Adam Przepiórkowski, Institute of Computer Science, Polish Academy of Sciences, Warsaw
Alexandr Rosen, Institute of Theoretical and Computational Linguistics, Charles University, Prague

# On the Case of Predicative Complements in Czech Infinitival Clauses[1]

## 1 Introduction

To the ear of a Czech speaker, examples in (1) sound perfectly acceptable:

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

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

Yet their theoretical account in some more prominent modern syntactic theories, including Principles and Parameters (P&P; Chomsky 1995 and references therein) and Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag 1987, 1994), seems to be problematic.

Both theories assume a clear distinction between raising to subject (SR) and so-called Exceptional Case Marking (ECM; here called raising to object, or OR) constructions, as in (2), on the one hand, and subject control (SC) and object control (OC) constructions, as in (3) and (1) above, on the other hand.

(2) a. John seems to like Mary. (SR)

b. Honza přestal mít rád Marii. (SR)
Honza-NOM stopped have-INF glad Marie-ACC
'Honza stopped liking Maria.'

c. I expect John to help Mary. (OR)

d. Viděl jsem Honzu líbat Marii. (OR)
seen AUX-1SG Honza-ACC kiss-INF Marie-ACC
'I saw Honza kissing Marie.'

(3) a. John tries to help Mary. (SC)

b. Honza zkouší pomáhat Marii. (SC)
Honza-NOM tries help-INF Marie-ACC
'Honza tries to help Marie.'

c. I ordered John to help Mary. (OC)

d. Nařídil jsem Honzovi pomáhat Marii. (OC)
ordered AUX-1SG Honza-DAT help-INF Marie-ACC
'I ordered Honza to help Marie'

[1]The authors would like to thank the audience at FDSL-5 for helpful comments. The work was partially supported by the Grant Agency of the Czech Republic, grant no. 405/03/0913, and by the governments of Poland and the Czech Republic within the programme KONTAKT (Czech reg. no. 23/2004, Polish no. 25/2004/CZ). See Przepiórkowski and Rosen 2004 for an extended version of this paper, based on parallel Czech and Polish data.

Two of the most robust cross-linguistic tests distinguishing raising and control involve passivisation (e.g., Pollard and Sag 1994 and, for Czech, Skoumalová 2002) and idiom chunks (e.g., Postal 1974): 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., (4)–(5), and ii) chunks of sentential idioms can be raised arguments, but they cannot be controllers, e.g., (6)–(7).

(4) a. Mary seems to be liked by John. (SR; $\approx$(2a))

b. I expect Mary to be liked by John. (OR; $\approx$(2c))

(5) a. Mary tries to be liked by John. (SC; $\not\approx$(3a))

b. I ordered Mary to be liked by John. (OC; $\not\approx$(3c))

(6) a. The cat seems to be out of the bag. (SR)

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

(7) a. *The cat tries to be out of the bag. (SC)

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

In both P&P and HPSG, two systematic differences between raising and control are postulated: i) semantically, raising verbs have one argument fewer than the corresponding control verbs, e.g., *seem* has one semantic argument, while *try* has two; 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 controllee (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 $\alpha$ 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.

The reason why examples in (1) are problematic is that this correlation is missing: the controllee's case is shown overtly as the case of the predicative complement, agreeing obligatorily with its subject (the controllee). Thus, the controller seems to share its case value with that of the controllee, contrary to the expectation that the controllee receives its case independently.

In the following, we use Czech data to show that the correlation between semantic and structural properties of raising and control does not hold and, thus, refute the Theta criterion of P&P. In doing this, we follow Hudson 1998, 2003, who postulates a similar dissociation of raising/control and SS on the basis of Icelandic, Ancient Greek, and — in the latter article — Russian, and Przepiórkowski 1999, 2004b, who confirms Hudson's conclusions using Polish data and provides an analysis in HPSG.

The next section, §2, presents the relevant case transmission data from Czech. Section 3 contains an HPSG analysis of the Czech data in terms of dissociating SS from raising. Finally, §4 summarises the main conclusions.

## 2 The data

### 2.1 Data on Subject Control/Raising

Consider the case transmission examples (8)–(9) (example (9b) is repeated from (1a)).

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

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

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

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

Both in SR examples, (8), and in SC examples, (9), 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) agreement with numeral subjects in examples (8b) and (9b).[2]

Examples (8) 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 (9) 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.

Very similar case transmission data in SR/OR constructions were previously discussed for Polish by Franks 1995 and Przepiórkowski 1999, 2004b, the only significant distinction being the availability of accusative case with numeral subjects (10).[3]

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

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

In order to analyse the Polish data, Franks 1995 argues that PRO must bear case, contrary to P&P assumptions, and sketches a few possible solutions, noting their various drawbacks.

[2]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 2000).

(i) Pět poslanců bylo neklidných. (P.)
five-ACC MPs-GEN were uneasy-GEN
'Five MPs were uneasy.'

[3]For the arguments that Polish numeral subjects are in fact accusative, see, e.g., Przepiórkowski 1999, 2004a. For the purpose of this article, we assume that numeral subjects are accusative also in Czech, although for Czech the arguments are less convincing.

Przepiórkowski 1999, 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. On the other hand, such SS must be forbidden in OC, as there is no case transmission in Polish OC and the predicative adjective occurs in the 'default' instrumental case, so-called 'instrumental of predication'.

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.

### 2.2 Data on Object Control/Raising

Czech OC constructions 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 the instrumental acceptable to some extent only with copula, as in (11).

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

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

The nominative as the non-agreeing predicative case is present also in Czech OC with dative objects (12) — there is no case transmission.[4]

(12) a. Marie nařídila Honzovi přijít střízlivý / *střízlivému. (OC)
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)
'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 (13), the case of the predicative complement can be either nominative (no case transmission, as in the dative object case) or accusative (case transmission).[5]

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

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

[4]Examples (12) with predicative complements in the dative are acceptable when predicative complements are understood as complementing the higher verb.

[5]Similarly as examples (12) above, examples (13) in the version with the predicative complement in the accusative can have the meaning 'While he was sober, she taught him to come home'. Nevertheless, the glossed reading is the preferred option. Example (13b) is repeated from (1b).

However, when the object is a numeral NP, as in (14), the preferred case of the predicative complement seems to be genitive, i.e., case transmission takes place.

(14) a. Marie přiměla pět pacientů přijít svlečených / Marie made five-ACC patients-GEN come-INF undressed-GEN ??svlečení / ??svlečené. (OC) undressed-NOM 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) 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,[6] seem to behave as expected, i.e., case transmission takes place.

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

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

As the tentative acceptability ratings of the examples above show, more representative and reliable data are needed, especially for 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.[7] Visitors of the site were asked to rate 26 Czech sentences, exemplifying the presence or absence of SS by means of the presence or absence of case transmission from the matrix object to the predicative complement of the embedded infinitive. Table 1 gives the number of sentences in the survey listed according to the type of the matrix verb (control or raising), the case of the *anchor* (the controller or the raised element)[8] and the absence or presence of SS (as shown by the case of the predicative complement).

Two kinds of accusative controllers were tested: plain and numeral NPs, in order to verify the hypothesis that the latter support SS 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).

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 SS. Such factors may include: the linear distance between the matrix object and the predicative complement (the acceptability

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

[7] The second co-author wishes to thank Hana Skoumalová and Johanka Doležalová for gladly provided technical assistance.

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

| | 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

of the non-sharing option increases with longer distance), lexical setting (e.g., infinitival copula tends to make non-sharing option more acceptable), idiomatic nature of the embedded predicative, verbal aspect of both verbs, and the often reported difficulty of judgment due to stylistic preference for alternative syntactic structures (finite clauses instead of infinitival clauses, adverbs instead of predicative adjectives).

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 for multiple examples of the same type.

| 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

The results of the survey support the following conclusions:

1. For ordinary accusative NPs as OC objects, SS is slightly preferred, as in (13).
2. For numeral accusative NPs as OC objects, SS is strongly preferred, cf. (14).
3. In OR, SS is the strongly preferred option, cf. (15).

### 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 by Polish and Czech.[9] Czech data are presented in more detail in Table 4 in parallel with comparable Polish data. The comparison shows the striking difference between SC and SR constructions on the one

[9] *$\theta$-marked* anchors are controllers, non-*$\theta$-marked* anchors are raised elements.

hand, where Czech and Polish behave in a similar way (except for the fact that with numeral subjects Polish allows for two or even three options), and OC and OR constructions on the other hand, where OC constructions with the predicative complement in the instrumental are the only possibility in Polish. In current Polish, there are no OR/ECM verbs.

| Anchor | | Structure (Structure Sharing or PRO) | | | | |
|---|---|---|---|---|---|---|
| $\theta$-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

| **Verb** | **Anchor** | | **Czech pred.** | | | **Polish pred.** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **subj** | **obj** | nom | gen | acc | nom | gen | acc | ins |
| **SC** | nom | | ● | | | ● | | | |
| | № | | | ● | | | ● | ● | |
| **SR** | nom | | ● | | | ● | | | |
| | № | | | ● | | | ● | ● | |
| **OC** | | dat | ● | | | | | | |
| | | acc | ○ | | ● | | | | ● |
| | | acc № | | ● | | | | | |
| **OR** | | acc | | | ● | | | | |
| | | acc № | | ● | | | | | |

● OK ○ acceptable № numeral phrase

● structure sharing ● unexpected structure sharing

Table 4: Czech and Polish raising/control in detail

## 3 Analysis in HPSG

The analysis of case transmission in Czech reflects the following two observations made on the basis of the data considered above, as well as the data presented in Timberlake 1988 and Hudson 1998, 2003: (i) in many languages (Icelandic, Ancient Greek, Lithuanian, Polish, Russian, Czech…) there are cases of semantic control involving case transmission, i.e., involving SS; and (ii] 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 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 case transmission, §3.3. The final subsection, §3.4, contains a discussion of the interaction of HPSG Raising Principle with the analysis proposed.

### 3.1 Control and Raising in HPSG

In standard HPSG, raising verbs are assumed to have lexical entries like the one schematically presented in (16).[10]

(16) A Subject Raising verb (e.g., Czech *začít* 'begin')

$$\begin{bmatrix} \textit{word} & \\ \text{ARG-ST} & \begin{bmatrix} \text{SUBJ} \langle \boxed{0} \rangle & \\ \text{COMPS} \langle \text{VP} \begin{bmatrix} \text{DEPS} \mid \text{SUBJ} \langle \boxed{0} \rangle \\ \text{CONT} \ \boxed{1} \end{bmatrix} \rangle & \end{bmatrix} \\ \text{CONT} & \mathbf{P}(\boxed{1}) \end{bmatrix}$$

In (16) 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 surface subject of the VP (0) is structure shared with the deep subject of the SR verb. The boxed numbers (e.g., 1) are variables whose values are feature structures; multiple occurrences of the same variable in a feature structure denote structure sharing.

In contrast to raising verbs, control verbs have lexical entries like (17).

(17) A Subject Control verb (e.g., Czech *zkusit*, 'try')

$$\begin{bmatrix} \textit{word} & \\ \text{ARG-ST} & \begin{bmatrix} \text{SUBJ} \langle \text{NP} \begin{bmatrix} \text{CONT} \ \boxed{1} \end{bmatrix} \rangle & \\ \text{COMPS} \langle \text{VP} \begin{bmatrix} \text{DEPS} \mid \text{SUBJ} \langle [\ \text{CONT} \ \boxed{1}\ ] \rangle \\ \text{CONT} \ \boxed{2} \end{bmatrix} \rangle & \end{bmatrix} \\ \text{CONT} & \mathbf{P}(\boxed{1},\boxed{2}) \end{bmatrix}$$

A SC verb, as in (17), is a (semantically) 2-argument verb: the first semantic argument (i.e., 1 in **P**(1,2)) is the content of its deep subject (1), while the second semantic argument is the content of the infinitival VP (cf. 2). The lexical entry does not specify full SS between the deep

[10] The attributes ARG-ST and DEPS reflect the assumption that there are at least two levels of argument structure: a 'deep' argument structure, constant for all forms of a given lexeme, and a 'surface' argument structure, which depends on the voice of a given word form, the two arguments structures being systematically related via general grammatical principles (cf., e.g., Manning and Sag 1998, 1999, Bouma et al. 2001, Avgustinova 2001). Moreover, DEPS is the locus of syntactic case assignment, as well as — at least in languages such as Czech and Polish — binding.

subject of the verb and the surface subject of the VP complement; instead, the two subjects are co-indexed (cf. 1), i.e., they share their semantics.[11] A related difference between raising and control verbs is that a raising verb does not specify the morhosyntactic makeup of its raised argument, it takes whatever subject is required by its VP complement, while a control verb specifies its controller argument as an NP.

Lexical entries of OR and OC verbs are analogous to the lexical entries of SR and SC verbs above, and they display analogous differences. An entry for OR verbs is shown below.

(18) An Object Raising (ECM) verb (e.g., Czech *vidět*, 'see')

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

According to (18), the first semantic argument of an OR verb is the content of its deep subject (cf. 1), while the second semantic argument is the content of the infinitival VP (cf. 2). The surface subject of the VP complement is structure shared with (raised to) the deep object of the OR verb (cf. 0).

Finally, OC verbs have three semantic arguments and their objects must be co-indexed with the subjects of their VP complements. In (19), as in (18), the first semantic argument of the OC verb is the content of its deep subject (cf. 1) and, again as in case of OR verbs, semantic argument of the OC verb is the content of the infinitival VP (cf. 3). However, the surface subject of the VP is only co-indexed (cf. 0) with the deep object of an OC verb and the content of that object (cf. 2) is the second semantic argument of the OC verb:

(19) An Object Control verb (e.g., Czech *přikázat*, 'order')

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

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, 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 parochial principles introduced in §3.3.

On the other hand, in case of raising predicates, the correlation between raising as a phenomenon of syntax (SS of arguments) and semantics (no semantic role assigned to the raised argument) has been ensured by the Raising Principle, schematically quoted below.

[11] 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.

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

$$\left[\begin{array}{ll} \textit{word} & \\ \textsc{arg-st} & \left[\begin{array}{ll} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \ldots [\ \textsc{deps} \mid \textsc{subj}\ \boxed{1}\ ] \ldots \rangle \end{array}\right] \end{array}\right]$$

or

$$\left[\begin{array}{ll} \textit{word} & \\ \textsc{arg-st} & \left[\textsc{comps} \quad \langle \ldots \boxed{1} \ldots [\ \textsc{deps} \mid \textsc{subj}\ \boxed{1}\ ] \ldots \rangle\right] \end{array}\right],$$

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 $E$).[12]

Note that the Raising Principle corresponds to the Theta criterion of P&P:[13]

> 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 the dissociation of raising/control and structure sharing proposed in §3.3 does not violate the Raising Principle.

### 3.2 Case Assignment and Case Agreement in HPSG

We follow the general approach of Przepiórkowski 1999 and assume the structural/lexical case dichotomy and the following principles of structural case assignment:[14]

(20) structural case is assigned (checked) at the level of the surface argument structure (DEPS) of words;

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

(22) assignment of structural case in Czech is guided by the following principles :

a. surface subjects of both finite and non-finite verbs are assigned the nominative;

b. surface complements of verbs are assigned the accusative;

Thus, if an argument occurs on several argument structures, its case is checked only on the highest of these. This way there is no case clash in, e.g., OR, where the lower subject is raised to the higher object position: it is assigned case according to its highest position, i.e., it receives the accusative (cf. (22b)), and no attempt is made to assign the nominative to its lower 'copy' (cf. (21)). On the other hand, in cases of object control, where only the INDEX value is structure-shared, the controller and the controlled unexpressed subject are two different arguments, so

[12] 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.

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

[14] See Przepiórkowski 1999 for formalisation and discussion.

they are both assigned case: the object receives the accusative (cf. (22b)), and the controlled subject receives the nominative (cf. (22a)).

Furthermore, we assume that the noun phrase modified by a predicative complement is subcategorized for by the complement as its subject. The modified noun phrase is also an argument of a verb, which serves as the head of the predicative complement. If this verb is non-finite, the argument can be controlled or raised. Then the subject of the predicative complement is overtly realized higher up, possibly in the matrix clause.

A predicative complement may share morphological case with its subject, or may be assigned the nominative, which is captured by the following implicational principle:[15]

$$(23) \qquad \begin{bmatrix} \text{SUBJ} & \langle \text{XP}[\text{CASE } \boxed{1}] \rangle \\ \text{HEAD} & \begin{bmatrix} \text{PRD} & + \\ \text{CASE} & \boxed{2} \end{bmatrix} \end{bmatrix} \rightarrow (\boxed{1} = \boxed{2}) \vee (\boxed{2} = \textit{ins})$$

According to this principle, for any cased ([CASE 2]) predicate ([PRD +]) subcategorising for a cased subject ([ SUBJ ⟨XP[CASE 1]⟩ ]), the case of the subject must agree with the case of the predicate (1 = 2), or else the case of the predicate must be instrumental (2 = *ins*). In Czech, the instrumental option seems to be further restricted, but we refrain from any attempt to formalize this restriction here.

### 3.3 Case Transmission

#### 3.3.1 Subject Raising and Object Raising

Nothing needs to be added to the HPSG assumptions discussed above to account for the case transmission in SR examples such as (8a). In such cases, the subject occurs on three argument structures, i.e., those of the finite raising verb, the infinitive copula and the predicative adjective. Since (21) says that case is assigned on the highest argument structure, the nominative is assigned via (22a). Hence, the subject of the predicative adjective is nominative, and, in order to satisfy (23), the adjective itself must be either nominative or (less likely in Czech) instrumental.

In case of SR examples involving numeral anchors, as in (8b), the reasoning would be essentially the same, although the predicative case marking principle (23) would have to take into account the quirky predicative agreement with numeral phrases; cf. fn. 15.

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 Subject Control

On the other hand, it seems that the above assumptions lead to blatantly wrong predictions in case of SC, which in Czech (as well as in Polish) involves obligatory SS, cf. (9a). The subject of the infinitive 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 infinitive (and, hence, with the subject of the predicate). The problem shows up with numeral subjects (cf. (9b)), which should agree with a predicative complement in the genitive. However, the unrealised subject of

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

the non-finite verb in control constructions is assigned the nominative case, cf. (22a). Then, the predicative complement must be in the nominative (or instrumental), contrary to the data.

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 SS of complete controller and controllee.[16]

So, SC verbs, whose structure is repeated in (17), are in principle ambiguous between a non-SS interpretation, in which $\boxed{3} \neq \boxed{4}$, and a SS interpretation, where $\boxed{3} = \boxed{4}$.

(17) A Subject Control verb

$$\begin{bmatrix} \textit{word} & \\ \text{ARG-ST} & \begin{bmatrix} \text{SUBJ} \langle \boxed{3}\text{NP}\left[\text{CONT } \boxed{1}\right]\rangle \\ \text{COMPS} \langle \text{VP}\begin{bmatrix}\text{DEPS} \mid \text{SUBJ} \langle \boxed{4}[\text{ CONT } \boxed{1}\ ]\rangle \\ \text{CONT } \boxed{2}\end{bmatrix}\rangle \end{bmatrix} \\ \text{CONT} & \mathbf{P}(\boxed{1},\boxed{2}) \end{bmatrix}$$

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 (nominative or instrumental instead of agreement), we face the problem of overgeneration (nominative or instrumental, 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):

(24) 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.*

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

### 3.3.3 Object Control

Czech data concerning OC are considerably more complex than those on SC or SR: case transmission is forbidden with dative controllers, but it is optional with accusative controllers, unless the accusative controller is a numeral phrase, where case transmission is actually obligatory.

For the dative controllers, a Czech-specific principle prohibiting SS must be stipulated:[17]

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

[17] A similar principle without the restriction to dative NPs covers all OC constructions in Polish.

(25) Dative Object Control without SS

*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.*

$$\left[\begin{array}{l}\textit{word}\\ \textsc{arg-st}\,|\,\textsc{comps}\ \left\langle \boxed{1}\mathrm{NP}\left[\begin{array}{l}\textsc{case}\ \textit{dat}\\ \textsc{cont}\ \boxed{0}\end{array}\right],\ \mathrm{VP}\left[\textsc{deps}\,|\,\textsc{subj}\ \langle \boxed{2}[\ \textsc{cont}\ \boxed{0}\ ]\rangle\right]\right\rangle\end{array}\right] \rightarrow \quad \boxed{1} \neq \boxed{2}$$

Similarly, for accusative numeral objects, SS may be enforced as in SC:

(26) Accusative Numeral Object Control with SS

*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.*

$$\left[\begin{array}{l}\textit{word}\\ \textsc{arg-st}\,|\,\textsc{comps}\ \left\langle \boxed{1}\mathrm{NP}\left[\begin{array}{l}\textit{numeral}\\ \textsc{case}\ \textit{acc}\\ \textsc{index}\ \boxed{0}\end{array}\right],\ \mathrm{VP}\left[\textsc{deps}\,|\,\textsc{subj}\ \langle \boxed{2}[\ \textsc{index}\ \boxed{0}\ ]\rangle\right]\right\rangle\end{array}\right] \rightarrow \quad \boxed{1} = \boxed{2}$$

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.

### 3.3.4 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 constructions are summarised in Table 5. Again, Polish is included for comparison.

| **verb** | **controller** | **Czech** | **Polish** |
|---|---|---|---|
| **SC** | | *force SS* (24) | |
| **OC** | **dat** | *disallow SS* (25) | *disallow SS* |
| | **acc** | | |
| | **acc №** | *force SS* (26) | |

Table 5: Parochial principles for Czech and Polish

### 3.4 Raising Principle Revisited

The HPSG Raising Principle quoted in §3.1 seems to be violated by SS in control environments, pivotal for the analysis presented in this paper: 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, as the only principle of this sort given in Pollard and Sag 1994, 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. 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.

## 4 Conclusion

Czech turns out to be one of the languages that confirm Hudson's (1998, 2003) claim that it is an empirical issue whether control structures involve SS or not. In Czech, subject control always involves SS, while in various types of OC constructions SS is obligatory, optional or forbidden.

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. 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. The analysis of complex case transmission facts in Czech consists in a small number of simple constraints on SS of arguments in control constructions.

This type of analysis is 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 is in contrast with the widely different structures of OC and ECM (OR) verbs in 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 are valid and, if so, whether they can be accommodated by other syntactic theories.

## References


Avgustinova, T. (2001). Arguments, grammatical relations, and diathetic paradigm. In Flickinger, D. and Kathol, A., editors, *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, pages 23–42. CSLI Publications, Stanford, CA.

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

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

Franks, S. (1998). *Parameters of Slavic Morphosyntax* revisited: A minimalist perspective. In Bošković, Ž., Franks, S., and Snyder, W., 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.

Hudson, R. (1998). Functional control with and without structure-sharing. In Siewierska, A. and Song, J. J., 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.

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, G., Koenig, J.-P., and Kathol, A., editors, *Lexical and Constructional Aspects of Linguistic Explanation*, pages 63–78. CSLI Publications, Stanford, CA.

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

Przepiórkowski, A. (2000). Predicative case agreement with Quantifier Phrases in Polish. In Okrent, A. and Boyle, J., 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. (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. and Rosen, A. (2004). Czech and Polish raising/control with or without structure sharing. *Research in Language*. Forthcoming.

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 Barlow, M. and Ferguson, C. A., 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 Koster, J. and Reuland, E., editors, *Long-Distance Anaphora*, pages 151–170. Cambridge University Press, Cambridge.