS,PRO

Table 3: Table I in Hudson 2003, extended with Polish and Czech

3 An HPSG Account

The analysis of case transmission in Czech and Polish reflects the following observations made on the basis of the data considered above, as well as the data presented in Timberlake 1988 and Hudson 1998,

---

[11]The second co-author wishes to thank Hana Skoumalová and Johanka Doležalová for their valuable technical assistance.

[12]We borrow the term *anchor*, as used in this context, from Hudson 2003.

[13]Similarly as in Polish SC and SR examples, where instrumental case of the embedded predicative complement is more acceptable with the growing distance from the matrix subject, cf. Przepiórkowski 1999a.

[14]θ-*marked* anchors are controllers, non-θ-*marked* anchors are raised elements.

Verb	Anchor		Czech pred.			Polish pred.			
	subj	**obj**	nom	gen	acc	nom	gen	acc	ins
SC	(nom) NP		●(SS)			●(SS)			
	(nom/acc) №			●(SS)			●(SS)	●(SS)	○
SR	(nom) NP		●(SS)			●(SS)			
	(nom/acc) №			●(SS)			●(SS)	●(SS)	
OC		dat	●						
		acc NP	○		●(SS)				●
		acc №		●(SS)					
OR		acc NP			●(SS)				
		acc №		●(SS)					

● acceptable (no SS) ●(SS) acceptable (SS)

○ marginally acceptable (no SS)

NP nominal phrase № numeral phrase

Table 4: Czech and Polish raising/control in detail

2003:

- in many languages (Icelandic, Ancient Greek, Lithuanian, Polish, Russian, Czech. . .) there are cases of semantic control involving case transmission, i.e., involving structure sharing,

- there are idiosyncratic differences between languages in the specific control environments which allow or force SS.

Hence, the analysis should follow as much as possible from general and independently needed principles, but the need for additional parochial principles should not be surprising.

The following subsections present an HPSG account of case transmission in Czech and Polish control constructions based on, and extending, the analysis for Polish presented in Przepiórkowski 2004b. We start with various independently needed HPSG assumptions concerning control, raising, and case assignment, §§3.1–3.2, and then we present the additional parochial principles needed to account for Czech and Polish case transmission, §3.3. The final subsection, §3.4, contains the justification for some of the technical details of the current account, as well as a discussion of possible revisions of this account.

3.1 Control and Raising in HPSG

In standard HPSG, raising verbs are assumed to have lexical entries like the one schematically presented in (17).

(17)   A Subject Raising verb (e.g., Cz. *začít* or P. *zacząć*, 'begin')

$$
\begin{bmatrix}
word \\
\text{SUBJ} & \langle \boxed{0} \rangle \\
\text{COMPS} & \left\langle \text{VP} \begin{bmatrix} \text{SUBJ} \langle \boxed{0} \rangle \\ \text{CONT} \boxed{1} \end{bmatrix} \right\rangle \\
\text{CONT} & \mathbf{P}(\boxed{1})
\end{bmatrix}
$$

In (17) the content of the infinitival VP (cf. [1]) is identified with the only semantic argument of the SR verb (cf. **P**([1])), where **P** stands for the semantic predicate expressed by the verb), while the subject of the VP ([0]) is structure shared with the subject of the SR verb. Note that boxed numbers (e.g., [1]) are variables whose values are feature structures, and multiple occurrences of the same variable in a feature structure denote structure sharing.

In fact, HPSG, just as other modern syntactic theories, makes the distinction between 'surface' arguments and 'deep' arguments, e.g., in passive sentences, the 'deep' object is the 'surface' subject, while the 'deep' subject may be realised as an oblique 'surface' argument. It is assumed that 'deep' argument structure is constant for all forms of a given lexeme, while the 'surface' argument structure may vary with, e.g., the voice of a given word form. The two argument structures are represented with the attributes ARG-ST (i.e., ARGUMENT-STRUCTURE, for 'deep' arguments) and DEPS (i.e., DEPENDENTS, for 'surface' arguments). Obviously, the values of these two attributes are systematically related via general grammatical principles (cf., e.g., Manning and Sag 1998, 1999, Bouma *et al.* 2001, Avgustinova 2001, Przepiórkowski 2004c).

Given this distinction, a more verbose, but still schematic, lexical entry for a subject raising verb is presented in (18). That lexical entry makes it clear that, in SR verbs, it is the 'deep' subject of the raising verb and the 'surface' subject of the lower verb that are co-indexed.

(18)   A Subject Raising verb (revised)

$$
\begin{bmatrix}
word \\
\text{ARG-ST} \quad
\begin{bmatrix}
\text{SUBJ} \langle [0] \rangle \\
\text{COMPS} \left\langle \text{VP}
\begin{bmatrix}
\text{DEPS} \mid \text{SUBJ} \langle [0] \rangle \\
\text{CONT} [1]
\end{bmatrix}
\right\rangle
\end{bmatrix} \\
\text{CONT} \quad \mathbf{P}([1])
\end{bmatrix}
$$

Lexical entries such as (18) will give rise to syntactic structures such as the structure in Figure 1, for Polish *Janek zaczął myśleć* 'Janek started to think'. This structure illustrates also a number of HPSG
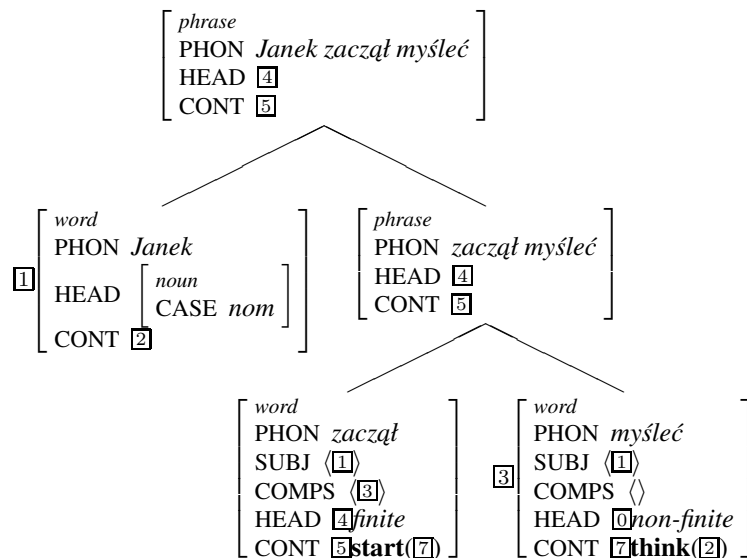
Figure 1: The preliminary structure of P. *Janek zaczął myśleć* (SR).

assumptions, including the following:

- there is an attribute whose values represent the phonological makeup of a given expression (word or phrase), PHON; for the purpose of this article, the values of this attribute are approximated orthographically,

- there is an attribute whose values represent the morphosyntax of a given expression, namely, HEAD,

- the values of HEAD are structure shared along the head projection, e.g., in Figure 1, subtrees corresponding to *zaczął*, *zaczął myśleć* and *Janek zaczął myśleć* have the same value of HEAD, ④,

- similarly, the values of CONT are normally structure shared along the head projection, cf. ⑤.

Throughout this paper, features which are not immediately relevant will often be ignored in such figures; e.g., SUBJ and COMPS are omitted from the structure corresponding to the word *Janek* (their values are empty lists here).

In contrast to raising verbs, control verbs have lexical entries like (19), giving rise to structures as in Figure 2, for Polish *Janek próbował myśleć* 'Janek tried to think'.

$$
\begin{bmatrix}
phrase \\
\text{PHON } \textit{Janek próbował myśleć} \\
\text{HEAD } ④ \\
\text{CONT } ⑤
\end{bmatrix}
$$

①
$$
\begin{bmatrix}
word \\
\text{PHON } \textit{Janek} \\
\text{HEAD } \begin{bmatrix} noun \\ \text{CASE } nom \end{bmatrix} \\
\text{CONT } ②
\end{bmatrix}
$$
$$
\begin{bmatrix}
phrase \\
\text{PHON } \textit{próbował myśleć} \\
\text{HEAD } ④ \\
\text{CONT } ⑤
\end{bmatrix}
$$

$$
\begin{bmatrix}
word \\
\text{PHON } \textit{próbował} \\
\text{SUBJ } \langle ① \rangle \\
\text{COMPS } \langle ③ \rangle \\
\text{HEAD } ④\textit{finite} \\
\text{CONT } ⑤\textbf{try}(②,⑦)
\end{bmatrix}
$$
③
$$
\begin{bmatrix}
word \\
\text{PHON } \textit{myśleć} \\
\text{SUBJ } \langle \text{NP[CONT }②] \rangle \\
\text{COMPS } \langle \rangle \\
\text{HEAD } ⓪\textit{non-finite} \\
\text{CONT } ⑦\textbf{think}(②)
\end{bmatrix}
$$
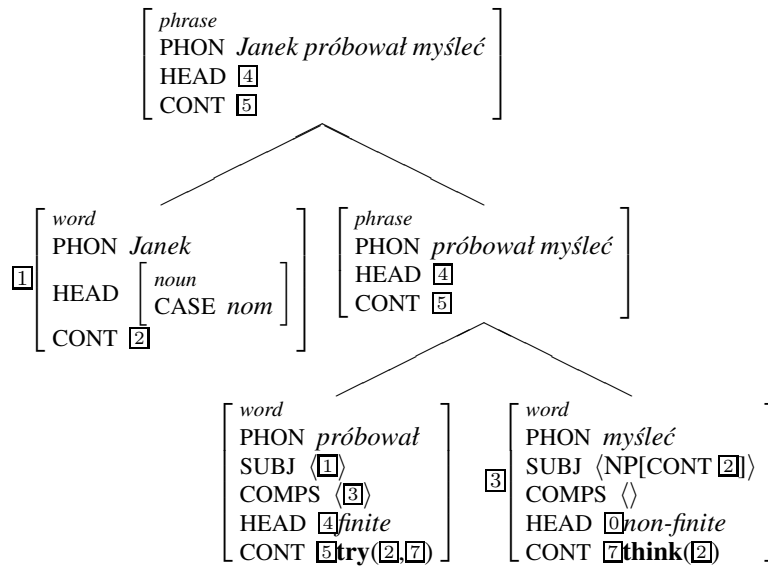
Figure 2: The preliminary structure of P. *Janek próbował myśleć* (SC).

(19)   A Subject Control verb (e.g., Cz. *zkusit* or P. *próbować*, 'try')

$$
\begin{bmatrix}
word \\
\text{ARG-ST } \begin{bmatrix} \text{SUBJ } \langle \text{NP}[\text{CONT } ①] \rangle \\ \text{COMPS } \left\langle \text{VP} \begin{bmatrix} \text{DEPS} \mid \text{SUBJ } \langle [\text{ CONT } ①] \rangle \\ \text{CONT } ② \end{bmatrix} \right\rangle \end{bmatrix} \\
\text{CONT } \quad \textbf{P}(①,②)
\end{bmatrix}
$$
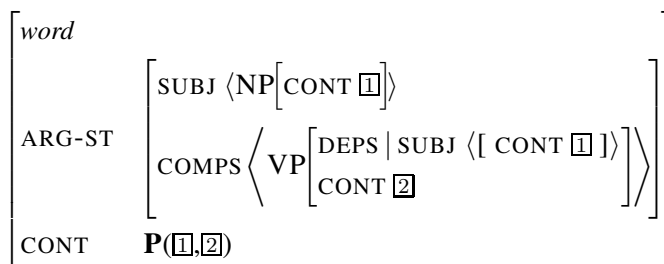
A subject control verb, as in (19), is a (semantically) 2-argument verb: the first semantic argument (i.e., ① in **P**(①,②)) is the content of its subject (①), while the second semantic argument is the content of the infinitival VP (cf. ②). Moreover, this schematic lexical entry does not specify full structure sharing between the subject of the verb and the subject of the VP complement; instead, the two subjects are co-indexed (cf. ①), i.e., they share their semantics.[15] A related difference between lexical entries of subject raising verbs and subject control verbs that should be noted is that raising verbs do not specify the morphosyntactic makeup of their subjects, they take whatever subjects are required by their VP complements, while control verbs specify their subjects as NPs.

---

[15]For the sake of cross-theoretical readability, throughout the paper we present a simplified version of HPSG structures and analyses. In particular, in standard HPSG, only parts of the values of CONT are structure shared in control constructions.

Lexical entries of ECM verbs and object control verbs are analogous to the lexical entries of raising to subject and subject control verbs above, and they display analogous differences. A schematic lexical entry for ECM verbs, treated in HPSG as raising to object verbs, is shown below.

(20)   An Object Raising (ECM) verb (e.g., Cz. *vidět*, 'see')

$$
\begin{bmatrix}
word \\
\text{ARG-ST} \begin{bmatrix} \text{SUBJ} \ \langle \text{NP}\begin{bmatrix}\text{CONT} \ \boxed{1}\end{bmatrix}\rangle \\ \text{COMPS} \ \left\langle \boxed{0}, \text{VP}\begin{bmatrix}\text{DEPS}\,|\,\text{SUBJ} \ \langle\boxed{0}\rangle \\ \text{CONT} \ \boxed{2}\end{bmatrix}\right\rangle \end{bmatrix} \\
\text{CONT} \quad \mathbf{P}(\boxed{1},\boxed{2})
\end{bmatrix}
$$

According to (20), the first semantic argument of an OR verb is the semantic content of the subject of that verb (cf. $\boxed{1}$), while the second semantic argument is the content of the infinitival VP (cf. $\boxed{2}$). Also, the subject of the VP complement is structure shared with (raised to) the object of the OR verb (cf. $\boxed{0}$).

Finally, according to lexical entries of object control verbs, such verbs have three semantic arguments and their objects must be co-indexed with the subjects of their VP complements:

(21)   An Object Control verb (e.g., Cz. *přikázat* or P. *kazać*, 'order')

$$
\begin{bmatrix}
word \\
\text{ARG-ST} \begin{bmatrix} \text{SUBJ} \ \langle \text{NP}\begin{bmatrix}\text{CONT} \ \boxed{1}\end{bmatrix}\rangle \\ \text{COMPS} \ \left\langle \text{NP}\begin{bmatrix}\text{CONT} \ \boxed{2}\end{bmatrix}, \text{VP}\begin{bmatrix}\text{DEPS}\,|\,\text{SUBJ} \ \langle[\ \text{CONT} \ \boxed{2}\ ]\rangle \\ \text{CONT} \ \boxed{3}\end{bmatrix}\right\rangle \end{bmatrix} \\
\text{CONT} \quad \mathbf{P}(\boxed{1},\boxed{2},\boxed{3})
\end{bmatrix}
$$

In (21), as in (20), the first semantic argument of the OC verb is the content of its subject (cf. $\boxed{1}$) and, again as in case of OR verbs, semantic argument of the OC verb is the content of the infinitival VP (cf. $\boxed{3}$). However, the subject of the VP is only co-indexed (cf. $\boxed{2}$) with the object of an OC verb and the content of that object (cf. $\boxed{2}$ again) is the second semantic argument of the OC verb.

Now, HPSG control theory as formulated in Pollard and Sag 1994 assumes that controlled subjects are really anaphors and that their distribution and reference is accounted for by the binding theory. This does not seem plausible for Czech or Polish, where anaphors are strictly bound by subjects, while controlled elements may be controlled by either subjects or complements, so we assume that such a theory is at best a parochial principle for English, corresponding to Czech and Polish parochial principles introduced in §3.3.

On the other hand, in case of raising predicates, the correlation between the syntactic aspects of raising (structure sharing of arguments) and the semantic aspects of raising (no semantic role assigned to the raised argument) has been ensured by the Raising Principle, schematically quoted below.

In every lexical entry E in which an argument is structure shared with another argument's subject, i.e., in every lexical entry E of the form

$$
\begin{bmatrix}
word \\
\text{ARG-ST} \begin{bmatrix} \text{SUBJ} & \langle\boxed{1}\rangle \\ \text{COMPS} & \langle... [\ \text{DEPS}\,|\,\text{SUBJ} \ \boxed{1}\ ] ...\rangle \end{bmatrix}
\end{bmatrix}
$$

or

$$
\begin{bmatrix}
word \\
\text{ARG-ST} \begin{bmatrix} \text{COMPS} & \langle... \boxed{1} ... [\ \text{DEPS}\,|\,\text{SUBJ} \ \boxed{1}\ ] ...\rangle \end{bmatrix}
\end{bmatrix},
$$

the element $\boxed{1}$ is assigned no semantic role in E (i.e., the CONT value of $\boxed{1}$ is not a semantic argument of the predicate expressed in the CONT value of E).[16]

Note that the Raising Principle corresponds to the Theta criterion of P&P:[17]

Each argument bears one and only one theta-role, and each theta-role is assigned to one and only one argument.

We will see in §3.4 that, perhaps contrary to expectations, the dissociation of raising/control and SS proposed in §3.3 does not violate the Raising Principle.

3.2 Case Assignment and Case Agreement in HPSG

Together with Przepiórkowski (1999b,a, 2004b), we assume the following principles of grammatical case assignment:[18]

(22)   grammatical case is assigned (checked) at the level of the 'surface' argument structure of words, i.e., at the level of DEPS (a possibly universal principle);

(23)   in case of raising, when an argument occurs at a number of argument structures of different verbs, case is assigned at the highest argument structure on which the argument occurs (possibly universal);

(24)   the following syntactic case assignment principles (among others) hold for Polish:

    a.   for subjects of finite verbs:

        i.   assign the nominative to NPs (P.);

        ii.   assign the accusative to Numeral Phrases (P.);

    b.   for subjects of non-finite verbs:

        i.   assign the null case (cf. Chomsky and Lasnik 1995; P.);

(25)   null case cannot be morphologically realised (universal).

For the purpose of this article, we extend these principles to Czech in the following way:

(24′)   the following syntactic case assignment principles (among others) hold for Czech:

    a.   for subjects of finite verbs:[19]

        i.   assign the nominative to NP and NumP phrases (Cz.);

    b.   for subjects of non-finite verbs:[20]

        i.   assign the nominative case (Cz.).

---

[16]In fact, this principle was originally formulated in terms of a single attribute SUBCAT, whose value is the list of all arguments, subjects and complements alike.

[17]This formulation is taken from Haegeman 1991, p. 63.

[18]For the sake of brevity, we will not for formalise these assumptions here. For discussion, justification and formalisation, the reader is referred to Przepiórkowski 1999a.

[19]Note that it would be possible to analyse Czech numeral subjects as accusative, just as in Polish, but, unlike in Polish, there are no strong arguments for doing so, so we respect the tradition here; cf. fn. 7.

[20]Note that, while in Polish infinitival copula constructions the predicate must occur in the instrumental case, cf. (i) below, in Czech it normally occurs in the nominative case, with the instrumental as a very restricted option, cf. (12) on p. 5.

(i)   a. Być   pijanym   to być   głupim.   (P.)
         be-INF drunk-INS is be-INF stupid-INS
         'To be drunk is to be stupid.'
      b. *Być   pijany   to być   głupi.   (P.)
         be-INF drunk-NOM is be-INF stupid-NOM

12

Let us now consider how these principles apply to the SR sentence *Janek zaczął myśleć* 'Janek started to think' of Figure 1. There are three words in that syntactic structure, two of which, *zaczął* 'started' and *myśleć* 'think', have non-empty argument structures. In fact, both verbs have the same argument structures, both containing just the subject *Janek*. If case were assigned at the level of the lower non-finite verb *myśleć*, the null case would be assigned, courtesy of principle (24bi). On the other hand, if case were assigned at the level of the higher finite verb *zaczął*, the subject would be marked as nominative, in accordance with (24ai). Since both principles would be assigning case to the same element, i.e., the raised subject, case clash would result and the whole sentence would be ungrammatical. This does not happen because principle (23) states that case is assigned only at the highest verb, i.e., in this case, at the level of *zaczął*, and — hence — it is nominative.

On the other hand, in case of the SC sentence *Janek próbował myśleć* 'Janek tried to think' of Figure 2, the argument structures of the two verbs are different: the two subjects are not structure shared, they only share their CONT values. That means that both case assignment principles, (24bi) and (24ai) are applicable, assigning the null case to the (unrealised) subject of the lower verb, and the nominative to the realised subject of the higher verb.

Let us now consider the case of adjectival or nominal predicates. According to standard HPSG assumptions, predicates (incl. predicative adjectives) have the following schematic lexical structure:

(26)
$$\begin{bmatrix} word \\ \text{ARG-ST|SUBJ} & \langle \text{XP[CONT } \boxed{0}] \rangle \\ \text{HEAD} & [\text{PRD } +] \\ \text{CONT} & \mathbf{P}(\boxed{0},...) \end{bmatrix}$$

That is, predicates are lexical structures (structures of type *word*) marked as predicative ([PRD +]), subcategorising for a subject (XP) and introducing a semantic predicate (**P**) whose argument (perhaps one of several arguments, in case the predicate subcategorises for complements) is the semantic content of that subject (cf. the variable $\boxed{0}$).

For example, a simplified lexical entry for the P. nominative adjectival wordform *miły* 'nice' is given below:

(27)
$$\begin{bmatrix} word \\ \text{PHON miły} \\ \text{ARG-ST|SUBJ} & \langle \text{NP[CONT } \boxed{0}] \rangle \\ \text{HEAD} & \begin{bmatrix} adjective \\ \text{CASE } nom \\ \text{PRD } + \end{bmatrix} \\ \text{CONT} & \mathbf{nice}(\boxed{0}) \end{bmatrix}$$

Now, we assume that, in the typical case, predicative case marking in Czech and Polish consists in agreement with the cased phrase modified by the predicate, or else, assignment of the instrumental case, the latter option, common in Slavic languages, being restricted in various ways.[21] That is, we assume the following principle:[22]

(28)
$$\begin{bmatrix} \text{ARG-ST|SUBJ} & \langle \text{XP[CASE } \boxed{1}] \rangle \\ \text{HEAD} & \begin{bmatrix} \text{PRD } + \\ \text{CASE } \boxed{2} \end{bmatrix} \end{bmatrix} \rightarrow (\boxed{1} = \boxed{2}) \vee (\boxed{2} = ins)$$

According to this implicational principle, for any cased ([CASE $\boxed{2}$]) predicate ([PRD +]) subcategorising for a cased subject ([ARG-ST|SUBJ $\langle$XP[CASE $\boxed{1}$]$\rangle$]), the case of the subject must agree with the case of the predicate ($\boxed{1} = \boxed{2}$), or else the case of the predicate must be instrumental ($\boxed{2} = ins$).

The wordform *miły*, whose lexical entry is given in (27) above, actually must — in order to satisfy the above principle — have the following structure:

---

[21]See the references cited in fn. 5, as well as the discussion in Przepiórkowski 1999a. In Cz., this option seems to be restricted to predicative complements of the copula-like verbs.

[22]This is a simplified version of the principle, which does not take into consideration quirky (possibly genitive in P., obligatorily genitive in Cz.) case agreement with numeral phrases; see Przepiórkowski 1999a, 2000 for the full version.

(29)
$$
\begin{bmatrix}
\textit{word} \\
\text{PHON miły} \\
\text{ARG-ST|SUBJ } \langle \text{NP} \begin{bmatrix} \text{CASE } \textit{nom} \\ \text{CONT } \boxed{0} \end{bmatrix} \rangle \\
\text{HEAD } \begin{bmatrix} \textit{adjective} \\ \text{CASE } \textit{nom} \\ \text{PRD } + \end{bmatrix} \\
\text{CONT } \mathbf{nice(\boxed{0})}
\end{bmatrix}
$$

Moreover, the predicative form *miły* does not take any complements, and its 'surface' arguments are equal to its 'deep' arguments:

(30)
$$
\begin{bmatrix}
\textit{word} \\
\text{PHON miły} \\
\text{ARG-ST } \begin{bmatrix} \text{SUBJ } \langle \boxed{1}\text{NP} \begin{bmatrix} \text{CASE } \textit{nom} \\ \text{CONT } \boxed{0} \end{bmatrix} \rangle \\ \text{COMPS } \langle \rangle \end{bmatrix} \\
\text{DEPS } \begin{bmatrix} \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \langle \rangle \end{bmatrix} \\
\text{HEAD } \begin{bmatrix} \textit{adjective} \\ \text{CASE } \textit{nom} \\ \text{PRD } + \end{bmatrix} \\
\text{CONT } \mathbf{nice(\boxed{0})}
\end{bmatrix}
$$

Before we can illustrate the background HPSG assumptions concerning predicative case marking, we need to have a look at the lexical entry for the predicative copula:

(31)
$$
\begin{bmatrix}
\textit{word} \\
\text{PHON być} \\
\text{ARG-ST } \begin{bmatrix} \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \langle \begin{bmatrix} \text{DEPS|SUBJ } \langle \boxed{1} \rangle \\ \text{CONT } \boxed{2} \\ \text{HEAD|PRD } + \end{bmatrix} \rangle \end{bmatrix} \\
\text{HEAD } \textit{non-finite} \\
\text{CONT } \boxed{2}
\end{bmatrix}
$$

According to this lexical entry, the infinitival form of the predicative copula *być* 'to be' is really a raising verb (cf. (18) in §3.1): its 'deep' subject (\boxed{1}) is structure shared with the 'surface' subject of its complement (also \boxed{1}). Moreover, the predicative copula is a semantically vacuous verb: its semantic content (\boxed{2}) is just the content of its predicative complement (again, \boxed{2}).

These background assumptions are illustrated with Figure 3, showing the syntactic structure of sentence (32):

(32) Janek   jest miły. (P.)
     Janek-NOM is   nice-NOM

     'Janek is nice'.

Note that the subject *Janek* occurs on two argument structures here: it is the subject of the finite verb *jest* and the subject of the predicative adjective *miły*. However, in accordance with (23), case is assigned at the level of *jest*, so it must be nominative (cf. (24ai)). On the other hand, the predicative complement of the copula must satisfy (28), so it must either agree with the nominative subject, or else occur in the instrumental case. Figure 3 illustrates the former possibility, the latter option being very restricted in Polish and often considered obsolete.

3.3 Case Transmission

3.3.1 SR and OR

It should be clear that nothing needs to be added to the HPSG assumptions discussed above to account for the case transmission in SR examples such as P. (7a) or Cz. (10a). In such cases:
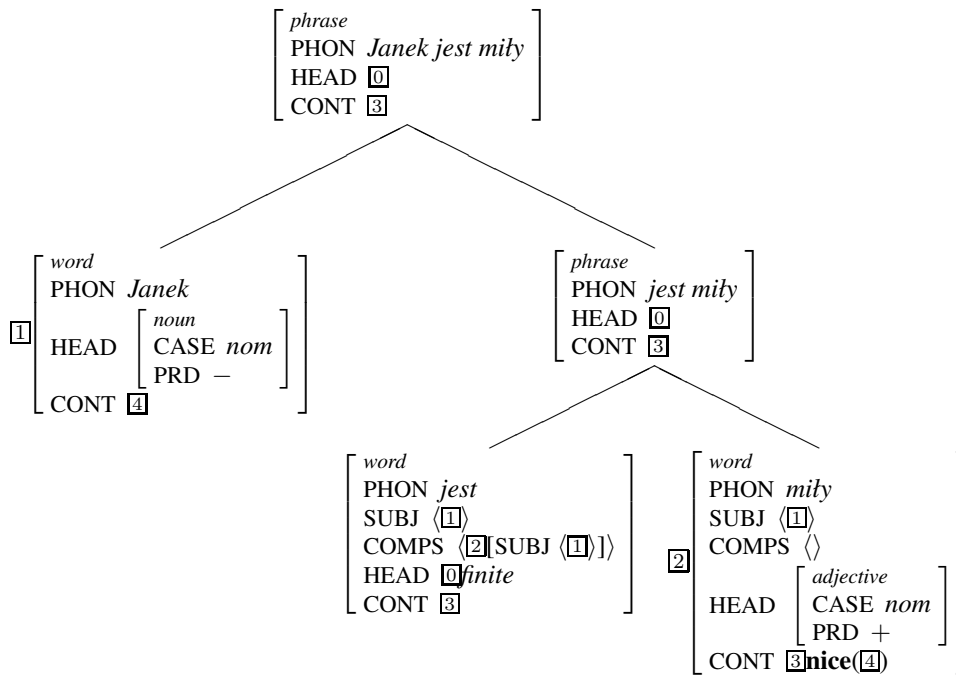
$$
\begin{bmatrix}
phrase \\
\text{PHON } \textit{Janek jest miły} \\
\text{HEAD } \boxed{0} \\
\text{CONT } \boxed{3}
\end{bmatrix}
$$

$$
\boxed{1}
\begin{bmatrix}
word \\
\text{PHON } \textit{Janek} \\
\text{HEAD }
\begin{bmatrix}
noun \\
\text{CASE } nom \\
\text{PRD } -
\end{bmatrix} \\
\text{CONT } \boxed{4}
\end{bmatrix}
\qquad
\begin{bmatrix}
phrase \\
\text{PHON } \textit{jest miły} \\
\text{HEAD } \boxed{0} \\
\text{CONT } \boxed{3}
\end{bmatrix}
$$

$$
\begin{bmatrix}
word \\
\text{PHON } \textit{jest} \\
\text{SUBJ } \langle \boxed{1} \rangle \\
\text{COMPS } \langle \boxed{2}[\text{SUBJ } \langle \boxed{1} \rangle] \rangle \\
\text{HEAD } \boxed{0}\textit{finite} \\
\text{CONT } \boxed{3}
\end{bmatrix}
\qquad
\boxed{2}
\begin{bmatrix}
word \\
\text{PHON } \textit{miły} \\
\text{SUBJ } \langle \boxed{1} \rangle \\
\text{COMPS } \langle \rangle \\
\text{HEAD }
\begin{bmatrix}
adjective \\
\text{CASE } nom \\
\text{PRD } +
\end{bmatrix} \\
\text{CONT } \boxed{3}\textbf{nice}(\boxed{4})
\end{bmatrix}
$$

Figure 3: The structure of (32).

- the subject occurs on three argument structures, i.e., those of the finite raising verb, the copula and the predicative adjective;

- (23) says that case is assigned on the highest argument structure, i.e.,

- the nominative is assigned via (24ai) for P. and (24′ai) for Cz.,

- hence, the subject of the predicative adjective is nominative,

- and, in order to satisfy (28), the adjective itself must be either nominative or instrumental, the latter option heavily restricted in case of Cz.

In Polish, (7a) illustrates the former possibility, with the instrumental option also (marginally) possible, cf. (33):

(33) ?Janek   wydawał się być   niespokojnym.
      Janek-NOM seemed   be-INF uneasy-INS

In case of SR examples involving numeral anchors, as in P. (7b) or Cz. (10b), the reasoning would be essentially the same, although the predicative case marking principle (28) would have to take into account the quirky predicative agreement with numeral phrases; cf. fn. 22.

A similar reasoning can be applied to Czech ECM constructions to explain the agreement between the predicative adjective and the object of the ECM verb.

3.3.2 SC

On the other hand, it seems that the above assumptions lead to blatantly wrong predictions in case of SC, which involves obligatory SS both in Cz. and in P., cf. (11a) and (8a):

- the subject of the copula is, again, structure shared with the subject of the predicative adjective,

- but the subject of the control verb is only co-indexed with the subject of the copula (and, hence, with the subject of the predicate), so

- although the nominative case is assigned to the subject of the finite verb just as in case of SR verbs,

- the unrealised subject of the non-finite verb receives the null case in P., cf. (24bi), or the nominative case in Cz., cf. (24′bi),

- so the adjective must be in the instrumental case in Polish, via (28) (it cannot agree with the null-cased subject because null case cannot be realised morphologically, cf. (25)), or in the nominative (or perhaps also instrumental) case in Czech.

This last prediction is not fulfilled for either language: in both languages the predicative adjective agrees with the subject of the SC verb (or, marginally, occurs in the instrumental).

The matter is not that simple, though. HPSG is a declarative (non-transformational) constraint-based theory, i.e., a structure is grammatical if and only if it satisfies all principles (constraints) of the grammar. That is, structures which are not explicitly forbidden by the grammar are licensed. Now, lexical entries for control verbs require that the controller and the controllee be co-indexed, i.e., that they share (parts of) their semantics, but nothing in the grammar actually forbids full structure sharing of complete controller and controllee.[23] So, in fact, SC verbs, whose structure is repeated in (19), are in principle ambiguous between a non-SS interpretation, in which $\boxed{3} \neq \boxed{4}$, and a SS interpretation, where $\boxed{3} = \boxed{4}$.

(19)   A Subject Control verb

$$
\begin{bmatrix}
\textit{word} \\
\text{ARG-ST} \quad
\begin{bmatrix}
\text{SUBJ} \ \langle \boxed{3}\text{NP}\big[\text{CONT} \ \boxed{1}\big]\rangle \\
\text{COMPS} \ \Big\langle \text{VP}\begin{bmatrix}\text{DEPS} \,|\, \text{SUBJ} \ \langle \boxed{4}[\ \text{CONT} \ \boxed{1}\ ]\rangle \\ \text{CONT} \ \boxed{2}\end{bmatrix}\Big\rangle
\end{bmatrix} \\
\text{CONT} \qquad \mathbf{P}(\boxed{1},\boxed{2})
\end{bmatrix}
$$

The reasoning about case non-transmission in SC constructions assumed the non-identity of the controller and the controllee ($\boxed{3} \neq \boxed{4}$); in case of the identity ($\boxed{3} = \boxed{4}$), the reasoning is the same as for the raising verbs in §3.3.1, i.e., case transmission is predicted. This means that, instead of facing the problem of wrong predictions (instrumental or, in Cz., nominative instead of agreement), we face the problem of overgeneration (instrumental or nominative, as well as agreement). That is, what is needed is a principle *forcing* the SS interpretation of control verbs. Such a principle is presented below (in words and as a formal HPSG principle):

(34)   Subject Control with SS

Whenever the deep subject of X is co-indexed with the surface subject of a VP complement of X, the two subjects are the same element.

$$
\begin{bmatrix}
\textit{word} \\
\text{ARG-ST} \quad
\begin{bmatrix}
\text{SUBJ} & \langle \boxed{1} [\ \text{CONT} \ \boxed{0}\ ]\rangle \\
\text{COMPS} & \langle \text{VP}\big[\text{DEPS} \,|\, \text{SUBJ} \ \langle \boxed{2} [\ \text{CONT} \ \boxed{0}\ ]\rangle\big]\rangle
\end{bmatrix}
\end{bmatrix} \rightarrow \boxed{1} = \boxed{2}
$$

3.3.3   OC in Polish

In case of P. OC, as in (9) above, we are faced with a similar overgeneration problem:

---

[23]The possibility of such an analysis of the Icelandic data is already suggested in Pollard and Sag 1994, p. 140, fn. 40.

- if the controller (the object of the OC verb) and the controllee are 'accidentally' structure shared, then the case of the controller is 'transmitted' all the way to the subject of the predicative adjective, in a way similar to OR cases,

- on the other hand, if only the CONT values (or, possibly, also other components but not the full structures) are shared between the controlling object and the controlled subject, then the null case assigned to the controllee and, according to the same reasoning as in the case of SC in §3.3.2, the adjective must bear the instrumental case in P., or the nominative case in Cz.

However, in case of P. OC, case is not 'transmitted' to the predicative adjective, so structure sharing must in fact be forbidden:

(35)   Object Control without SS (Polish only)

Whenever a deep NP complement of X is co-indexed with the surface subject of a VP complement of X, the NP and the VP's subject are different elements.

$$\begin{bmatrix} word \\ \text{ARG-ST} \mid \text{COMPS} \quad \left\langle \begin{array}{l} \boxed{1}\text{NP}[\ \text{CONT}\ \boxed{0}\ ], \\ \text{VP}\Big[\text{DEPS} \mid \text{SUBJ}\ \langle\boxed{2}[\ \text{CONT}\ \boxed{0}\ ]\rangle\Big] \end{array} \right\rangle \end{bmatrix} \rightarrow \boxed{1} \neq \boxed{2}$$

3.3.4   OC in Czech

As discussed in §2.2, Czech OC data are considerably more complex than in Polish: case transmission is forbidden with dative controllers, just as in Polish, but it is optional with accusative controllers, unless the accusative controller is a numeral phrase, where case transmission is actually obligatory.

For the dative controller case, a principle similar to (35), but appropriately restricted to *dative* controllers, needs to be present in the grammar of Czech:

(36)   Dative Object Control without SS (Czech only)

Whenever a dative NP as a deep complement of X is co-indexed with the surface subject of a VP complement of X, the NP and the VP's subject are different elements.

$$\begin{bmatrix} word \\ \\ \text{ARG-ST} \mid \text{COMPS} \quad \left\langle \begin{array}{l} \boxed{1}\text{NP}\begin{bmatrix} \text{CASE}\ dat \\ \text{CONT}\ \boxed{0} \end{bmatrix}, \\ \\ \text{VP}\Big[\text{DEPS} \mid \text{SUBJ}\ \langle\boxed{2}[\ \text{CONT}\ \boxed{0}\ ]\rangle\Big] \end{array} \right\rangle \end{bmatrix} \rightarrow \boxed{1} \neq \boxed{2}$$

Similarly, for accusative numeral objects, structure sharing may be enforced in a way similar to enforcing the obligatory structure sharing in SC:

(37)   Accusative Numeral Object Control with SS (Czech only)

Whenever an accusative numeral NP as a deep complement of X is co-indexed with the surface subject of a VP complement of X, the NP and the VP's subject are the same element.

$$\begin{bmatrix} word \\ \\ \\ \text{ARG-ST} \mid \text{COMPS} \quad \left\langle \begin{array}{l} \boxed{1}\text{NP}\begin{bmatrix} numeral \\ \text{CASE}\ acc \\ \text{INDEX}\ \boxed{0} \end{bmatrix}, \\ \\ \text{VP}\Big[\text{DEPS} \mid \text{SUBJ}\ \langle\boxed{2}[\ \text{INDEX}\ \boxed{0}\ ]\rangle\Big] \end{array} \right\rangle \end{bmatrix} \rightarrow \boxed{1} = \boxed{2}$$

Finally, in order to account for the optionality of case 'transmission' with other accusative controllers, we need to say... nothing! As discussed in §3.3.2, such SS / no SS ambiguity follows from the independently needed HPSG accounts of control and case marking.

3.3.5 Summary

The parochial principles which had to be added to the general HPSG analyses of control and case assignment in order to account for the distribution of case transmission in Czech and Polish control constructions are summarised in Table 5.

verb	controller	Czech	Polish
SC		*force SS* (34)	
OC	**dat**	*disallow SS* (36)	*disallow SS* (35)
	acc NP		
	acc №	*force SS* (37)	

Table 5: Parochial principles for Czech and Polish

3.4 Further Considerations

So far, we have assumed that anchors should be characterised in terms of the 'deep' argument structure (ARG-ST), while the raised or controlled elements are 'surface' (DEPS) subjects. Subsections 3.4.1–3.4.2 justify this position. Then, §3.4.3 briefly discusses the role of the HPSG Raising Principle in the current account, while the final subsection of this section, §3.4.4, considers — on the basis of apparently marginally acceptable Czech examples of raising without case transmission — the possibility of even further dissociation between raising and SS.

3.4.1 Deep Anchors

It is known cross-linguistically that, e.g., *object* in 'object control' or 'raising to object' should be understood as *deep object* (cf., e.g., Růžička 1999). For example, in (38), the anchor is the subject of a passive verb.

(38) a. John was expected to help Mary. (OR; E.)

      b. John was ordered to help Mary. (OC; E.)

Similar examples can be found in Czech:

(39) a. Šéf   nutil Marii   vstávat   brzy.   (OC; Cz.)
          boss-NOM urged Marie-ACC get up-INF early
          'The boss urged Marie to get up early.'

      b. Marie   byla nucena vstávat   brzy.   (OC; Cz.)
          Marie-NOM was urged   get up-INF early
          'Marie was urged to get up early.'

Also in Polish, there is at least one OC verb which involves an accusative controller and which can be passivised:

(40) a. Janek   nauczył dziecko myć   ręce przed jedzeniem.   (OC; P.)
          Janek-NOM taught   child-ACC wash-INF hands before eating

‘Janek taught the child to wash hands before eating.’

   b.   Dziecko   zostało nauczone myć   ręce   przed jedzeniem.   (OC; P.)
        child-NOM was   taught   wash-INF hands before eating
        ‘The child was taught to wash hands before eating.’

Now, if SUBJ and COMPS in the Subject Control with SS principle (34) referred to 'surface' subject and complements, then this principle would enforce the SS interpretation in the passive examples (39b) and (40b). However, as (41) below shows, such constructions do not involve structure sharing in Polish, case transmission is not observed, while, as (42b) illustrates, similar constructions in Czech involve optional structure sharing, with some preference for structure sharing, just as their active voice counterparts do (cf. (14)–(15) in §2.2.2).

(41)   a.   Dziecko   zostało nauczone być   miłym / ?*miłe.   (OC; P.)
            child-NOM was   taught   be-INF nice-INS   nice-NOM
            ‘The child was taught to be nice.’

       b.   Pięcioro dzieci   zostało nauczone być   miłymi / *miłych   / ?*miłe.
            five-ACC children-GEN was   taught   be-INF nice-INS   nice-GEN   nice-ACC
                                                                                    (OC; P.)
            ‘Five children were taught to be nice.’

(42)   a.   Marie   byla nucena vstávat   nevyspalá.   (OC; Cz.)
            Marie-NOM was urged   get up-INF sleepy-NOM
            ‘Marie was urged to get up without proper sleep.’

       b.   Sedm   trpaslíků   bylo nuceno vstávat   ?nevyspalí / nevyspalých. (OC; Cz.)
            seven-NOM dwarfs-GEN were urged   get up-INF sleepy-NOM   sleepy-GEN
            ‘The seven dwarfs were urged to get up without proper sleep.’

This justifies our analysis of anchors in terms of 'deep' arguments.

3.4.2 Surface Controllees

Similarly, it can be shown that the elements targeted in raising and control are 'surface', not 'deep', subjects. Consider, e.g., the following two sentences.

(43)   Janek   chciał zapamiętać   Tomka.   (SC; P.)
       Janek-NOM wanted remember-INF Tomek-ACC
       ‘Janek wanted to remember Tomek.’

(44)   Tomek   chciał zostać   zapamiętany przez Janka.   (SC; P.)
       Tomek-ACC wanted become-INF remembered by   Janek
       ‘Tomek wanted to be remembered by Janek.’

Example (43) shows that the element controlled by the subject of *chciał* 'wanted' is a subject, but it does not make it clear, whether it is the 'deep' subject or the 'surface' subject. However, in (44) the controller is understood as co-referential with the 'deep' object, i.e., as the 'surface' subject of the passive participle *zapamiętany*. This minimally justifies the claim that the raising auxiliary *zostać* targets the 'surface' subject of the passive participle and, in the absence of arguments to the contrary, we can (defeasibly) infer that raising predicates target 'surface' subjects.

Similarly, constructions such as the following, involving the object control of subjects of active and passive participles, show that at least in some control constructions, it is clearly the 'surface' subject that is the controllee.

19

(45)   Pamiętam   rodzinę   porzucającą   go.   (P.)
              remember-1.SG family-ACC abandoning-ACC him-ACC
              'I remember the family abandoning him.'

(46)   Pamiętam   go   porzuconego   przez rodzinę.   (P.)
              remember-1.SG him-ACC abandoned-ACC by   family
              'I remember him abandoned by the family.'

Thus, we conclude that the targets of raising and control should be characterised in terms of the 'surface' argument structure, i.e., in terms of the value of DEPS.

3.4.3   Raising Principle

Let us reconsider the Raising Principle quoted in §3.1 above.

> In every lexical entry E in which an argument is structure shared with another argument's subject, i.e., in every lexical entry E of the form
>
> $$\begin{bmatrix} word \\ \text{ARG-ST} \quad \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle ... [\, \text{DEPS} \,|\, \text{SUBJ} \; \boxed{1} \,] \,...\rangle \end{bmatrix} \end{bmatrix}$$
>
> or
>
> $$\begin{bmatrix} word \\ \text{ARG-ST} \quad \begin{bmatrix} \text{COMPS} & \langle ... \boxed{1} ... [\, \text{DEPS} \,|\, \text{SUBJ} \; \boxed{1} \,] \,...\rangle \end{bmatrix} \end{bmatrix},$$
>
> the element $\boxed{1}$ is assigned no semantic role in E (i.e., the CONT value of $\boxed{1}$ is not a semantic argument of the predicate expressed in the CONT value of E).[24]

Structure sharing in control environments, pivotal for the analysis presented in this paper, seems to violate this principle: the element which occurs simultaneously on the higher ARG-ST and on the lower DEPS|SUBJ does have a semantic role assigned by the higher verb, apparently contrary to the Raising Principle.

However, it should be noted that the Raising Principle is formulated as a constraint on lexical entries (i.e., on descriptions), not on linguistic entities — it constrains the way that words can be described in the lexicon. This is the only principle of this sort given in Pollard and Sag 1994, it is not formalisable in RSRL (Richter 2000), the most comprehensive mathematical formalism for HPSG, and so it has been criticised by some HPSG practitioners. But it is exactly because of that lexical nature of the Raising Principle that the present account does not violate it: lexical entries for control and raising verbs assumed in this analysis (discussed in §3.1) are the kinds of lexical entries usually given for control and raising verbs, i.e., lexical entries for control verbs do not specify full SS, while lexical entries for raising verbs, which do specify full SS, do not assign a semantic role to the raised argument. To put it succinctly, according to the account proposed here, control verbs receive the SS interpretation outside the realm of the lexicon, i.e., outside the scope of the Raising Principle.

3.4.4   Raising *without* Structure Sharing?

The account presented above is based on the observation that, while raising always involves full structure sharing, including the sharing of case values, control in principle may or may not involve structure sharing of morphosyntactic information. That is, although control constructions may or may not exhibit case transmission, raising always involves case transmission.

---

[24]In fact, this principle was originally formulated in terms of a single attribute SUBCAT, whose value is the list of all arguments, subjects and complements alike.

From this perspective, it is worrisome that, in the survey mentioned above, some speakers of Czech (about 12%) accepted the following raising construction without case transmission:[25]

(47) ??Bylo vidět dost   lidí   vracet se   od   okénka   nespokojení.
      was   seen   enough-NOM   people-GEN   return-INF from   counter dissatisfied-NOM
                                                                          (OR, no SS; Cz.)

      'One could see quite a few people return from the counter dissatisfied.'

The acceptability of nominative predicatives seems to improve with the distance between object and predicative, or when the infinitive is copula, as in (48).

(48) ??Marie viděla Honzu   být   střízlivý.   (OR, no SS; Cz.)
      Marie saw   Honza-ACC be-INF   sober-NOM.MASC
      'Mary saw John being sober.'

According to Hudson 2003, cases of OR without case transmission exist also in Ancient Greek. Example (49), where SS is optional, originally appears in Lacarme 1978, p. 107.

(49) sunoidá   soi   eu   poiésanti   / poiésanta   (OR, SS / no SS; A.G.)
      know-1.SG you-DAT   well having-done-DAT   having-done-ACC
      'I know you have done well.'

In both cases, the status of such data is not clear. Ideally, psycholinguistic experiments should be carried out to determine the acceptability status of Czech sentences such as (47)–(48). However, assuming that such sentences should be permitted by the grammar, how can we account for them?

The most obvious solution, namely, treating raising verbs in the lexicon just the same way as control verbs and then perhaps forcing structure sharing (and, hence, case transmission) whenever it occurs via principles similar to those in §3.3 is technically problematic on standard HPSG assumptions. According to those assumptions, control involves co-indexation, i.e., structure sharing of values of the attribute INDEX, which is part of CONT values of referential nominal expressions. A corollary of this position is that control is a relation between two referential nominal expressions. However, raising may involve other kinds of arguments, e.g., expletives or clausal subjects. That means that, if raising verbs had lexical entries just like those of control verbs, i.e., involving structure sharing of indices, then only raising of referential nominal phrases would be allowed, contrary to facts.

Note, however, that in an effort to simplify the background assumptions for the benefit of a reader less familiar with HPSG, we assumed throughout the paper that co-indexation involves structure sharing of complete CONT values, regardless of whether they contain the nominal INDEX attribute or not. Given that the traditional HPSG approach to semantics has repeatedly been questioned in recent literature and a number of alternative approaches have been proposed (Nerbonne 1992, Copestake *et al.* 1997, Richter and Sailer 1999b,a, 2001, 2003), this approach might turn out to be viable. One empirical question that must be answered first, though, is what are the exact conditions on the violation of SS in raising constructions.

We leave this issue for further research.

4 Conclusion

Czech and Polish overwhelmingly confirm Hudson's (1998, 2003) claim that it is an empirical issue whether control structures involve structure sharing or not. Polish is a particularly simple case: subject control always involves SS, while object control never does. Czech shows the fuller range of possibilities: obligatory SS in SC, and obligatory, optional or forbidden SS in various types of OC constructions.

---

[25]Recall that, in Czech, it is the genitive predicates that agree with nominative and accusative numeral phrases, cf. fn. 6, and for this reason we may speak of lack of case transmission in (47), where both the numeral phrase and the predicative adjective are in the nominative!

The account presented above is based on the standard HPSG assumptions about the difference between control and raising, and on an earlier HPSG analysis of syntactic case assignment. It turns out that no modifications of these background assumptions were necessary, with the exception of relegating HPSG's assumption that controllees are anaphors to the status of a parochial principle for English, and that the analysis of complex case transmission facts in Czech and Polish consists in a small number of simple constraints on the distribution of structure sharing of arguments in control constructions. The analysis is empirically superior, both, to the standard P&P account of control in terms of PRO and Theta criterion, which is at odds with case transmission facts witnessed in some control constructions, and to the more recent Relational Grammar and Minimalist analyses of Dziwirek 1994, 1998 and Hornstein 1999, which incorrectly predict case transmission in all obligatory control constructions.

It should be noted that this type of analysis has been made possible because of two crucial traits of HPSG. First, on the linguistic side, raising constructions have the same constituent structure as corresponding control constructions. This should be contrasted with the widely different structures of object control and ECM (here treated as raising to object) verbs in standard P&P. Second, on the formal side, HPSG is a constraint-based formalism, where any structure not explicitly forbidden by a grammatical principle is licensed. Again, this should be contrasted with formalisms in which structures are licensed via explicit structure-building operations. It remains to be seen to what extent the intuitions embodied in this analysis can be accommodated by other syntactic theories.

References

Avgustinova, T. (2001). Arguments, grammatical relations, and diathetic paradigm. In Flickinger and Kathol (2001), pages 23–42.

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

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. (1986). *Barriers*, volume 13 of *Linguistic Inquiry Monographs*. The MIT Press, Cambridge, MA.

Chomsky, N. (1995). *The Minimalist Program*, volume 28 of *Current Studies in Linguistics*. The MIT Press, Cambridge, MA.

Chomsky, N. and Lasnik, H. (1995). Principles and parameters. In Chomsky (1995), pages 13–127.

Copestake, A., Flickinger, D., and Sag, I. A. (1997). Minimal recursion semantics: An introduction. Ms., Stanford University, available from ftp://csli-ftp.stanford.edu/linguistics/sag/mrs.ps.

Culicover, P. W. and Jackendoff, R. (2001). Control is not movement. *Linguistic Inquiry*, **32**, 493–512.

Daneš, F., Grepl, M., and Hlavsa, Z., editors (1987). *Mluvnice češtiny*, volume 3. Academia, Praha.

Dziwirek, K. (1994). *Polish Subjects*. Garland, New York.

Dziwirek, K. (1998). Reduced construction in universal grammar: Evidence from the Polish object control construction. *Natural Language and Linguistic Theory*, **16**, 53–99.

Filip, H. (2001). The semantics of case in Russian secondary predication. In R. Hastings, B. Jackson, and Z. Zvolenszky, editors, *Proceedings of Semantics and Linguistic Theory (SALT) XI*, pages 192–211, Ithaca, NY. Department of Linguistics, Cornell University, CLC Publications.

Flickinger, D. and Kathol, A., editors (2001). *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*. CSLI Publications, Stanford, CA.

Franks, S. (1995). *Parameters of Slavic Morphosyntax*. Oxford University Press, New York.

Franks, S. (1998). *Parameters of Slavic Morphosyntax* revisited: A minimalist perspective. In Ž. Bošković, S. Franks, and W. Snyder, editors, *Annual Workshop on Formal Approaches to Slavic Linguistics: The Connecticut Meeting 1997*, volume 43 of *Michigan Slavic Materials*, pages 134–165, Ann Arbor. Michigan Slavic Publications.

Haegeman, L. (1991). *Introduction to Government and Binding Theory*. Blackwell Textbooks in Linguistics. Blackwell, Oxford, first edition.

Hornstein, N. (1999). Movement and control. *Linguistic Inquiry*, **30**, 69–96.

Hudson, R. (1998). Functional control with and without structure-sharing. In A. Siewierska and J. J. Song, editors, *Case, Typology and Grammar*, volume 38 of *Typological Studies in Language*, pages 151–169. Benjamins, Amsterdam.

Hudson, R. (2003). Case agreement, PRO and structure sharing. *Research in Language*, **1**. Forthcoming.

Ionin, T. and Matushansky, O. (2002). Encasing the time: Temporal effects on Russian predicative case. NELS 33, MIT.

Lacarme, J. (1978). *Aspects syntaxiques des complétives du Grec*. Ph. D. dissertation, University of Montreal, Montreal.

Manning, C. D. and Sag, I. A. (1998). Argument structure, valence, and binding. *Nordic Journal of Linguistics*, **21**(2), 107–144.

Manning, C. D. and Sag, I. A. (1999). Dissociations between argument structure and grammatical relations. In Webelhuth *et al.* (1999), pages 63–78.

Nerbonne, J. (1992). Constraint-based semantics. In P. Dekker and M. Stokhof, editors, *Proceedings of the Eighth Amsterdam Colloquium*, pages 425–444, Amsterdam. Institute for Logic, Language and Information.

Nichols, J. (1981). *Predicate Nominals: A Partial Surface Syntax of Russian*. University of California Press, Berkeley.

Pisarkowa, K. (1965). *Predykatywność określeń w polskim zdaniu*. Zakład narodowy im. Ossolińskich, Wrocław.

Pollard, C. and Sag, I. A. (1987). *Information-Based Syntax and Semantics, Volume 1: Fundamentals*. Number 13 in CSLI Lecture Notes. CSLI Publications, Stanford, CA.

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

Postal, P. M. (1974). *On Raising*. Massachusetts Institute of Technology, Cambridge, MA.

Przepiórkowski, A. (1999a). *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph. D. dissertation, Universität Tübingen.

Przepiórkowski, A. (1999b). On case assignment and 'adjuncts as complements'. In Webelhuth *et al.* (1999), pages 231–245.

Przepiórkowski, A. (2000). Predicative case agreement with Quantifier Phrases in Polish. In A. Okrent and J. Boyle, editors, *The Proceedings from the Main Session of the Chicago Linguistic Society's Thirty-sixth Meeting*, volume 36-1, pages 343–354. Chicago Linguistic Society, Chicago, IL.

Przepiórkowski, A. (2001). ARG-ST on phrases: Evidence from Polish. In Flickinger and Kathol (2001), pages 267–284.

Przepiórkowski, A. (2004a). O wartości przypadka podmiotów liczebnikowych. *Biuletyn Polskiego Towarzystwa Językoznawczego*, **LX**. Forthcoming.

Przepiórkowski, A. (2004b). On case transmission in Polish control and raising constructions. *Papers and Studies in Contrastive Linguistics*. Forthcoming.

Przepiórkowski, A. (2004c). Towards the design of a syntactico-semantic lexicon for Polish. In *Proceedings of IIS:IIPWM 2004*.

Richter, F. (2000). *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Ph. D. dissertation, Universität Tübingen.

Richter, F. and Sailer, M. (1999a). LF conditions on expressions of Ty2: An HPSG analysis of Negative Concord in Polish. In R. D. Borsley and A. Przepiórkowski, editors, *Slavic in Head-Driven Phrase*

Structure Grammar, pages 247–282. CSLI Publications, Stanford, CA.

Richter, F. and Sailer, M. (1999b). Underspecified semantics in HPSG. In H. C. Bunt and R. Muskens, editors, *Computing Meaning*, pages 95–112. Kluwer, Dordrecht.

Richter, F. and Sailer, M. (2001). Polish negation and Lexical Resource Semantics. *Electronic Notes in Theoretical Computer Science*, **53**.

Richter, F. and Sailer, M. (2003). Basic concepts of Lexical Resource Semantics. Ms., Universität Tübingen.

Růžička, R. (1999). *Control in Grammar and Pragmatics: A cross-linguistic study*. Benjamins, Amsterdam.

Skoumalová, H. (2002). Verb frames extracted from dictionaries. *The Prague Bulletin of Mathematical Linguistics*, **77**, 19–62.

Timberlake, A. (1988). Case agreement in Lithuanian. In M. Barlow and C. A. Ferguson, editors, *Agreement in Natural Language: Approaches, Theories, Descriptions*, pages 181–199. CSLI Publications, Stanford, CA.

Toman, J. (1991). Anaphors in binary trees: an analysis of Czech reflexives. In J. Koster and E. Reuland, editors, *Long-Distance Anaphora*, pages 151–170. Cambridge University Press, Cambridge.

Webelhuth, G., Koenig, J.-P., and Kathol, A., editors (1999). *Lexical and Constructional Aspects of Linguistic Explanation*. CSLI Publications, Stanford, CA.

Zabrocki, T. (1981). *Lexical Rules of Semantic Interpretation: Control and NP Movement in English and Polish*. Number 14 in Filologia Angielska. Adam Mickiewicz University, Poznań.

Authors' addresses:

Adam Przepiórkowski
Instytut Podstaw Informatyki PAN
ul. Ordona 21
01-237 Warszawa
Poland

tel.:   (+48 22) 8362841
fax: (+48 22) 8376564
e-mail: `adamp@ipipan.waw.pl`
URL: `http://dach.ipipan.waw.pl/~adamp/`

Alexandr Rosen
Institute of Theoretical and Computational Linguistics
Faculty of Arts, Charles University
Celetná 13
110 00 Praha 1
Czech Republic

tel.:   (+420) 224491858
fax: (+420) 224491875
e-mail: `alexandr.rosen@ff.cuni.cz`
URL: `http://utkl.ff.cuni.cz/~rosen`