Czech in Generative Grammar

Editors:
Mojmír Dočekal
Petr Karlík
Jana Zmrzlíková

Contents

Petr Biskup: Sentence-final sentence adverbs in the phase model 7

Pavel Caha: A Note About *A Note About Nothing* 17

Markéta Ceplová: Infinitives under 'have'/'be' in Czech 31

Mojmír Dočekal: *Only*, bound variables and VP ellipsis in Czech 47

Jakub Dotlačil: Clitic omission in Czech as across-the-Board extraction 65

Joseph Emonds: Czech Cases and the Syntacticon:
Poznámky k, o, okolo, nad něčím a pro někoho 81

Petr Karlík: Mixed Nominals in Czech 105

Lucie Medová and Tarald Taraldsen: 1, 2, SE 119

Radek Šimík: The Czech invariant demonstrative *to* is a Foc head 139

Hana Skrabalova: Wh-questions with conjoined wh-words 161

Andrea Volencová: Reflexive verbal forms in Czech
from the Romance perspective 175

Markéta Ziková: Why Czech case markers sometimes get lost 193

Preface

This festschrift is the collective work of people who believe that transformational generative grammar is empirically the strongest linguistic theory, as it offers a restrictive and yet sufficiently robust apparatus for the description of any natural language, while offering mechanisms for evaluating each newly arising hypothesis. All the authors are united as well either by their knowledge of the Czech language, or by their use of Czech as a data source for more general understanding. As far as we know, ours is the first work with such an ambition.

But more important, all the authors are united in their wish to give pleasure to the First Lady of Czech linguistics, who introduced modern generative grammar to the Czech Republic, and who taught and still teaches generative grammar to many of us directly or indirectly, our collaborator in the past and the present: Lída Veselovská. Without Lída generative grammar would not have arisen in the Czech Republic, just as it would not have existed in the world without Chomsky's book which gave it birth fifty years ago.

Sentence-final sentence adverbs in the phase model

Petr Biskup

biskup@rz.uni-leipzig.de

1 Introduction

It has been argued that sentence adverbs cannot occur in the sentence-final position, unless they are separated by a comma intonation. See, for example, Jackendoff (1972) for English (1), Belletti (1990) for Italian (2) and French (3), and Alexiadou (1997) for Greek (4).

(1) *Horatio has lost his mind evidently/probably.
(Jackendoff 1972, 50 (3.9))

(2) *Gianni partirà probabilmente.
'Gianni will leave probably'
(Belletti 1990, 53 (53))

(3) *Jean partira probablement.
'Jean will leave probably.'
(Belletti 1990, 53 (53))

(4) *O Janis tha figi pithanos.
The John$_{\text{NOM}}$ FUT go$_{\text{3SG}}$ probably
(Alexiadou 1997, 157 (80))

The following examples show that the same holds for Czech; sentence adverbs such as the epistemic *možná* 'possibly' or *pravděpodobně* 'probably' are not allowed in the sentence-final position.

(5) *Tu knihu pošle Jirkovi Pavel pravděpodobně.
the book$_{\text{ACC}}$ sends Jirka$_{\text{DAT}}$ Pavel$_{\text{NOM}}$ probably
'The book, Pavel will probably send to Jirka.'

(6) *Tu knihu pošle Jirkovi Pavel možná.
the book$_{\text{ACC}}$ sends Jirka$_{\text{DAT}}$ Pavel$_{\text{NOM}}$ possibly
'The book, Pavel will possibly send to Jirka.'

This seems to be in accordance with the claim that sentence adverbs occupy a pretty high position in the clausal structure (Cinque 1999, Laenzlinger 2002). There is also an alternative explanation why the adverbs in (1)-(6) are impossible in the sentence-final position. According to Lang (1979), sentence adverbs are focus sensitive operators and they themselves cannot be focalized. Compare also Hajičová, Partee & Sgall (1998a,b) or Koktová (1987, 1999), who

argue that the prototypical position of focus sensitive adverbs is at the boundary between the background and the focus, and Krifka (1992) or Jacobs (1986, 1988) in Czech who argue that focusing adverbs must c-command their focus.

However, there are sentence adverbs that can occur in the sentence-final position in Czech and Italian; consider examples (7) and (8).

(7) Tu knihu pošle Jirkovi Pavel určitě.
the book_{ACC} sends $\text{Jirka}_{\text{DAT}}$ $\text{Pavel}_{\text{NOM}}$ certainly
'The book, Pavel will certainly send to Jirka.'

(8) Gianni lo merita sicuramente / di sicuro.
'Gianni deserves it surely.'
(Cinque 1999, 180 note 80))

These examples pose a problem for the argument that the source of the ungrammaticality of sentence adverbs in the sentence-final position is the height of their structural position and they show that the ungrammaticality of (1)-(6) is not due to the general impossibility of sentence adverbs to appear in the sentence-final position and be accented there.

Therefore Cinque (1999, 180 note 80) proposes that adverbs like *sicuramente* belong to the class of (realis) mood adverbs that can be used as focusing adverbs - they are heads taking their modifees as complements - and allow their complements to move across them. On the other hand, Lang (1979) distinguishes three classes of German sentence adverbs. Class A adverbs are, for example, *wahrscheinlich* 'probably' or *möglicherweise* 'possibly'; class B adverbs *bedauerlicherweise* 'unfortunately' or *überraschenderweise* 'surprisingly'; and class C adverbs *tatsächlich* 'really' or *in der Tat* 'certainly'. According to him, class A adverbs, that is the adverbs in (5) and (6), differ from class C adverbs, that is, the adverbs in (7) or (8), in that they cannot be accented and focalized.

However, both proposals have a problem with cases like (9) because the sentence-final adverb *možná* 'possibly' is epistemic (that is, Lang's class A), is accented and represents the focus itself; it is associated with the focus-sensitive adverbial *jenom* 'only'. Since non-clausal non-selected adverbials are merged to the left in Czech (see Biskup in prep.), the grammaticality of (9) cannot be accounted for through right adjunction of *možná*.

(9) Tu knihu pošle Jirkovi Pavel jenom možná.
the book_{ACC} sends $\text{Jirka}_{\text{DAT}}$ $\text{Pavel}_{\text{NOM}}$ only possibly
'It is only possible that Pavel will send the book to Jirka.'

The remainder of this paper is organized as follows. In section 2, I will discuss veridicality and a downward-monotonicity approach to sentence adverbs and on the basis of empirical arguments I will argue that these approaches are not appropriate. In section 3, I will argue that sentence adverbs generally can be merged in the vP phase and that the (un)grammaticality of certain sentence adverbs in the sentence-final position depends on the (non-)interpretability of a given adverb in the vP position and not on its syntactic position. I will propose an analysis in a model with a correlation between the phase structure, tripartite quantificational structure and information structure and will argue that sentence adverbs can appear in the sentence-final position only if they represent the extreme value with respect to the set of focus alternatives. Conclusions will be drawn in section 4.

2 Other proposals: veridicality and downward-monotonicity

Let us begin with the difference between the adverbs *určitě* 'certainly' and *možná* 'possibly'. It is easy to show that these adverbs have different lexicosemantic properties. The epistemic adverb *možná* is excluded from contexts licensing negative polarity items. More specifically, it is degraded, for example, under the question operator (10) or the imperative operator (11).

(10) Pošle Pavel (*možná) Jirkovi tu knihu (*možná)?
sends $\text{Pavel}_{\text{NOM}}$ possibly $\text{Jirka}_{\text{DAT}}$ the book_{ACC} possibly
'Will Pavel possibly send the book to Jirka?'

(11) Pošli Jirkovi (*možná) tu knihu (*možná)!
send $\text{Jirka}_{\text{DAT}}$ possibly the book_{ACC} possibly
'Possibly send the book to Jirka.'

In contrast, the (realis) mood adverb *určitě* can appear in these environments, as demonstrated by question (12) and the imperative sentence in (13).

(12) Pošle Pavel (určitě) Jirkovi tu knihu (určitě)?
sends $\text{Pavel}_{\text{NOM}}$ certainly $\text{Jirka}_{\text{DAT}}$ the book_{ACC} certainly
'Will Pavel certainly send the book to Jirka?'

(13) Pošli Jirkovi (určitě) tu knihu (určitě)!
send $\text{Jirka}_{\text{DAT}}$ certainly the book_{ACC} certainly
'Certainly send the book to Jirka.'

In the recent literature, there are two interesting approaches to this issue. First, let us look at the veridicality approach.

2.1 Veridicality

It has been argued that questions and imperatives are nonveridical environments, see, for example, Giannakidou (1999, 2002). Giannakidou (2002, 5) defines (non)veridicality for propositional operators as follows:

> *(Non)veridicality for propositional operators*
> (i) A propositional operator F is veridical iff Fp entails p: $Fp \rightarrow p$; otherwise F is nonveridical.
> (ii) A nonveridical operator F is antiveridical iff Fp entails *not* p: $Fp \rightarrow \neg p$.

Given the different behavior of the adverbs in (10)-(13), one might suggest that the reason why (5) and (6) are ungrammatical is that the appropriate sentence adverbs are in the scope of a nonveridical operator. Since there is no other overt operator in the sentence, it could be something like Jacobs's (1988) assertion operator (ASSERT). It is feasible to suggest that ASSERT(p) does not entail p. A potential problem is that one cannot assume that the truth of a sentence is always evaluated with respect to a certain epistemic model because in a speaker's belief model, the asserted proposition can be true (see discussion in Giannakidou 1999). However, there is also an empirical argument against this analysis. When the sentence adverb *možná* occurs in the background domain of a declarative sentence (as evidenced by the congruent context (14a)), the sentence is grammatical; compare (14b) with example (6).

(14) a. Komu pošle Pavel tu knihu? (To whom will Pavel send the book?)

b. Tu knihu pošle možná Pavel Jirkovi.
the book_{ACC} sends possibly $\text{Pavel}_{\text{NOM}}$ $\text{Jirka}_{\text{DAT}}$
'The book, Pavel will possibly send to Jirka.'

Given the grammaticality of (14b) and the fact that the assertion operator has in its scope the whole sentence, one needs to use a tripartite structure for the operator. This is in line with Jacobs (1988), according to whom all sentences have a focus-sensitive element (either overt or covert) and all illocutionary operators can be focus sensitive and can participate in a tripartite structure. Thus, if one wants to keep the idea that it is the nonveridicality of the assertion operator to which the sentence adverb is sensitive, the assertion operator would have to be veridical in its restrictor (the background domain of the sentence) and nonveridical in its nuclear scope (the focus domain). However, there is a reason to think that nonveridicality of the focalized position is not the right issue here. According to Zwart (1995), the dyadic operator *or* (both exlusive and nonexclusive) is nonveridical in both conjuncts. (Non)veridicality for dyadic operators is defined as follows (Zwart 1995, 288):

> *(Non)veridicality for dyadic operators*
> Let C be a dyadic truth-functional connective. C is said to be *veridical* with respect to p [q] just in case pCq ⇒ p [pCq⇒ q] is logically valid.
> If C is not veridical with respect to p [q], then C is *nonveridical* with respect to p [q].

This means that the sentence adverb *možná* should be bad in this environment. This is indeed so, as shown by the following example.

(15) ??Buď sousedka možná vaří, nebo soused lakuje auto.
Either $\text{neighbor}_{\text{FEM}}$ possibly cooks or $\text{neighbor}_{\text{MASC}}$ paints car
'Either the neighbor possibly is cooking or her husband is painting his car.'

So far so good. However, according to Czech speakers, there is a clear contrast in judgments of (15) with *možná* in the background position and (16) with *možná* in the focus position. This unexpected since the focus position of the adverb should also be excluded due to nonveridicality.

(16) *Buď sousedka vaří možná, nebo soused lakuje auto.
Either $\text{neighbor}_{\text{FEM}}$ cooks possibly or $\text{neighbor}_{\text{MASC}}$ paints car
'Either the neighbor possibly is cooking or her husband is painting his car.'

Since the explanation in terms of nonveridicality of the focalized position does not work, let us turn to the downward-monotonicity approach.

2.2 Downward-monotonicity

Nilsen (2003) argues that sentence adverbs like *možná* are positive polarity items and as such are excluded from environments licensing negative polarity items. More specifically, he argues that they are excluded from downward-entailing environments. Downward-entailing operators reverse the direction of entailment. This means that in downward-entailing environments truth is preserved when the predicate is replaced by a stronger (subset) predicate. See the definition below, taken from Nilsen (2003, 41).

DE function
A function f is downward-entailing iff whenever a is semantically stronger than b, it holds that *f(b)* is semantically stronger than *f(a)*.

Thus, for example, (17a) entails (17b); the replacement of *knihu* with the stronger predicate *zajímavou knihu* under negation preserves the truth.

(17) a. Pavel nepošle Jirkovi knihu.
Pavel$_{\text{NOM}}$ $_{\text{NEG}}$sends Jirka$_{\text{DAT}}$ book$_{\text{ACC}}$
'Pavel will not send a book to Jirka.'

b. Pavel nepošle Jirkovi zajímavou knihu.
Pavel$_{\text{NOM}}$ $_{\text{NEG}}$sends Jirka$_{\text{DAT}}$ interesting book$_{\text{ACC}}$
'Pavel will not send an interesting book to Jirka.'

Again, given the grammaticality of sentence (14b) with *možná* in the background domain, the covert assertion operator should be downward-entailing in its nuclear scope (the focus of the sentence) but not in its restrictor (the background domain of the sentence). This does not go through because example (18) - with the appropriate context (18a) - shows that the focus of the assertion operator is upward-entailing; sentence (16b) entails (16c).

(18) a. Co pošle Pavel Jirkovi?
'What will Pavel send to Jirka?'

b. Pavel pošle Jirkovi zajímavou knihu.
Pavel$_{\text{NOM}}$ sends Jirka$_{\text{DAT}}$ interesting book$_{\text{ACC}}$
'Pavel will send an interesting book to Jirka.'

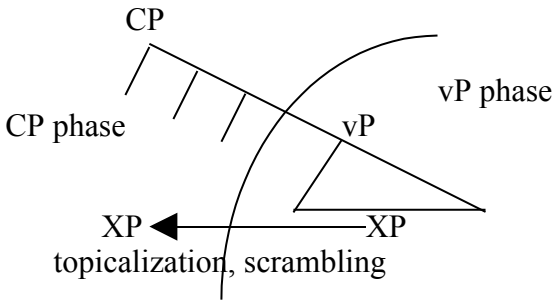
c. Pavel pošle Jirkovi knihu.
Pavel$_{\text{NOM}}$ sends Jirka$_{\text{DAT}}$ book$_{\text{ACC}}$
'Pavel will send a book to Jirka.'

Thus, the conclusion drawn from this section is that the explanation in terms of downward monotonicity does not work either. If it is not nonveridicality or downward-monotonicity that makes the sentence adverbs like *možná* bad in the sentence-final position in declarative sentences, it is necessary to find an alternative analysis.

3 The analysis

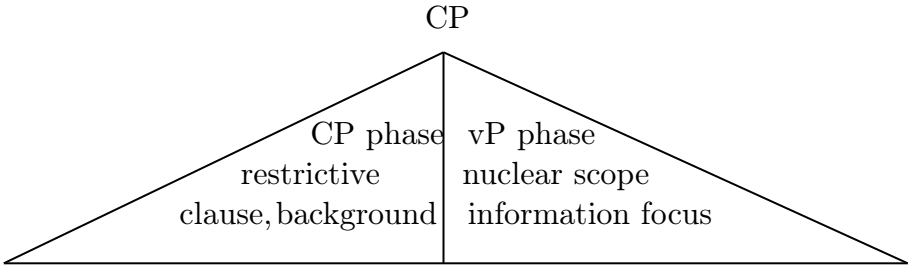
According to Chomsky (2001), every phase has its own subarray that chooses its elements from the numeration. Therefore it seems natural that one and the same adverb can appear in both the vP subarray and the CP subarray, and can be merged in both the vP phase and the CP phase. In Biskup (2006), I show that phrases moved to the CP phase – that is, scrambled or topicalized - are backgrounded and get a specific interpretation, as illustrated in (19). The specific interpretation can be epistemic, partitive, or generic. I argue that this is driven by the grammar requirement that backgrounded specific elements are to be linearized and interpreted in the CP phase (the left part of sentences) in scrambling languages like Czech.

(19)



Building on Partee (1992), Diesing (1992), and Chomsky (2000, 2001, 2005, 2006), I propose that there is a correlation between the phase structure, tripartite quantificational structure and information structure of the sentence. This is illustrated in (20). This means that at the semantic interface, the vP phase (the elements in the phase) is interpreted as the nuclear scope of the quantificational structure and the information focus. The CP phase is interpreted as the restrictive clause and the domain of background.

(20)



Then, without going into details, examples (6), (7) and (9) look like (21), (22), and (23).

(21) *[$_{CP}$ [$_{XP}$ Tu knihu pošle [$_{TP}$ Jirkovi Pavel [$_{vP}$ možná]]]].
the book$_{ACC}$ sends Jirka$_{DAT}$ Pavel$_{NOM}$ possibly
'The book, Pavel will possibly send to Jirka'

(22) [$_{CP}$ [$_{XP}$ Tu knihu pošle [$_{TP}$ Jirkovi Pavel [$_{vP}$ určitě]]]].
the book$_{ACC}$ sends Jirka$_{DAT}$ Pavel$_{NOM}$ certainly
'The book, Pavel will certainly send to Jirka.'

(23) [$_{CP}$ [$_{XP}$ Tu knihu pošle [$_{TP}$ Jirkovi Pavel [$_{vP}$ jenom možná]]]].
the book$_{ACC}$ sends Jirka$_{DAT}$ Pavel$_{NOM}$ only possibly
'It is only possible that the book Pavel will send to Jirka.'

I make the standard assumption that every sentence has a focus. Then, given the proposed model, sentence adverbs in the sentence-final position in vP are necessarily interpreted as focalized at the semantic interface and introduce a set of alternatives. To account for the behavior of sentence adverbs in the sentence-final position and the difference in behavior between *možná* and *určitě*, I will make use of Krifka's (1995) notion of 'extreme value'.[1] Krifka argues that polarity items as well as focalized elements introduce alternatives and that the alternatives are ordered according to semantic strength (set relations). And that the set of alternatives can be represented by values on a scale.

Here, I propose that a sentence adverb can appear in the sentence-final position in the vP phase and be focalized there only if it represents the extreme value with respect to the set of

[1] It has been suggested by Piñón (2006) that Krifka's approach (1995) might be used for the analysis of sentence adverbs. By 'extreme value', I mean the extremity with respect to particular focus alternatives, not with respect to all the alternatives together, for details, see Krifka (1995).

focus alternatives. Epistemic adverbials can be taken to correspond to certain values on an epistemic scale. It is a lexicosemantic property of the (realis) mood adverbial *určitě* that it corresponds to the highest value on the epistemic scale. Thus, it represents the extreme value with respect to the focus alternatives because all other alternatives are, of course, lower on the epistemic scale. Therefore *určitě* can occur in the focalized sentence-final position and in this way sentences with (realis) mood adverbials like *určitě* get the verum-focus interpretation; see (7)=(22) again.

In contrast, *možná*, given its lexicosemantic properties, does not represent an extreme value on the epistemic scale, therefore it cannot serve as the asserted focus alternative. Hence, if *možná* appears in the sentence-final position in vP, the sentence becomes ungrammatical, as demonstrated by (6)=(21). However, this changes when the exclusive particle *jenom* 'only' is added. I demonstrated by example (9)=(23) that the sentence adverbial *možná* can in fact occur in the sentence-final position. It is well-known that the exclusive adverb *only* is focus sensitive, see, for example, Rooth (1985), Beaver & Clark (2003) or Krifka (2006). In sentence (9)=(23), *jenom* associates with the focalized sentence adverbial *možná* and excludes all other alternatives. The sentence then gets the interpretation according to which it is only possible – not, for example, probable or certain, which are possible alternatives to *možná* here - that Pavel will send the book to Jirka. Thus, *jenom* makes from *možná* the lowest alternative on the epistemic scale, which means that this alternative has the required extreme value; but in this case, it is extreme low probability. So, sentence (9)=(23) shows that sentence adverbs other than (realis) mood adverbs can also merge within vP and appear in the sentence-final position, and that the appropriate adverbs do not have to be used as focusing adverbs in that position.

Similarly, the epistemic adverbial *pravděpodobně* can also occur in the sentence-final position if it represents the extreme value in the set of focus alternatives. This is shown in example (24).

(24) a. *[$_{CP}$ [$_{XP}$ Pavel přijde [$_{vP}$ pravděpodobně]]].
Pavel$_{NOM}$ comes probably
'Pavel will probably come.'

b. [$_{CP}$ [$_{XP}$ Pavel přijde [$_{vP}$ nanejvýš pravděpodobně]]].
Pavel$_{NOM}$ comes most highly probably
'It is at most probable that Pavel will come."
'It is highly probable that Pavel will come.'

Sentence (24a) demonstrates that under usual circumstances, *pravděpodobně* is ungrammatical in the sentence-final position. As shown by (24b), after adding of the adverb *nanejvýš* 'most highly', the sentence becomes grammatical. This case is especially interesting because (24b) allows the extreme value to appear on either end of the epistemic scale. The first interpretation is the extremely-low-value interpretation, according to which there is no focus alternative lower than the *pravděpodobně* value on the epistemic scale. In this case, sentence (24b) can be continued, for example, by the following sentence: *Určitě ne najisto* 'Certainly not beyond doubt'.

According to the second interpretation, the probability of Pavel's coming is highest, it is in fact certain that Pavel will come. Hence, in this case, the asserted alternative represents the extremely high value on the epistemic scale. As in the case of sentence (9)=(23), example (24b) demonstrates that although sentence adverbs like *pravděpodobně* cannot occur in the sentence-final position by themselves, they can occur there and be interpreted if they get help through another element. Hence, the problem does not lie in their low syntactic position, counter to

the accounts mentioned in the introduction.

In certain contexts, *pravděpodobně* associated with the focus particle *dokonce* 'even' can also represent the extreme value. Consider example (25).

(25) $[_{CP}\ [_{XP}$ Pavel přijde $[_{vP}$ dokonce pravděpodobně $]]]$.
Pavel$_{NOM}$ comes even probably
'It is even probable that Pavel will come.'

Imagine a situation where somebody is asking a few friends whether they will come to his party. All guys before Pavel say that they possibly will come. In this context, the value of *pravděpodobně* is extreme because it is unexpectedly high on the probability scale with respect to the other alternatives.

4 Summary

In this paper, I have argued that sentence adverbs can be merged in the vP phase, that they can occur in the sentence-final position, be accented there and focalized. Depending on their lexicosemantic properties and on the properties of the appropriate sentence, either they can be interpreted in the vP position or they cannot. This means that the (un)grammaticality of the appropriate sentence depends on the (non-)interpretability of the adverb in the vP position and not on the syntactic position of the adverb. I have argued that the ungrammaticality of certain sentence adverbs in the sentence-final position in the vP phase is due to the fact that they cannot serve as the asserted alternative in the focus in the appropriate sentence. Sentence adverbs can appear in the sentence-final position only if they represent the extreme value with respect to the set of focus alternatives.

References

Alexiadou, Artemis. 1997. *Adverb placement: A case study in antisymmetric syntax.* Amsterdam: John Benjamins.

Beaver, David and Brady Clark. 2003. Always and Only: Why not all Focus Sensitive Operators are Alike. *Natural Language Semantics* 11:323–362.

Belletti, Adriana. 1990. *Generalized Verb Movement. Aspects of Verb Syntax.* Turin: Rosenberg & Sellier.

Biskup, Petr. 2006. Scrambling in Czech: syntax, semantics and information structure. Seok Koon, Chin & Fujimori Atsushi, (eds.). *Proceedings of the 21st North West Linguistics Conference. UBC Occasional Papers of Linguistics* 1:1–15.

--. in prep. *The Phase Model and Adverbials.* PhD. dissertation. Universität Leipzig.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. Martin, R., D. Michaels & J. Uriagereka (eds.). *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik*: 89–156. Cambridge, Mass.: MIT Press.

--. 2001. Derivation by Phase. Kenstowicz, M. (ed.). *Ken Hale: A Life in Language*: 1–52. Cambridge, Mass.: MIT Press.

--. 2005. *On Phases.* Ms. MIT.

--. 2006. *Approaching UG from below.* Ms. MIT.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Giannakidou, Anastasia. 1999. Affective dependencies. *Linguistics and Philosophy* 22:367–421.

--. 2002. Licensing and Sensitivity in Polarity Items: from Downward Entailment to (Non)Veridicality. *Chicago Linguistic Society* 39:29–53.

Hajičová, Eva, Barbara Partee and Petr Sgall. 1998a. Focus, topic, and semantics. Benedicto, E., M. Romero & S. Tomioka (eds.). *UMOP 21: Proceedings of workshop on focus*: 101–124. University of Massachusetts, Amherst: GLSA.

--. 1998b. *Topic-Focus Articulation, tripartite structures, and semantic content*. Dordrecht: Kluwer.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Jacobs, Joachim. 1986. The Syntax of Focus and Adverbials in German. Abraham, W. & S. de Meij (eds.). *Topic, Focus and Configurationality*: 103–127. Amsterdam: Benjamins.

--. 1988. Fokus-Hintergrung-Gliederung und Grammatik. Altmann, H. (ed.). *Intonations-forschungen*: 89–134. Tübingen: Narr.

Koktová, Eva. 1987. On the scoping properties of negation, focusing particles and sentence adverbials. *Theoretical Linguistics* 14:173–226.

--. 1999. *Word-order Based Grammar*. Berlin: Mouton de Gruyter.

Krifka, Manfred. 1992. A Compositional Semantics for Multiple Focus Construction. Jacobs, J. (ed.). *Informationsstruktur und Grammatik*: 17–53. Opladen: Westdeutscher Verlag.

--. 1995. The Semantics and Pragmatics of Polarity Items. *Linguistic Analysis* 25:1–49.

--. 2006. Association with focus phrases. Molnar, V. & S. Winkler (eds.). *The Architecture of Focus*: 105–136. Berlin, New York: Mouton de Gruyter.

Laenzlinger, Christopher. 2002. A Feature-Based Theory of Adverb Syntax. *GG@G* 3:67–106.

Lang, Ewald. 1979. Zum Status der Satzadverbiale. *Slovo a slovesnost* 40:200–213.

Nilsen, Øysten. 2003. *Eliminating Positions*. Ph.D. dissertation, Utrecht University.

Partee, Barbara. 1992. Topic, focus, and quantification. *Proceeding of the First Conference on Semantic and Linguistic Theory* (SALT 1):159–187.

Piñón, Christopher. 2006. Modal Adverbs. Gärtner, H.M., S. Beck et al. (eds.). *40-60 puzzles for Manfred Krifka*. Berlin. http://www.zas.gwz-berlin.de/40-60-puzzles-for-krifka/

Rooth, Math. 1985. *Association with Focus*. Ph.D. dissertation. University of Massachusetts, Amherst.

Zwart, Frans. 1995. Nonveridical Contexts. *Linguistic Analysis* 25 (3/4):286–312.

A Note About *A Note About Nothing*[*]

Pavel Caha

Pavel.Caha@hum.uit.no

1 Introduction

I start off by reporting some observations made by Lída in her paper *A Note About Nothing* (Veselovská, 2003) concerning Case-assignment asymmetries in Czech DPs with inanimate pronouns (like *nic* 'nothing'). These data show that there is a clear asymmetry between Structural (nominative, accusative) and Oblique Cases, which Lída accounts for in terms of S-structure vs. D-structure asymmetry. Since D-structure as an explanatory tool has been abandoned in recent minimalist theorizing (see McCawley, 1968 for an early proposal along these lines), I try to recast the distinction in different terms. Specifically, I look at pairs of Czech constructions that exhibit a systematic alternations between the instrumental (INS, an Oblique Case) and accusative (ACC, a Structural Case), and I propose the following. (i) these constructions are transformationally related, (ii) the DPs are base-generated bearing INS, and (iii) they become ACC by movement. I derive (ii) and (iii) from the assumption that INS contains two case projections, K_1 and K_2, whereas ACC only K_2, reconstructing the S-structure vs. D-structure distinction in terms of the amount of functional structure. I further follow Starke (2005) by having the projection of K_2 raise from within K_1P. I suggest to extend this approach to Case marking in general, and I lay out a first step of the program: I propose a hierarchy of Czech Case.

2 Structural Cases vs. Oblique Cases

Veselovská (2001) observes that DPs headed by inanimate pronouns behave differently in Structural and Oblique Case environments. Specifically, adjectives following the pronoun bear genitive when the pronoun is nominative or accusative, and thus their Case differs from that of the pronoun, as demonstrated in (1). However, they are marked instrumental when the pronouns appears in instrumental, thus agreeing with the Case of the pronoun, as can be observed in (2).

(1) *něco* *takového* *velkého*
nothing$_{NOM/ACC}$ such$_{GEN}$ big$_{GEN}$ (Veselovská, 2003:ex.17a)

(2) *s* *ničím* *takovým* *velkým*
with nothing$_{INS}$ such$_{INS}$ big$_{INS}$ (Veselovská, 2003:ex.20b)

[*]I am grateful to the following people: Marcel den Dikken, Gillian Ramchand, Björn Lundquist, Lucie Medová, Marina Pantcheva, Michal Starke, Peter Svenonius, Tarald Taraldsen.

The first step of the analysis that Veselovská (2003) puts forth in order to understand the data theoretically, is to ground the distinction between Structural and Oblique Cases in terms of Case Hierarchy, understood (along the lines of Franks (1995)) in terms of different timing of operations. Specifically, Structural Case is assumed to be assigned at S-structure, whereas Oblique Case is assigned at D-structure.

(3) Case Hierarchy
Lexical [+Oblique] Case is assigned at D-Structure
Structural Case is assigned at S-Structure (Veselovská, 2003:ex.22)

The problem I address here is what to do with the distinction of Structural vs. Oblique Case when the distinction between D-structure and S-structure is abandoned (a point argued for in McCawley, 1968, or much later in Chomsky, 1995).

Apart from general considerations presented in the works cited, there is in fact empirical evidence against the implementation of Case Theory given in (3), which comes from the domain of Case itself. Specifically, the formulation in (3) stems from the conviction that while Structural Cases are purely structural and devoid of any special semantic contribution, Oblique Cases are "Lexical" in the sense that they satisfy both semantic and syntactic selectional requirements of items that they co-occur with.

Here, I argue against both of these propositions. On one hand, I present (additional) evidence that Structural Cases are relevant for semantic information (and here I follow much of previous work). On the other hand, I try to show that there is a great deal of specifically syntactic factors that contribute to the distribution of Oblique Case.[1]

Hence, I try to re-establish the distinction in completely different terms: in terms of syntactic structure.[2] I start by looking at Czech instrumental and accusative alternations and see what kind of theory is able to capture their systematic relation.

3 Instrumental and Accusative

3.1 Raising-to-Object Constructions

Consider the pair (4) and (5).

(4) *Petr* *nechal* [*televizi* *opravit* *odborníkem*]
Peter let television $\text{repair}_{\text{INF}}$ $\text{professional}_{\text{INS}}$
'Peter had the TV repaired by a professional'

(5) *Petr* *nechal* [*odborníka* *opravit* *televizi*]
Peter let $\text{professional}_{\text{ACC}}$ $\text{repair}_{\text{INF}}$ television
'Peter had the professional repair the TV'

The DP *odborník* 'professional' takes INS in (4), but ACC in (5). First thing to note is that the verb *nechat* 'let/have' is a restructuring verb (for restructuring in Czech see Dotlačil, 2004), and takes a clause headed by the infinitive as its complement.

[1] This conclusion in fact does not make sense if syntax and compositional semantics are identical, as suggested in some recent works (e.g. Ramchand, 2007).

[2] For proposals along similar lines see e.g. Bayer et. al. (2001) and Asbury (2006), as well as references therein.

The infinitival clause is passive in (4), with the internal argument (IA) of the infinitive raised to the object of the matrix verb 'let', and the external argument EA (of the infinitive) expressed by INS (as is the standard for EAs in Czech passives). In (5), the infinitive is active, and it is the infinitive's EA that becomes the object of the matrix verb.

That this analysis is on the right track, is confirmed by the (im)possibility of so called long passive (see e.g. Wurmbrand, 2001). Specifically, the reflexive passive of the matrix verb is able to promote the IA of the infinitive in the case of (4), but not in the case of (5), as shown in (6).

(6) *Televize se nechala opravit odborník-em / *-a*
Television $_{\text{REFL}}$ let repair$_{\text{INF}}$ professional$_{\text{INS / *ACC}}$
'The TV was repaired by a professional'

On standard interpretation, the EA blocks the long passive (by relativized minimality) in (5), but not in (4), suggesting that the IA must have moved across the EA in the course of deriving (4).

The conclusion is the following: the EA and the IA are both generated in the infinitival clause, but only one of them can become the object of the matrix verb. If the IA manages to do so (as in (4)), the EA is marked INS. If the EA becomes the object (blocking the long passive), it is marked ACC.

3.2 Spray/Load Alternation

The Czech spray/load alternation is presented in (7) and (8).

(7) *Naložil* [*vůz* *senem*]
loaded$_{3.\text{PL}}$ truck hay$_{\text{INS}}$
'They loaded the truck with hay.'

(8) *Naložil* [*seno* *na vůz*]
loaded$_{3.\text{PL}}$ hay$_{\text{ACC}}$ on truck
'They loaded the hay on the truck.'

In (7), the DP 'hay' is marked INS, but it shows ACC in (8). Again, the question arises how to capture the case marking of the DP in question. It is theoretically attractive to understand this pattern in terms of the preceding one. Specifically, the bracketed structure in (7) is to be understood as a passive version of the bracketed structure in (8), on analogy with (4) and (5). The same conclusion is reached by Romanova (2004, 2007) and Ramchand and Svenonius (2004). The works cited propose that the 'hay' and the 'truck' are to be understood as the EA and the IA, respectively, of a spatial relation denoted by the preposition 'on'.

In (7), the IA of the preposition moves to the object position of the verb 'load', and the EA is marked by INS. In (8), it is the EA that moves to the object position of the matrix verb, and surfaces as ACC.

The conclusion here can thus be very similar to the one arrived at in the first subsection: the EA and the IA of the P are both generated in a non-finite clause, but only one of them can become the object of the matrix verb 'load'. If the IA manages to do so (as in (7)), the EA is marked INS. If the EA becomes the object, it is marked ACC.

3.3 EAs of PPs

So far, there is no clear indication of how to analyze the relation between INS and ACC, nor how to address the problem of case in these constructions in general. One option is to take INS for an Oblique Case that serves the purpose of marking demoted EAs (of both Vs and Ps), and ACC to be assigned by the matrix verb in a designated structural position, keeping to the version of Case Theory presented in (3). However, a clearer picture (as for which way not to go) emerges, if we take a closer look on the ACC/INS alternation of the P's EA.

(9) *Petr bodnul nožem do chleba*
Peter stuck knife$_{INS}$ in bread
'Petr stuck the knife in the bread and pulled it out.'

(10) *Petr bodnul nůž do chleba*
Peter stuck knife$_{ACC}$ in bread
'Petr stuck the knife in the bread and the knife stayed there.'

As in the previous section, I consider the 'knife' in both examples to be the EA of the spatial relation denoted by the directional P *do*. The difference between the INS and ACC here does not have to do with active or passive, since the IA of the P stays the same in both examples and never moves across the EA 'knife'. Rather, the difference here is best characterized in semantic terms: either the event expressed by the V is followed by a state resulting from that event (10), or not (9). We can clearly see here that (contrary to expectations based on the version of Case Theory in (3)), it is a Structural Case that brings in a semantic load.

One way how to understand the semantic (result state vs. pure process) and syntactic (ACC vs. INS) difference theoretically, is to have the resulting state (knife in the bread) represented and projected as a phrase in syntax (resP), along the lines of Ramchand (2007). Ramchand further proposes that the head of the resP is a complement of the process (proc) denoted by the verb, which again she takes to be represented in syntax by a projection she calls procP. In these terms, the constructions under consideration differ in whether or not the syntactic structure contains the resP. Specifically, the PP (['knife' [into 'bread']]) is a complement of proc in (9), but a complement of res in (10).

(11) [proc [EA [P IA]]] no entailment of a state resulting from the process
(12) [proc [res [EA [P IA]]]] entailment of a state resulting from the process

If this is a correct account of the semantic difference, it is tempting to understand the case marking pattern in terms of the res head. A straightforward way how to relate it to the structures proposed, is to connect the ACC with the EA raised to Spec,resP, and to see the INS as arising if the EA stays in situ. The generalization, then, is that if the P's EA raises, it surfaces as ACC, whereas if it stays in situ, it bears INS.

It is not difficult to see how the solution proposed for these examples generalizes to the case of spray/load alternation. First thing to note is that in the case of 'load', there is always a resulting state implied, hence both versions have the structure in (12). The difference is whether it is the 'hay', or the 'truck' that ends up in the state of being loaded. In terms of the decomposition adopted, the difference is whether the IA or the EA ends up in Spec,resP. If the EA stays in situ, the IA becomes the holder of the resulting state (Resultee), and the EA is marked INS, in line with the observations made for the verb 'stick'. If the EA moves, it becomes the Resultee, and it surfaces as ACC.

(13) [proc [res [EA [P IA]]]] EA stays in situ --> INS

(14) [proc [res [EA [P IA]]]] EA moves --> ACC

A similar story can be told for the cases we have observed at the beginning, replacing res in (13) and (14) for the relevant head corresponding to the restructuring verb, and P for the embedded verb.[3]

An intermediate conclusion is that whatever it is that drives the movement of EA in (14), it is not the need to be assigned case, since instrumental case is available already in Spec,P. This is not as surprising as it might seem, since a similar conclusion can be drawn for pseudo-passive sentences (*John was yelled at John*), where the DP (*John*) raises from a Case position (unless there is a way how to connect the passive marking on the verb with the P's ability to assign accusative Case).

Put even strongly, the DPs in the examples discussed not only do not move for reasons of Case, they move from one Case position to another. Hence, the reasons for DP movement are rather to be sought in the fact that in each movement step, there is a new predication relation is established between the moved DP and the target of movement. So in the examples at hand, the P's EA does not move for Case reasons, it moves in order to be interpreted as the subject of the result phrase.

3.4 ias of Ps

With a partial solution in place, we want to know more on how case marking and case shifting happens. Is INS assigned to every DP that happens to be in Spec,P at the end of the day? Is all ACC marking connected to Spec,Res? Let us see if further patterns help to deepen our understanding and sharpen our conclusions. Let's start off with the complements of locative and directional PPs and case shifting that happens there.

(15) *nad* / *pod* /*před* / *za* / *mezi* *něčím*
above / under / in front of/ behind / between something$_{INS}$ LOCATION

(16) *nad* / *pod* /*před* / *za* / *mezi* *něc-o*
above / under / in front of/ behind / between something$_{ACC}$ DIRECTION

All the PPs in (15) denote simple locations, whereas those in (16) denote directions ('to above / to under' etc.). The change from locative to directional meaning is accompanied by the change of case on the P's object from INS to ACC.

It is immediately clear that INS here cannot be viewed as a case that marks EAs in situ, since here INS is associated with the IA of P. It is also clear that ACC can't be considered the case assigned by V (or res) to the direct object either: all the PPs in (16) can in fact be embedded under certain types of Ns (like 'path' or 'jump') without any effect on the case of the P's object.

Still there is a way how to connect the pattern at hand with the one arrived at before. In order to see that, let me first present some basic observations concerning the syntax of locative and directional PPs. A body of literature (van Riemsdijk, 1990, Koopman, 1997, Helmantel, 2002, den Dikken, 2003, Svenonius, 2004 on the syntactic side, and also Jackendoff, 1990, Zwarts and

[3] See Collins (2005) for a way how to get the IA across the EA. See Taraldsen (2006) for an alternative. See Hoekstra (2004:ch.1) for an interesting proposal in the same spirit as presented here.

Winter, 2000, and Kracht, 2002 on the semantic side) has established that directional PPs are built on top of locative PPs. The basic structures, taken from Svenonius (2004), are given below.

(17) [Place [DP]] locative adposition
(18) [Path [Place [DP]]] directional adposition

Again, we face the problem of how to connect the structures with Case. Making profit of the Path head in (18), it is plausible to view it as the source of ACC in directional PPs. Let me therefore propose that the DP has to raise to Spec,PathP (see (19)), which is followed by a remnant movement of the PlaceP to a yet higher position. Before I turn to some evidence for the movement, let us see where this account brings us wrt Case.

(19) [Path [Place [DP]]] ACC
(20) [Place [DP]] INS

In (20), I repeat the structure for locative PPs with the DP in INS. As with the previous constructions, we can see the INS arising as a result of a DP having stayed in situ. In directional PPs, the DP moves to Spec,Path, where it surfaces as ACC. Granted the assumptions I have made, the following unified view is achieved:

(21) DPs in situ are INS, DPs in derived position are ACC.
(22) Case marking correlates with movement, oblique marked DPs are in situ, structural case arises in derived positions.

Let me call (22) the Case-Movement Conjecture (CMC), which is meant to be an update of (3). Now back to what justifies the movement in (19). Consider the following examples from Dutch:

(23) *de weg op de heuvel* locative
the road up the hill

(24) *de weg de heuvel op* directional
the road the hill up (Dutch, den Dikken, 2006)

The Czech alternation in Case is replicated in Dutch by alternation in word order, where in directional contexts, the DP precedes the P for exactly the same reason as the Czech DPs are marked ACC, in accordance with CMC.[4]

3.5 Affected Instrumentals Surface as Accusatives

Let me continue with the idea that some DPs take INS, when in situ, and switch to ACC when they move up. The following alternation then falls in place:

[4] The reason is movement of the P's IA to Spec,PathP in directional contexts.

(25) *Petr* *prošel* *lesem*
Peter went$_{VIA}$ wood$_{INS}$
'Peter went through the wood'

(26) *Petr* *prošel* *les*
Peter went$_{VIA}$ wood$_{ACC}$
'Peter walked through the whole wood/ Peter searched the wood through'

Let me first present data that shed light on the INS in (25). As shown in (27), bare INS on a DP can be used in Czech to denote 'through'. The 'path' here is unbounded, which means that it can be fully contained in the 'wood.'

(27) *cesta lesem*
path wood$_{INS}$
'a path through the wood'

The sentence in (25) exploits this usage, with the additional factor that the prefix (*pro-*) introduces a bounded interpretation of the path Peter moves along, so the interpretation is that he enters the wood on one side and leaves it on another side. For examples like (25), I propose that the INS DP, denoting an unbounded path 'through', is a structural complement of the prefix *pro-* that introduces the bounded interpretation of the 'through' path. The PrefixP (headed by *pro-*) is further embedded under the proc head, which gives the right interpretation: there is a process of Peter walking that proceeds along the bounded path through the wood.

(28) [proc (*walk*) [*pro-* [DP]]] INS

Now the change from INS to ACC brings about a kind of "affectedness" of the DP 'the wood': it has to be affected somehow by Peter having walked through it. This gives rise to a slight shift in meaning, and the sentence roughly means that Peter has searched the wood through.

Theoretically speaking, the sentence with ACC brings about a state resulting from the process of walking. Again, this is to be captured by the presence of res in (26). The case marking then complies with the CMC: once the resP is present, the DP moves to its Spec and surfaces with ACC.

(29) [proc (*walk*) [res [*pro-* [DP]]]] ACC

3.6 Interim Summary

Before we go on, let us see where we are. I have started with the observations concerning active and passive non-finite complements to verbs, concentrating on the case marking of the EA. The initial observation (for passive VPs and PPs) was that INS marks EAs that are demoted. This had to be changed for EAs of Ps embedded under the verb 'stick', where the same alternation was due to semantic factors (having a resulting state or not). The conclusion which emerged was that EAs in situ are marked INS, whereas moved EAs receive ACC. For these reasons, I have suggested that DP movement is to be dissociated from Case licensing per se, and that DPs are free to originate in one Case position and move to another.

The investigation of case alternations in PPs brought us further to the conclusion that there is no designated projection for either INS or ACC. It was concluded that a unified view on the alternations can only be achieved if CMC is adopted: DPs are marked by Oblique Case in situ, and they receive ACC once moved to a derived position. The formal mechanism underlying CMC is presented in the next section.

4 Peeling

The technical problem I want to find a solution for in this section is how to make one Case (INS, associated with the base position) change into another (ACC) in the derived position. There seem to be two easy options: the addition of a feature (e.g. along the lines of Sportiche, 2005, or Koopman, 2005), or the loss of a feature (along the lines of Sportiche, 1988). Let Case morphology be our guide here.

Table I: Subset-superset relations of instrumental and accusative in Coll. Czech

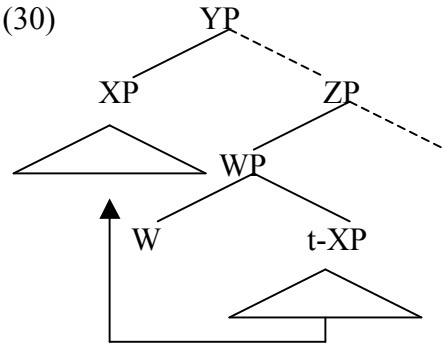
	muž ‚man', pl	*kuře*, ‚chicken', pl	*kost*, ‚bone', pl	*stavení*, ‚building', pl	*dobrý*, adj.,pl.
NOMINATIVE	muž-i	kuřat-a	myš-i	staven-í	dobr-ý
ACCUSATIVE	muž-e	kuřat-a	myš-i	staven-í	dobr-ý
INSTRUMENTAL	muž-e-ma	kuřat-a-ma	myš-i-ma	staven-í-ma	dobr-ý-ma
GENITIVE	muž-ů	kuřat-Ø	myš-í	staven-í	dobr-ý-ch
DATIVE	muž-ů-m	kuřat-ů-m	myš-í-m	staven-í-m	dobr-ý-m
LOCATIVE	muž-í-ch	kuřat-e-ch	myš-í-ch	staven-í-ch	dobr-ý-ch

It can be easily observed that morphology tells us to choose the latter way and treat the case shifting patterns discussed as a feature loss. The reason is that INS in Czech tends to be morphologically more complex than ACC, which is shown in the paradigms above. Here, the INS actually seems to be built on top of the ACC, by the addition of *-ma*. In other words, the change from INS to ACC is to be understood as stranding of *-ma*.

4.1 How to Lose Features?

The scenario of feature loss is given below in (30). This tree encodes the way DPs move from one case position to another in the constructions discussed. Starke (2005) generalizes this scenario for all DP movement (with slight qualifications concerning successive cyclic movement).

(30)



Starke (2005) calls the W in (30) the Peel, and the movement of XP is called Peeling. The adoption of Starke's Peeling proposal for the cases at hand, leads to the conclusion that INS

is structurally more complex than ACC, and that, in fact, accusative is contained inside the INS. The following representations for INS and ACC encode this, with INS containing two Case heads, K_1 and K_2, as opposed to ACC with only one Case head.

(31) [K_1 [K_2 [XP]]] INS
(32) [K_2 [XP]] ACC

In these representations, ACC is a proper subset of INS, reflecting directly the morphology of Table I. In syntax, DPs are base generated as INS (at least in the positions discussed), and when they move, they move by Peeling.[5] This means that when they undergo movement for reasons of extra predication, they strand the highest shell, K_1, in situ, and surface as K_2Ps, that is as ACCs.

5 Nominative

If (30) is the way Case movement is to be represented in the case of ACC and INS, a natural hypothesis is that (30) holds for all Case driven movement. In particular, pairs of active and passive sentences (illustrated in (33) and (34)) seem to be easily susceptible for such an approach.

(33) *Vypil becherovku*
drank becherovka$_{ACC}$
'He drank becherovka.'

(34) *Becherovka byla vypita*
becherovka$_{NOM}$ was drunk
'Becherovka was drunk.'

In present terms, when the DP 'becherovka' moves from the object position to the subject position, it changes case from ACC to nominative (NOM). The way this can be understood under the Peeling theory of Case assignment, is to assume (30) together with the representation of NOM and ACC depicted below:

(35) [K_2 [K_3P]] ACC
(36) [K_3P] NOM

On analogy with the previous example, NOM is a syntactic subset of ACC. When an ACC DP undergoes passivization, it moves to Spec,TP and since it moves by Peeling (30), it strands the projection of accusative in situ and surfaces as NOM in the derived position.

6 The Case Hierarchy

The observations above suggest that Cases are neither primitive units, nor structurally identical to one another. In fact, it appears that each Case is a unique collection of functional heads, and that individual Cases stand in structural subset - superset relations, as determined by the collection of heads they are the manifestation of. Specifically, I have argued that if we want

[5] See Medová (to appear) for the same approach.

to capture the systematic relationship between ACC and INS in Czech, we can consider ACC to be a structural subset of INS, which is directly reflected by morphology of certain paradigms presented in Table I.

In this section, I build on cross-linguistic comparison of Case systems conducted by Blake (1994) and I propose a universal hierarchy of Cases. I show what predictions this hierarchy (understood in terms of syntactic structure) makes with respect to syncretism patterns that are expected, and I demonstrate the validity of the predictions on Czech Case morphology.

6.1 Blake's (1994) Case Sequence (with a Czech Twist)

Blake (1994) discusses Case systems of the languages of the world, observing the number and kind of cases each language distinguishes. Among others, he mentions the following types of languages (with languages exemplifying these types in brackets):[6]

(37) NOM – ACC/OBL (Chemehuevi)
(38) NOM – ERG/OBL (Kabardian)
(39) NOM – ACC – GEN/OBL (Modern Greek, Semitic, Nubian)
(40) NOM – ACC – GEN – DAT/OBL (Icelandic, German, Nilo-Saharan, Yaqui, Ancient Greek)
(41) NOM – ACC – GEN – DAT – LOC/OBL (Latin)
(42) NOM – ACC – GEN – DAT – LOC – INSTR/OBL (Slavic)
(43) NOM – ACC – GEN – DAT – LOC – ABL/OBL (Turkish)
(44) NOM – ACC – GEN – DAT – LOC – INSTR – ABL (Classical Armenian)
(45) NOM – ACC – GEN – DAT – LOC – INSTR – ABL - COM (Tamil)

What can be observed is that there is a clear hierarchy in which Cases tend to appear in the languages of the world. The hierarchy is given below:

(46) NOM > ACC > GEN > DAT > LOC > INSTR > ABL > COM

Interpreting this hierarchy in terms of syntactic structure means that not only is accusative formed on top of nominative by the addition of a feature (as already argued), but genitive is formed on top of accusative in a similar fashion (and mutatis mutandis for other Cases). Having rebuilt Blake's hierarchy this way in the syntax, we in fact achieve the result that the hierarchy is derived, granted (in addition) the standard assumption that if in a particular language, the head noun is able to move to the left of, say, dative (so that dative surfaces as a suffix), the DP can also move to the left of all cases lower on the hierarchy (and so they also end up as suffixes).

Another consequence of the Blake's hierarchy (given its syntactic implementation) can be brought to light under particular assumptions concerning Case syncretism. Specifically, Bobaljik (2007) argues that only heads that are adjacent in the syntactic structure can be syncretic.[7] It follows that if two non-adjacent Cases of the Blake's hierarchy are syncretic, all cases that intervene have to be syncretic with these two Cases. This is what Bobaljik refers to as *A-B-A.

This kind of reasoning applied to Czech, however, reveals that a twist in the hierarchy is needed, and that its correct rendering for Czech implies the order LOC > DAT (47) (opposed

[6] See also a summary of the Moving Right Along seminar held in Tromsø on on June 13, 2006 by Pantcheva, available at http://www.hum.uit.no/mra/.

[7] See Bobaljik's work for one possible reasoning that derives this fact, but see also the implementation in Caha (2007), which derives the same result under different assumptions concerning lexical access.

to Blake's DAT > LOC). There are two reasons. First, it is sometimes the case that locative is syncretic with genitive, with dative not participating in the syncretism (see the first two lines below (47)). Second, dative can be syncretic with instrumental, with locative standing aside (a syncretism characteristic of the old dual). Syntactically, the reason that Czech requires the order LOC > DAT is that the Czech Locative never conveys locational meaning on its own, but always requires the help of a preposition. Hence, it does not count as a locative case in a strict sense, and the fact that it is called locative is rather misleading.

(47)	[Ins	[Dat	[Loc	[Gen	[Acc	[Nom ....
'we'	ná-m-i	ná-m	ná-s	ná-s	ná-s	my
'Adj., pl.'	A-ý-m-i	A-ý-m	A-ý-ch	A-ý-ch	A-ý	A-ý
'both'	ob-ě-m-a	ob-ě-m-a	ob-ou	ob-ou	ob-a	ob-a

Space limitations prevent me from demonstrating Czech Case morphology in greater detail. However (once certain intervening factors are controlled for), there are only two exponents whose distribution presents a counterexample to the hierarchy (47) coupled with the *A-B-A reasoning. The first problematic exponent is acc.sg. *-i* in the paradigm *růže*, which is syncretic with locative and dative, but accusative is (unexpectedly) *-e*.[8] Similarly, *-ou* is accusative and instrumental of feminine adjectives, but the intervening Cases have a different ending. I suppose that these are examples of homonymy, and not systematic syncretism.

(47), then, is the Case hierarchy that seems motivated on both morphological and typological grounds, underlying the syntax of Czech Case.

7 Conclusions

The proposal I have put forth here is that the distinction between Structural and Oblique Cases is a a difference in the amount of structure. Specifically, I have proposed that Oblique Cases have a large number of functional projections and they in fact contain the Structural Cases inside them. This proposal has a direct bearing on several questions that have been discussed. First, the fact that Oblique Cases tend to be expressed by more morphology, and that Oblique Cases can be missing from a given language (Blake's hierarchy), are both in direct correlation with the proposed hierarchy of Cases. Second, a mechanism for Case shifting, namely Peeling, has been proposed, following the guidelines of Starke (2005).

References

Asbury, Anna. 2006. Towards a typology of morphological case. In: Svenonius, P. (ed.). *Nordlyd: Tromsø Working Papers in Linguistics, special issue on Adpositions* 33.1:120–148.

Bayer, Joseph, Markus Bader, and Michael Meng. 2001. Morphological underspecification meets oblique case: Syntactic and processing effects in German. *Lingua* 111:465–514.

Blake, Barry. 1994. *Case*. Cambridge: Cambridge University Press.

Bobaljik, Jonathan. 2007. *ABCs of comparative suppletion*. Ms., University of Connecticut.

Caha, Pavel. 2007. *The Superset Principle*. Ms., University of Tromsø.

Chomsky, Noam. 1995. *The minimalist programme*. Cambridge, MA: MIT Press.

[8] Incidentally, the accusative sg. *-i* also gives rise to A-B-A pattern of another exponent in the paradigm *růže*, namely *-e*, which appears in nominative and genitive.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8:81-120.

den Dikken, Marcel. 2006. *On the functional structure of locative and directional PPs.* Ms., CUNY.

Dotlačil, Jakub. 2004. *The syntax of infinitives in Czech.* Master's thesis, University of Tromsø.

Fong, Viviane. 1997. *The Order of Things: What Directional Locatives Denote.* PhD thesis, Stanford University.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax.* Oxford, New York: Oxford University Press.

Helmantel, Marijon. 2002. *Interactions in the Dutch adpositional domain.* Utrecht: LOT.

Hoekstra, Teun. 2004. *Arguments and Structure. Studies on the Architecture of the Sentence.* Sybesma, R. et al. (eds.). Berlin: Mouton de Gruyter.

Jackendoff, Ray. 1990. *Semantic Structures.* Cambridge, MA: MIT Press.

Koopman, Hilda. 1997. *Prepositions, postpositions, circumpositions and particles: The structure of Dutch PPs.* Ms. UCLA. http://www.linguistics. ucla.edu/people/ koopman/papers/pp.pdf

Koopman, Hilda. 2005. Korean (and Japanese) Morphology from a Syntactic Perspective. *Linguistic Inquiry* 36:601-633.

Kracht, Markus. 2002. On the semantics of locatives. *Linguistics and Philosophy* 25:157–232.

McCawley, John. 1968. Lexical insertion in a grammar without Deep Structure. In: Darden, B. J., C.-J. N. Bailey & A. Davison (eds.): *Papers from the fourth regional meeting of the Chicago Linguistic Society*, Chicago.

Medová, Lucie. 2007. Reflexive clitics, Movement, and Layered Case. To appear in: *Proceedings of FDSL 6.5.* Frankfurt: Peter Lang.

Ramchand, Gillian. 2007. *Verb Meaning and Lexicon.* Ms. Tromsø, lingBuzz/000346.

Ramchand, Gillian and Peter Svenonius. 2004. *Prepositions and External Argument Demotion.* Handout of a talk in Oslo. Downloadable at www.hum.uit.no/a/svenonius/papers/RamchandSvenonius04Oslo.pdf

Riemsdijk, Henk van. 1990. Functional prepositions. In: Pinkster, H. & I. Genee (eds.). *Unity in diversity: Papers presented to Simon Dik on his 50th birthday*: 229–241. Dordrecht: Foris.

Romanova, Eugenia. 2004. *Transitive, intransitive and passivizing... prefixes.* Handout of a talk at the Argument Structure Workshop, November, Tromsø.

Romanova, Eugenia. 2007. *Constructing Perfectivity in Russian.* Ms. Tromsø.

Sportiche, Dominique. 1988. A Theory of Floating Quantifiers and its Corollaries for Constituent Structure. *Linguistic Inquiry* 19:425-451.

Sportiche, Dominique. 2005. *Division of Labor between Merge and Move: Strict Locality of Selection and Apparent Reconstruction Paradoxes.* Ms. UCLA. lingBuzz/000163

Svenonius, Peter. 2004. *Spatial P in English.* Ms. University of Tromsø.

Starke, Michal. 2005. Lecture in Tromsø, Fall.

Taraldsen, Tarald. 2006. *Resumptive pronouns and stranded Case in the Bodø dialect.* Talk at NORMS Workshop, October, Tromsø.

Veselovská, Ludmila. 2003. *A note about nothing.* Ms., Palacký University, Olomouc.

Wurmbrand, Susanne. 2001. *Infinitives: restructuring and clause structure.* Berlin, New York: Mouton de Gruyter.

Zwarts, Joost and Yoad, Winter. 2000. Vector Space Semantics: A Model-Theoretic Analysis of Locative Prepositions. *Journal of Logic, Language, and Information* 9:169–211.

Infinitives under 'have'/'be' in Czech

Markéta Ceplová

ceplova@mit.edu

1 Introduction

The verb 'have' in Czech can show up in three different versions: 1) possessive 'have', 2) obligational 'have', and 3) wh-existential 'have', as seen in the following example:

(1) a. Anna měla kočku.
Anna$_{nom.sg}$ have$_{3sgf.past}$ cat$_{acc.sg}$
'Anna had a cat.'

b. Anna měla spát.
Anna$_{nom.sg}$ $_{3sgf.past}$ sleep$_{inf}$
'Anna was supposed to sleep.'

c. Anna měla kde spát.
Anna$_{nom.sg}$ have$_{3sgf.past}$ where sleep$_{inf}$
'Anna had a place where she could have slept.'

Parallel to the verb 'have', similar options, with the exception of the obligational construction, are also available to the verb 'be':

(2) a. Na zahradě byla kočka.
in garden be$_{3sgf.past}$ cat$_{nom.sg}$
'There was a cat in the garden.'

b. * Na zahradě bylo spát.
in garden be$_{3sgn.past}$ sleep$_{inf}$

c. Na zahradě bylo kde spát.
in garden be$_{3sgn.past}$ where sleep$_{inf}$
'There was a place in the garden where we could have slept.'

The three options vary significantly in several factors. The first difference is what complement the 'have'/'be' verbs take, a DP/NP ('cat') in the first case, an infinitival verb ('sleep') in the second (this option is not a possible complement for the verb 'be'), or a wh-complement with a non-finite (infinitival or subjunctive) verb in the third ('where to sleep').

Parallel to these three options of complements is the interpretation of the whole clause. While the first case receives a possessive/existential interpretation, the second, as well as the third one, receives a modal interpretation. The difference between the last two options is in part in the specific flavor and force of the modality. A bare infinitival verb, as in (1b)), receives an

obligational modality (a standard 'have-to' interpretation, as in e. g. Bhatt, 1998), the modal force in the third version of the verb 'have' receives a modality of 'existential possibility.'

The constructions along the ones in (1c) and (2c) are available not only in Czech, but also in other Slavic, Romance, as well as in other languages[1].

These constructions have been given several different analyses and different names in the literature: irrealis free relatives (Grosu and Landman, 1998), indefinite free relatives (Caponigro, 2001), indefinites in disguise (Rappaport, 1986; Rudin, 1986), infinitival existential sentences (Babby, 2000) or wh-existential constructions (Izvorski, 1998; Pancheva Izvorski, 2000). I will henceforth call them 'wh-existential constructions,' or WHECs for short.

In this paper I look mainly at the wh-complements of the wh-existential constructions. But before that, let us look at the constructions as whole. WHECs show several notable properties, that will not be a subject of this paper, but are still worth setting clear up-front and that distinguish them from other cases that could look similar.

For one thing, as already mentioned above, the constructions obligatorily receive a modal interpretation – one that is parallel to the English case in (3), namely the "could" interpretation that is in (4a), but not a "should" interpretation as in (4b) that is an alternative interpretation of the English sentence here[2]:

(3) I have something to read.

(4) a. I have something that I *can* read.
b. I have something that I *have to* read.

To be somewhat more precise about it, the available modality is claimed by Pancheva Izvorski (2000) to be "associated with a particular modal interpretation – an existential modal restricted by a bouletic accessibility relation" (Pancheva Izvorski, 2000, p. 24), or "the root modal of circumstantial possibility" (p. 36).

Another point worth mentioning is that the number of possible matrix verbs that can be involved in WHECs is argued to be strictly limited by Pancheva Izvorski (2000). Besides 'be' and 'have' there are (in some languages) also other ones available in WHECs, namely Pancheva Izvorski (2000) includes 'find', 'look for', 'choose' in the languages she looks at, see her ex. (39) from Russian on page 35. However, neither of these alternative matrix verbs allows properties of the wh-clause parallel to the ones allowed by 'be' and 'have' in Czech (if available at all).

(5) On isčet [s kem poexat'].
he look-for$_{3sg.pres}$ [with whom go$_{inf}$]
'He is looking for someone to go with.'

For the purposes of this paper, I am going to concentrate on the matrix verbs 'be' and 'have'.

The properties of the wh-complement, which is the main focus of this paper, are as follows: I first show that in general and more specifically in Czech 1) the verb needs to be non-finite, namely infinitival, or subjunctive (see section 2); 2) it needs to involve one or more wh-words (see section 3); 3) the size of the complement has to be small enough to enable restructuring, at least optionally (see section 4; 4) the (understood) Subject of the wh-clause is obligatorily coreferential with the Subject of the matrix verb (where available).

[1] Pancheva Izvorski (2000) includes in the list also Greek, Hebrew, dialects of Arabic, and Yiddish.

[2] For arguments and treatment that exclude the "should" modality reading in WHECs see (Pancheva Izvorski, 2000, p. 27n).

2 Non-finiteness restriction on the wh-complement

Besides other restrictions, there is a strict restriction on the verb in the wh-complement of the matrix verb. Specifically, the verb has to be either infinitival, or have subjunctive[3] tense, no finite verb is allowed:

(6) a. *infinitive*

Petr neměl koho pozvat.
Petr$_{nom.sg}$ not-have$_{3sg.past}$ who$_{acc.sg}$ invite$_{inf.}$
'Petr didn't have anyone he could invite.'

b. *subjunctive*

Petr neměl koho by pozval.
Petr$_{nom.sg}$ not-have$_{3sg.past}$ who$_{acc.sg}$ aux$_{subj.-cl}$ invite$_{pprt.}$

c. *finite verb*

* Petr neměl koho pozval.
Petr$_{nom.sg}$ not-have$_{3sg.past}$ who$_{acc.sg}$ invite$_{3sg.past.}$

In general, the subjunctive form of the Czech verb is a compound form, including an auxiliary (derived from 'be') that behaves as a clitic and agrees in person/number with the Subject, and a past participle of the verb that agrees with the Subject in person/number/gender (marked here as 'pprt.' in the glosses).

The subjunctive verb, unlike the infinitival one, can also take a Subject, marked with Nominative Case in most instances, though it can leave it null (*pro*, or PRO):

(7) Petr by přišel rád.
Petr$_{nom.sg}$ aux$_{3sg}$ come$_{pprt.3sgm}$ glad$_{3sgm}$
'Petr would be glad to come.'

(8) Přišel by rád.
come$_{pprt.3sgm}$ aux$_{3sg}$ glad$_{3sgm}$
'He would be glad to come.'

(9) Doporučil jí, aby přišla brzo.
recommend$_{past.3sgm}$ she$_{dat.sg}$ that$_{subj.3sg}$ come$_{past.3sgf}$ early
'He recommended to her to come early.'

(10) * Petr přijít rád.
Petr$_{nom.sg}$ come$_{inf}$ glad$_{3sgm}$

When a subjunctive form of the verb is used in the wh-complement, it obligatorily agrees with the (matrix) subject, in person, number, and gender, to the extent the agreement has effects on the morphological form of the verb. Since infinitives in Czech do not show any agreement, nothing similar can be observed in those cases:

(11) *subjunctive*

[3] This form of the verb is usually described in the Czech literature as 'conjunctive'. To the extent that 'conjunctive' ('kondicionál' in Czech) can be taken as another name for 'subjunctive' it is fair to talk about it this way. Junghanns (1999) also gives what he calls 'subjunctive markers' for Czech, equating 'subjunctive' to the standard Czech 'kondicionál'. The 'markers' that he provides are the same ones I have been looking at here.

a. Petr měl, co by mu koupil.
Petr$_{nom.sg.m}$ have$_{past.3sgm}$ what$_{acc.sg}$ aux$_{subj.3sg}$ he$_{dat.sg}$
'Petr could buy him something.'

b. Marie měla, co by mu koupila.
Marie$_{nom.sg.f}$ have$_{past.3sgf}$ what$_{acc.sg}$ aux$_{subj.3sg}$ he$_{dat.sg}$ buy$_{past.sgf}$
'Marie had something she could buy him.'

c. Máme, co bychom mu koupili.
have$_{pres.1pl}$ what$_{acc.sg}$ aux$_{subj.1pl}$ he$_{dat.sg}$ buy$_{past.pl}$
'We have something we could buy him.'

d. Mám, co bych mu koupila.
have$_{pres.1sg}$ what$_{acc.sg}$ aux$_{subj.1sg}$ he$_{dat.sg}$ buy$_{past.sgf}$
'I have something I could buy him.'

(12) *infinitive*

Měl (jsem) / měla (jsme) / měli (jsme) / mám co
have$_{past.3sgm}$ (aux$_{1sg}$) / have$_{past.3sgf}$ (aux$_{1sg}$) / have$_{past.3pl}$ (aux$_{1pl}$) / have$_{pres.1sg}$ what
mu koupit.
he$_{dat.sg}$ buy$_{inf.}$

'He/she/we/I had/have something he could buy him.'

Another difference between the two verbal forms also includes the fact that unlike the infinitival wh-complements that allow restructuring (and thus clitic climbing), wh-complements with a subjunctive verb do not have this option.

In the rest of this paper I look primarily at cases with the infinitival version of the verb, noting in passing where relevant or important differences obtain without necessarily providing a detailed account of them.

3 Wh-words in the wh-complement

In this section I look at what wh-words can be present in the wh-complement and what determines their exact status and shape. I first go over the ways the form of the wh-words is determined, then specifically looking at those that come with a Nominative Case and restrictions on those, based on that I also look at where the wh-words start and how far they can go.

3.1 Restricting the form of the wh-words

I first show that the syntactic restrictions (if any) the wh-words show come from the wh-complement itself (specifically its verb). The restrictions that come from the matrix clause/verb (no which- or how-many-constructions can appear in the wh-complement), are mostly of semantic nature. Wh-words with '-ever' (as the English 'whatever', 'whoever', 'wherever'...) are prohibited in the wh-existential construction:

(13) Kalina má koho(*koli) pozvat.
Kalina$_{nom.sg}$ have$_{pres.3sg}$ who(*ever)$_{acc.sg}$ invite$_{inf.}$
'There is someone Kalina can invite.'

This clear impossibility of adding -ever to the wh-words is a strong argument against treating the wh-complement as a Free Relative. Pancheva Izvorski (2000) argues that based on the difference in the (non)availability of -ever with the wh-words, as well as other properties of the wh-clauses, the wh-clauses pattern with interrogative clauses, rather than Free Relatives. The (lack of) availability of wh-pronouns with -ever in English, as in Czech, strongly support that conclusion, as seen below:

(14) a. Kdo(koli) přijde, bude považován za pozvaného.
who(ever)$_{nom.sg}$ come$_{3sg.pres}$ be$_{3sg.fut}$ considered$_{nom.sgm}$ for invited$_{acc.sg}$
'Who(ever) comes will be considered as having been invited.'

b. Vím, kdo(*koli) přišel.
know$_{1sg.pres}$ who(*ever)$_{nom.sg}$ come$_{3sgm.past}$
'I know who came.'

(15) a. Who(ever) came was sent to the kitchen.

b. I wonder who(*ever) came.

The wh-existential construction is restricted to clauses that include at least one wh-word in the complement of "be"/"have". Without it, the sentence receives an entirely different interpretation that is not related to the construction under disscussion, as seen below:

(16) a. Kalina má někoho pozvat.
Kalina$_{nom.sg}$ have$_{3sg.pres}$ someone$_{acc.sg}$ invite$_{inf.}$
'Kalina is supposed to invite someone.'

b. Kalina má koho pozvat.
Kalina$_{nom.sg}$ have$_{3sg.pres}$ who$_{acc.sg}$ invite$_{inf.}$
'There is someone who Kalina can invite.'

The wh-existential constructions allow almost any wh-word in the wh-complement, both nominal as well as adverbial ones, as seen in (17):

(17) a. Josef má (určitě) koho pozvat.
Josef$_{nom.sg}$ have$_{3sg.pres}$ (certainly) who$_{acc.sg}$ invite$_{inf.}$
'Josef certainly has someone he can invite.'

b. Josef má kam Annu pozvat.
Josef$_{nom.sg}$ have$_{3sg.pres}$ where Anna$_{acc.sg}$ invite$_{inf.}$
'Josef has a place he can invite Anna to.'

c. Josef (opravdu) nemá proč tam chodit.
Josef (really) neg-have$_{3sg.pres}$ why there go$_{inf.}$
'Josef does not have a(ny) reason to go there.'

The wh-complement allows more than one wh-word:

(18) a. *(a case in which everyone has to keep introducing people to other people, but Josef refuses to continue and a friend is trying to defend him)*

Josef už opravdu nemá koho komu představit.
Josef$_{nom.sg}$ already really not-have$_{3sg.pres}$ who$_{acc.sg}$ who$_{dat.sg}$ introduce$_{inf}$

'Josef is done with all introductions.'

b. *(A statement of an usherer in a completely full space:)*

Už nemám kam koho posadit.
already not-have$_{1sg.pres}$ where who$_{acc.sg}$ seat$_{inf}$

'No-one else can be seated here.'

Yet a further note that is needed here is a note about nominative wh-words. While infinitives in Czech cannot assign a Nominative Case, or take any overt Subject, see e. g. (19), there still are some cases of nominative/Subject wh-words in the WHECs, as e. g. in (10) (repeated here as (20)):

(19) * (Petr) přijít rád.
Petr$_{nom.sg}$ come$_{inf}$ glad$_{3sgm}$

(20) Neměl kdo přijít.
not-have$_{3sgm.past}$ who$_{nom}$ come$_{inf.}$
'There wasn't anyone who could come.'

At this point, the only place the Nominative Case could be assigned from is the matrix verb, provided it itself is able to do that. I will return to these cases, as well as to the discussion of the Subject and its interpretation and other cases of nouns with Nominative Case in the wh-complement, in section 5.

3.2 Wh-movement

The wh-words also have the option to move higher, out of the wh-complement, e. g. in cases the whole sentence is turned into a wh-question (showing that the wh-complement is not an island[4]; compare the following cases with a declarative sentence in (18a)):

(21) a. Má Josef koho komu představit?
have$_{3sg.pres}$ Josef$_{nom.sg}$ who$_{acc.sg}$ who$_{dat.sg}$ introduce$_{inf.}$
'Can Josef still make more introductions?'

b. Koho má (ještě) Josef komu představit?
who$_{acc.sg}$ have$_{3sg.pres}$ (still) Josef$_{nom.sg}$ who$_{dat.sg}$ introduce$_{inf.}$
'Who could Josef still introduce to someone?'
'Who is Josef (still) supposed to introduce to someone?'

c. Komu má (ještě) Josef koho představit?
whom$_{dat.sg}$ have$_{3sg.pres}$ (still) Josef$_{nom.sg}$ who$_{acc.sg}$ introduce$_{inf}$
'Who could Josef (still) introduce someone to him?'
'To whom is Josef (still) supposed to introduce someone?'

Three things need to be added to the examples in (21). One is that in this case, either of the two wh-words can move higher, without any clear preference, as seen in (21b) and (21c).

3.2.1 Alternative interpretations

The second thing to note is the fact that the sentences with one of the wh-words moved out of the wh-complement (above the verb 'have') allow two possible interpretations, as long as at least one wh-word stays in the wh-complement (see the translations in (21)). One of these two possible interpretations is the one that wh-existential constructions bring (i. e. asserting

[4] I would like to thank David Pesetsky for pointing out the islandhood status to me.

the existence of the entity that the wh-word stands for, and its availability to the subject of the matrix clause, together with an existential force of a modality; for more on that, see the introduction in section 1). The other one is a "have-to" interpretation that "have" plus an infinitive can bring in Czech independently (in a 'have-to' construction):

(22) a. Josef má přijít zítra.
Josef$_{nom.sg}$ have$_{3sg.pres}$ come$_{inf.}$ tomorrow
'Josef should (is supposed/expected to) come tomorrow.'

b. Kdy má Josef přijít?
when have$_{3sg.pres}$ Josef$_{nom.sg}$ come$_{inf.}$
'When should Josef (is J. supposed to) come?'

However, when all wh-words move higher, only the "have-to" interpretation remains, and the one that is a part of the wh-existential construction is not available at all:

(23) Koho komu má Josef představit?
who$_{acc.sg}$ who$_{dat.sg}$ have$_{3sg.pres}$ Josef$_{nom.sg}$ introduce$_{inf.}$
'Who is Josef supposed to introduce to someone?'

Another take on this argument is also available from cases that do not show any alternatives of interpretation when the wh-words are moved out of the wh-complement (as in sentences with subjunctive verbs in the wh-complements):

(24) * Koho komu má Josef by představil?
who$_{acc.sg}$ who$_{dat.sg}$ have$_{3sg.pres}$ Josef$_{nom.sg}$ aux$_{subj-cl}$ introduce$_{pprt.3sgm}$

In cases with subjunctive verbs in the wh-complement, when all the wh-words are moved out of the wh-complement, the sentences then become ungrammatical, as expected.

4 Size of the wh-complement

The size of the wh-complement has standardly been taken to be a CP, see e. g. Babby (2000); Pancheva Izvorski (2000). However, Czech WHECs have some properties that argue against this treatment. Namely clitic climbing that is completely grammatical in WHECs, while it has been argued to be impossible out of all finite, as well as some infinitival clauses, e. g. non-finite wh-complements of interrogative verbs (such as 'know' or 'ask'). The clitic climbing itself then strongly suggests the size of the wh-complements of WHECs to be smaller than CP.

(25) a. Petr má kam ho pozvat.
Petr$_{nom.sg}$ have$_{3sg.pres}$ where he$_{acc.sg-cl.}$ invite$_{inf.}$
'Petr has a place where he could invite him.'

b. Petr ho má kam pozvat.
Petr$_{nom.sg}$ he$_{acc.sg-cl.}$ have$_{3sg.pres}$ where invite$_{inf.}$

It has been argued in the literature (Rezac, ms.; Dotlačil, in progress) that clitics in Czech cannot climb out of a CP, not even out of a non-finite one, as seen in the following examples:

(26) a. * Ale nevím ho$_i$ opravdu jak zapisovat t$_i$.
but not-know him really how record$_{inf.}$
'But I really do not know how to record it.' [Lenertová (2004), taken from Dotlačil (in progress)]

b. * Petr mu_i nevěděl koho doporučit t_i
Petr$_{nom.sg}$ he$_{dat.sg-cl.}$ not-know$_{3sg.past}$ who$_{acc.sg}$ recommend$_{inf.}$
'Petr did not know who to recommend to him.'

Furthermore, Rezac (ms.) argues that clitics can only climb out of *restructuring* infinitives (based on looking at the restructuring infinitives within the framework of Aissen and Perlmutter, 1983).

Based on some tests, and adopting conclusions from Wurmbrand (1998), Rezac concludes that the infinitive that allows clitic climbing has to be a restructuring one, and more specifically of the size of a bare VP (while a non-restructuring infinitive as a complement is larger).

For Dotlačil, the restructuring infinitive is (or at least can be) slightly bigger than just a bare VP. Based on Lenertová (Lenertová, 2004)[5], "clitic climbing is not degraded if the clause hosts a temporal adverb." Thus "the infinitival clause must be realizable with T node but without PRO" which Dotlačil argues is the same as that for long-distance agreement in Czech.

Based on the previous section, the possibilities of the size of the wh-complement in WHECs are either bare VP, or vP. I suggest treating the wh-complements of WHECs as an vP based on the wh-nature of the complement.

5 Subject of the wh-complement: Control or Raising?

The embedded non-finite (especially the infinitival) wh-clause does not allow (with some exceptions to which I return in section 5.2) an overt subject to be present. It has been commonly understood that the understood Subject of the infinitive has to be coreferential with the Subject of the matrix verb. Thus the following sentence in (27a) has only the Subject interpretation that is provided, with no other Subject interpretation, as e. g. in (27b), possible.

(27) a. Josef má koho Marii představit.
Josef$_{nom.sg}$ have$_{3sg.pres}$ who$_{acc.sg}$ Marie$_{dat.sg}$ introduce$_{inf.}$
'Josef$_i$ has someone who he$_i$ can introduce to Marie.'

b. 'Josef$_i$ has someone who he$_j$ can introduce to Marie.'

There are several options for capturing this strong coreference. One option, that looks very tempting, would be to consider this as a case of Raising. Another option, considered and argued for in (Pancheva Izvorski, 2000), treats the WHECs, and the coreference between the two Subjects, as a case of Obligatory (Exhaustive) Control.

In this section I argue that the Subject of the embedded verb (in the wh-complement) is best treated as a case of Raising, rather than Control, contra Pancheva Izvorski (2000)'s arguments. I also provide reasons that the Raising treatment for the Subject provides the right treatment for arguments with Nominative Case in the wh-complement, be it overt Subjects, or nominative forms of 'who'.

5.1 Pancheva Izvorski's arguments against Raising

Here I briefly look at Pancheva Izvorski's arguments against considering Raising in WHECs. It is based on several parts: 1) arguing that "A-movement out of the *wh*-clauses is theoretically problematic", 2) trying to account for the "lack of agreement between *Mary* and *be*", and 3) the

[5] This quotation is taken from Dotlačil (2005).

fact that "*Mary* cannot appear in the nominative when the matrix predicate is *be*" (p. 65)[6] To these three arguments, she then adds the lack of availability of expletives in WHECs (standardly associated with cases of Raising), and the lack of truth-functional equivalence between active and passive sentences.

I first look at Pancheva's argument from the lack of Nominative Case and agreement, then I look at the lack of expletives and truth-functional equivalence between active and passive sentences, and then at the argument about the "theoretically problematic" A-movement out of the wh-complements.

In the end, those arguments, as they stand, turn out not to be tenable, or relevant, for the cases of WHECs in Czech that are considered in this paper. I provide the way to treat (overt) Subjects in WHECs in section 5.2.

5.1.1 'Be' taking more than one (nominal) argument?

I look first at the arguments 2) and 3), i. e. the (im)possibility of 'be' assigning Nominative Case together with showing agreement (or, taken together, taking an overt Subject in addition to the wh-constituent) in WHECs. I show that the cases that Pancheva Izvorski (2000) considers are not surprising at all, given the properties of the matrix verb 'be' and thus cannot be used as good arguments against Raising.

It is worth noting that the verb 'be', unlike the verb 'have', can only take one (nominal) argument in general[7], as well as in WHECs:

(28) a. Byla tam hruška.
be$_{3sgf.past}$ there pear$_{nom.sg}$
'There was a pear there.'

b. * Josef byl/-a tam hruška/hrušku.
Josef$_{nom.sg}$ be$_{3sgm/f.past}$ there pear$_{nom.sg}$/pear$_{acc.sg}$

c. Josef měl hrušku.
Josef$_{nom.sg}$ have$_{3sgm.past}$ pear$_{acc.sg}$
'Josef had a pear.'

(29) a. Bylo tam co jíst.
be$_{3sgn.past}$ there what eat$_{inf}$
'There was something that one can eat/that can be eaten.'

b. * Josef byl/-o (tam) co jíst.
Josef$_{nom.sg}$ be$_{3sgm/n.past}$ what eat$_{inf}$

c. Josef měl co jíst.
Josef$_{nom.sg}$ have$_{3sgm.past}$ what eat$_{inf}$
'Josef had something that he could eat.'

Based on that property of the matrix verb 'be', the absence of Nominative Case on *Mary* is accounted for because 'be' can only assign Case to one argument and in WHECs it, by definition, takes the WH-constituent as its argument. To the extent that the wh-constituent as an argument needs a Case, there isn't any other Case that 'be' could assign (and, as shown in (10), infinitives in Czech do not assign Nominative Case either).

[6] As much as one would like to see an example of either one, or both, 2) and 3), there is none, not even ones showing their non-existence or ungrammaticality, provided in Pancheva Izvorski (2000).

[7] The existential verb 'be' in Czech, as well as in e. g. English, takes a locative (adverbial) argument as well. For the purposes of this paper I am going to put that issue aside.

Parallel to the inability of 'be' to assign Nominative Case to a 'Subject', the lack of agreement is not surprising since assigning Case is standardly taken to go hand-in-hand with Agreement (both happening through the checking of features through the operation Agree for Chomsky, 2000). Thus neither the lack of a Nominative Case, nor the lack of agreement with the matrix verb 'be' argues against Raising in WHECs.

5.1.2 Expletives

Pancheva Izvorski (2000) provides the prohibition of weather-subjects, as well as expletives as a support for her argument in favor of Control, rather than Raising as the basis for the coreference of the matrix and embedded Subjects.

The cases of impossible "weather-subjects" for Pancheva Izvorski include the following ones from Russian (Pancheva Izvorski, 2000, p. 65, ex. 91):

(30) * Est' kogda idti dožd / doždju'.
be$_{3sg}$ when go$_{inf}$ rain$_{nom}$ / rain$_{dat}$
'There is a time such that it can rain then.'

This case is bad in Czech too, but there still are other potential cases to warrant the other conclusion (arguing for Raising rather than Control):

(31) a. * Je / není kdy pršet.
be$_{3sg.pres}$ / not-be$_{3sg.pres}$ when rain$_{inf}$

b. ? Má / nemá už kam pršet.
have$_{3sg.pres}$ / not-have$_{3sg.pres}$ already where-to rain$_{inf}$
'There is/isn't already place that could take rain.'

c. Je / není co jíst.
be$_{3sg.pres}$ / not-be$_{3sg.pres}$ what eat$_{inf}$
'There isn't anything one could eat.'

What is even more important for the case provided here in (31c) is that a version of that is also possible with an expletive[8]:

(32) (V)ono nebylo co jíst.
it$_{nom.sg-expl.}$ not-be$_{3sgn.past}$ what eat$_{inf}$
'There wasn't anything one could eat.'

Thus the argument from the lack of possibility of expletives, as well as no "weather-subjects" is not as clear-cut as one could hope for.

Another comment worth making about the possibility of Raising from the (infinitival) wh-complement is that the matrix verb does play a role here. Given that Raising is a name for movement for Case, it can only happen with verbs that can possibly assign a Nominative Case. That is exactly where 'be' and 'have' as matrix verbs differ in Czech, as already argued above in section 5.1.1. Thus, Raising is only a matter to be considered with the matrix verb 'have', not 'be'.[9]

[8] For more on the expletive "(v)ono" in Czech see Rezac (2004).

[9] A rather interesting comment on this topic was made to me by Norvin Richards (p.c.), on the note that 'have' could be treated as 'be' plus Raising. I leave that comment, while interesting, aside for now.

5.1.3 No A-movement out of wh-clauses?

Here I will look at what "theoretical" evidence against Raising out of *wh*-clauses there is that warrants Pancheva Izvorski's conclusion and give evidence that A-movement out of the *wh*-constituents in WHECs is possible (parallel to evidence of such movement in Russian provided by Babby (2000)).

To start with, it has been generally accepted that A-movement out of CPs is theoretically problematic. However, I have already argued (and shown) in section 4 that the wh-constituent (from which the Raising would need to occur here) does need to be smaller than a CP since it independently shows other properties that are restricted to infinitival complements smaller than a CP (e.g., clitic climbing). If the wh-constituent is indeed smaller than a CP, it is worth looking at the possibility of Raising in WHECs again.

On the same note are arguments by Babby (2000), that show that the "understood" Subject of the infinitive is in fact a true Subject of the infinitive that raises higher and ends up being in a position where one might expect the Subject of the matrix verb 'be'.

The cases that Babby (2000) is looking at are Russian WHEC cases, called by Babby Infinitival Existential Sentences, or IES; examples are given in (33), (Babby, 2000, page 1, ex. (1)).

(33) a. Nam est' gde spat'.
us$_{dat}$ there-is where to-sleep
'There is somewhere for us to sleep.'

b. Nam negde spat'.
us$_{dat}$ nowhere to-sleep
'There is nowhere for us to sleep.'

Babby's arguments go along the following lines. First of all, Babby adopts the statement that "infinitive clauses in Russian all have a dative subject, overt or null" (Babby, 2000, p.3). If the Subject "is overt, it is normally fronted, occupying the Spec position on the sentence's highest functional projection" (*ibid*).
Based on that Babby (*ibid*) argues that the structure in (34b) is preferable over the structure in (34a):

(34) a. Nam$_i$ [$_{IP}$ [$_{VP}$ [$_V$ est'] [$_{CP}$ gde$_j$ [$_{IP}$ t$_i$ spat' t$_j$]]]]
b. [$_{IP}$ [$_{VP}$ [$_V$ est'] [$_{CP}$ gde$_j$ [$_{IP}$ PRO$_i$ spat' t$_j$]]]]

Further "reasons for assuming that the dative in (1) [(33) here] is in fact the subject of the infinitive" (*ibid*). First reason is that "pronouns other than the dative pronouns systematically prepose out of the infinitival clause" – as e. g. in (35), (Babby, 2000, p. 4, ex. (6), (7)), which would result in having to claim that the verb 'be' in Russian can take Subjects marked with different Cases (at least accusative and dative ones).

(35) a. Menja$_i$ nekomu bylo [$_{CP}$ vstrečat' t$_i$]
me$_{acc}$ no-one$_{dat}$ was$_{3sgn}$ to-meet
'There was no one to meet me.'

b. Ne dumaju čtoby teper' nas$_i$ bylo [$_{CP}$ za čto upreknut' t$_i$].
'I don't think that there is anything to rebuke us for now.'

Another argument for Babby is that when the wh-word (K-word in Babby (2000)) is itself the dative subject of the infinitive, no other (non-wh) dative argument of *est'* can occur. The

last argument Babby (*ibid*) gives concerns the properties of the matrix verb 'be' in Russian negative WHECs where it occurs in the copula version which cannot assign theta roles. Thus the only verb that could assign a theta role to the dative arguments is the infinitival verb here.

Thus for Russian, Babby concludes, the most reasonable way of treating the dative subjects is treating them as subjects of the infinitival verb in the complement of the existential verb, and systematically preposing them outside of the wh-complement.

5.1.4 Conclusion

While Pancheva Izvorski (2000) presents a fairly convincing argument for Control in WHECs for Bulgarian, the conclusion is still somewhat half-way for Czech based on the sections above. In the next section, I provide some more arguments for Raising (parallel to some extent to the Russian cases provided by Babby, 2000) in Czech WHECs and a study of arguments with a Nominative Case in the wh-complements in section 5.2.

5.2 Nominative arguments/wh-words in the (non-finite) wh-complement

Even though the infinitival verb does not allow an overt Subject in Czech (as noted in section 2), the wh-existential construction, in certain and restricted cases, does allow an overt Subject in the wh-complement. Both a nominative wh-word (who$_{nom}$, which is clearly morphologically marked as nominative), as well as a name, or other nominals, are available.

(36) a. Nemá kdo přijít včas.
not-have$_{3sg.pres}$ who$_{nom}$ come$_{inf}$ on-time
'There isn't anyone to/who would come on time.'

b. (Budou hrát,) dokud (se) bude mít s kým (se)
be$_{fut.3pl}$ play$_{inf.}$ until (refl$_{acc-cl.}$) be$_{fut.3sg}$ have$_{inf.}$ with who$_{inst.sg}$ (refl$_{acc-cl.}$)
Petr utkat.
fight$_{inf.}$
'(They will play) as long as Petr will have someone to play with.'

While it seems that both 'be' and 'have' are available and can occur with a Nominative argument in the wh-complement, the situation is not that simple. While cases with 'be' as the matrix verb and an infinitival form of the verb in the wh-complement are fine (with either a nominal or adjectival wh-word), they do not allow a Nominative Case to be assigned to an argument (compare the ungrammatical cases in (37) with parallel ones with the matrix verb 'have' in (38)):

(37) a. Tady (si) není kam (si) sednout.
here refl$_{dat-cl.}$ not-be$_{pres.3sg}$ where refl$_{dat-cl.}$ sit-down$_{inf.}$
'There is no place to sit down here.'

b. Sem není koho pozvat.
here not-be$_{pres.3sg}$ who$_{acc}$ invite$_{inf.}$
'There is nobody to/we could invite here.'

c. * Není kdo přijít brzo.
not-be$_{3sg.pres}$ who$_{nom}$ come$_{inf.}$ early

d. ?? Tady (to) není kdo (to) uklidit.
here (it$_{acc}$) not-be$_{3sg.pres}$ who$_{nom}$ (it$_{acc}$) clean$_{inf.}$

e. (Budou hrát) dokud bude s kým se (*Petr) utkat.
be$_{fut.3pl}$ play$_{inf.}$ until be$_{fut.3sg}$ with who$_{inst.}$ refl$_{acc-cl}$ Petr$_{nom.sg}$ fight$_{inf.}$
'They will play as long as there will be someone one could fight with.'

(38) a. ? Nemá kdo přijít brzo.
not-have$_{3sg.pres}$ who$_{nom.sg}$ come$_{inf}$ early
'There isn't anyone who could come early.'

b. Tady (to) nemá (to) kdo uklidit.
here (it$_{acc}$) not-have$_{3sg.pres}$ (it$_{acc}$) who$_{nom}$ clean$_{inf}$
'There isn't anyone who could clean it here.'

c. (Budou hrát) dokud (se) bude mít s kým (se)
be$_{fut.3pl}$ play$_{inf.}$ until (refl$_{acc-cl}$) be$_{fut.3sg}$ have$_{inf}$ with who$_{inst.}$ (refl$_{acc-cl}$)
Petr utkat.
Petr$_{nom.sg}$ fight$_{inf.}$
'They will play as long as Petr will have someone he could fight with.'

The question that this situation raises then is what assigns the Nominative Case, since it has already been shown that the infinitival verb itself cannot do that. The options are basically two: either it is provided from the matrix verb, or some other null verb (e.g., a modal which needs to be present in the construction anyway, see the initial overview of WHECs in section 1).

One way of checking these two options is trying to check whether the tense of the matrix verb has any effect on the status of the sentence, since turning the matrix verb into an infinitival form should strip it of the ability of assigning Nominative Case.[10] As the next case shows, if we try to put an infinitival matrix 'have' in the construction, the Nominative Case is no longer available, while other Cases do stay available[11]:

(39) a. Že by měl kdo přijít (se mi zdálo
that aux$_{subj.3sg}$ have$_{pprt.3sgm}$ who$_{nom}$ come$_{inf.}$ refl$_{acc-cl.}$ I$_{dat-cl.}$ seem$_{past.3sgn}$
nepravděpodobné).
not-likely
'It didn't seem to me to be likely for anybody to come.'

b. * Mít kdo přijít se mi zdálo nepravděpodobné.
have$_{inf.}$ who$_{nom}$ come$_{inf.}$ refl$_{acc-cl.}$ I$_{dat-cl.}$ seem$_{past.3sgn}$ not-likely

c. Mít s kým si povídat je jedna z nejdůležitějších
have$_{inf}$ with who$_{inst}$ refl$_{dat-cl}$ talk be$_{3sg.pres}$ one from most-important$_{gen}$
věcí v životě.
thing$_{gen.pl}$ in life
'To have someone one can talk to is one of the most important things in life.'

Another possible argument for claiming that it is the matrix verb 'have' that assigns the Nominative Case would be looking at the possibility of another nominative argument/Subject of a finite 'have'. Parallel to the cases of Russian dative subjects, mentioned in section 5.1.3, a nominative form of *who* in the wh-complement blocks any other nominative Subject in WHECs:

[10] I would like to thank David Pesetsky (p.c.) for suggesting this to me.

[11] The point of the other Cases is to show that the infinitive form of the matrix verb 'have' does not interfere with other properties of the construction.

(40) * Josef neměl kdo přijít.
Josef$_{nom.sg}$ not-have$_{3sgm.past}$ who$_{nom}$ come$_{inf}$

Even though (or because of the case that) the sentences are grammatical, provided the matrix verb ('have') is finite and there isn't any other argument requiring a Nominative Case, it still needs to be specified how the Nominative Case can be assigned into the wh-complement (where the Subject is located in the examples given above).

The fact that the nominative arguments can be found in the wh-complement further strengthens the case for Raising in Czech, given that if the Subject of the infinitival verb was just a PRO, there would be no reasonable way of lowering the argument/wh-word to the place where it is found. I argue that the Case in these examples is assigned in situ, without the noun/wh-word raising higher to be assigned the Case.

6 Conclusion

In this paper I have shown that the wh-existential constructions in Czech have two options of the non-finite form of the verb in the wh-complements of matrix verbs 'be' or 'have', namely they can appear in subjunctive or infinitival forms.

I have argued that the wh-words that have to be present in the wh-existential constructions are determined in their form by the verb in the wh-complement; the wh-word are also allowed to long-distance wh-move out of the wh-complement, provided that at least one wh-word is left inside the wh-complement; when one or more wh-words moves out of the wh-complement with an infinitival form of the verb, the sentence gains another alternative of interpretation – one of a have-to construction.

I have further argued that the wh-complement in wh-existential constructions has to have a structure smaller than CP, since it allows restructuring that is the prerequisite of clitic climbing which is freely available in wh-existential constructions.

Based on the previous conclusions I have also shown that the Subjects of the matrix and the embedded (non-finite) verb are related through Raising, rather than Control.

References

Aissen, Judith L. and David M. Perlmutter. 1983. Clause reduction in spanish. Perlmutter, D. M. (ed.). *Studies in Relational Grammar* 1:360–403. Chicago: University of Chicago Press.

Babby, Leonard. 2000. Infinitival existential sentences in Russian: A case of syntactic suppletion. King, T. H. and I. A. Sekerina (eds.). *Formal Approaches to Slavic Linguistics: The Philadelphia Meeting 1999*:1–23. Ann Arbor: Michigan Slavic Publications.

Bhatt, Rajesh. 1998. Obligation and possession. Harley, H. (ed.). *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect. MIT Working Papers in Linguistics* 32:21–40. Cambridge, Mass: MITWPL.

Caponigro, Ivano. 2001. On the semantics of indefinite free relatives. van Koppen, M., J. Sio & M. de Vos (eds.). *Proceedings of ConSOLE X*:49–62. Leiden: SOLE.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. Martin, R., D. Michaels & J. Uriagereka (eds.). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*: 89–155. Cambridge, MA: The MIT Press.

Dotlačil, Jakub. 2005. Non-local binding in Slavic languages and restructuring. Blaho, S., L. Vicente & E. Schoorlemmer (eds.). *Proceedings of ConSOLE XIII*:1–16. Leiden, Netherlands: SOLE.

--. In progress. *Why clitics cannot climb out of CP*.

Grosu, Alexander and Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6:125–170.

Izvorski, Roumyana. 1998. Non-indicative wh-complements of possessive and existential predicates. Tamanji, P. N. & K. Kusumoto (eds.). *Proceedings of NELS* 28:159–173. Amherst, MA: GLSA.

Junghanns, Uwe. 1999. Generative Beschreibung periphrastische Konstruktionen des Tschechischen. Anstatt, T., R. Meyer & E. Seitz (eds.). *Linguistische Beiträge zur Slavistik aus Deutschland und Österreich. VII. JungslavistInnen-Treffen. Tübingen/Blaubeuren 1998*: 133–165. München: Sagner.

Lenertová, Denisa. 2004. Czech pronominal clitics. *Journal of Slavic Linguistics* 12: 135–171.

Pancheva Izvorski, Roumyana. 2000. Free relatives and related matters. Ph.D. thesis. University of Pensylvania.

Rappaport, Gilbert. 1986. On a persistent problem of Russian syntax: Sentences of the type *Mne Negde Spat'*. *Russian Linguistics* 10:1–31.

Rezac, Milan. 2004. *Elements of cyclic syntax: Agree and Merge*. Ph.D. thesis. University of Toronto, Toronto.

--. *The syntax of clitic climbing in Czech*. http://minimalism.linguistics.arizona.edu/AMSA/PDF/AMSA-202-0602.pdf.

Rudin, Catherine. 1986. *Aspects of Bulgarian Syntax: Complementizers and* Wh-*Constructions*. Columbus, OH: Slavica.

Wurmbrand, Susi. 1998. *Infinitives*. Ph.D. thesis, MIT.

Only, bound variables and VP ellipsis in Czech*

Mojmír Dočekal

docekal@phil.muni.cz

1 Introduction

This paper presents evidence that a possessive anaphor *svůj* in Czech can be interpreted either as a bound variable (which is unproblematic) or as a covaluated pronoun.[1] The distinction between bound variables and covaluated expressions surfaces in Czech sentences with the focus sensitive particle *jen/pouze* (which corresponds to the famous *only* in English), but disappears in the VP ellipsis contexts. This is a bit puzzling, because VP ellipsis and *only* are both contexts where the distinction between bound variables and covaluated pronouns can be detected, at least in English.

In the first sections I will examine the syntax and semantics of VP ellipsis and then I will investigate the behavior of Czech pronouns and anaphors in the context of VP ellipsis and the *jen/pouze* expressions.

Let us introduce basic data. Czech does not have the equivalent of VP-ellipsis (VPE) in English. VPE with past tense auxiliaries and conditional auxiliaries is ungrammatical in Czech:

(1) a. Mary is leaving and John is $[_{VP}$ e] too.
b. Mary hasn't left, and John has $[_{VP}$ e].

(2) a. *Past tense auxiliaries*
*Já jsem odešel a ty jsi taky.
I AUX leave and you AUX too
'I left and you did too.'
b. *Conditional auxiliaries*
*Já bych odešel a ty bys taky.
I COND leave and you COND too
'I would leave and you would too.'

VP-ellipsis in Czech is grammatical only with lexical or semi-lexical verbs, like a future auxiliary, or a copula:

(3) a. *Lexical verb:*

*The present study is a part of the project GAČR 405/07/P252, the goal of which is an analysis of the Czech negation and negative polarity items.

[1]See Reinhart (2000) for the distinction between bound variables and covaluated pronouns.

Karel umí odejít, ale Petr neumí.
Karel knows leave but Petr neg-knows
'Karel know how to leave, but Petr does not'

b. *Future auxiliary:*
Karel bude odcházet, ale Petr nebude.
Karel will leave but Petr neg-will
'Karel will leave, but Petr will not.'

c. *Copula:*
Karel je učitel, ale Petr není.
Karel is teacher but Petr neg-is
'Karel is a teacher, but Petr is not.'

This is exactly the difference between *high verbal forms* (conditional auxiliaries, past tense auxiliaries) and *low verbal forms* (main finite verbs, future auxiliaries), as proposed in Veselovská (2004).

So Czech is somewhere between English, with full VPE, and Romance languages, where VPE is not possible with any auxiliary (see Zagona (1982) and Busquets 2005).

2 Syntax of VPE: stripping and VPE

According to Hankamer and Sag (1976: 409) *stripping* is a rule that deletes everything in a clause under identity with corresponding parts of the preceding clause (*the correlate*), except for one constituent (*the remnant*). The distinction between VPE and stripping is that VPE is the ellipsis of the whole verb phrase, which includes the verb and any objects or modifiers it might take. VP ellipsis is licensed by the by the immediately preceding auxiliary (in English) or by the polarity particles (*ne*, *taky*, *ano* in Czech). (4) shows examples of stripping and (5) shows examples of VPE.

(4) a. Petr dal knihu Karlovi, ale taky Marii.
Petr gave book Karel but too Mary
'Petr gave a book to Karel, but also to Mary.'

b. Petr políbil Marii, ale ne Karla.
Petr kissed Mary but not Karel
'Petr kissed Mary, but not Karel.'

(5) a. Petr dal knihu Karlovi, ale Marie ne.
Petr gave book Karel but Mary not
'Petr gave a book to Karel but Mary did not.'

b. Petr políbil Marii a Karel taky.
Petr kissed Mary and Karel too
'Petr kissed Mary and Karel did too.'

The fundamental properties of stripping in Czech are the following (see McShane 2000 and Busquets 2005):

1. Unlike VPE, stripping is not allowed in subordinate clauses:

(6) *Stripping:*
a. *Petr dal knihu Karlovi, protože ne Kláře.
Petr gave book Karel because not Klara

'Petr gave the book to Karel because not to Klara.'

b. *Petr políbil Marii, protože ne Karla.
Petr kissed Mary because not Karel
'Petr kissed Mary because not Karel.'

(7) *VPE*:

a. Petr dal knihu Karlovi, protože Klára taky.
Petr gave book Karel because Klara too
'Petr gave the book to Karel because Klara did too.'

b. Petr políbil Marii, protože Karel taky.
Petr kissed Mary because Karel too
'Petr kissed Mary because Karel did too.'

2. VPE is a constituent operation, stripping is not.

3. Unlike VPE, stripping does not appear to conform to the backward anaphora constraint (i. e. the remnant cannot precede the correlate):

(8) a. *VPE:*
Přestože Karel neumí, Petr umí hrát na piáno.
Although Karel neg-knows Petr knows play on piano
'Although Karel does not know, Petr does know how to play the piano.'

b. *Stripping*:
*Ale ne Kláře, Petr dal knihu Karlovi.
But not Klara, Petr gave book Karel
'Although not to Klara, Petr gave the book to Karel.'

4. According to Busqutes (2005), stripping is constrained by its interaction with focus, but this is not the case for VPE:

(9) a. *Stripping:*
Petr dal květiny [$_F$ Kláře], ale ne Karlovi.
Petr gave flowers Klara but not Karel
'Petr gave the flowers to Klara but not to Karel.'

b. #Petr dal květiny [$_F$ Kláře], ale ne knihu.
Petr gave flowers Klara but not book
'Petr gave flowers to Klara, but not the book.'

(10) *VPE:*
Petr dal květiny [$_F$ Kláře], ale Karel ne.
Petr gave flowers Klara but Karel not
'Petr gave the flowers to Klara, but Karel did not.'

This is not exactly true in Czech, as we will see later. But the restriction, which was formulated in Busquets (2005: 14) for Catalan, holds also in Czech: "The focused expression constructs a set of alternatives C, if the remnant does not belong to C, then the sentence is infelicitous. By contrast, the position of focus in TP-Ellipsis is not a necessary condition in order to recover the missing VP. "

In Czech, as in Romance languages, the VPE can be licensed by polarity particles. For Catalan Busquets 1997 points to following examples:

(11) En Pere anirà a Paris, i la Marie també.
'Pere will go to Paris and Maria (will) too.'

(12) El Rafel va convidar el Ricard al teatre, però al cinema no
'Rafel invited Ricard to the theatre, but (he didn't) to the movies not.'

In Czech polarity particles similar to Catalan *també, no* can be found. Their role is analogous to English auxiliaries in VPE:

(13) a. Petr přišel, ale Karel ne.
Petr came but Karel not
'Petr came, but Karel did not.'
b. Petr přišel a Karel taky.
Petr came and Karel too
'Petr came and Karel did too.'

There is sometimes proposed a dedicated functional projection for the polarity particles (see Laka 1990) which is also used for focus licensing. Laka (1990) postulates two different *Polarity phrases* (ΣP), one between VP and TP and one between TP and CP. If we assume that Czech polarity particles in VPE are located in the lower ΣP, then we predict the difference between high verbal forms and low verbal forms in their ability to occur in VPE. High verbal forms are base generated in T and low verbal forms are base generated somewhere lower, e.g.. in v (future auxiliary) and in V (main verbs). The exact location of low verbal forms is irrelevant for the purposes of this paper; it is sufficient to assume that low verbal forms are base generated below ΣP.

This is of course the same phenomena as the negation test used in Veselovská (2004). The negation test distinguishes between low verbal forms and high verbal forms – the later can bear negation, the former cannot.

(14) a. *Já nejsem přišel.
I neg-AUX came
'I did not come.'
b. Já jsem nepřišel.
I did neg-came
'I did not come.'

(15) a. *Já nebych přišel.
I neg-AUX-COND came
'I did not come.'
b. Já bych nepřišel.
I AUX-COND neg-came
'I would not come.'

(16) a. Já nebudu chodit.
I neg-AUX come
I will not come.
b. *Já budu nechodit.
I AUX neg-come
'I will not come.'[2]

[2]This is ungrammatical in the sentential negation reading.

(17) a. Já nejsem student.
I neg-am student
'I am not a student.'
b. *Já jsem nestudent.
I am neg-student
'I am not a student.'

If we assume head movement through ΣP for negated verbs, then high verbal forms cannot be negated, because the T head is above Σ head thoughout a derivation.

3 VPE

Czech pronouns displays a strict/sloppy identity ambiguity in VPE. Compare the standard case shown in (18), which can be interpreted in one of two ways: either John's mother went to Peter's house (strict) or she went to John's house (sloppy). On the other hand, Czech anaphors (possessive or not) usually involve only sloppy identity. Compare the data in (19).

(18) Petrova matka šla do jeho domu a Honzova taky.
Peter's mother went to his house and John's too
'Peter's mother went to his house and John's did too.'

(19) Petrova matka šla do svého domu a Honzova taky.
Peter's mother$_i$ went to her$_i$ house and John's too
'Peter's mother$_i$ went to her$_i$ house, and John's did too.' [sloppy/*strict identity]

This is quite unproblematic if we assume two distinct procedures for pronoun resolution: binding and covaluation (see Reinhart 2000). In the first, we close the property: a common technical implementation is that the variable gets bound by the λ-operator. In the second, the free variable is assigned a value, say, from the discourse storage.

Bound variables in natural language must be c-commanded by their antecedents, but covaluation is a pragmatic phenomena, so it is constrained by saliency and other pragmatic notions. From this it follows that a intersentential pronoun dependency must always be covaluation, but in one sentence there are two possibilities: either binding or covaluation. (20) is ungrammatical, because quantifiers cannot refer, so covaluation is forbidden, but binding is forbidden also, because there is a sentence boundary between the quantifier and the pronoun. In contrast, (21) is grammatical, because the quantifier c-commands the pronoun.

(20) [Každý člověk]$_1$ je smrtelný. *On$_1$ je také racionální.
'[Every man]$_1$ is mortal. *He$_1$ is also rational.'

(21) [Každý člověk]$_1$ doufá, že ho$_1$ ostatní chápou.
'[Every man]$_1$ hopes that the others understand him$_1$.'

The pronoun *ho* in (21) must be interpreted as a bound variable, because the quantifier cannot refer, so the covaluation option is impossible. In contrast, the pronoun *on* in (22) must be interpreted as a covaluation for the exactly opposite reasons.

(22) [Ten člověk]$_1$ je smrtelný. On$_1$ je také racionální.
'[The man]$_1$ is mortal. He$_1$ is also rational.'

4 Possessives

A well known disambiguator for the contrast between binding and covaluation is the focus particle *only* as in (23), where distinct truth conditions result from the two interpretations of pronouns. Intuitively (23-a) means that Karel is the only person in the universe of discourse who thinks that his dog is amazing. In contrast, (23-b) means that Karel is the only person in the universe of discourse who thinks that Karel's dog is amazing.

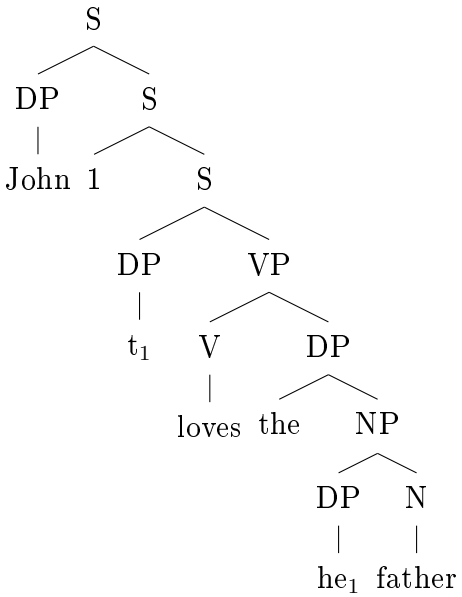
(23) Jen Karel si myslí, že jeho pes je úžasný.
Only Karel thinks that his dog is amazing
a. binding: Only Karel (λx (x think, that x's dog is amazing));
b. covaluation: Only Karel (λx (x think, that y's dog is amazing) & y = Karel)

In English the pronoun *his* can also have the mentioned interpretations. According to Heim & Kratzer (24) has three following structures. (25) is a tree which corresponds the binding semantics for *his*: *John* after quantifier raising yields the index 1, which is the H&K system translatable as the λ abstraction with the variable corresponding to the index 1 and because *he* has the same index, the semantics is straightforward: John (λx (x loves x's father)).

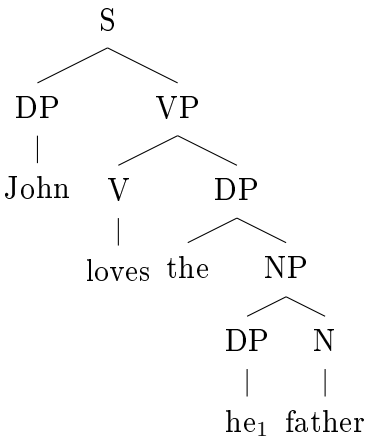
For the covaluation option there are two possibilities: (i) either *John* does not undergo quantifier raising (QR), so there is no λ abstraction and the pronoun remains free – this is (26), or (ii) *John* undergoes QR, but the indexes are different – as in the tree (27). Both trees correspond to the covaluation reading: John (λx (x loves y's father) & y = John), where the "& y = John" part of the formula is supplied from the pragmatics, but not from the semantics.

(24) John loves his father. Heim & Kratzer (1998: 246)

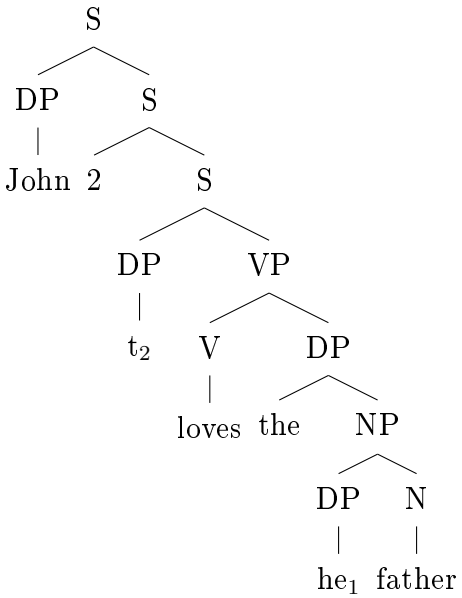
(25)



(26)

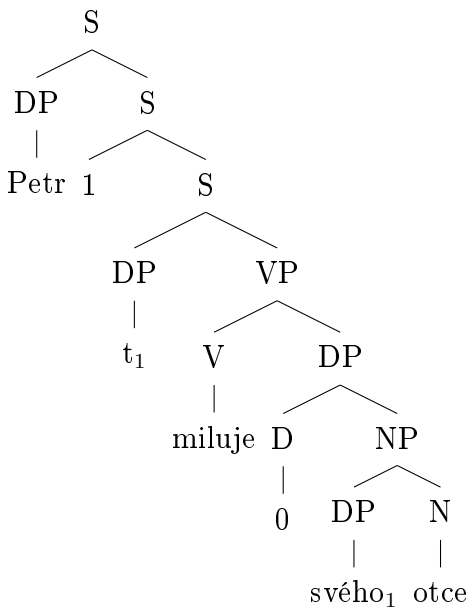


(27)



Now the hypothesis is that for Czech possessive anaphors, only the tree (28) is viable, because the binding theory demands it – after all *svůj* is an anaphor, so there is nothing really unexpected. English *his* is a pronoun, thus can take its antecedent from the discourse storage, which is not an option for the anaphor.

(28)



Now let us return back to VPE. We can account for the data in (18) and (19), if we assume that Czech pronouns can be interpreted either as bound variables or as covaluated free variables; but Czech anaphors can be interpreted only as bound variables. (18) can have three interpretations:

1. strict identity to Peter (covaluation);
2. strict identity to John (covaluation);
3. sloppy identity (binding).

The fourth interpretation of (18) (Peter's mother went to John's house and John's mother went to Peter's house) is ungrammatical. This follows from any theory of VP ellipsis which states that the ellipsis is dependent on identity in LF, where identity means either identity of binding or covaluation (if the antecedent is interpreted as bound, then the elided constituent is too and the same for the covaluation) – see Fiengo & May (1994) and Busquets (1997).

The LF identity of elided material with its antecedent can be seen independently from (29), where the non-elided VP has two possible interpretations: (29-a) or (29-b), but if we chose (29-a), then we must chose the wide scope interpretation of DP *nějaký dárek* in the elided VP also. And the same holds for (29-b).

(29) Klára dala nějaký dárek každému návštěvníkovi a Bára taky.
'Klara gave a present to every visitor and Bara did too.'
a. $\exists$x(present'(x) $\supset\forall$y(visitor'(y) $\supset$ give'(Klara',x,y)))
b. $\forall$y(visitor'(y) $\supset \exists$x(gift'(x) $\supset$ give'(Klara',x,y)))

The account of the LF identity of the elided VP with its antecedent holds also for the bound/covaluation distinction. This is confirmed by the following data. I have asked 20 Czech native speakers for their grammaticality judgments for (30) and the results are the following:

1. Strict reading, denotation of both pronouns is Karel: 50% of participants.
2. Strict reading, denotation of both pronouns is Petr: everybody agreed that it is a possible interpretation, but for nobody was it the first reading.
3. Sloppy reading: 50% of participants.
4. Strict reading, denotation of the non-elided pronoun as Petr and denotation of the elided pronoun as Karel: impossible reading for everybody.

(30) Karlova matka uklízí jeho knihovnu, ale Petrova ne.
'Karel's mother cleans his library, but Peter's does not.'

The explanation is straightforward, if ellipsis must preserve the binding/covaluation distinction:

1. For the covaluation reading the referential indexes must be identical (reading 1 and 2); this is translated as the identity of variables:
   a) reading 1:
   Karel (λy (y's mother (λx (x cleans z's library) & z = Karel))) & Petr (λy (y's mother (λx (x ¬cleans z's library) & z = Karel)))
   b) reading 2:
   Karel (λy (y's mother (λx (x cleans z's library) & z = Petr))) & Petr (λy (y's mother (λx (x ¬cleans z's library) & z = Petr)))
2. For the binding reading (reading 3) the indexes can be different, but the λ abstraction of the elided VP must be the alphabetical variant of its antecedent:
   Karel (λy (y's mother (λx (x cleans y's library)))) & Petr (λw (w's mother (λz (z ¬cleans w's library))))
3. The fourth reading is forbidden, because the covaluation is not possible and binding is not possible either, since the DP *Petrova matka* does not c-command the first possessive pronoun.

Discourse Representation Theory (DRT, see Kamp & Reyle 1993) offers quite sensible formalization for these facts. VPE is naturally treated here as a subtype of anaphora (see Klein 1987), where VP as a predicate is represented by a separate subDRS, where a distinguished discourse referent represents λ-bound variable. This anaphoric treatment of VPE is quite natural, because it predicts some of the properties of VPE:

1. Impossibility of VPE in subordinate sentences: this follows from the anaphoricity of VPE, because this would lead to inappropriate self-embedding (the antecedent would contain the elided constituent).
2. Pragmatic summation of two or more VP into antecedent for elided VP: *Petr became ill. He was taken to hospital. Karel did too.*

In this DRT theory of VPE, for example (30), can be represented as (31), where $[x_4]$ is a distinguished discourse referent, P is the subDRS, where $[x_4]$ is the distinguished discourse referent. Conditions (i) and (ii) represent excluding alternatives – only one of them can be chosen, if (i) is chosen, then the sloppy reading results, if (ii) is chosen, then the strict reading emerges.

(31)

$x\ x_1\ y\ P\ Q$
karel(x)
mother(y)
r(x,y)
P(x)
P: $\boxed{\begin{array}{\|c\|}\hline [x_4]\ x_3\ x_2 \\ \hline \text{library}(x_3) \\ \text{r}(x_3,x_2) \\ \text{clean}(x_4,x_3) \\ \text{(i) } x_2{=}x_4 \\ \text{(ii) } x_2 = x \\ \hline \end{array}}$
petr(x_1)
$\neg$ $\boxed{\begin{array}{\|c\|}\hline \text{do}(x_1,Q) \\ Q{=}P \\ \hline \end{array}}$

Realizations for each individual case:

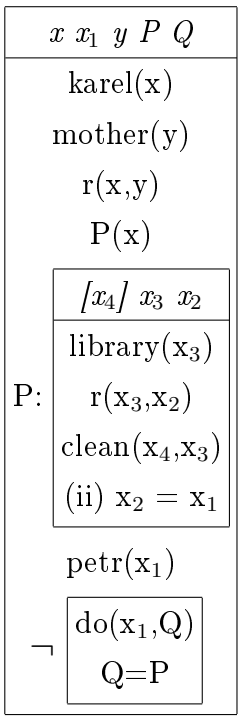
1. The first reading (strict denotation to Karel):

(32)

$x\ x_1\ y\ P\ Q$
karel(x)
mother(y)
r(x,y)
P(x)
P: $\boxed{\begin{array}{\|c\|}\hline [x_4]\ x_3\ x_2 \\ \hline \text{library}(x_3) \\ \text{r}(x_3,x_2) \\ \text{clean}(x_4,x_3) \\ \text{(ii) } x_2 = x \\ \hline \end{array}}$
petr(x_1)
$\neg$ $\boxed{\begin{array}{\|c\|}\hline \text{do}(x_1,Q) \\ Q{=}P \\ \hline \end{array}}$

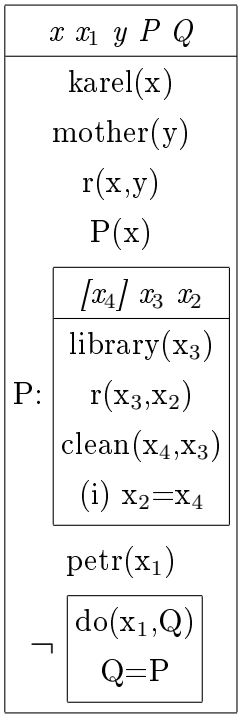
2. The second reading (strict denotation to Petr):

(33)



3. The third, sloppy reading:

(34)



Interestingly sometimes even Czech possessive anaphors can behave as covaluated, as can be seen from (35) and its two possible continuations. *Only* in Czech also seems to change the argument structure: *-able* adjectives are usually taken as agent absorbing, but with *only* the agent is again possible – see (36-b).

(35) Jen prof. Novák nedoporučuje své knihy, ale ostatní profesoři své/Novákovy
Only prof. Novak neg-recommends his books, but other professors their/Novak's
knihy doporučují.
books recommend
'Only prof. Novak does not recommend his books, but other professors do.'

(36) a. *Ten brouk je pozorovatelný Karlem.
The beetle is observable by Karel
b. Ten brouk je pozorovatelný jen Karlem.
The beetle is observable only by Karel

The central idea is that *only* in Czech triggers a background-presupposition rule which states that whenever focusing gives rise to a background λx.ϕ(x), there is a presupposition to the effect that λx.ϕ(x) holds of some individual. This can explain (36-b) (if we take *only* A B as $\neg\exists x[\neg Ax \wedge Bx]$, where B is presupposed).

But for the purposes of this article the most important point is that (35) shows that *svůj* is interpretable either as a bound variable or as a covaluated expression. This is quite unexpected, because *svůj* is an anaphor. I assume that the covaluation reading of the Czech possessive anaphor stems from its smaller ϕ-deficiency (in the sense of Reuland & Reinhart 1995) in comparison with the ϕ-deficiency of other Czech anaphors.

Other Czech anaphors cannot be interpreted as covaluated expressions – anaphors *sobě/se* in (37) and (38) can be interpreted only as bound variables.

(37) Jen prof. Novák vykládá pořád o sobě.
'Only prof. Novak always talks about himself.'
a. [...] ale ostatní profesoři vykládají i o lingvistice.
'[...] but other professors talk also about linguistics.'
b. #[...] ale ostatní profesoři nevykládají o prof. Novákovi.
'[...] but other professors do not talk about prof. Novak.'

(38) Jen prof. Novák se pořád chválí.
'Only prof. Novak praises himself.'
a. [...] ale ostatní profesoři chválí i jiné lingvisty.
'[...] but other professors praise also other linguists.'
b. #[...] ale ostatní profesoři nechválí prof. Nováka.
'[...] but other professors do not praise prof. Novak.'

5 *Only*

I suppose that the distinguishing property of *only* can be traced back to its semantics. This section will be dedicated to an incomplete but sufficient (for our purposes) story of *only*.

I assume the familiar alternative semantics framework (Rooth 1992 and many others). *Only* is the focus sensitive particle which can (as other focus sensitive particles) alter the truth conditions of a sentence with the presentational focus. Compare (39-a) and (39-b) – only the later sentence has other truth conditions than its non-focused counterpart. In the (39-a) focus evokes the set of alternatives to Mary, but its truth-conditions are the same as its non-focused counterpart. But (39-b) means that Mary was the only girl (from the set of alternatives), which was kissed by Peter.

(39) Petr políbil Marii.
'Petr kissed Mary.'

a. Petr políbil [$_F$ Marii].
'Petr kissed Mary.'

b. Petr políbil pouze [$_F$ Marii].
'Petr kissed only Mary.'

I do not discuss contrastive focus like (40), where it is sometimes claimed that similar truth-condition effects which in (39) stem from the focus sensitive particle are caused solely by the contrastive focus.

(40) [$_F$ Marii] políbil Petr.
'Mary, kissed Petr.'

The classical semantics for the *only* particle is the reverse of implication or inclusion.

(41) Karel will come, if it will rain.
p ... it will rain, q ... karel will come
$p \rightarrow q$

(42) Karel will come, only if it will rain.
$q \rightarrow p$

(43) Teachers read poetry.
Ax ... teachers, Bx ... poetry readers
$\forall x[Ax \rightarrow Bx]$

(44) Only teachers read poetry.
$\forall x[Bx \rightarrow Ax]$

But this is not the whole story. Van der Sandt & Geurts (2004) claim that *only* should be analyzed as a weak quantifier along the lines of (45).

(45) *only A are B* $= \neg\exists x[\neg Ax \wedge Bx]$

This is of course truth-conditionally equivalent to the strong-quantifier analysis, but it makes some interesting predictions. The first is the scope of negation: in the weak quantifier analysis the negation has scope over the whole formula, so it should license the NPIs in the whole sentence, unlike in the analysis with the strong quantifier, where the downward entailing ability of the strong quantifier is limited to its first argument.[3] This is corroborated by the Czech data in (46)– the NPIs *vůbec* and *kdy* are in boldface.

(46) a. Všichni, kdo **vůbec kdy** navštívili Brno, ho milují.
Everybody who at_all ever visited Brno it love
'Everybody who ever visited Brno loves it.'

b. *Všichni, **kdo** milují Brno, navštívili **vůbec kdy** jiné město.
Everybody who love Brno visited at_all ever another town
'Everybody, who loves Brno, ever visited another town.'

(47) a. Pouze ti, kdo **vůbec kdy** navštívili Brno, ho milují.
Only the_people who at_all ever visited Brno it love

[3] $\forall$ is downward entailing (entailing goes from set to its subsets) on its first argument and upward entailing (entailing goes from set to its supersets) on its second argument – see (46).

'Only the people who ever visited Brno love it.'

b. Pouze ti, kdo milují Brno, navštívili **vůbec kdy** Horní Heršpice.
Only the_people who love Brno visited at_all ever Horni Herspice
'Only the people who love Brno ever visited Horni Herspice.'

However, if the weak quantifier analysis is on the right track, then there is a problem with the presupposition of sentences like (48). (48) presupposes that there was somebody else than Petr, who visited Brno. From the strong-quantifier analysis of *only* this presupposition follows straightforwardly – if the strong quantifier presupposes (as is commonly assumed) existential quantification of its first argument (see (49)). (48-a) is the most salient reading of (48), where the negation scopes over the focused element, and in the strong-quantifier analysis it correctly predicts the existential presupposition of the set of Brno visitors. But (48-b), where the first negation before *Petr(x)* comes from the negation and the second from the rule for *only*, does not presuppose the non-emptiness of the set of Brno visitors.

(48) Brno did not visit only [$_F$ Petr].

a. $\forall x[BrnoVisitor(x) \rightarrow \neg Petr(x)]$

b. $\neg\exists x[\neg\neg Petr(x) \wedge BrnoVisitor(x)]$

(49) All students came.

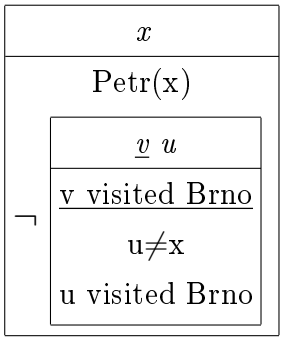
For these purposes we can use the Background/Presupposition Rule from Geurts & van der Sandt 2004 as can be seen in (50).

(50) The Background/Presupposition Rule (BPR)
Whenever focusing gives rise to a background $\lambda x.\phi(x)$, there is a presupposition to the effect that $\lambda x.\phi(x)$ holds of some individual.

In (51) the *Only Petr did not visit* is a background: $\lambda x.\neg visit(x)$, and then (50) predicts that the existential presupposition will arise.

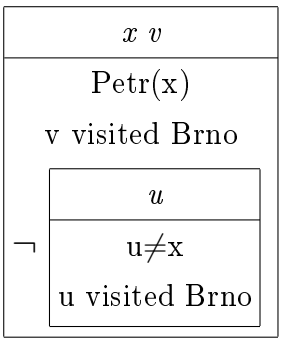
(51) Only Petr did not visit [$_F$ Brno].

(52)



After the global accommodation of the presupposition we have the following DRS, which is equivalent to the predicate logic formula in (54). For details of presupposition projections see van der Sandt & Geurts (2004).

(53)



(54) $\exists xv[Karel(x) \wedge VisitedBrno(v) \wedge \neg\exists u[u \neq x \wedge VisitedBrno(u)]]$

Now let us return to (35), repeated below as (55). For this example the presupposition analysis means that there are two options for the accommodation of the presupposition:

1. Presupposition λx(x recommends Novak's books), holding of some individual, is accommodated globally – this presupposition is then picked up in the second sentence in its reading where other professors recommend Novak's books.

2. Presupposition λx(x recommends Novak's books), holding of some individual, is accommodated locally – in the scope of negation, so there is no presupposition that somebody recommends Novak's books, so therefore the sloppy reading in the second sentence arises.

(55) Jen prof. Novák nedoporučuje své knihy, ale ostatní profesoři své/Novákovy
Only prof. Novak not-recommends his books, but other professors their/Novak's
knihy doporučují.
books recommend

6 Solution

In this section I will propose the solution for the puzzle of the interpretation of bound variables in Czech VPE. I will summarize the facts: the Czech anaphor *svůj* can be interpreted either as a bound variable or as a covaluated expression; this was shown with the help of the focus sensitive particle *jen* ('only'). The covaluation reading disappears in the VPE context, but shows up again if the VPE contains *jen*.

The solution can be found in the alternative semantics framework (Rooth 1992). The idea of my proposal is to compare the entailing capability of the discussed sentences. (56-a) is stronger than (56-b), which means that (56-b) implies (56-a). In other words: the set of those people who love somebody's library is the super-set of those people who love their own library: if you love your own library, then you love somebody's library, but not the other way around.

This pattern of reasoning is an example of Gricean quantity implicature and can be generalized using the notion of a scale of alternative assertions. Since my loving my library implies my loving somebody's library, but not conversely, we can set up a partially ordered set of two open propositions ordered by entailment: $\{love(x, x's\, library), love(x, y's\, library)\}$, the ordering relation $\geq$ is entailment.

(56) Petr miluje svou knihovnu.
Petr loves his library
a. Petr (λx (x loves x's library))
b. Petr (λx (x loves y's library) & y = ???)

If we look at the focusing pattern in VPE, we see that there is focus on the polarity particle, which is sometimes dubbed polarity focus. The focus of this type is motivated by the suggested contrast, for instance between Karel and Petr. I follow Rooth (1992) with his rule applying to contrastive focus – see (57).

(57) **Contrasting Phrases**. Construe a phrase α as contrasting with a phrase β, if $\|\beta\|^o \in \|\alpha\|^f$.

$\|\beta\|^o$ denotes ordinary semantic value of β (if β is *Petr*, then $\|\beta\|^o$ is Petr, if β is *love* then $\|\beta\|^o$ is the appropriate relation, ...), $\|\alpha\|^f$ denotes a focus semantic value for the α, which in the case of a sentence is a set of propositions. For instance, the focus semantic value for the sentence *[Karel]$_F$ loves Petr* is the set of propositions of the form 'x loves Petr', while the focus semantic value for *Karel loves [Petr]$_F$* is the set of propositions of the form 'Karel loves x'. For (58) the Contrasting Phrases Rule, β is the elided VP, α is its antecedent VP and the Contrasting Phrases Rule states that $\|Karel(\lambda x(x\,loves\,x's\,library)\|^o \in \|\lambda x(x\,loves\,x's\,library\|^f$, which is of course fulfilled.

(58) a. [Petr]$_F$ miluje svou knihovnu a Karel [taky]$_F$.
Petr loves his library and Karel too
b. [Petr]$_F$ miluje svou knihovnu, ale Karel [ne]$_F$.
Petr loves his library but Karel not

This rule has the right predictions, because in (59) we do not have a sloppy reading in the second conjunct – this sentence must mean that Karel does not love Mary's library. This follows from the (57), because $\|Karel(\lambda x(x\,loves\,x's\,library)\|^o \notin \|\lambda x(x\,loves\,Mary's\,library\|^f$, but $\|Karel(\lambda x(x\,loves\,Mary's\,library)\|^o \in \|\lambda x(x\,loves\,Mary's\,library\|^f$.

(59) [Petr]$_F$ miluje Mariinu knihovnu, ale Karel [ne]$_F$.
Petr loves Mary's library but Karel not

But (57) cannot by itself predict why (60-a) has only the sloppy reading. For this we need something stronger. The basic idea is that the listener selects the lowest element from the scale of the open propositions.

(60) a. Petr miluje svou knihovnu a Karel taky.
Petr loves his library and Karel too
Petr (λx (x loves x's library)) & Karel (λy (y loves y's library))
b. Petr (λx (x loves x's library)) & Karel (λy (y loves y's library))

In discussing example (61), Rooth (1992) says that asserting an element ϕ of C implicates the negation of any higher element in the scale, that is any ψ such that $\psi \geq_C \phi$ and $\phi \neq \psi$. So in this case, if the underlying set C is partially ordered as $\{ace(m), pass(m)\}$, then asserting 'Mats passed' implicates the negation of 'Mats aced'.

(61) Well, I [passed]$_F$.

In the (60-a) the scale is $\{love(x, x's\,library), love(x, y's\,library)\}$ and the process goes the other way round than in Rooth's example. Asserting polarity of VP implicates the negation of any lower element of the scale, which in this example is $love(x, y's\,library)$.

On the other hand *svůj* in (62) can have both covaluated and bound interpretations. This follows from the possible interpretations for the first sentence – see the last section – and either

of the two can be picked up by the following VPE. This is also circumstantial evidence for the pragmatic restrictions on VPE.

(62) Pouze Karel miluje svou knihovnu, ale Petr ne.
Only Karel loves his library but Peter not

Summary: the Czech anaphor *svůj* can be interpreted either as a bound variable or as a covaluated expression. But the covaluation reading disappears in the VPE context. This follows from the Contrasting Phrases Rule and from the scale of the open propositions, which emerges as a consequence of focusing the antecedent VP for VPE. From this scale the listener chooses the highest element, which in this case is the bound variable reading.

References

Asher, Nicholas. 1993. *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer.

Busquets, Joan. 1997. *A Discourse-Based Approach to VPE : A Survey*. Research report RI IRIT/97-48-R, 1997, IRIT Université Paul Sabatier, Toulouse.

Busquets, Joan. 2005. *Stripping vs VP-Ellipsis in Catalan*. Research report RR5616, 2005, INRIA, Université Bordeaux-1, LaBri.

Fiengo, Robert and Robert May 1994. *Indices and Identity*. Cambridge, MA: MIT Press.

Geurts, Bart. 1999. *Presuppositions and Pronouns*. Oxford: Elsevier.

Geurts, Bart and Rob van der Sandt. 2004. Interpreting Focus. *Theoretical Linguistics* 30:1–44.

Hankamer, Jorge and Ivan Sag. 1976. Deep and Surface Anaphora. *Linguistic Inquiry* 7(3):391–426.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic*. Dordrecht: Kluwer.

Klein, Ewan. 1987. VP ellipsis in DR theory. Groenendijk, J. & D. de Jongh & M. Stokhof (eds). *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*: 161–187. Dordrecht: Foris Publications.

Kučerová, Ivona. 2005. The T-Extension Condition. J. Alderete et. al. (eds.): *Proceedings of the 24th West Coast Conference on Formal Linguistics*: 227–235. Somerville, MA: Cascadilla Proceedings Project.

McShane, Marjorie J. 2000. Verbal ellipsis in Russian, Polish, and Czech. *Slavic and East European Journal* 44(2): 195–233.

Reuland, Eric and Tanya Reinhart. 1995. Pronouns, Anaphors and Case. Haider, H., S. Olsen & S. Vikner (eds.): *Studies in Comparative Germanic Syntax*: 241–268. Dordrecht: Kluwer.

Reinhart, Tanya. 1983: *Anaphora and Semantic Interpretation*. London: Croom Helm.

Reinhart, Tanya. 2000. Strategies of Anaphora Resolution. Bennis, H., M. Everaert & E. Reuland (eds.): *Interface Strategies*: 295–324. Amsterdam: North Holland Amsterdam.

Veselovská, Ludmila. 2004. Rozšířená verbální projekce v češtině: tři druhy slovesa *být*. Hladká, Z. & P. Karlík (eds.): *Čeština – univerzália a specifika* 5:203–213. Praha: Nakladatelství Lidové noviny.

Zagona, Karen. 1982. *Government and Proper Government of Verbal Projections*. Doctoral dissertation, University of Washington.

Clitic omission in Czech as across-the-Board extraction

Jakub Dotlačil

Jakub.Dotlacil@let.uu.nl

1 Introduction

In the tradition of Slavic generative linguistics, one of the main discussions concentrate on the second position clitics and a question of how to account for their position. Should we blame syntax, phonology, or a mixture of the two?

The battlefield of various approaches to the second position of clitics was mostly Serbian or Croatian (see, for example, Schutze 1994 for a phonological approach, Progovac 1996 for a syntactic approach, or Bošković 2001 for a mixture of the two). On the other hand, the literature on similar issues in Czech was surprisingly unequivocal. As far as I know, linguists applied a purely syntactic approach to the second position in Czech (see, for example, Rivero 1991, Veselovská 1995, Lenertová 2004).

In this paper, I am going to discuss an argument from clitic omission in conjunction which goes against a purely syntactic account of clitic placement in Czech and supports Bošković's analysis of clitic positioning: namely, their surfacing in the second position is a result of an interplay of both, syntax (which takes care of clitics' movement) and phonology (which takes care of pronunciation of the right copy) (see Bošković 2001).

To have a taste of the argument, take a look at the following examples. In coordination, Czech allows for clitic omission in one of the conjuncts, as shown on two examples: (1-a), in which the clitics *jsem* and *ho* are omitted in the second conjunct, and (1-b), in which the clitic *se* is omitted in the second conjunct.[1]

(1) a. Já **jsem** **ho** zavolal a představil známým.
I past-aux$_{1sg}$ him$_{acc}$ called and introduced friends
'I called him and introduced him to friends.'

b. Petr **se**$_i$ umyl a oholil.
Petr refl washed and shaved
'Petr washed and shaved.'

As I am going to argue, cases like (1-a) and (1-b) should be analyzed as cases of Across-the-Board (ATB) extraction.

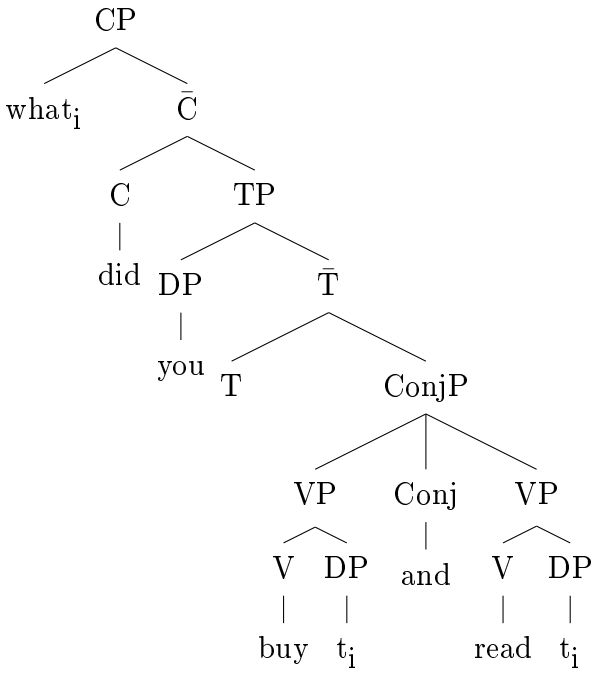
ATB extraction is, descriptively speaking, movement of like-phrases *out* of the coordination. (2) is an example of ATB wh-movement.

[1] All clitics are boldfaced. All glossed examples are from Czech.

(2) What did you buy and read?

The wh-word *what* starts as an argument of both *buy* and *read*. Later on, it moves out of the VP coordination (i.e., undergoes ATB-extraction) into its specifier position (Spec, CP). The final structure with traces after movement is sketched in (3):

(3)



Similarly to the wh-word in example (2), I am going to argue that in examples (1-a) and (1-b), clitics originate in both conjuncts. Later on, they move out of the coordination (i.e., undergoing ATB extraction) into their specifier position.

Somewhat surprisingly, I will show that we can find cases similar to (1-a) and (1-b) (i.e., cases of ATB extraction of clitics), in which, crucially, clitics are pronounced not outside of the coordination but inside the first conjunct. In other words, we will deal with a special case of ATB extraction: one in which ATB extracted constituents do not surface *outside* of the coordination but *inside* the first conjunct. This, I am going to argue, provides an argument for a particular view on the second position clitic placement. In particular, it is compatible with Bošković's analysis of clitic placement but problematic for other, purely syntactic, accounts.

The paper is organized in the following way. Section 2 discusses arguments for ATB extraction analysis of (1-a) and (1-b). Section 3 discusses cases of ATB extraction of clitics in which clitics surface inside the first conjunct. Section 4 focuses on some problems with the presented analysis of clitic placement in Czech. Section 5 concludes.

2 Clitic omission as ATB movement

In Czech, clitics can be omitted in one of the conjuncts, as shown in (4), repeated from above:

(4) Já **jsem** **ho** zavolal a představil známým.
I past-aux$_{1sg}$ him$_{acc}$ called and introduced friends
'I called him and introduced him to friends.'

The clitic omission in the second conjunct, one might suggest, is a result of an ellipsis that targets clitic clusters.

However, an ellipsis analysis runs into immediate problems. First, notice that clitic omission is possible only in coordinations. It is ungrammatical in a non-coordinated sequence of clauses (5-a). Furthermore, clitic omission is impossible in the matrix clause if the anteceding clitics appear in a subordinate clause (5-b) or if the anteceding clitics appear in the matrix clause and the clitic cluster is deleted in a subordinate clause (5-c).

(5) a. *Já **jsem** **ho** zavolal. Potom představil známým.
I past-aux$_{1sg}$ him$_{acc}$ called. Then introduced friends
'I called him. Then I introduced him to friends.'
b. *Poté co **jsem** **ho** zavolal představil známým.
After what past-aux$_{1sg}$ him$_{acc}$ called introduced friends
'After I had called him I introduced him to friends.'
c. *Představil **jsem** **ho** známým poté co zavolal.
introduced past-aux$_{1sg}$ him$_{acc}$ friends after what called
'I introduced him to friends after I had called him.'

The fact that clitic omission is licensed only in coordination would seem arbitrary and surprising if we assumed that clitic omission is a result of an ellipsis of clitics.[2] However, if we assumed that clitic omission is a result of ATB extraction of clitics the difference between (4) and (5-a)-(5-c) would fall out automatically. This is so because ATB extraction (a label for the phenomenon in which a constituent that is understood to originate in all the conjuncts surfaces outside of the conjunction (see (2))) occurs only in coordinations.

Thus, based on the difference between (4) and (5-a)-(5-c) I suggest that (4) should be analyzed as (6):

(6) I past-aux$_i$ him$_j$ ... [$_{ConjP}$[$_{XP}$ called t$_i$ t$_j$] and [$_{XP}$ introduced t$_i$ t$_j$ friends]]

The next three subsections are going to present an additional support for the analysis of clitic omission as ATB extraction.

[2] VP ellipsis, sluicing, or NP ellipsis can normally be licensed in contexts similar to (5-a)-(5-c). Gapping (verb deletion) is a surprising case: it can be licensed only in coordinations (cf. (i) and (ii)):

(i) Marie políbila Petra a Táňa Tomáše.
Marie kissed Petr and Táňa Tomáš
'Marie kissed Petr and Tanya - Tom.'

(ii) ??Marie políbila Petra předtím než Táňa Tomáše.
Marie kissed Petr before than Táňa Tomáš
'*Marie kissed Petr before Tanya - Tom.'

In fact, this has been used as an argument to treat gapping not as an ellipsis (contra Ross 1967) but as ATB extraction (Johnson, 2006).

2.1 Case matching

It has been noted (see Dyla 1984, Franks 1993) that the constituent that undergoes ATB extraction must be assigned cases in conjuncts which have the same overt manifestation.

To see the validity of this generalization, let us turn for a moment to ATB wh-movement.

The following table summarizes nominative, genitive and accusative forms for Czech *kdo* 'who' and *co* 'what'. Notice that 'who' has a syncretic form in genitive and accusative, whereas 'what' has syncretic forms in nominative and accusative:

(7) Table 1: Case of *who* and *what*

Case	who	what
nominative	kdo	*co*
genitive	*koho*	čeho
accusative	*koho*	*co*

Syncretism in forms corresponds to acceptability of ATB extraction. In case the first conjunct assigns nominative and the second conjunct assigns accusative, only 'what' is (marginally) acceptable (8-a). Similar examples with 'who' are ungrammatical ((8-b)-(8-c)).

(8) 1st conjunct - nominative, 2nd conjunct - accusative

a. ?Co tam leželo a ty jsi sebral?
what$_{\text{gen-acc}}$ there lay and you past-aux picked
'What lay there and you picked it?'

b. *Kdo tam ležel a ty jsi sebral?
who$_{\text{nom}}$ there lay and you past-aux picked
'Who lay there and you picked him?'

c. *Koho tam ležel a ty jsi sebral?
who$_{\text{acc}}$ there lay and you past-aux picked
'Who lay there and you picked him?'

This exemplifies the descriptive generalization, mentioned above: only 'what' can undergo ATB extraction in (8-a) because it has one and the same form for nominative and accusative. 'who' cannot undergo ATB extraction in this case because its forms for nominative and accusative differ.

This descriptive generalization can be made sense of if we follow approaches to the architecture of the language which assume that syntax feeds lexical insertion.[3] In the examples (8-a) to (8-c), the wh-words are assigned two cases in syntax (either by agreement projections inside each conjunct, or by the verb itself, or by combination of the two). When lexical insertion applies, the lexicon is searched for an item that matches the requirement of being an exponent of both cases (nominative and accusative). The paradigm of 'what' has such a lexical item (namely, *co*), unlike the paradigm of 'who'. Thus, in case of 'who', whatever lexical item is inserted it cannot satisfy morphological requirement and the ungrammaticality of (8-b) and (8-c) follows.[4]

[3] Generative semantics was one such an approach. Recently, the same idea has been revived in Distributed Morphology.

[4] This reasoning only works if we do not assume that morphological forms in paradigms are always underspecified.

Notice that there is no reason to expect a similar requirement for ellipses. This is so because whereas in cases of ATB movement one and the same object originates in both conjuncts and is assigned two cases in each of the conjuncts, in cases of ellipses, we deal with two different objects: the antecedent and the ellided constituent. And, in fact, the head that assigns case to the antecedent and the head that assigns case to the ellided constituent do not need to assign case with the same overt manifestation. For example, a subject pro-drop can differ from its antecedent (9-a). The same holds for noun ellipses (9-b).[5]

(9) the antecedent - accusative, the ellipsis - nominative

a. Petr měl rád Marii$_{\text{i}}$. Hlavně když _$_{\text{i}}$ mu vařila.
Petr had like Marie$_{\text{acc}}$. especially when _ him cooked
'Petr liked Marie. Especially when she cooked for him.'

b. Marie měla ráda černovlasé kluky$_{\text{i}}$. Táně se líbili blondátí _$_{\text{i}}$
Marie had like black-hair guys$_{\text{acc}}$. Tanya refl appealed blond$_{\text{nom}}$ _
'Marie liked blackhair guys. The blond ones appealed to Tanya.'

Crucially, clitic omission patterns with ATB-wh-movement and unlike ellipsis.

The following table shows that he$_{\text{cl}}$ has a syncretic form for genitive and accusative, unlike they$_{\text{cl}}$.

Table 2: Case of *he$_{cl}$* and *they$_{cl}$*

(10)

Case	he	they
genitive	*ho*	jich
accusative	*ho*	je

As predicted, in case one conjunct assigns genitive and the other conjunct assigns accusative, only 'he$_{\text{cl}}$' is acceptable:

(11) 1st conjunct genitive, 2nd conjunct accusative

a. ?Já **se ho** bojím a nenávidím.
I refl him$_{\text{gen-acc}}$ be-afraid and hate
'I am afraid of him and loathe him.'

b. *Já **se jich** bojím a nenávidím.
I refl them$_{\text{gen}}$ be-afraid and hate
'I am afraid of them and loathe them.'

c. *Já **se je** bojím a nenávidím.
I refl them$_{\text{acc}}$ be-afraid and hate
'I am afraid of them and loathe them.'

2.2 Coordination with more than two conjuncts

As discovered by Ross (1967), conjuncts are islands (i.e., structures which block movement out of them). This descriptive generalization is known as the Coordinate Structure Constraint:

(12) Coordinate Structure Constraint (CSC): In a coordinate structure, no element contained in a conjunct may be moved out of that conjunct.

[5]Ellipses in these two examples are marked by the underline _ coindexed with the ellipsis' antecedent.

There is a well-known exception to that generalization, an exception that is a topic of this paper: ATB extraction. This has also been noticed by Ross in his seminal thesis:

(13) There is an important class of rules to which (12) does not apply. These are rule schemata which move a constituent out of all conjuncts in a coordinate structure. (Ross, 1967, page 107)

Thus, whereas (14-a) is ungrammatical (violation of the CSC), (14-b) is fine (ATB extraction).

(14) a. *Co jsi koupil a odešel?
what past-aux bought and left
'*What did you buy and left?'
b. Co jsi koupil a prodal?
what past-aux bought and sold
'What did you buy and sell?'

So far, we have seen examples of clitic omission in which clitics originated in both conjuncts and surfaced outside of the coordination. This confirms the description as given in (13). If clitics originated only in some of the conjuncts we should expect ungrammaticality because clitic omission is a case of ATB movement which is movement out of *all* conjuncts. If clitics moved from some, but not all, conjuncts, the Coordinate Structure Constraint should be violated, which should cause ungrammaticality.

This prediction is borne out. (15) shows a case of coordination of two conjuncts. Clitics undergo ATB extraction out of both conjuncts:

(15) Petr **se** umyl a oholil.
Petr refl washed and shaved
'Petr washed and shaved.'

Coordination of more than two conjuncts is also possible if clitics undergo ATB extraction out of all conjuncts:

(16) Petr **se** myl, česal a holil.
Petr refl washed, combed and shaved
'Petr washed, combed and shaved.'

However, it is crucial that clitics end up outside of the coordination and are omitted in *all* conjuncts:

(17) *Petr **se** myl, čistil si zuby a holil.
Petr refl washed, brushed refl teeth and shaved
'Petr washed, brushed his teeth and shaved.'

As said above, the ungrammaticality of (17) follows since the CSC is violated in this case.

On the other hand, if clitic omission was a case of ellipsis, the ungrammaticality of (17) would come as a surprise. Simply put, there is no reason why ellipses should be all-or-nothing phenomenon. In other words, there is no reason to expect that ellipses can either appear in all conjuncts (apart from the first one), or in none of them.

2.3 When ATB extraction is blocked

The final argument that I believe supports my analysis of clitic omission as ATB extraction comes from a surprising piece of data. As pointed out to me by an anonymous reviewer of Formal Approaches to Slavic Linguistics 2007, not all coordinations allow for clitic omission within their conjuncts. According to the reviewer (18), in which the clitic *mu* 'him$_{dat}$' is omitted, is ungrammatical (I myself do not find it completely ungrammatical, but agree that the sentence is degraded):

(18) *Už **se** **mu** ulevilo a je líp
already refl him$_{him}$ relieve and is better
'He was relieved of pain and feels better.'

The ungrammaticality of (18) surely is surprising. The example represents a coordinated structure and therefore ATB extraction of clitics should be licensed.

Fortunately, independent reasons have been found which suggest that (18) might not be a coordinated structure after all. As discussed at length by Postal (1998), coordinations in which the first conjunct serves as a cause of the second conjunct show different behavior than standard coordinations. For example, they allow for violation of the Coordinate Structure Constraint (19), as noticed by Lakoff (1986):

(19) The stuff which$_i$ the guys in the Caucasus drink t$_i$ and live to be 100.

Postal (1998) suggests that these constructions should not be analyzed as coordinate structures but as subordinations. If his analysis is on the right track, we might have an explanation for the ungrammaticality of (18). Notice first that (18) is also a case in which the first conjunct is a cause of the second conjunct. Now, if, following Postal (1998), we do not analyze (18) as a coordinate structure, we actually expect ATB extraction to be ungrammatical. Notice that ATB extraction of a wh-word is also impossible in the same coordination which supports our conclusion that (18) is ungrammatical because ATB extraction is impossible in this case:

(20) *Komu$_i$ se už ulevilo a je líp?
who already refl relieve and is better
'Who was relieved of pain and feels better?'

This concludes arguments for analysing clitic omission as ATB extraction. From now on, I am going to assume that this analysis is correct. In the next section I am going to concentrate on a surprising case of clitic omission: one in which clitics arguably do not move out of the coordination but stay inside the first conjunct.

3 ATB extraction with seemingly no extraction

3.1 Problematic examples

Take a look at the following example:

(21) Zavolal **jsem** **ho** a představil známým.
called past-aux$_{1sg}$ him$_{acc}$ and introduced friends
'I called him and introduced him to friends.'

Let us go through this example more carefully. First, notice that the clitics *jsem* and *ho* are omitted in the second conjunct. Clearly we deal with clitic omission, which, as I argued, should be analyzed as an ATB extraction of clitics. Thus, both clitics *jsem* and *ho* originate as an auxiliary and an argument in both conjuncts. Later on, they both move out of the coordination.

Notice that both clitics are linearly ordered at the end of the first conjunct. Since clitics undergo ATB extraction out of the coordination the verb preceding them must undergo movement out of the coordination as well. Thus, we end up with the structure (22) for example (21).

(22) [called$_i$ [past-aux$_j$ him$_k$ [$_{ConjP}$ [$_{VP}$ t_i t_j t_k] [$_{VP}$ introduced t_j t_k to friends]]]]

The problem is that in this structure the verb underwent movement from the first conjunct. But this movement violates Coordinate Structure Constraint and thus should render (21) ungrammatical! Furthermore, notice that after ATB extraction of clitics the first conjunct consists only of 'called'. Its movement out of the coordination violates another well-known constraint, the Conjunct Constraint (Grosu, 1981), which states that no conjunct may be moved.

We have already seen (in Section 2.2) that the Coordinate Structure Constraint applies in Czech. (23) shows that the Conjunct Constraint applies in Czech, as well:

(23) *Chlapec kterého jsem potkal a Petra.
boy which past-aux met and Petr
'*The boy which I met and Petr.'

In short, we have ended up in a paradox. If we assumed that clitic omission is a case of ellipsis which targets clitic cluster we would have no explanation for the data discussed in the previous section (Section 2). But if we assume that clitic omission is a case of ATB extraction we expect the movement of the verb 'called' in (21) to violate the Coordinate Structure Constraint and the Conjunct Constraint and, therefore, we expect (21) to be ungrammatical, contrary to the facts.

There are, as far as I can see, two possible ways out of this paradox. The first one is to assume that the Coordinate Structure Constraint and the Conjunct Constraint are structure-specific. They do apply in cases like (23) but are not applicable to (21). The second way out of the paradox is to assume that there is something special about ATB extraction of clitics. Something that enables them not to be pronounced outside of the coordination in cases like (21). In that case, the verb 'called' do not need to move out of the coordination either, and, therefore, no violation of the Coordinate Structure Constraint and the Conjunct Constraint takes place.

I am going to follow the second route.

3.2 Clitic placement at the syntax-phonology interface

One of the many issues connected to the study of clitics is a question of how to account for their placement. Is their second position a result of a phonological requirement, syntactic requirement, or both?

In his recent work, Željko Bošković (Bošković, 2001) follows Franks (2000) and suggests that we should consider clitic placement an interface phenomenon. It is a result of an interplay between syntax and phonology.

Following the Minimalist Program (see, especially, Chomsky 1995), Bošković assumes that when a constituent moves it leaves a copy of itself in the original position. The two copies (one in the base position and the other in the target position of the movement) are indistinguishable

from each other. However, they do behave differently with respect to pronunciation. As has been noted many times in literature, it is usually only the highest copy that is pronounced. All the lower copies are deleted (see Nunes 2001 and Nunes 2005 for an extensive discussion and interesting explanation of why this should be so). It is this last point that, according to Bošković, differentiates clitics from other moved constituents.

Clitics enter computation with syntactic requirements. For the sake of argument let us assume that all clitics in Czech need to move via head movement to C.[6] However, unlike most other words, clitics also come with phonological requirements that must be satisfied. Following Bošković I assume that clitics in Czech need to lean on a host to their left.[7]

Recall that it is normally the highest copy of a moved constituent that is pronounced. Thus, we would expect clitics to be pronounced in their highest position, the head of C. However, this copy might violate clitics' phonological requirements: if there is no material higher in the sentence, clitics cannot lean on any host to their left and the sentence is ungrammatical. Bošković (2001) suggests that in this scenario, a lower copy of clitics is pronounced: the one that satisfies phonological requirements.

Let us go through one example:

(24) Zavolal **jsem** **ho**.
called past-aux$_{1sg}$ him$_{acc}$
'I called him.'

Let us assume that the clitic 'him' is merged as a direct object of the verb and later on moves as a head to the projection which hosts the auxiliary. Furthermore, let us assume (following Veselovská 2004) that past auxiliaries are located in the head T. Thus, at the level of $\bar{T}$, the structure looks followingly:

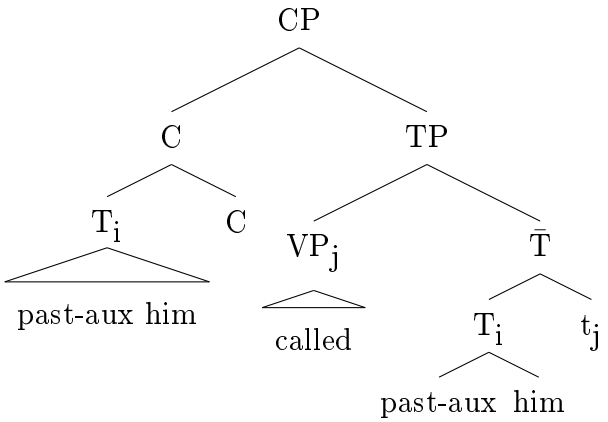
(25) $[_{T'}\ [_{T}$ past-aux him$_i$ $]\ [_{vP}\ [$v $]\ [_{VP}\ [$v called $]$ t$_i$ $]\ ]\ ]$

The T projects further up. Following Migdalski (2006), let us assume that to satisfy EPP requirements of the T, the vP moves into its specifier. After that, C selects for the TP and clitics head move via the T into the head C. This is the final structure:

[6] This goes contra arguments that Bošković explicitly makes for Serbian, namely that clitics do not all move into the same position. However, the arguments on which this conclusion is based do not work in Czech. Thus, I assume that unlike in Serbian, all clitics in Czech do target the same position. If this turns out to be incorrect, the analysis that I am going to propose can still be maintained, with appropriate modifications.

[7] There are problems with this assumption for Czech because clitics in Czech can be either enclitics or proclitics. I will turn to this issue in Section 4.

(26)



However, in this case the highest copy of the clitics cannot be pronounced because it would violate phonological requirements. Therefore, a lower copy is selected: the one in which clitics sit in the T. The pronunciation of the lower copy gives us the correct word order for (24).

3.3 ATB extraction with seemingly no extraction explained

Armed with an analysis of how clitics' syntactic and phonological requirements together derive their position let us move to the example discussed in Section 3.1.

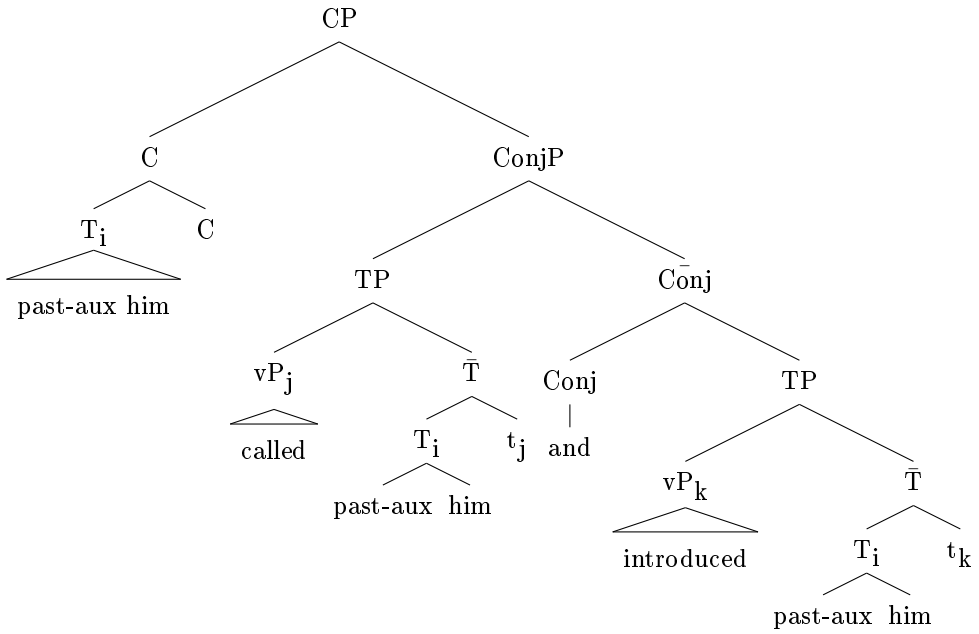
(27) Zavolal **jsem** **ho** a představil známým.
called past-aux$_{1sg}$ him$_{acc}$ and introduced friends
'I called him and introduced him to friends.'

Recall what the problem was. Example (27) is a case of ATB extraction of clitics. Therefore, clitics must have moved out of the coordination. But if they did (so the reasoning went) then the verb 'called' must have moved out of the coordination, too, in violation of the Coordinate Structure Constraint and the Conjunct Constraint.

There is a flaw in this reasoning, and, I believe, it should be obvious by now what the flaw is. In Section 3.1, I have assumed without discussion that what is pronounced is the highest copy of a moved constituent. In fact, this assumption is natural if one follows syntactic accounts to the second position of clitics and assumes that the second position is derived solely in syntax. However, we have seen in the previous section that this does not need to be so. More concretely, I presented an approach to clitic placement which argues that the second position is a result of a phonological requirement which leads to the pronunciation of a lower copy.

This is one possible analysis of (27) which avoids violation of the Coordinate Structure Constraint and the Conjunct Constraint: first, the object is merged as the object of the verbs in both conjuncts and it adjoins to the auxiliary in the T, still separately in the two conjuncts. The T in both conjuncts project. The T in the first conjunct attracts the verb phrase consisting of 'called' into its specifier. The T in the second conjunct attracts the verb phrase consisting of 'introduced' into its specifier. Now, the two TPs are conjoined and C takes this coordination as its complement. The clitics *jsem* and *ho* undergo ATB extraction into the C. This is the final structure:

(28)



However, if the highest copy of the clitics was pronounced the clitics' phonological requirement would not be satisfied. Therefore, a lower copy must be pronounced. Counting the number of nodes, we arrive at the second highest copy: the one in which the clitics sit in the T in the first conjunct. Thus, this copy of clitics is pronounced and we end up with the correct word order without violating the Coordinate Structure Constraint or the Conjunct Constraint.

There are at least two predictions that this approach makes. First, notice that full DPs should behave differently than clitics do with respect to examples like (27). This is so because full DPs do not have a phonological requirement that would force pronunciation of a lower copy. Since it is always the case that the highest copy of these DPs is pronounced examples like (27) with DPs in the place of the clitics should be ungrammatical.

This is correct. Notice first that full DPs can license what looks like an object drop:

(29) Petra jsem zavolal a představil známým.
Petr$_{acc}$ past-aux called and introduced friends
'I called Petr and introduced him to friends.'

Since there is no object drop in Czech, I conclude that example (29) is a case of ATB extraction: 'Petr$_{acc}$' starts as an argument of the verbs in both conjuncts and undergoes ATB extraction out of the coordination.

Thus, example (29) shows us that ATB extraction of full DPs is possible.

Interestingly, an example parallel to (27), in which the clitic *ho* 'him' is substituted by the full DP 'Petr$_{acc}$' is ungrammatical:

(30) *Zavolal jsem Petra a představil známým
called past-aux$_{1sg}$ Petr$_{acc}$ and introduced friends
'I called Honza and introduced him to friends.'

As discussed above, this follows from the presented analysis. 'Petr$_{acc}$' comes with no phonological requirements that would force pronunciation of a lower copy in this example. Therefore, the highest copy of 'Petr$_{acc}$' has been pronounced in (30). Since the highest copy must be outside of the coordination (otherwise, there is no way to license object drop in this coordination), the verb 'called' must have moved out of the coordination, as well. But this movement of the verb violates the Coordinate Structure Constraint and the Conjunct Constraint!

The second prediction that the presented analysis makes concerns interaction of clitic omission with the size of conjuncts. To license clitic omission in (27), repeated here as (31), clitics have to move out of coordination even though this movement is masked by the fact that the copy outside of the coordination is not pronounced.

(31) Zavolal **jsem** **ho** a představil známým.
called past-aux$_{1sg}$ him$_{acc}$ and introduced friends
'I called him and introduced him to friends.'

Crucially, clitics must be able to move out of the coordination, otherwise the structure is illicit. If we dealt with a coordination of conjuncts which are islands for clitic movement (i.e., if we dealt with a structure bigger than the TP) clitics should not be able to undergo ATB extraction out of the coordination and clitic omission should consequently be ungrammatical. This should be so even though clitics are not pronounced outside of the coordination.

This prediction is also borne out. Notice first that clitic omission is possible in conjoined TPs in embedded contexts:

(32) Petr řekl že **jsem** **ho** zavolal a představil známým.
Petr said that past-aux$_{1sg}$ him$_{acc}$ called and introduced friends
'Petr said that I had called him and had introduced him to friends.'

However, once we conjoin CPs clitic omission is impossible:

(33) *Honza řekl že **jsem** **ho** zavolal a že představil známým.
Honza said that past-aux$_{1sg}$ him$_{acc}$ called and that introduced friends
'Honza said that I had called him and that I had introduced him to friends.'

As is well-known, clitics cannot move out of CP (see Progovac 1993, Veselovská 1995, Rezac 2005, or Dotlačil 2007 for different explanations of this fact). In (33), clitics have to undergo ATB extraction out of the coordination. However, since clitics cannot move out of CP this movement is illicit, and since ATB extraction is impossible, clitic omission cannot take place in (33).

In conclusion, the grammaticality of clitic omission in (31) supports the analysis of clitic placement in Czech along Franks (2000) / Bošković (2001) line. We have seen an argument for the language architecture in which phonology overrides syntax in deciding which copy should be pronounced (see Nunes 2005 for an approach to the pronunciation of copies which is compatible with this view).

The next section is going to discuss one problem ignored so far: do we really have any reasons to assume that clitics come with phonological requirements?

4 Czech clitics and their phonological requirements

As discussed by Fried (1994), Toman (1996), and many others, clitics in Czech do not need to lean on a host to their left. Clitics in Czech can be at the beginning of an intonational phrase

and lean on a host to the right (i.e., behave as proclitics) (example (34-a)). They can, in fact, even be at the beginning of a clause in colloquial Czech (example (34-b)) (see Lenertová 2004 for more examples and discussion).

(34) a. Já # tvoje máma # **jsem ti** slíbila hračku.
I # your mother # aux you promised toy
'I, your mother, promised you a toy.'
b. **Se mi** včera narodil syn. (ok only in colloquial Czech)
refl me yesterday born son
'My son was born yesterday.'

But is the fact that Czech clitics can be either proclitics or enclitics important for my analysis of examples like (35)?

(35) Zavolal **jsem** **ho**.
called past-aux$_{1sg}$ him$_{acc}$
'I called him.'

In fact, it is. Following Bošković (2001), I analyzed (35) as a case in which phonology forces pronunciation of a lower copy of clitics. But if clitics can also be proclitics there is no reason to expect that phonology should force the pronunciation of a lower copy in this case. The highest copy would violate no phonological requirements after all. Since phonologically, clitics can be both proclitics and enclitics, shouldn't it be only syntax that derives clitics' second position (in standard Czech), with no help from phonology?

One way out from this problem that I would like to suggest is that clitics are not underspecified for proclitization and encliticization. Instead, they are ambiguous. In (35) enclitics were chosen in the lexicon: since these need to lean on their hosts to the left, the pronunciation of a lower copy is forced. If proclitics were used instead the highest copy could be pronounced (for reasons unclear to me, this latter option is possible only in colloquial Czech).

Surprisingly, (34-a) does not allow for the possibility in which clitics become enclitics (and end up not in the second, but third position in the clause):

(36) ??Já # tvoje máma # slíbila **jsem ti** hračku.
I # your mother # promised aux you toy
'I, your mother, promised you a toy.'

This is surprising if clitics were ambiguous: why should we not be able to retrieve to enclitics in (36) which would force the pronunciation of a copy below the predicate?

I tentatively assume that there is nothing wrong with choosing enclitics in (36). What goes wrong with this example is the predicate movement. The predicate 'promised' cannot move above clitics to support them. Notice that if the predicate cannot move from its base position both clitics *jsem* and *ho* must end up at the left edge of the intonational phrase.[8]

To be sure there are much more cases that one should analyze before jumping at the conclusion that the hypothesis of clitics being ambiguous between enclitics and proclitics really is viable. Unfortunately, since this would lead me too far afield from the topic of this paper, I have to leave them aside.

[8]I assume that the base position of *ho* cannot be pronounced for independent reasons - see Moro (2000)

5 Conclusion

This paper has discussed cases of clitic omission as in (37):

(37) Já **jsem** **ho** zavolal a představil známým.
I past-aux$_{1sg}$ him$_{acc}$ called and introduced friends
'I called him and introduced him to friends.'

I argued that these cases should be analyzed as ATB extraction of clitics. However, this conclusion turned out to be problematic in cases of examples in which clitics seemed to stay inside the first conjunct:

(38) Zavolal **jsem** **ho** a představil známým.
called past-aux$_{1sg}$ him$_{acc}$ and introduced friends
'I called him and introduced him to friends.'

This, I argued, should be seen as evidence that clitics might be pronounced in their non-highest position, which is in line with Bošković's analysis of clitic placement in Serbian (Bošković, 2001). Based on the examples like (38), I argued that clitics' position in Czech is not solely derived in the syntactic part of the computation but is a result of an interplay between syntax and phonology. If this analysis is on the right track cases like (38) supports a particular view on the syntax-phonology interface. Under this view, phonology can alter word order that is derived in syntax by deciding which copy of movement done in syntax should be chosen for pronunciation.

References

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: cliticization and related phenomena*. Oxford: Elsevier Science.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.

Dotlačil, Jakub. 2007. Why clitics cannot climb out of CP: a discourse approach. Compton, R. et al. (eds.). *Formal Approaches to Slavic Linguistics: The Toronto Meeting 2006*. Ann Arbor: Michigan Slavic Publications 15:76–93.

Dyła, Stefan. 1984. Across-the-board dependencies and Case in Polish. *Linguistic Inquiry* 15:701–705.

Franks, Steven. 2000. Clitics at the interface. Beukema, F. & M. den Dikken (eds.). *Clitic phenomena in European languages*. 1-46. Amsterdam, Philadelphia: John Benjamins.

--. 1993. On parallellism in across-the-board dependencies. *Linguistics Inquiry* 24:509–529.

Fried, Mirjam. 1994. Seccond-position clitics in Czech: Syntactic or phonological? *Lingua* 94:155–175.

Grosu, Alexander. 1981. *Approaches to island phenomena*. Amsterdam, New York: North-Holland.

Johnson, K. 2006. *Gapping is not (VP) ellipsis*. Unpublished, ms.

Lakoff, George. 1986. Frame semantic control of the Coordinate Structure Constraint. Farley, A. M. (ed.). *Papers from the Parasession on Pragmatics and Grammatical Theory*. Chicago: University of Chicago.

Lenertová, Denisa. 2004. Czech pronominal clitics. *Journal of Slavic linguistics* 12:135–171.

Migdalski, Krzysztof. 2006. *The Syntax of Compound Tenses in Slavic*. PhD dissertation, Tilburg University, Tilburg.

Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, Massachusetts: MIT Press.

Nunes, Jairo. 2005. *Linearization of Chains and Sideward Movement*. Cambridge, Massachusetts, London, England: MIT Press.

--. 2001. Sideward movement. *Linguistic Inquiry* 31:303–344.

Postal, Paul M. 1998. *Three investigations on extraction*. Cambridge, Massachusetts: MIT Press.

Progovac, Ljiljana. 1996. 1993. Locality and Subjunctive-like Complements in Serbo-Croatian. *Journal of Slavic Linguistics* 1:116–144.

--. Clitics in Serbian/Croatian: Comp as the second position. Halpren, A. & A. M. Zwicky (eds.). *Approaching second: Second position clitics and related phenomena*: 411–428. Calif: CSLI.

Rezac, Milan. 2005. The syntax of clitic climbing in Czech. Heggie, L. & F. Ordøñez (eds.). *Clitic and Affix Combinations*. Amsterdam: Philadelpia: John Benjamins.

Rivero, Mária Luisa. 1991. Long head movement and negation: Serbo-Croat vs. Slovak and Czech. *The Linguistics Review* 8:319–351.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. PhD dissertation, MIT, Cambridge, Massachusetts.

Schutze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. *MIT Working Papers in Linguistics* 21:373–473.

Toman, Jindřich. 1996. A note on clitics and prosody. Halpren, A. & A. M. Zwicky (eds.). *Approaching second: Second position clitics and related phenomena*: 505–510. Stanford: CSLI.

Veselovská, Ludmila. 1995. *Phrasal Movement and X-Morphology: Word Order Parallels in Czech and English Nominal and Verbal Projections*. PhD dissertation, Palacký University, Olomouc.

--. 2004. The Extended verbal projection in Czech: Three Variants of the verb be. *FDSL 5*. Leipzig.

Czech Cases and the Syntacticon: Poznámky k, o, okolo, nad něčím a pro někoho

Joseph Emonds

jeemonds@hotmail.com

Czechs and non-Czechs alike generally feel that the Czech morphological case system is quite complex, especially as it manifests itself with prepositions or without them, but in phrases which correspond to PPs in say English.[1]

It is the purpose of this study to show that Czech cases in PP structures are more than "orderly"; they are in fact, with literally one clear exception, entirely regular and predictable.[2] Of course, in order to do this, one cannot just accept every truism that has been handed down through traditional grammar. Moreover, it is necessary to crucially use a few highly motivated but innovative constructs formulated within generative grammar. If such moves were not necessary, the regularity of the Czech Case system would have become apparent before now. Generally, structuralist and generativist authors alike are overly reliant on traditional grammar's truisms. Even generativists seem to feel that these truisms are somehow "inherent in the data" and can be used at no cost. So in practice they prefer to perpetuate them rather than utilize theoretical constructs not yet approved in their perception by "current theory."

Non-Slavic speaking linguists acquainted with Indo-European (I-E) morphological case soon sense that many uses of Czech's six (non-vocative) cases are far from unfamiliar.[3] Nominative and accusative have the same form for inanimates, those termed both "neuter" and "masculine." Most predicate nominals and adjectives agree in case with the noun phrase they modify, whether in primary or secondary predication. Partitives are expressed in the genitive. Noun phrase objects of prepositions expressing motion toward a Goal are in the accusative.

What seems bewildering though are the Czech cases used with Ps (or to express some English PPs without a P): accusatives, genitives, datives, locatives and instrumentals all find their places in its system. Compared with Classical Greek, Standard German, Icelandic, Latin and

[1] I have had fruitful and enlightening discussions on this topic with both L. Veselovská and M. Martinková. I thank them both, and am very grateful to M. Martinková for carefully reading and commenting on various drafts of this paper and for supplying numerous useful references. Of course, neither is responsible for my errors of fact and interpretation.

[2] At least among prepositions that are not highly lexicalized and idiomatic. An entertaining exercise for Czech readers may be to guess which grammatical P is irrevocably exceptional in my system.

[3] Czech vocatives also have a familiar Indo-European feel, i.e., akin to those in Latin. In almost all paradigms, a vowel is added to nominatives of proper nouns: a high vowel with masculine stems and a mid-vowel with feminine stems. As vocatives are not related to sentence structures involving other categories, they don't properly speaking constitute a separate case. Compare them to the case structures discussed below in section 1.

Old English, in which 3 cases can appear after Ps,[4] Czech cases in PP structures seem more arbitrary and less predictable.

1 Against the traditional definition of "dative" for the third case

To my mind, the earliest generative attempts around 1980 to make sense of morphological case incorporated two insightful idealizations, but at the same time formulated them in terms of an unperceptive and unreflective lapse. The good news is (i) that cross-linguistically each basic case results from a *characteristic case-assigning category*, and (ii) these basic case-assigning categories are few in number, basically four:

(1) a. "Nominative" is assigned by I to the DP closest to it.
b. "Accusative" is assigned by V to the DP closest to it.
c. "Dative" or "oblique" is assigned by P to the DP closest to it.
d. "Genitive" is assigned by some category in N-projections to the closest DP.

The bad news is that the field retains to this day these "names" of these cases as if they are somehow categories in their own right. Perhaps this gives the impression that generative grammar explains these traditional concepts, but in most cases it simply uncritically accepts them, e.g. even elaborating on the timeworn distinction between "structural" and "semantic" case.[5] Emonds (1985: Ch. 1 and 5), argues in detail that no special case categories should appear in grammatical representations; "nominative" is nothing other than the appearance of I as a feature or index on nominal projections and likewise for the other cases. Consequently, the case features should be represented essentially as follows, where the traditional names for them have absolutely no theoretical status.

(2) a. Nominal projections D^k in the nominative are notated D_I.
b. Nominal projections D^k in the accusative are notated D_V.
c. Nominal projections D^k in the oblique case are notated D_P.
d. Nominal projections D^k in the genitive are notated D_Q. [6]

These notations involve more than terminological parsimony, even though that alone would justify them. They also remove prejudices that impede good analyses, e.g. the instinctive feel of many linguists (including generativists) that traditional terms like e.g. "datives" or "genitives" capture similarities across languages independent of actual empirical cross-linguistic

[4] Latin grammars always remark that no P can ever be followed by its dative case, leaving only genitives, accusatives, and ablatives that Ps can introduce.

[5] "Semantic" cases are those assigned before movement of object nominals in passives, while "structural" cases are those assigned after it. This leads to a problem with genitives, which government and binding treatments turned a blind eye to. As observed in Veselovska (1997), genitive phrases act "structurally" inside noun phrases, but "semantically" when they are complements to verbs.

The minimalist recognition (Chomsky, 1995) that trees are constructed from the bottom up removes any problem, as well as the need for the semantic/structural distinction. Once assigned, case cannot change. DPs inside PP domains receive case before the PPs merge into larger structures ("semantic case"), while DPs not inside PPs merge in IP or larger DP domains without case ("structural case"). If such IPs or DPs then merge further, the cases on the DPs they properly contain cannot change, exactly as with "semantic case."

[6] Veselovská (2001) argues, to my mind convincingly, that the case assigning category inside nominal projections is Q, for (existential) "quantifiers" including, in Czech, high numerals. I adopt this here throughout.

justifications. This harmful hidden assumption puts meaningless burdens of proof on any analyses in which Case X in L_1 is the basic equivalent of some otherwise named Case Y in L_2, rather than what tradition calls Case X in L_2.

In addition, reducing "cases" to features or indices as in (2) (i.e. an accusative is a D or N projection with a V feature or index) makes possible some revealing structural analyses of how predicate nominals and adjectives receive morphological cases, including those in Czech (Emonds 2000: Ch. 8). For example, the Case feature D_V on a direct object assigns case to a following secondary predicate nominal. Such analyses cannot even be envisioned if one insists on the traditional case names formally unrelated to the case assigners. Although most work on case is unaware of such accounts, they nonetheless constitute unanswered and empirically based arguments in favour of eliminating, as in (2), traditional or any other ad hoc names for cases.

Perhaps the most striking innovation in the generative approach to case is embodied in (1c) and (2c). Traditional grammar always uses the terms nominative, accusative and dative for the morphological cases of respectively the subject, direct object and indirect object of typical simple three place predicates. It furthermore treats the three cases as entirely unrelated. However, massive cross-linguistic syncretism between nominatives and accusatives (throughout I-E and not least in Czech) shows that the first two cases are indeed subcases of some more general verb-related "archi-case."[7] But related to each other or not, both "nominative" and "accusative" noun phrases are associated in (1a-b) with structural configurations closely connected with I and with V. Hence the generative treatment of these cases is not so far from their traditional associations with subjects and direct objects.

For "datives," the situation is different. And indeed, the early generative proposal for case in e.g., Chomsky (1981) does *not* define datives as the "case of the indirect object." Rather, the generative "dative" was (and is) the case structurally associated with *the structure of PP.* Indeed, the surface form of indirect objects in many systems is *not* associated with PP at all. For example, in several languages lacking morphological case, benefactive applicative phrases, as studied in e.g. Baker (1988), typically surface in direct object position and are arguably accusative, sometimes by morphological as well as syntactic criteria. In Modern Standard Arabic, indirect objects also appear as accusatives, i.e. with the same case as direct objects.[8]

Pursuing this line, it seems like the "third case" or the "P-assigned case" in (1c) should be identified with the case used for expressing *locations* of a verb's action or state. In Classical Greek, Standard German, Icelandic and Old English, this case happens to also be that of the indirect object (the so-called "dative"). However, Latin's locative case is its ablative, not its dative.[9] The third Czech case, judging from how many locative Ps govern it, is neither what is called the dative or the locative, *but is rather the "instrumental."* Such are the cases that express, in each of these languages, the Ps or PP structures translating *under, over, between, in front of, behind, through* and *with,* when these are conceived of as pure locations unassociated

[7] Generative treatments have not recognized this. I pursue this topic elsewhere; it does not interact with this essay's subject matter except in one particular. The fact that nominatives and accusatives are partly "the same" can explain why Ps in I-E systems never "take a nominative." Traditional grammars, if they even note this fact, provide no explanation for it.

[8] There is a underlying universal relation between indirect objects and PP structure, but it is not a surface relation. This position is cross-linguistically supported in Emonds (1993).

[9] The Latin "dative" is a variant on the ablative and has different forms only in (most but not all) singulars; Latin plural datives and ablatives *always have the same form.* Moreover, in all paradigms where it differs from the ablative, the dative singular has a *single* form *–í.* Emonds (1985: Ch. 5) argues that this special allomorph appears in Latin PPs if and only if an introductory [P, GOAL] is ∅. As discussed in section 7, the status of the Czech dative resembles to some extent that of the Latin dative.

with Goals or Sources of movements or transitions.

Reinforcing this first conclusion for Czech, we find this language has a sometimes specially remarked use of the instrumental, the bare (preposition-less) "instrumental expressing place," exemplified in bold here with some examples from Holá (2000: 235):

(3) Musíte jít **podchodem** a pak doleva.
'You must go **through the underpass** and then left.'
Šli jsme **jinou ulicí.**
'We went **by the other street**.'
Vlak projel **malým nádražím.**
'The train went **past a small station**.'
Řeka Vltava teče **Prahou.**
'The Vltava River flows **through Prague**.'

I will further be developing the idea that Czech system is most closely related, in terms of other I-E systems, to that of Classical Greek (the latter's nominative, genitive, dative and accusative). In Czech, the closest structural relative to the Classical Greek dative is the instrumental. Of course, the Czech dative and locative cases will also be analyzed; in fact, the patterns surrounding these "additional" cases will be a crucial testing ground for refining and developing a theory of the lexicon that I have proposed in earlier work, so these recalcitrant cases are in fact the main motivation for pursuing this topic.[10]

This study will incorporate the "good news" on morphological case from Government and Binding (four basic cases each associated with a universal structure) and discard the "bad news." The names of the cases are just mnemonics for DPs structurally marked as occurring in certain constructions. I use the neutral descriptive term "oblique" for the case generally associated with a P of location, i.e. the Classical Greek dative, the Czech instrumental and the Latin ablative.

2 Probable universal syntactic features of Czech Ps

Several works by Jackendoff (1977, 1983, 1990) and others have established the need to clearly distinguish noun phrases denoting a LOCATION of a verb's action or state from those denoting the GOAL of a verb's action. Only the latter are typically introduced with a counterpart to English *to*. This work has also shown that noun phrase SOURCEs of a verb's action, indicated in English with *from,* share some similarities with GOALs, which he proposes to capture by treating Ps that introduce both types as "Prepositions of PATH".

The Path Ps that introduce GOALs and SOURCE DPs differ in their grammar from the Place Ps that introduce LOCATIONs. Many relevant syntactic analyses by van Riemsdijk (1978, 1988, 2002) have brought out further differences between these two classes of Ps, which he differentiates by means of a feature that we can identify with ±PATH.[11]

Van Riemsdijk's work on complex case systems of e.g, some Caucasian languages, as well as

[10] Case systems outside Indo-European exhibit little if any relation between expressing simple locations with Ps and the case of indirect objects. Locations with P use the nominative in Turkish and the genitive in Modern Standard Arabic. Descriptions of Turkish follow common practice and call the case of its indirect objects "dative".

[11] Van Riemsdijk and Huijbregts (1998) show that Ps of PATH often introduce PPs of PLACE as their complements, i.e., two Ps can succeed each other: *from behind the barn, onto (⇐ to + on) the table,* French [PATH *en*] [PLACE *dessous*] *de la table* 'to underneath (of) the table'. We are not concerned here with sequence of Ps, but only with those Ps closest to DPs that assign them case.

traditional analyses of somewhat more familiar grammars of e.g. Sanskrit and Finnish, suggest that the basic features on P further divide [P, PATH] into those expressing "motion toward", "motion up to", and "motion into", in both directions (*to* and *from*). It appears that Czech cases do not grammatically distinguish these three subcategories of path, but only two in each direction, as in (4). We can specify those of the left column as +CONTACT and those of the right column as –CONTACT.

(4) Czech grammatical P of +CONTACT | Czech grammatical P of –CONTACT

motion up to or into: P = *do*+ genitive	motion toward: P = *k(e)*+ dative
motion off or out of: P = *z(e)* + genitive	motion away from: P = *od* + genitive

The grammatical features that seem to cross-classify Czech prepositions thus include at least LOCATION, GOAL, SOURCE and CONTACT. The feature -PATH stands for the conjunction [-GOAL, -SOURCE], but I am not sure if as a feature on D it has a theoretical status or is just shorthand.

In terms of these features, we can cross-classify Czech Ps in a table, indicating at the same time the cases typically exhibited by their object nominals.

(5) Table of Czech cases on DPs objects of Ps

Czech features on Ps that have DP objects	Grammatical entries for P = Ø or P ≠ Ø P, ±F, -f$_j$, where F$_I$ =±GOAL, ±LOCATION, ±CONTACT	Open Class Dictionary (P ≠ Ø): P, +LOCATION, ±GOAL, +f$_j$ Positions of: x = modified XP, s = speaker, o = object of P
F$_i$ = -LOCATION -GOAL, -SOURCE	P,-F = Ø 'with' (instrument) = *s(e)* 'with'	Not defined, since Dictionary Ps must be +LOCATION.
F$_i$ = +LOCATION -GOAL, -SOURCE In G-B terms, this case is "oblique." The Czech term is **instrumental,** and is in **red** font here. In Latin, this is "ablative."	P,-F = Ø 'through, by' (bare locational DP)	s x o *před* 'in front of' s o x *za* 'beyond' x o x *mezi* 'between' x o *nad* 'above' s (or other reference point) o x *pod* 'below' s (or other reference point)
	P,+F$_i$ = *o* 'about' = *v(e)* 'at, in' = *při* 'near, during' = *po* 'over, along' = *na* 'on' This square to be analysed in section 5. (This will be the Czech "locative square.")	
F$_i$ = +LOCATION +GOAL -SOURCE The default is the super-case assigned by V, namely the Czech **accusative**, in **blue** font here.	P = Ø (indirect objects) P = *(až) k(e)* 'toward' This square to be analysed in section 6. (This will be the Czech "dative square.") Exception: *do* + genitive: '(up) to'	The same case as assigned by V, namely accusative. *o* '(to) against, than' *přes* '(to) across' *před* '(to) in front of', *mezi* '(to) between', *nad* '(to) above', etc. *na* 'onto', purpose, duration *mimo* 'outside, besides' *pro* 'for' *po* 'up to, during' *v(e)* 'on' (time), some idioms *za* 'for' (in exchanges), 'as'
F$_i$ = +LOCATION +SOURCE -GOAL The Czech **genitive** case is in **green** font here. The genitive here is similar to that in Classical Greek. It is unclear whether +SOURCE should be +GOAL or –GOAL.	P = Ø (partitives) P = *od(e)* '(away) from' *z(e)* 'out of, off'	*bez* 'without' *blízko* 'near from' *nedaleko* 'not far from' *okolo* 'around' *vedle* 'next from' *kromě* 'except' *vyjma* 'except' (?)Lexical items with [$_{P,SOURCE}$Ø]: *během* 'during' *koncem* 'at the end of' *podél* 'along' *u* 'at the place of, near' *včetně* 'inclusive of'

Some one way semantic implications related to case can be appreciated immediately. For example, the genitive expresses motion away from while the accusative expresses motion into. Another tendency evident in Table (5) is that the majority of Ps take objects in the accusative, genitive or instrumental cases, while not many Ps take datives or locatives. That is, in the presence of P, not all cases are equal, an asymmetry which requires an account.

[11] Emonds (2000, Ch. 2) argues that no actual categorial difference separates P and the feature LOCATION. The interpretation "spatial or temporal location" is nothing more than the presence of P at Logical Form ("LF"). For a mechanism by which P can be marked as *not* interpreted as LOCATION in LF, see section 8.

3 Correspondences between Czech and Classical Greek Prepositional Cases

This section will not yet approach the difference between what in Table (5) are termed "grammatical" and "open class" prepositions, in part because neither traditional grammar, nor structuralism, nor generative grammar provides any clear criteria to work with. This section focuses only on *the rows* of Table (5), making no distinction between its columns. This step leads us to investigate three basic case types among Czech objects of P: those traditionally called the instrumental/ locative (rows 2 and 3), the dative/ accusative (row 4), and the genitive (row 5). Prior to specifying a case-assignment theory, these labels have no content other than descriptive mnemonics for phonological spellings.

Table (6) expresses two pivotal correspondences between Czech and its ancient Indo-European cousin Classical Greek. Rows 1-3 in Table (6) correspond to rows 3-5 in Table (5). We for the moment ignore case usages lacking overt Ps, such as indirect objects.

(6) Cases on objects of overt Ps expressing location (*not* extended to temporal P):

+LOCATION	Czech	Classical Greek
-GOAL (pure location)	instrumentals/ locatives	datives
+GOAL, -SOURCE	datives/ accusatives	accusatives
+SOURCE	genitives	genitives

Two traditionally named cases in Table (6) have corresponding uses: accusatives and genitives. I turn first to an important characteristic common to the two systems (in row 3), the fact that genitive case is used for "motion away from" and also for "privative" prepositions expressing concepts such as *instead of, except, without* and *not far from.*

3.1 An extended I-E use of genitives common to Classical Greek and Czech

The use of the genitive to express separation from a SOURCE is not "natural", that is, a widespread pattern imposed by Universal Grammar. A number of non-I-E languages with a clear genitive case mark, perhaps the best studied being Japanese, *entirely reserve the genitive for use within noun phrase projections.* That is, whatever the complexities of Japanese structures, one thing is for sure, if one sees the genitive *no* 'of', then one is *inside* a nominal structure. A Japanese genitive phrase with or without an immediately dominating PP can never be a constituent of a verbal projection.

In contrast, both the Romance and Germanic I-E families have an impressive range of genitive constituents of verb phrases, such as introduced by e.g. English *of*: *His talk was of no interest; we think of her as intelligent; they heard/ spoke/ knew of that fact; John is of a different opinion.* The French genitive constructions with *de* 'of' in verbal projections differ from the English ones, but are equally if not more varied. Moreover, I take it the English genitive constructions are typical within Germanic and those of French within Romance.

Languages in these same families with morphological genitives (Latin, German, Icelandic) also exhibit such genitives in various verbal constructions (unlike Japanese). However, one pattern conspicuously absent in both Romance and Germanic is any systematic reflex of genitive case for objects of Ps expressing motion away from. This particular property of early I-E (characteristic

of Sanskrit) is preserved in only some descendent families, but nonetheless is clear in both Czech and Classical Greek.

Thus, the treatment of Greek Ps of locations and the required cases of their objects in Niederle, Niederle and Varel (1991: 214-223) is organized around a Table on their page 213 with the following columns:[12]

Col. 1 Ps that take accusatives, express goals, and answer "where to?"

Col. 2 Ps that take genitives, express "point of departure," and answer "where from?"

Col. 3 Ps that take datives, express "location" and "(at rest) contact", and answer "where?"

They also refer to such genitives as "genitives of separation." Clearly, their Table's columns correspond respectively to rows 4, 5 and 3 of the Table (5) above of Czech Ps.

Thus, Greek translations of Czech Ps of motion away from and some other Greek Ps indicating motion away from take object nominals marked with genitive case. The Roman spellings here don't indicate the accents on the final vowels of Greek Ps.

	apo '(away) from', Czech *od*	ek/ ex 'out of', Czech *z(e)*
(7)	hyper 'from above downwards'	hypo 'from below upwards'[13]
	kata 'from above'	para 'from close to'[14]

These are only some of the most studied Greek Ps with genitives, but they suffice to establish that using this case to express "motion away from" is systematic in Greek, parallel to the Czech Ps in the last row of Table (5). And also as in Czech, several Greek Ps take genitives simply by lexical stipulation, independent of their meanings: *anti* 'against', *dia* 'through' and *pro* 'in front of'.[15]

We can thus propose that Czech and Classical Greek share the essential aspect of how genitive case is spelled out in PPs. The symbol π_{genitive} represents the phonological forms of the genitive inflection, which are too varied to concern us here.

(8) Lexical entry for **Genitive Case Marking** (Czech and Classical Greek):

$$<[\ +\text{N}\ ]\_\_\_>,\ \{\ \text{D}_\text{Q}\ /\ [\ \text{D}_\text{P},\ \text{SOURCE}\ ]\ \},\ \pi_{genitive}$$

I here explain the form of lexical entries that appear in this study, based on the lexical theory developed in Emonds (2000, 2003, 2005a). The material inside $<\ldots>$ is a word-internal subcategorization frame, i.e. the D_i are suffixes on a stem of category +N, where +N encompasses the nominal categories N, A and D, precisely those that exhibit morphological case. The idea that case suffixes are of category D is justified in Veselovská (2001); cf. note 5. Since a noun-internal location is not an interpretable position in universal grammar for any of D_i, PLURAL, or SOURCE, these latter categories can be spelled out as suffixes only by virtue of some LF-independent licensing device that constrains the mappings from syntax to phonological form (PF). In this framework, the general device employed for this is "Alternative Realization" (18), formulated and exemplified in section 5.3 below.

[12]German Ps that take genitives are a group of lexical items without this semantic basis.

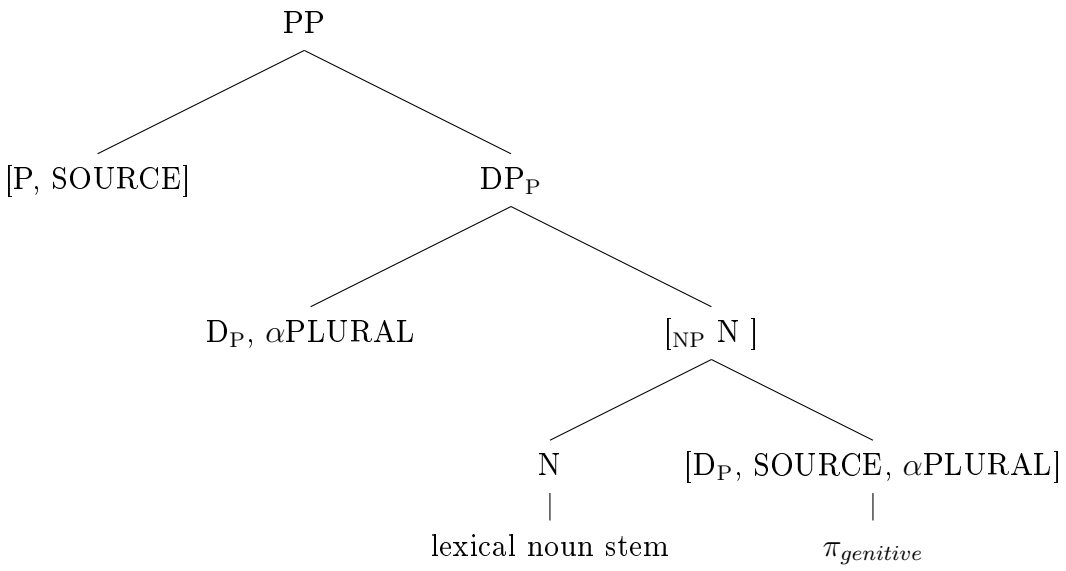
[13]These two Ps correspond to Czech *pod* and *nad*. Their objects receive different cases according to the scheme in Table (6). The glosses in (7) are relevant for *genitive* objects.

[14]Greek *para* is like Czech *při*. The gloss in the text of 'from' is again only relevant for genitive objects. The grammatical item is –CONTACT; the possibility of three different cases after it suggests that this feature is grammatical in Greek.

[15]Niederle, Niederle and Varel (1991) discuss some further Greek Ps whose uses with genitives remain unclear to me: *amphi, epi, meta, peri* and *pros*.

Consequently, a "mini-tree" associated e.g. with a genitive expressing motion away from is (9), where the higher P, SOURCE, D_P, and PLURAL are all in their canonical (= LF-interpretable) positions.

(9)



In lexical entries such as (8), the syntactic specifications appear to the left and the phonological ones to the right. This reflects a central claim that I argue for in Emonds (2003), to the effect that grammatical items are lexically identified (i.e. accessible in processing) by their unique *grammatical context and category feature combinations.*[16]

Finally, we can now also state how to specify Ps that take genitives lexically. For example, the fact that Czech *do* 'to' is [+GOAL, -SOURCE] and yet takes a genitive object is clearly exceptional in Table (5). Consequently *do* is treated as a lexical exception throughout this study. Using the widespread generative term for item-specific lexical case marking, the object of *do* is thus a "quirky genitive."[17]

(10) **Quirky Case.** P, +GOAL, +CONTACT, <D_Q>, *do*

When no blank occurs inside the context specification <F_i>, an item is a free morpheme and occurs with a full phrase complement, whose lexical head is required to have the feature F_i.

[16] In contrast, I claim that open class items are uniquely identified by their phonology, and moreover subgrouped into searchable "cohorts" by their initial consonant clusters (Marslen-Wilson 1987, 1990). Thus, a single open class Dictionary item such as *open* can be an A, V, or even N (a golf tournament), but a suffix such as the English causative *–en* is unique by virtue of its syntax; *-en* can be grouped as a single lexical item with phonologically distinct causative suffixes *–ify, ize,* and ∅, but not with its homophone, the passive suffix *–en.*

It is then correctly predicted that free morphemes with different initial consonants can constitute a single lexical entry only if they are grammatical (*bad/worse; go/went; she/her*).

[17] M. Martinková (pers. comm.) informs me that the Russian translation *v(o)* of Czech *do* is not exceptional or "quirky" with respect to Table (6); its object DP is accusative. Not unexpectedly, some Czech verbs also take "quirky case" complements, e.g. *ptát se* 'ask (of someone)' and *bát se* 'fear' take quirky genitives; *věřit* 'believe' and *rozumět* 'understand' take quirky datives; *plýtvat* 'waste' and *opovrhovat* 'despise' take quirky instrumentals.

3.2 Further properties of the system of overt Ps that express LOCATION

Generally, one may ask whether and how the rows of Table (6) should be grouped. There are three plausible possibilities: one is that GOAL and SOURCE Ps have a common PATH feature (+SOURCE would also be +GOAL); a second is that Universal Grammar does not group them (+SOURCE would then be -GOAL); and a third is that pure LOCATION should be grouped with SOURCE but not GOAL. The least option might seem unmotivated semantically, but in fact the Latin system, which conflates both in a single "ablative case," seems to suggest this. For exposition, I treat Ps of GOAL and SOURCE as not sharing a feature, but do not investigate this question further here.

The previous subsection has tried to capture the common expression of SOURCE Ps in Czech and Classical Greek, schematized in Table (6), row 3. The next section will turn to row 2, which proposes a general equivalence of Czech instrumentals and locatives with Classical Greek datives for expressing LOCATIONS. Here the traditional case names obscure the parallelisms, which is why I employ a neutral yet traditional term "oblique case."

Finally, let us turn to the use of the accusative for "prepositions of motion toward," as in row 1 of Table (6). Both Czech and Classical Greek robustly exhibit a widespread property of Indo-European morphological case systems, namely the use of the accusative for their objects. (The Czech variant called the dative is treated in section 7.) Focusing here on the accusative, the issue arises, is this use of it stipulative? Does it arise from lexical marking on various and sundry Ps, or is there something more general behind this uniform use of accusatives for Goal DPs?

4 The I-E accusative as a default morphological realization of Case

The fact is, the I-E morphological accusative has an essentially default function in several case-marking contexts (i.e. embedding in syntactic structures, as opposed to dislocated topics, vocatives, citation forms, etc.). The same I-E languages in which (i) "accusatives are used for motion to" also use them in nominative positions for all nouns of "neuter gender". For example, the nominative of Latin neuters like *bellum* 'war' has a superficial accusative suffix *–m*. The following traditional truism expresses this, but provides no explanation: (ii) "For neuter nouns, nominative and accusative are always the same".

In fact, it seems that the Indo-European accusative is simply the general mark of constructional case, i.e. the case assigned when case must be assigned, but when no other case is available. As a third example, when the finiteness constituent I is unspecified for features and hence cannot assign nominative to its subject, what emerges in Latin and other languages is a proposition whose form is, using another unexplained shibboleth of traditional grammar, (iii) "accusative with infinitive."

In light of the accusative usages (i)-(iii), we need not say that the GOAL prepositions of Czech, Classical Greek or any other I-E language are lexically stipulated to take accusatives. Rather they are *unspecified* for requiring any case. This perspective has been suggested in Veselovská (2001), who proposes that accusative is *an unmarked structural case* on DPs. Then, because (some version of) the Classical Case Filter of Rouveret and Vergnaud (1980) requires DPs to have case generally, if no other Case assignment applies, Case Marking is supplied by

the following highly general default lexical entry.[18]

(11) Lexical entry for **Accusative Case Marking** (Indo-European):
[+N]___, D_X, π *accusative*

This perspective accounts for why accusative case appears after Ps of such *varied meanings* in (5), not just those of spatial GOALS. Consider the range of functions of the bold Ps with italicized accusative objects in (12), from Naughton (1999: 301-304).

(12)

To je **pro** *mě.*	'This is for me.'
Jdu **pro** *mléko.*	'I go for milk.'
Čekám **na** *tramvaj.*	'I wait for a tram.'
Jsem tu jen **na** *týden.*	'I'm here only for a week.'
Jsi **o** *rok* starší než já.	'You're (by) a year older than me.'
Pršelo **(po)** *celý den.*	'It rained (for) all day.'
Ve *středu*	'On Wednesday'
Nevěříš **v** *Boha*?	'You don't believe in God?'
Koupil to **za** *pět korun.*	'He bought it for five crowns.'
Udělám to **za** *hodinu.*	'I accomplish it (with)in an hour.'

Certainly these usages cannot result from some extension or contortion of a special rule involving motion toward. Alternatively, to claim that so many Ps take the accusative via lexical specifications would amount to a loving embrace of irregularity. In contrast, the simple claim here involves no lexical or grammatical stipulations; the objects of these Ps are accusatives merely by default. These D projections must have case, yet no rule spells out case on them other than (11).

When (11) applies e.g. to an oblique D_P in a Goal phrase or to the nominative D_I of a neuter gender lexical subject, their syntactic case features in LF nonetheless remain respectively, D_P and D_I, even though the morphological realizations are phonologically accusative. As a default, rule (11) *follows* the other more specific Case Marking rules, i.e. those in Czech that spell out Ds as Instrumentals (i.e. the Oblique Case D_P), Genitives, and (Animate Gender) Nominatives. However, (11) never "overwrites" such spell outs, since it applies to only those D_x that have not been assigned phonological content.[19] It moreover does not conflict with a requirement that distinct complement XPs of a single predicate exhibit *different* syntactic case features in Logical Form.[20]

[18] Thus, there is no special Accusative Marking statement that specifies the category V or D_V. The complex phonological alterations among accusative allomorphs, represented in (11) by $\pi_{\text{accusative}}$, are not treated here.

[19] One type of phonological content that can be specified in a lexical entry is a null allomorph, such as the Czech feminine and neuter genitive plurals. Such null morphemes must be distinguished from categories that are null for theoretical reasons. See Emonds (2005).

[20] This very general and intriguing pattern is called the Logical Form Case Filter in Emonds (2000: Ch. 8). It explains several previously mysterious or unremarked restrictions on combinations of XP complement types. One of the possibilities is having "no case" in Logical Form, a situation that arises with clausal complements and predicate attributes.

5 Locative vs. Instrumental: boundary and properties of Syntacticon items

5.1 A lexical entry for Czech instrumental endings

For discussion, I repeat in (13) the two upper rows of Table (5):

(13) Czech Prepositions not related to GOALs or SOURCEs

F_i = -LOCATION -GOAL, -SOURCE	P,-F = Ø 'with' (instrument) = *s(e)* 'with'	Not defined, since Dictionary Ps must be +LOCATION.
F_i = +LOCATION -GOAL, -SOURCE In G-B terms, this case is "oblique." The Czech term is **instrumental,** and is in **red** font here. In Latin, this is "ablative."	P,-F = Ø 'through, by' (bare locational DP)	s x o *před* s o x *za* x o x *mezi* x o *nad* s (or other reference point) o x *pod* s (or other reference point)
	P,+F_i = *o* 'about' = *v(e)* 'at, in' = *při* 'near, during' = *po* 'over, along' = *na* 'on' (This is the Czech "locative square.")	

As observed in textbooks of Czech, the instrumental and locative cases have in common that they "can denote a fixed location". In this way they contrast with genitives and datives, which "can denote a change of place" (Naughton 1999: 66).

Interestingly, both the least semantically specific, *s(e)* and Ø 'with', and the most semantically specific of the P in (13) select object DPs in the instrumental. This is a typical "elsewhere" distribution. That is, a special statement is needed for the Ps in violet font (section 5.3 below), and then a simple default statement assigns instrumental case to the rest of the Czech oblique objects which are neither GOALS nor SOURCES, indicated here shorthand with –PATH.

(14) Lexical entry for **Instrumental Case Marking** on Czech nouns, singular and plural:

$$<[\ +\mathrm{N}\ ]\ \_\_\_\_\ (\mathrm{PLUR})>,\ [\ \mathrm{D_P},\ \text{-PATH}\ ]\ \{\ m,\ \acute{\imath},\ \emptyset\ \}^{21}$$

Any phonological specifications in the entries of this study are meant to suggest only the essential aspects of the various suffixes, ignoring the phonetic detail. Thus, (14) doesn't attempt to specify the phonological realizations of the three instrumental allomorphs or the conditions on choosing among them.[22]

[21] In proposing his Mirror Principle relating morphological and syntactic positions, Baker (1985) observes that case affixes generally *follow* plural affixes if the two are distinguishable. However, as (14) stipulates, the Czech instrumental suffix, whose most common realization is *–m-*, seemingly *precedes* its plural *–y/í*; adjectival inflection directly reflects this. That is, Standard Czech instrumental plurals always end in *–y/í* (a rare exception being *dvěma* 'two').

[22] Nonetheless, certain aspects of the morphophonology of Czech instrumental suffixes stand out: (i) The feminine singular allomorph seems to be a high vowel whose value for ±FRONT is determined by the stem-final segment, and then *...au⇒...ou, and* (ii) the allomorph is otherwise–*m* or Ø. The zero allomorph blocks palatalization of final non-palatalized consonants in front of the plural *–y/i* (nom. plur. *páni* 'sirs' vs. instr. plur. *pány*).

5.2 The nature of the instrumental/ locative distinction

We next examine a difference in (13) between what I claim are open class Dictionary items (in column 3) and "grammatical" items (in column 2). In several studies, I have claimed that the defining difference between two entirely different types of lexical items resides in that fact that open class items are always associated with some *purely semantic features f not used in syntax proper*. Before going into this difference more, let us ask: Are there actually purely semantic features needed for the Czech Ps in the third column in (13)? I believe there are.

The meanings of these latter P are indeed more complex than those of the P in column 2. This is obvious for *mezi* 'between'. The other Ps in column 3 require some kind of frame of reference, such as an observer or gravity (to determine up vs. down). Consider: if two rocky masses are floating close to each other in deep space with neither observer, star nor the earth as a known reference point, one could not be said to be "beyond/ in front of/ above/ below" the other (column 3). But they could be said to be "on/ in/ near/ with" each other (column 2). These reference points and an orientation toward them are thus part of the lexical meaning of only the Ps in column 3 and are indicated by purely semantic features f_j.[23]

This conclusion is not far from that reached in an extensive discussion of the semantics of Place in Leech (1969: esp. sections 8.1-8.5). The English Ps translating Czech Ps that take locatives are said to involve only "dimensionality", while those Ps whose Czech counterparts take instrumentals express "relative position" and "orientation". For these latter notions, he introduces a number of features in addition to those needed for simple locations, which I take to be examples of semantic f_j.[24]

Now exactly which features F are syntactic and which f are semantic is not obvious, prior to syntactic investigation or to constructing a lexical theory. But *given* the well justified open vs. closed class distinction in grammatical phenomena far from prepositional meanings, and *given* a methodological commitment to the idea that syntax is not arbitrary, we look for and actually find a needed distinction. In particular, the need for "orientation features" f_j for the P in column 3 in Table (13) sheds light on what has previously looked like a semantic morass. That is, the different cases assigned by Czech Ps *reveal the boundary* between the syntactic feature complex [P, -PATH] and some further Ps with purely semantic orientation features f_j.

In several works, most collected in Emonds (2000; 2007), I have developed the idea that the timeworn distinction between open and closed classes is not some kind of continuum or some vague property of lexicons without central implications for syntax. In fact, I find that all of syntactic theory turns crucially on the different behaviors of these two kinds of lexical classes. The open classes are strictly limited to four lexical categories, N, V, A, and P, which I term the Dictionary. The closed classes include all other categories, and *crucially* closed subsets of the lexical categories, and this component I call the Syntacticon. In terms of this section, both *s(e)* and the locative-assigning Ps of Czech are in its Syntacticon, while its other Instrumental-assigning Ps are in its Dictionary.

The limited scope of this study does not allow for any more than a summary of the fundamen-

[23] Since the locative P *u* 'at the house of' seems to have a semantic feature f indicating something like "the area/ house of", we might expect it to take instrumental case. But it takes a genitive, which I assume is "quirky", like the genitive with *do* in (10). Compare the French P *chez*, which has *only* the "house of" meaning.

[24] Leech's intuitive theory nonetheless fails to predict certain aspects of Czech case distribution. First, his semantic dichotomy would lead to expecting locative case for the bare "instrumentals of location" in (3), since these do not express orientation or relative position. Second, whenever a P of relative position implies further some kind of separation of two objects (*bez* 'without', *kromě* 'except', *koncem* 'at the end of', *nedaleko* 'not far from'), Czech uses a genitive rather than an instrumental, suggesting a role for the feature SOURCE, which in Leech's system has no special status. I conclude that syntactic features and the structure of the lexicon make better case predictions than his purely semantic investigation.

tal differences between Dictionaries and Syntacticons, which I reproduce with minor changes from Emonds (2005a). Not all the terms in (15) are discussed in this study.

(15) Dictionary Syntacticon

Defining property: Items with purely semantic features *f*	YES	NO
a. Syntactic properties:		
i. Grammatical categories in the inventory	N,V,A,P	ALL
ii. "Late insertion" possible during syntax and at PF	NO	possible
iii. Items with "alternatively realized" features; cf. (18) below	NO	possible
iv. Full suppletion inside paradigms (*go/ went;*); cf. note 16	NO	possible
b. Phonological properties:		
i. Items conform phonologically to "primary vocabulary"	possible	YES
ii. Bound items heading compounds have relatively less stress	NO	YES
iii. Phonetically zero morphemes (see section 8)	NO	possible
c. Intermodal and processing properties:		
i. Open classes; adults can coin neologisms	YES	NO
ii. Interface with non-linguistic memory and culture	YES	NO
iii. Processing look-up in terms of initial consonant cluster	YES	NO
iv. Processing look-up in terms of syntactic addressing	NO	YES
v. (?) Limited to Broca's area of the brain	NO	YES

5.3 The Czech "locative square": a lexical entry for the locative suffixes

As seen in the previous section, the relevant P (*o, v(e), při, po, na*) requiring locative endings that "replace" the instrumental ones should be specified as [P, F_i, -f_j], i.e. they lack any semantically specific features f characteristic of open class Dictionary items. Yet in order to get locative case, P must have some contentful syntactic feature(s) F_i other than P itself. For when P has no such additional feature, it is realized as *s(e)* or ∅, giving rise to Instrumental case spell outs in (14). So we need to now determine what this F for locatives is.

The key to the nature of F is a curious gap in the gamut of Czech prepositions. L. Veselovská (pers. comm.) has pointed out that *at* presents a difficulty for Czech-speakers learning English. This item is taught to Czech learners as conveying the sense that its object is "dimensionless" or a "point" or an "abstraction" lacking internal structure.

(16) We should stop { at/ in } Barcelona for petrol.
Cf. We should spend our vacation { in/ ??at } Barcelona.
He won't stay at the University for long.
The people at the pub were noisy.
I saw a police car at the last stop light.

Translations of these PPs into Czech must use prepositions that imply some kind of dimension or relative position; see for example Peprník (1995) who typically uses Czech *u*. On the possible "dimensionality" and quirky case of *u*, see note 23; English *at* has no such implications. That is, both *at* and *with* seem to lack a feature that Leech associates with (positive) "dimensionality". For this concept of conveying that an object DP has spatial dimensionality, I propose a syntactic feature of P labeled EXTENSION. Except for the completely unmarked *s(e),* all closed class Czech Ps that are –PATH are then +EXTENSION: *o* 'about', *v(e)* 'in', *při* 'near', *po* 'over, along', *na* 'on'. From this perspective, *with* is a pure exemplum of [P, -PATH] with no further specification. [25]

These considerations allow us to write a lexical entry for the Czech locative affixes. The following entry does not reflect the similar morphological shapes of locative and dative singulars in Czech; section 7.2 returns to the possible significance of this.

(17) Lexical entry for Czech **Locative Marking**:
<[+N]___>, EXTENSION, $\pi_{locative}$

We now come to the crux of the puzzling "locative square" in (5) and (13). Why does (17) never spell out locative case on objects of Czech *open class Ps* of *EXTENSION,* for example those meaning 'on top of' or 'in front of' or 'alongside'? A revealing answer apparently requires further restricting a fundamental notion that distinguishes Syntacticon from Dictionary items. The relevant principle, which I and others have justified in too many works to list, is given in Emonds (2000; 125): [26]

(18) **Alternative Realization ("AR")**. A syntactic feature F associated in UG with category B can be alternatively realized in a *closed class grammatical morpheme* under X^0, provided that X^0 is the lexical head of a sister of B^j.

Now Locative Marking (17) is clearly a case of AR, where F is EXTENSION, B (and B^j) is P^0, and X^0 is $[+N]^0$. The closed class grammatical morphemes are the D_P spelled as $\pi_{locative.}$ Note that , as indicated in (15a-iii), open class morphemes *never* exemplify AR.

However, the earlier formulated Czech case assignments for genitives (8), accusatives (11) and instrumentals (14) are also legitimate and in fact typical instances of AR under heads $[+N]^0$. In all three entries, P is again at least one of the values of B, while the copied features F of P here include *P itself* (that is, in all three rules P can be a value of X in the case feature D_X). Therefore, up to this point, the AR of Locative Marking presents no formal difference from the other case assignments. And yet it is empirically more restricted, since no open class Czech P can be spelled out as (i.e. assign) a locative case suffix. To express this difference, it appears that we must supplement AR with (19).

(19) **Restriction on open class AR**. Lexical entries can alternatively realize features of open class (Dictionary) items of category B *only if they spell out the category B itself.*

Though appearing at first glance ad hoc, this modification is an important and justified restriction. For example, it (apparently correctly) forbids say, different agreements on finite verbs that reflect only whether an open class subject noun is count or mass. And in the situation under discussion, since P is not mentioned in Locative Marking (17) and its appearance would

[25] In the terms developed here, we can probably maintain that Ps of PATH never have an EXTENSION feature. That is, PATH and EXTENSION are mutually exclusive.

[26] A summary of the many constructions from many different languages that have been or can be handled in terms of AR appears in Emonds (2000, section 4.4 and 4.5).

be superfluous, the restriction (19) derives the fact that (17) cannot apply to object nominals of open class Czech P.

6 The crucial characteristic of true case-marking

It should now be noted that Locative Marking (17) is formally not actually a "case-marking rule", because it specifies *no category with a case feature.* In this study Case Assignments, as seen in (8), (11) and (14) always have entries containing D_X where X is a case-assigning category. Leaving aside for the moment a possible relation to the dative, this lexical entry (17) for locative endings simply spells out special endings on nominal categories +N in the presence of *certain* prepositions. These spell outs amount to a more salient way to signal the presence of the special feature complex [P, EXTENSION, -f_j.].

For convenience, though the term has no real theoretical status, we can call such a pattern, in which AR spells out secondary features of a case-assigner on the assigner's object, a **"quasi-case"**. Thus, the Czech locative is a quasi-case, as is the Latin dative, compared to its true oblique case, the ablative (see note 9).

This asymmetric relation between Instrumental Case Marking (14) and Locative Marking (17) can explain why the latter takes preference over the former. A first way to explain this might invoke Pánini's and Kiparsky's "Elsewhere Principle", by which rules that apply in a more limited context, here (17), have precedence over those such as (14) that are more general.

Another line of thought may be more revealing. In a wide variety of syntactic patterns in terms of AR, the grammatical morphemes that alternatively realize canonical positions and features are not themselves part of LF. That is, such AR morphemes, for example those of quasi-cases, are not interpreted in their surface positions. They appear in these positions only subsequent to "Spell Out" and so are inserted in PF. However, Case Assignment rules such as (14) plausibly apply during a derivation *prior to Spell Out* (and are used in LF). If so, when Locative Case (17) applies in PF, it must "overwrite" the instrumental case spellings, thus rendering ungrammatical any instrumental case on objects of Czech grammatical Ps of EXTENSION. Such overwriting plausibly distinguishes true case lexical insertions used in LF *(they never overwrite each other),* from quasi-cases used only in PF *(they overwrite previous Spell Outs).*

In any case, whether due to the Elsewhere Condition or to a difference in component (narrow syntax vs. PF), Czech Locative Marking, linked to the specific feature EXTENSION on [P, -GOAL], pre-empts, overwrites, or otherwise precludes its Instrumental Case Marking.

Just as the Latin dative is not an autonomous case (Emonds 1985, Ch. 5), I conclude that there is no autonomous locative case in Czech. Its locative DPs are indeed oblique case complements of [P, -GOAL], but additionally alternatively realize the syntactic feature EXTENSION of this introductory P. The new Restriction on AR (19) predicts why only a small subset of grammatical [P, EXTENSION] can take such locative objects. This account thus reduces Czech non-vocative cases to five (nominative, genitive, oblique, accusative and dative), which is closer to but still one more than the four expected by the theoretical approach to case summarized in (1).

7 The Czech "dative square": a second quasi-case

7.1 Alternative Realization of the feature GOAL

Let us now focus on the remaining Czech case that falls outside of the restricted lists of (1) and (2), namely the dative. Datives with no overt P are used for indirect objects and also for "benefactive" nominals as in *uvařil* ***kamarádovi*** *oběd* 'made lunch for **a friend**'. Dative case is also always found with the grammatical P *k(e)* 'toward'. But overall, it is striking in how few structural contexts this case appears. Dative seems to simply spell out (alternatively realize) the feature GOAL of a P which has no purely semantic features f.[27]

[27] The preposition *ke* has the syntactic feature –CONTACT in many if not all of its uses. Although *do* 'to' expresses [GOAL, +CONTACT], its object is a lexically stipulated quirky genitive (10).

(20) Lexical entry for **Dative Marking** on Czech Nouns:[28]

$$<[+\mathrm{N}]\_\_\_>,\ \mathrm{GOAL},\left\{\begin{array}{c}\mathrm{PLUR},(vowel)\underline{m}\\ high\,vowel\end{array}\right\}$$

Besides the Syntacticon item *k(e)*, a very few Dictionary items such as *díky* 'thanks to', *vůči* 'compared to', *kvůli* 'because of', *(na) proti* 'against, opposite to' are generally listed as prepositions that take the dative. Given their lexical specificity, we can plausibly maintain that these few open class items take "quirky datives," consisting of lexical specifications for DP complements with a GOAL feature. Since the only way a DP (rather than P) can have such a non-canonical feature is via AR (18), these quirky datives are necessarily of the form [$_{\mathrm{PP}}$ [$_{\mathrm{P,\ GOAL}}$ $\emptyset$] - DP], which structure is identical to that of indirect objects and benefactives. Their dative inflection is then spelled out by (20).

Since the feature GOAL is a canonical feature of P, instances of GOAL being spelled out in dative inflections are further instances of AR. Again, it seems that the complex of P features copied in these inflections cannot have any purely semantic features f, i.e. Dative Marking (20) like Locative Marking (17) is limited to, and is therefore a test for determining, which Ps are in the Syntacticon. Since *do* 'to' takes quirky genitives, the remaining GOAL Ps are *k(e)* 'to/ toward' and $\emptyset$.

The Czech dative is thus not so much a case used in several structures as a means of morphologically realizing on DP the presence of a *minimally specified* P expressing GOAL. In this it completely resembles the Latin dative. The dative is thus also a quasi-case; whenever such Ps of GOAL have any further specification in terms of semantic features f, then in fact their objects appear with the *accusative* case, as seen clearly in the right column of Table (5). Therefore, there is no special oblique case for these GOALs, only the default "super case" of accusative, which is assigned by the general default entry (11).

7.2 Speculations on the "morphological (in)dependence" of quasi-cases

There is nonetheless an apparent difference between Latin and Czech datives with respect to their morphophonological realizations. The Latin dative is strikingly parasitic on the true oblique case, differing from the ablative only in the singular, and then only on some noun classes and some feminine agreeing adjectives. And in the patterns where a Latin dative singular differs from the ablative, it always spells out the same way, *-í*. That is, *a single morphophonological statement specifies completely the phonological differences between its datives and its ablatives* (Emonds 1985: Ch. 5). We can call this property the "morphological dependence" of the Latin dative quasi-case.

Several textbooks on Czech suggest that its locative is dependent (in the same sense as the Latin dative), though at first paradoxically, this dependence seems to partly be on the dative, itself a quasi-case. Specifically, the locative plural forms depend on genitive morphology and the singular forms on dative morphology. Hence, rather than being phonologically independent as presupposed in (17), perhaps Czech locative morphology somehow "piggybacks" on that of

[28] Much of the phonological detail, e.g. exactly which vowels appear in dative suffixes, is due to Czech (morpho-)phonology. Of interest is that the dative "high vowel" corresponding to *–a* in singulars is a palatalizing *–ě*, which perhaps should be represented as*–ye* with a high on-glide. In certain paradigms, stem-final unpalatalized dentals are separated from the dative suffix *–i* by means of epenthetic *–ov-*, possibly a "conspiracy" to remain unpalatalized (Czech datives don't seem to induce any palatalization that is not arguably part of the stem).

true cases. Here is a possible scenario.

Let us first look at the relation in plurals. Czech genitives independently require two allomorphic statements whose detailed phonology, as throughout, we must leave aside. Also as throughout, D_Q represents genitive case.[29]

(21) a. **Genitive Plural on Nouns.** <N_ _ _>, [D_Q, PLUR], { *u, í,* ∅ }
E.g. *muž-ů* 'of men', *radost-í* 'of joys', *měst-*∅ 'of towns'

b. **Elsewhere Genitive Plural.**<[+N]_ _ _>, [D_Q, PLUR], *(front vowel)ch*
Nouns of adjectival form such as *hajných* 'of gamekeeper's wives' and *vstupných* 'of entry fees' follow this pattern.

Now it turns out that (21b) is used only for genitive and locative, possessive pronouns, demonstratives, etc., but also for locative plural *nouns* N. So we need a statement (22) stating that in PF (not in LF) these nouns acquire a feature of an adjective [+N, +V] that contrasts with some feature of a noun [+N, -V].

(22) **Locative .** $D_P \Rightarrow$ [D_Q , +V] / [+PLUR, +EXTENSION, _ _ _]

Singular locatives are more closely related to datives, but nonetheless involve some rather minimal phonological differences:

(23) **Locative.** $D_P \Rightarrow$ GOAL / [-PLUR, +EXTENSION, _ _ _];
further, (i) delete *–u* following *m* (in modifiers), and (ii) *–u* ⇒ *-ě* in –ANIMATE nouns.

Keep in mind that if the "syntactically formulated" (22)-(23) replace the independent morphology statement for locatives in (17), they are nonetheless both instances of AR (18) and so have effect only in PF. That is, the syntactic features they introduce, namely D_Q, +V and GOAL respectively, have no relation to LF interpretations.

(22)-(23) express patterns that are remarked, even emphasized, in Czech textbooks. They give the impression that locative morphology pervasively depends on dative morphology. And in fact, since (22) contains only syntactic features and no phonology, (23) does constitute *a single morphophonological statement that specifies completely the phonological differences between Czech locatives and its genitives and datives.* Hence, exactly as the Latin dative is related to its ablative (oblique) case, the Czech locative quasi-case is morphologically dependent on its genitive and datives.

In view of this conclusion, it is tempting to ask whether even Czech datives might be dependent on its true oblique instrumental case. Probably only a "believer" in this study's approach, for example the author, might try. But it is curious how often the Czech datives can be derived from instrumentals by applying the following rough "switching" algorithm:

(24) Recipe for **Czech dative forms based on instrumentals.** (For this recipe, a final surface *–ou* must be treated as a form with an underlying final *-m*.)

a. If an instrumental feminine singular *noun* ends in a vowel, make it short.

b. In other instrumentals ending in a vowel, do the minimum to make them end in *–m*.

c. If an instrumental ends in *–m,* do the minimum to make it end in a vowel.

In this paper, I don't really stand by this recipe – or even by (22)-(23) – because they all depend

[29] My contention is that true cases are morphologically complex and thus independent. The fact that the morphology of a true case is irreducibly complex, as in (21), supports this view.

on unformulated "universal optimality phonology" that determines what constitutes "doing the minimum". But it seems natural that vocalizing a final *–m* consists in retaining its roundedness so that the corresponding vowel is *u*. In some perhaps exceptional situations, this recipe may falter or need modification. More tinkering is not the purpose of this study, but the very fact that a switching algorithm seems psychologically plausible supports (25).

(25) We *cannot be sure* whether Czech dative suffixes are morphologically independent, like those of a true case, or morphologically parasitic on the instrumentals.

In particular, since it is not implausible that Czech speakers might internalize a kind of "vowel vs. *–m* switching" to make datives, a rather strong hypothesis can at least be entertained:

(26) **Quasi-case affixes.** These differ from true cases (i) in spelling out only Syntacticon items (19) and additionally (ii) in being morphologically parasitic on true case affixes.

One can of course question whether Czech children can internalize three non-trivial algorithms such as some perfected versions of (22)-(24). The alternative is that they internalize the Declension Tables (spread out before me) that contain more than 100 allomorphs *for datives and locatives alone*, which are presented as somehow independent and moreover don't include clitics and all the possessive pronoun forms. Does the Czech brain learn 3 or so clever tricks at an early age, or is it doomed to plod through memorizing 100 plus allomorphs (for only 2 cases)? That is the question posed in this subsection.

8 Null vs. Phonological Spell Outs of minimally specified P

Table (5) of Czech prepositions and associated cases embodies the claim (15b-iii) that only Syntacticon items, and not open class items, can have null allomorphs. Some detailed implications of this claim are discussed in Emonds (2005b). Actually, most syntactic frameworks take for granted the possibility of phonologically null grammatical items, and do not propose null open class items. In view of this, a central question is, *under exactly what conditions can a grammatical P be null?* Let me state here a principle I have defended in many studies which answers this question. (Emonds 1987; 2000: 135)

(27) **Invisible Category Principle** (ICP). If all marked canonical features F on B are alternatively realized, except perhaps for B itself, then B may be empty.[30]

Now Case Marking lexical entries that spell out the various values of D_X as in (2), are a sort of formal prototype of Alternative Realization (see again section 5.3). Consequently, if all the *marked features* of such a P are spelled out in its object, P can be empty.[31]

We can see one effect of (27) with the use of dative case for expressing indirect objects. Cross-linguistically, in languages without morphological dative case, indirect objects can almost invariably and usually must be introduced with a minimal P expressing GOAL (English *to,* French *à,* Japanese *ni,* etc). Presumably, such P spell out the feature complex [P, GOAL, CONTACT]. Moreover, since *to* (French *à*) seems less marked than *toward* (French *vers*),

[30] As discussed in section 4.4.4 of the work cited, the fact that a category may be empty almost always leads to the situation where Economy requires that it must be empty.

[31] The Accusative Case Marking entry (11) does not alternatively realize the case category P nor the feature GOAL. Therefore, it cannot license an empty unmarked P of "motion toward".

+CONTACT is plausibly the *unmarked* value of this feature. Hence in order to satisfy the ICP (27) only the *marked* canonical feature GOAL of an indirect object P need be alternatively realized. Since the Czech entry (20) indeed alternatively realizes GOAL as dative inflections, this P introducing its indirect objects is ∅.

Minimally contrasting with this null P is the preposition *k(e)*, which spells out [P, GOAL, -CONTACT]. Here –CONTACT, a *marked* canonical feature of P, is *not* alternatively realized. Hence this P must be phonologically overt.

A similar dichotomy distinguishes the "bare instrumentals" of location in (3), from PPs introduced by the overt P *s(e)* 'with'. Keeping in mind that the LF interpretation of P is simply LOCATION (note 11), bare instrumentals express semantically a vague sense of the place of the action. Moreover, even the use of this case for which it is named, for expressing instruments, apparently has an unmarked P similar to that of indirect objects. That is, an instrument as in *jet autem* 'go by car' structurally indicates nothing more than a general *location* of the action; implying that "instrument" is not actually an explicit LF concept. The notion of "means" or "instrument" must be no more than pragmatic implication, due to cars being generally items designed for purposes of transport.[32]

Now the situation is different with accompaniment, *jet autem se studentem* 'go by car with a student'. The student is not in any sense the locus of the going, but in this phrase rather more like someone who is also going, a sort of secondary agent. That is, in "accompaniments" the sense of unmarked location is absent in LF. To express this, I use the following lexical formalism, developed in more detail elsewhere.

As in note 11, a P present in LF signifies LOCATION in the most general sense. Similarly, all other syntactic categories have such general senses: D has reference, A indicates properties, V indicates actions, etc. I claim that a marked syntactic "cancellation" feature +∅ allows the basic categories to lose their general interpretive sense: [D, +∅] loses reference (in expletives); [A, +∅] loses its property interpretation (e.g. in verbal passive participle affixes); [V, +∅] loses its activity interpretation (in stative verbs), etc. Similarly, those Ps which do not denote any type of location, such as English *of*, agentive *by,* the *with* of accompaniment, *despite,* etc. are all similarly lexically specified as [P, +∅].

Such is also the situation for Czech [$_P$ *s(e)*] 'with' when its object represents (a) accompaniment, (b) emotion or (c) mental perception or activity. The P carries no sense, even a vague one, of the action's physical location. Hence, the locational nature of P must be cancelled in LF by the specification [P, +∅]. This marked feature +∅ precludes a null P, so no *s(e)* in (28) can be omitted. (Examples of M. Martinková gratefully acknowledged.)

(28) a. Tom vešel s kamarádem dovnitř.
'Tom went with a friend inside.'
Objednali si margaritu se solí.
'They ordered a Margarita with salt.'

b. Šel do toho s velkým odhodláním.
'He went into it with great determination.'
Mluvil o ní s láskou.
'He spoke about her with love.'

c. Chci pokoj s výhledem na moře.
'I want a room with a view on the sea.'
S chutí začali pít, ale pili s mírou.

[32] Of course, when something is used as an instrument, there is no interest in more detail about its location with respect to the action, so no specific preposition is used.

'With gusto they started to drink, but they drank with restraint.'

There are no different "senses" of *with* here, only a general means of syntactic juxtaposition lacking any implication of location, in both Czech and English. Now, the cancellation feature $+\emptyset$ of *s(e)* and *with* is both marked and unrelated to any Alternative Realization. So if *s(e)* with the feature $+\emptyset$ were not spelled out, the ICP (27) would be violated.

The cancellation feature $+\emptyset$ thus explains the alternation pattern of *s(e)* with a null morpheme in the top row of Table (5). The ICP is an integral part of this explanation, as it is also in explaining the difference between Czech datives with and without *k(e)*.

9 Conclusion: Czech has the same four structural cases as other I-E languages.

We started by observing that Czech Ps seem to "take" five different morphological cases on their objects, counter to some more studied I-E systems which apparently allow only three such cases. But we have now seen that the Czech locatives and datives are not actually full-fledged cases. Like the latter, they alternatively realize the categories and features of introductory Ps. But unlike them, lexical entries for datives and locatives *don't spell out case features* of the form D_X as in (2), but rather only syntactic sub-categories of P: GOAL and EXTENSION. Consequently, the "fine-tuning" of Alternative Realization (18), by means of the new Restriction on AR (19), predicts that dative and locative assignment is permitted only in the presence of closed class P, that is, Ps that are in the Czech Syntacticon.

Turning around the perspective, studies of the Dictionary/Syntacticon dichotomy have not previously had clear criteria for assigning individual Ps to one or the other lexical component. The syntax of Czech cases, however, provides sharper ideas and criteria. For example, there are very few GOAL Ps in the Czech Syntacticon, only *k(e)* and $\emptyset$, while there are somewhat more LOCATION items, namely the Locative-assigning Ps that are core items expressing EXTENSION or what Leech (1969) calls "dimensionality."

The Czech special "quasi-cases," the Locative and the Dative, interestingly share many phonological spell outs (section 7), which further reinforces their similar theoretical status. Beyond this, however, the Czech syntactic system of morphological cases and of Ps does not differ in any essentials from those of its non-Slavic I-E relatives, especially Classical Greek, with which it shares the properties of (i) using Genitives for "motion away from" and (ii) expressing objects of Dictionary Ps, many expressing GOALs, with what is, following the suggestion of L. Veselovská, the unmarked structural case Accusative.

To me, the most intriguing remaining question is how to formally express in the lexicon or PF the similarities of Czech datives and locatives regularly remarked in books that teach Czech. These similarities, once they are captured, should shed some light on a notion of "morphological (in)dependence." That is, the allomorphs of the true case suffixes $\pi_{accusative}$, $\pi_{genitive}$ and $\pi_{oblique}$, seem to be significantly more diverse (and different from each other) than the quasi-case suffixes π_{dative} and $\pi_{locative}$. But for the moment, I cannot fully explain the phenomena, but only note that Czech textbooks suggest that it is a very real one.

References

Baker, Mark. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–416.

--. 1988. *Incorporation: A Theory of Grammatical Function Changing.* Chicago: University of Chicago Press.

Brame, Michael. 1984. The Head-Selector Theory of Lexical Specifications and the Non-existence of Coarse Categories. *Linguistic Analysis* 10:321–325.

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

--. 1995. *The Minimalist Program.* Cambridge: MIT Press.

Emonds, Joseph. 1985. *A Unified Theory of Syntactic Categories.* Dordrecht: Foris.

--. 1987. The Invisible Category Principle. *Linguistic Inquiry* 18:613–632.

--. 1993. Projecting Indirect Objects. *The Linguistic Review* 10:211–263.

--. 2000. *Lexicon and Grammar: the English Syntacticon.* Berlin: Mouton de Gruyter.

--. 2003. Two Types of Lexical Access: Why Syntacticon Items can be Null. *Theoretical and Applied Linguistics at Kobe Shoin* 6:1–24.

--. 2005a. The Computational Lexicon. *English Linguistics* 22:232–266.

--. 2005b. Syntactic Conditions on Phonetically Empty Morphemes. Broekhuis, H., N. Corver, R. Huybregts, U. Kleinhenz & J. Koster (eds.). *Organizing Grammar. Linguistic Studies in Honor of Henk van Riemsdijk*: 111–121. Berlin: Mouton de Gruyter.

--. 2007. *Discovering Syntax: Clause Structures of English, German and Romance.* Berlin: Mouton de Gruyter.

Holá, Lída. 2000. *Czech Step by Step.* Prague: Fragment Publications.

Jackendoff, Ray. 1977. *X-bar Syntax.* Cambridge: MIT Press.

--. 1983. *Semantics and Cognition.* Cambridge: MIT Press.

--. 1990. *Semantic Structures.* Cambridge: MIT Press.

Leech, Geoffrey. 1969. *Towards a Semantic Description of English.* London: Longman Group.

Marslen-Wilson, William D. 1987. Functional Parallelism in spoken word recognition. *Cognition* 25:71–102.

--. 1990. Activation, competition and frequency in lexical access. *Cognitive Models of Speech Processing.* G.T.M. Altmann, ed. London: MIT Press.

Naughton, James. 1999. *Colloquial Czech, 2^{nd} edition.* Oxon: Routledge.

Niederle, Jindřich, Václav Niederle, and Ladislav Varel. 1991. *Mluvnice řeckého jazyka.* Praha: Scriptum Publishers.

Peprník, Jaroslav. 1995. *Angličtina pro pokročilé.* Olomouc: FIN publishers.

Riemsdijk, Henk van. 1978. *A Case Study in Syntactic Markedness.* Dordrecht: Foris.

--. 1998. Categorial feature magnetism: the endocentricity and distribution of projections. *Journal of Comparative Germanic Linguistics* 2:1–48.

--. 2002. The unbearable lightness of GOing. *Journal of Comparative Germanic Linguistics* 5:143–196.

-- and Riny Huijbregts. 1998. *Interface in Space–How Natural Language Expresses Spatial Relations.* Paper presented at the Third International Symposium on Language, Logic and Computation. Batumi.

Rouveret, Alain and Jean-Roger Vergnaud. 1980. Specifying Reference to the Subject: French Causatives and Conditions on Representations. *Linguistic Inquiry* 11:97–201.

Veselovská, Ludmila. 1998. Possessive Movement in the Czech Nominal Phrase. *Journal of Slavic Linguistics* 6(2):255–300.

Veselovská, Ludmila. 2001. Agreement Patterns of Czech Group Nouns and Quantifiers. In: Corver, N. & H. van Riemsdijk (eds.). *Semi-Lexical Categories*: 273–320. Berlin: Mouton de Gruyter.

Mixed Nominals in Czech*

Petr Karlík

pkarlik@phil.muni.cz

1 Introduction

In the generative tradition, the mixed categories pose a problem from the very beginning. The thing is that regarding the postulate of endocentricity of syntactic constructions, in the lexical-based models the structures showing morphological and syntactic properties of two different "categories" X and Y (even though their lexical heads are single words, i. e. X or Y) present a potential problem. Standard generative insight into this type of mixed categories is that they have the external syntax of a X projection and the internal syntax of both a Y projection and of X or Z projections. For mixed categories, it is typical that a phrase, exclusively a constituent of XP, co-occurs with a constituent exclusively found in YP. Another specific property of mixed categories is their internal word structure the root of which is a lexical head of a mixed phrase. If, according to the lexicalist hypothesis, syntactic categories as V, N, A are related to morphological forms, mixed categories prototypically show anomalous morphological forms of words as noun, verb, adjective: it often happens that mixed category phrases are headed by words which appear to be morphophonologically ambiguous or neutral between the two categories of the mixed phrase.

A stereotyped example which complies with the above mentioned description of a mixed category are Italian phrases headed by the *infinito sostantivato* (1). The distribution of these phrases is identical with the distribution of DPs, see (1a) and (1a'), while their internal syntax is hybrid: In the internal structure, determiners, possessive and qualifying adjectives, i. e. expressions being typical constituents of a phrase with a lexical head N appear before *eseguire*. As for the postpositions, *eseguire* can be followed by a direct object and by adverbs, i. e. expressions being typical constituents of a phrase with a lexical head V:

(1) il suo continuo eseguire la canzone impeccabilimente
the his/her continual perform.INF the song impeccably
'his continually performing the song impeccably' (Zucchi 1993:55)

a [il suo continuo eseguire la canzone impeccabilimente] mi piaceva

a' [la cancone] mi piaceva

Another instance of a mixed category are English phrases headed by the *verbal gerund* (2). Identical distribution as with DP shows their external syntax of nominal projection, see (2) –

*The present study is a part of the project MSM 0021622435, the goal of which is an analysis of the Czech mixed categories from the diachronic point of view. The present-day Czech data and their analysis should be one of the conceptual and empirical diagnostics for this research.

(2a), their accusative case assignment and adverbial modification show their internal syntax of verbal projection, see (2) – (2b). At the same time, in the internal syntax of those phrases, expressions appear which are known as typical constituents of NPs, namely possessives (2'), but not determiners (2''). Further, there can be occurrences of expressions which are not constituents of noun phrases or unembedded vPs or higher projections containing the vP; e.g., in (2''') the subject receives accusative case. Morphonologically, *verbal gerunds* are ambiguous between the two categories of the mixed phrase as well ($_{\mathrm{V}}$*destroying the book* × $_{\mathrm{N}}$*destroying of the book*):

(2) [John deliberately destroying the book] annoyed everybody
(Alexiadou (2005), slightly modified by the author)

a. [my friend] annoyed everybody
b. John deliberately destroyed the book

(2') John's / his destroying the book
(2") *that / *the destroying the book
(2"') John / him destroying the book

In Czech, syntactic mixed categories carrying nominal and verbal properties similar to Italian phrases headed by the *infinito sostantivato* or English phrases headed by the *verbal gerund* apparently do not exist. On the other hand, phrases headed by words like (a) *stavění*, (b) *stavba* seem to be potential candidates for mixed categories, even though their properties are different from those defined as mixed category typical. The external distribution of these phrases is identical with DP (3) – (3a) and their internal structure shows prototypical morphosyntactic features of nominal phrases, i. e. no accusative case assignment but genitive case assignment (3b), and no adverbial modification but adjectival modification (3c). In common with DPs, possessives can occur within them (3d). Moreover, nominal properties can be seen in the inflection of *stavění* and *stavba* realizing the features Gender, Number and Case:

(3) [stavění / stavba domu] ho unavil-o/a

a. [Marie] ho unavila
b. staví dům × stavění / stavba *dům stavění / stavba domu
c. rychle staví × *rychle stavění / stavba rychlé stavění / stavba
d. Petrovo kolo Petrovo/a stavění / stavba × *Petrovo čte

The aforementioned properties of Czech phrases headed by *stavění* and *stavba* make them comparable to English phrases headed by the *nominal gerunds* (mixed nominalizations), as (4), and to phrases headed by the *derived nominals*, as (5), with both external and internal syntax of DP.[1] Those English phrases – unlike the phrases with verbal gerunds[2] – do not have accusative case assignment, but an of-PP one (see (4), (5) × (2)), they can have a determiner (while this is impossible for the verbal gerunds (see (5a) × (2''))), they have adjectival modification while the verbal gerunds have adverbial modification (see (5b) × (2)), etc.:

[1] In the generative tradition, the differences among verbal gerunds, nominal gerunds and derived nominals and phrases headed by them have been described innumerable times, from Lees (1960), Chomsky (1970), Abney (1987), Grimshaw (1990), Marantz (1997) to Harley & Noyer (2000), Borer (2003), Harley (in press) and others. The mentioned studies have been used by current authors as argumentation in favour of influential generative theories. It is not my intention to comment on all the mentioned analyses, nevertheless, I am going to use some of the data presented in them.

[2] I take into account only the differences relevant for the analysis of Czech data. Therefore, I omit the differences concerning particle shift, e. g., H. Harley & R. Noyer (1998).

(4) [John's destroying of the book] annoyed everybody

(5) [John's destruction of the book] annoyed everybody

a. the destroying / the destruction of the manuscript

b. the deliberate destroying / destruction of the manuscript

Verbal features are essential in order to consider the phrases headed by words of the type *stavění / stavba* with distribution identical with NP/DP instances of a mixed category. In both the external and in the internal syntax of the phrases sub (3) no verbal features can be seen so far; nevertheless, they can be expected at least in the internal structure of the words whose root is their lexical head: there is an apparent relation between *stavění / stavba* and *stavět* and, if there is (in the lexicalist framework) [$_{V}$stav-ě-t], there must be a derivation towards [$_{N}$[$_{N}$[$_{N}$[$_{V}$stav-ě] $_{N}$n] ?í] -inflection], [$_{N}$[$_{N}$[$_{V}$stav] $_{N}$b] -inflection].

Verbal gerunds having syntax properties similar to related VPs, while nominal gerunds and derived nominals have the properties of NP/DP, predict that an auxiliary can be present within phrases headed by the former (6) and, that this is not possible within phrases headed by the later (6a), a well known fact:

(6) John's having criticized the play annoyed us

a. *John's having criticized of the play annoyed us

For this paper, another well known fact is more important, namely that in auxiliary constructions aspect can be expressed on the verb, but not tense or mood:

(7) Mary's eating the last piece of cake
(Baker 2005:11)

a. Mary's having eaten the last piece of cake

b. *Mary's will eating / willing eat...

c. *Mary's might eating / mighting eat...

It is exactly the potential to express aspect (but not tense and mood) which lines up the Czech phrases headed by *stavění* with the English phrases headed by the verbal gerund: Part of the internal structure of the word *stavění* can be the same affixes which in the parallel Czech VPs realize either the aspect: prefixes (8) or suffixes (8'), or the so-called Aktionsarts: prefixes (9), or suffixes (10). On the other hand, in the Czech phrases headed by *stavba* the prefixes and suffixes are present in the internal structure of the phrase just exceptionally (which may be idiosyncratic), see (8) – (10) × (8a) – (10a). At the moment, we do not have any evidence to claim that in both types the affixes realize the aspect identically with VPs (to be mentioned furthermore):

(8)	stavění–dostavění	stavět$_{ipf}$–dostavět$_{pf}$
a.	stavba–dostavba	
(8')	dostavění–dostavování	dostavět$_{pf}$–dostavovat$_{ipf}$
a.	dostavba–∅	
(9)	střelení–zastřelení	střelit$_{pf}$–zastřelit$_{pf+feature\ x}$
a.	střelba–*zástřelba	
(10)	hraní–hrávání	hrát$_{ipf}$–hrávat$_{ipf+feature\ y}$
a.	hra–* ∅	

The present data provide some preliminary support for a hypothesis that phrases headed by the words of the type *stavění* are "more verbal" then the phrases headed by the words of the

type *stavba*. I will try to support this hypothesis by an analysis in the non-lexicalist theoretical framework because in the lexicalist analysis, similar conclusions were reached by L. Veselovská (2001). Since I am not a generativist, my argumentation will consist in gathering of the relevant data and in relating them to classical analyses of nominalizations.

2 Analysis

Starting from the standard non-lexicalist hypothesis (A. Marantz, 1997, and many others), according to which the nominalizations are Spell-Out categorially neutral $\sqrt{}$P in a D-context: An acategorial root becomes a nominal by a head raising of the structure containing it (and any features associated with typically verbal features) to a nominal head realised by a nominal suffix. Tentatively, it can be presupposed that if the structures of the type *stavění* and *stavba* are nominalizations, then they are structures on Spell-Out which are a result of a derivation: this derivation is entered by a categorially neutral lexical head root $\sqrt{\text{STAV}}$ and the categorial properties of the projected phrase are determined by the functional heads received during the syntactic derivation. It is obvious that structural portions of such a projected phrase must be smaller than TP because (among others) there is no Nominative in the phrases with *stavění* a *stavba*: *Petr čte* × **Petr četba / čtení*.

2.1

The first step of the analysis (with a direct empirical support) means to distinguish the nominalizations of the types *stavění* and *stavba*. I am using an independent observation of T. Scheer (2001): The quantity of a vowel in the prefix *za-* is determined by a [verbal] / [nominal] feature quality of the first suffix adjoined to a root. If the suffix bears a feature [nominal] the vowel is long, if the suffix bears a feature [verbal] the vowel is short. I have verified that this finding is generally valid for all the prefixes with a final vowel.[3] (The addition is that the second step of the Scheer's algorithm is the following: A long prefix gets reduced if the root contains a long vowel (the law of the three mores): *zá-stav Ø-a* × *za-stáv-k-a*.)

The data in (11) show that the prefixes in both types of nominalizations behave differently:

(11)

a.	pří-SPĚV-(e)k	vý-HR-Ø-a	ná-STAV-b-a	zá-KLOP-k-a
b.	při-SPÍV-á-n-í	vy-HR-á-n-í	na-STAV-e-n-í	za-KLOP-e-n-í
c.	při-SPÍV-á-m	vy-HR-a-(j)i	na-STAV-í-m	za-KLOP-í-m

The structures of the type *stavba* show at their prefix (which can be a part of them) that the first suffix merging to the right with the root bears the feature [nominal]; these are the *root-derivates*; see (a). On the other hand, the structures of the type *stavění* show on the prefix that the first suffix on the right merging with the root has the feature [verbal], see (b). It is confirmed by the examples in (c); therefore these cannot be *root-derivates*, but necessarily *stem-derivates*. It means (if this analysis is correct) that a mixed category, under standard concepts a phrase whose categorial features appear to change at some point in its projection line (P. Ackema & A. Neeleman (2004), among others), are only the nominalizations of the *stem-derivate* type because only those provide the change of the categorial features in the course of derivation: At first, in the syntax the acategorial root obtains verbal features through merging with a Theme-head and then, in the certain phase of derivation as a result of a merge

[3] A more complicated situation arises if together with the vowel quantity its quality changes: both *průřez* and *prořez*, but, expectedly, *průvod*, *průchod*... and, unexpectedly, *prodej*, *prohoz*.

of this "verbal" structure as a complement of a functional head little n, the categorial features of a phrase change.

2.2

I propose a common feature of all *root-derivates* is that they are licensed by a feature [event], which the root obtains from Encyclopedia: $\sqrt{}$PÁD – *pád-∅(-∅/...) (meteoritu)*, $\sqrt{}$JÁS – *jás-ot(-∅/...) (diváků)*, $\sqrt{}$VOL – *vol-b(-a/...) prezidenta* etc. This hypothesis is supported by the fact that the roots without an event structure, e. g., $\sqrt{}$BAGR, $\sqrt{}$SUCH, are from Encyclopedia furnished by the features making them ready to merge into a D-context: *bagr-∅(-∅/...)*, or into an A-context: *such-∅(-ý/...)*. If they are to be interpreted as mixed nominals, they need a larger portion of a verbal-associated structure, minimally a ThP, because the Theme-head is the node that furnishes the embedded structure with the features contributing to its event-interpretation, i. e. with the features obtained by the roots of the type $\sqrt{}$PÁD already in Encyclopedia, not in syntax. Therefore: LOV-∅(-∅/...) × *BAGR-∅(-∅/...), but : BAGR-o-vá-n-í(-∅/...); *SUCH-∅(-∅/...), but: SUŠ(-i)-e-n-í(-∅/...), or SCH(N)-u-t-í(-∅/...); as for the analysis of *root-derivates* see P. Karlík (in press).

2.3

On the other hand, even though the *stem-derivates* are furnished with event features of the root from Encyclopedia, they are not licensed by it. See the contrast: $\sqrt{}$CHOD$_{+\text{EVENT}}$: *chod-*$\emptyset_{\text{root-nom}}$ – *choz-e-n-í*$_{\text{stem-nom}}$ × $\sqrt{}$MLAD$_{-\text{EVENT}}$: $\emptyset_{\text{root-nom}}$ – *mlád-nu-t-í*$_{\text{stem-nom}}$ / *(o)mlaz-e-n-í*$_{\text{stem-nom}}$. Therefore, they can be derived from all the structures furnished (by the merge with the Theme-head in the syntax) by the feature [verbal]:[4] [*uč-e*]*-n-í*, [[*uč-i-tel*]*-ov-á*]*-n-í*.

It is not quite sure whether the portion of a structure which is merged in the D-context is smaller than vP or whether the vP is contained in it. I will try to summarize the data which can be a basis of an adequate analysis.

2.4

Observing the morphophonological structure of the *stem-derivates* represented by, e.g., (12),

(12) děl-á-n-í(-∅/...)

we can see that it contains minimally terminal nodes filled by Vocabulary Items (VI) -á-, -n-, -í and the inflexion material -∅/... (The structure (12) does not provide any evidence for terminal nodes filled by a ∅ element, nonetheless, they can be assumed, cf. *do-děl-á-vá-n-í-*). At this moment, I pass over the realization of features in the head Th (here, by VI -á-), supposing with a high level of probability that those features are of a verbal nature (supporting data see above). I consider important to find data which could be a base for reasoning what kind of features are realized by VI -n- (standing in complementary distribution with -t- (*mlaz-e-n-í / mlád(-n)-u-t-í*), uninteresting for the present paper) and, in what kind of a head? To compare (12) with (12a) and (12b) may help to find the solution:

(12) a. děl-á-n(-∅/...)

[4] In the first place, stem-derivates are not derived from non-lexical verbs which do not take aspectual prefixes but only suffixes: *muset/musívat (číst)* – ∅, *stát se/stávat se (učitelem)* – ∅, *nechat/nechával (ho zavřít)* – ∅.

b. děl-a-n(-ý/...)

There is a reason to ask a question whether the attachment of the identical material (-n-/-t-) to the stem of the stem-derivates (12), passive participles (12a) and passive adjectives (12b) is random or motivated. One of the options is that this suffix is inserted into a head responsible for a passive morphology. Stem-derivates, passive participles and adjectives share a part of the morphophonological form [děl-á-n-]. At the same time the head into which (in passive participles) VI -n- (or -t-) is inserted is, according to classical analyses, responsible for the passive nature of the phrase. If we accept Harley's (in press) hypothesis, that the analysis of structures proposed for a form must also be contained within the analysis of any structure derived from that form,[5] it necessarily follows that also the stem-derivates must be structures with a passive morphology. The idea that all the nominalizations (more exactly: all the process nominals), including structures without a visible passive morphology, are passive structures has been known since H. Borer (2001) and others. Borer's hypothesis has been used by M. Engelhardt & H. Trugman (1998) to explain (13) in Russian:

(13) konspekt lekcii brata × *konspektirovanie lekcii brata

According to this analysis, the nominal *konspekt* does not contain any VP: the genitive case of the subject *brata* is licensed by D, whereas the complement *lekcii* is inherently case-marked by the lexical head *konspekt*. But on the other hand, *konspektirovanie* contains a projection of V incorporated into the head noun: genitive internal argument occupies the [Spec,VP] position of the VP contained within the DP (hence it is subject) as a result of passive formation prior to the incorporation of V-to-N. Within process nominals, the internal argument being the subject, not the complement, the only source for the licensing of the genitive case is D. The proposed analysis claims that the stem-derivates contain a large portion of a verbal structure including a vP projection. The putative evidence supporting the analysis of process nominals as passive forms concerns: (a) the availability of instrumental agents (see further), (b) failure of the incorporated verbal head to assign accusative case to its internal argument etc. This analysis is not entirely unproblematic for the approaches cooperating with distributed morphology, because its part is not a syntax-based approach to morphology (e.g., no acategorial roots occur). For the Czech language, it is not even descriptively adequate, since Czech process nominals can take just one genitive case[6] (like in Russian), but, unlike in Russian, they allow just one genitive case at non-process nominals, too. (The other genitive surfaces as a Possessive, in both cases):

(14) *projekt reformy premiéra *projektování reformy premiéra
premiérův projekt reformy premiérovo projektování reformy

The empirical support disputing the hypothesis of the passive status of nominalizations comes from the data which show that nominalizations and passive participles face the mentioned (Harley's) requirement of containedness in such manner that they cannot be considered related. Data in (15) show a discrepancy at unaccusatives, unergatives, argumentless verbs and at non-

[5] The present hypothesis is necessary in the syntax-based morphology and is confirmed by the following data: At stem-derivates, the genitive DP can be interpreted both as A-1 and A-2, or, as Poss, respectively (*učení Pavla*$_{\text{A-1/A-2/Poss}}$). At action nominalization, the genitive DP is necessarily interpreted as A-2 or Poss (*učitel Pavla*$_{\text{A-2/Poss}}$) since A-1 is realized by the suffix -tel- (c.f. P. Caha – P. Karlík, 2005). Therefore, in the structure ?*naučitelnost Pavla (matematice)* the DP *Pavla* is necessarily interpreted as A-2 but in the structure with a stem-derivate containing the structure of action nominalization, there is another head licensing A-1 present, and so the genitive DP is necessarily interpretable as A-1 (*učitelování Pavla*$_{A\text{-}1}$).

[6] Two of the Genitive cases can possibly occur only if one of them is a lexical case: *zbavení ženy starostí* / *zbavit ženu starostí* (J. Panevová, 2000).

actional transitives, data in (16) at reflexives:

(15) blednutí, běžení, sněžení, dostání × *blednut, *běžen, *sněžen, *dostán

(16) umývání se × *umýván se

Next: If a stem-derivate really is a passive structure it should have in a DP (17) a genitive DP with the interpretation A-2 (an internal argument) in a subject position, i. e. in [spec, nP], analogically as is a DP with the A-2 interpretation in the subject position with a nominative morphology , i. e. in [spec, TP] in CP with passive participles (17). The contrast in (17a) exploiting the binding theory shows that a nominative has the properties of a subject while a genitive does not. It means that not even these data do not support the passive nature of stem-nominalizations:

(17) kritizování učitele žákem × učitel je kritizován žákem

a. *kritizování učitele$_i$ svým$_i$ žákem × učitel$_i$ je kritizován svým$_i$ žákem

Also, most of the native Czech speakers do not consider grammatical the structure with the Possessive$_{A\text{-}2}$ (17'), either:

(17') *učitelovo$_i$ kritizování svým$_i$ žákem

I take this evaluation for relevant because for many Czechs the nominal Possessive with the interpretation A-1 binds an anaphora (see (18) × (18a). It can mean that the pronominal Possessive$_{A\text{-}1}$ has the properties of a subject while a genitive does not (19). (This is not true for pronominal Possessives; see the contrast (19)):

(18) $_{A\text{-}1}$Petrovo$_i$ kritizování svého$_i$ učitele

a. *$_{A\text{-}2}$Petrovo$_i$ kritizování svým$_i$ učitelem

(19) ?Petrovo$_i$ pobíhání ve své$_i$ pracovně × *pobíhání Petra$_i$ ve své$_i$ pracovně

(19') ???moje pobíhání ve své pracovně

Another pair of examples show that a possible evaluation of a Possessive as a subject in the phrases with nominalizations contrasts with an unambiguous evaluation of a Possessive as a non-subject in the phrases with genuine nouns (cf. important example (20)):

(20) ?Petrovo$_i$ pobíhání ve své$_i$ pracovně × *Petrovo křeslo ve své pracovně

The presented data give a good reason to conclude that the analysis of the phrases headed by stem-derivates as passive structures is not compatible with the distributed morphology and it is not descriptively adequate, either. Therefore, I can announce another hypothesis: The feature contributed by a head into which VI -n-/-t- is inserted is [gender]. This hypothesis can be easily tested, using the already mentioned finding that a nominal suffix attached to the root as first prolongs the prefix with a final vowel. Data in (21), especially the contrast (21a) × (21b) plus comparison of (21a) and (22) show that the suffix -n- has not only a nominal feature but also a gender feature. Since the gender feature is not interpretable it must be eliminated during the derivation. It can be eliminated by checking the under Agree by a closest accessible node containing gender feature: the derivate then agrees (as an adjective) with an NP: [gender: masc. / fem. / neutrum]. Alternatively, this feature must be eliminated by Merge with a head containing the feature [gender: value] and, in this case, the VI *-í-* [gender: neutrum] is inserted and a derivate is a noun:

(21) (a) zá-chyt-n-ý × (b) za-chyt-i-t
zá-chyt-k-a

(22) za-chyc-e-n-ý dopis / za-chyc-e-n-á zpráva
za-chyc-e-n-í(-∅/...)

It seems that a structure [F_1 - F_n- Th - $\sqrt{}$][7] with the event and aspect interpretation generated by a syntactic process can be considered a universal structure which can be merged further: If it is merged as a complement of little n, the result is a NP/DP, if it is merged as a complement of little v, the result is vP, in the style of the distributed morphology. What follows from this analysis is the finding that at the stem-derivates the chunks of the structure merged with a D-context contain a VP, but don't contain a vP. The acceptance of this analysis, first proposed by Marantz (1997) and accepted by others, can be supported by the following data: The absence of little v brings about the fact that the complement is not assigned an accusative case (it correctly predicts the contrast: *čte knihu* × *čtení *knihu*). Further, the absence of little v causes that a [Spec, vP] position is not available, therefore, no external argument is licensed. This analysis correctly predicts the following facts: In a transitive sentence with a potential object *Petr napomíná,* the DP *Petr* is interpretable necessarily only as a A-1 („agent") since it has been inserted into the structure syntactically, through little v, and then moved from [Spec, vP] to [Spec, TP]. In the DP *Petrovo napomínání* is the Possessive *Petrovo* possessor (inserted either through the head little n or through the head Poss) and exactly only at the stem-derivates and picture nouns it can be reinterpreted – on the principle of coercion, e.g. – both as an „agent" and as a „patient" (*Petrovo*$_{\text{Poss} > \text{A-1/A-2}}$ *napomínání*). At the nouns not contaminated with verbal features it can be interpreted only structurally, as a possessor: *Petrovo*$_{\text{Poss}}$ *auto*). The same is true about the interpretation of a Genitive DP: *napomínání Petra*$_{\text{Poss} > \text{A-1/A-2}}$[8] × *letadlo prezidenta*$_{\text{Poss}}$. If in the structure of a phrase another DP or Possessive is present, their thematic interpretation is computable from UTAH: *Petrovo*$_{\text{A-1}}$ *napomínání Pavla*$_{\text{A-2}}$, *Petrovo*$_{\text{A-2}}$ *napomínání Pavlem*$_{\text{A-1}}$.

Also the analysis according to Marantz faces problems, e.g., with the already mentioned occurrence of an Instrumental adjunct. The instrumental presupposes a passive morphology (23).[9] Other problems pose the reflexives (see below), the occurrence of vP adverbs (in the post-subject position) (24) while the sentential adverbs in this position are unacceptable (24a).[10] In (25) we can see that those adverbs do not occur at the root-derivates either:

(23) *kritizuje Petrem × je kritizován Petrem
*poslední kritika románu Petrem × poslední kritizování románu Petrem

(24) vyčištění Vašeho oblečení rychle a spolehlivě (Vám nabízí čistírna ...)
rychlé a spolehlivé vyčištění vašeho obleku (Vám nabízí čistírna)

a. *vyčištění Vašeho oblečení pravděpodobně (Vám nabízí čistírna)
pravděpodobné vyčištění Vašeho oblečení (Vám nabízí čistírna)

(25) čtení té detektivky rychle (je skoro hřích) × *četba té detektivky rychle (je skoro

[7] F represents other heads of functional projections above Theme, e.g., two heads for a lexical aspect. The Theme can combine with a genuine $\sqrt{}$ from the Encyclopedia (UČ-e-n-í) or with a complex structure derived by a syntactic process from a $\sqrt{}$ (UČ-I-TEL-ov-á-n-í).

[8] As for the enforcement of the interpretation of the genitive DP as A-2 shown by the contrast of *napomínání Pavla*$_{\text{A-1/A-2}}$× *napomenutí Pavla*$_{\text{A-2}}$ at the stem-derivates with the feature [aspect: perfective] see below.

[9] A counterexample can be found in the Infinitival phrases in ECM-constructions: *Petr nechal* [*Marii přečíst básničku*] × *Petr nechal* [*básničku přečíst Marií*].

[10] A similar Czech example was brought by L. Veselovská (2001); as for the analysis of the phenomenon, see J. Fu – T. Roeper – H. Borer (2001).

hřích)

At the moment, we give up the solution because it is necessary to deal also with the properties of the head into which the suffix -í- is inserted. The example (26) brings an assumption that the Czech *s*tem-derivates are mass nouns:

(26) Balení těch dárků ho unavilo × *Dvě balení těch dárků ho unavila

Stem-derivates as mass nouns behave analogically with pluralia tantum, namely they co-occur with generic and not with cardinal numerals (26a); the predicate agreement confirms that they are in singular (in the same way as the pluralia tantum are in plural (27)):

(26) a. *Dvě balení těch dárků ho unavila × Dvojí balení těch dárků Petra unavil-o/*-a

(27) Jedny housle ležel-y/*-a na pohovce

It looks that all the stem-derivates are compositionally mass nouns, and therefore, in accordance with standard opinions (e.g., H. Borer (1999-2001), H. Harley (in press)) process nominals. If in general it is true that mass nouns can be turned into count nouns (e.g., genuine nouns in syntax with the help of classificators: *dobytek*$_{\text{mass}}$ > **dva dobytky* > *dva kusy dobytka*$_{\text{count}}$ × *pes*$_{\text{count}}$ - *dva psi* > **dva kusy psa*), it is necessary to expect that also the stem-derivates as the instances of mass nouns allow for this change. The shift of mass nouns into count nouns is here started by a context, as shown in (28). It manifests itself both in the fact that stem-derivates$_{\text{count}}$ can occur in singular and in plural (29), and in their idiomatic interpretation as result nominals denoting especially a result (*hlášení*, *pohoštění*, *čalounění* ...), a place (*propadání*, *stoupání* ...), means of activity (*krmení*, *osvětlení*, (*lyžařské*) *vázání*, *oblečení* ...) etc.:

(28) Balení$_{\text{mass}}$ ho unavuje × Balení$_{\text{count}}$ je vodotěsné.

(29) *Dvě balení$_{\text{mass}}$ těch léků ho unavila × Dvě balení$_{\text{count}}$ těch léků byla v ledničce

The presented data show that *-í-* could be VI manifesting the feature [mass]; therefore the feature [gender] has a value "neuter", cf. *stář-í*, *list-í* ..., *cukr-ov-í*, *kř-ov-í* ..., *pan-stv-í*, *sochař-stv-í* ..., *pod-hůř-í*, *ná-vrš-í*, *Po-lab-í*.[11]

The distinction process nominals (PN) × result nominals (RN) is relevant for the analysis of *stem-derivates* because each of the types shows a different syntactic behaviour (since J. Grimshaw, 1990) an obligatory component of all nominalization theories):

(a) Only subjects and complements of PN can be interpreted as arguments:

(30) Petrovo$_{\text{A-1}}$ oblečení$_{\text{PN}}$ Marie$_{\text{A-2}}$ trvalo dlouho / Petrovo$_{\text{A-2}}$ oblečení$_{\text{PN}}$ trvalo dlouho
Petrovo$_{\text{Poss}}$ oblečení$_{\text{RN}}$ se sušilo na dvoře

(b) PNs can compositionally express the values of the feature [aspect] (*psaní* – *napsání, opsání* – *opisování*)[12] and the values of the feature [quant] (*psaní* – *psávání*).
- RNs have only one aspect form (idiosyncratically, either perfective, or imperfective), but not an aspect meaning deductible from it:

[11] It is necessary, though, to explain a single exception, namely *paní* [Fem, Count], and numerous neuter nouns with the feature [count] (*pondělí*, *ústí* ...).

[12] I suppose that a good diagnostics can be seen in the preposition *během* selecting a complement with the feature [aspect: imperfective]: *během balení* × **během zabalení / během zabalování.*

(31) Balení$_{PN}$ / zabalení$_{PN}$ toho léku trvalo chvilku

a. Balení$_{RN}$ / *zabalení$_{RN}$ toho léku bylo poškozeno

(32) Jeho propadání$_{PN}$ / propadnutí$_{PN}$ už v základní škole nás mělo varovat

a. Navštívili jsme Rudické propadání$_{RN}$ / *propadnutí$_{RN}$

(c) PNs can be negated, RNs can not:

(33) Balení$_{PN}$ / nebalení $_{PN}$ svačiny do sáčku se mu vymstí

a. Balení$_{RN}$ / *nebalení$_{RN}$ léků bylo poškozeno vodou

(d) PNs can be modified by vP adjectives (and adverbs) marking the presence of an external argument, RNs can not:

(34) Rychlé / úmyslné balení$_{PN}$ těch léků ho překvapilo

a. *Rychlé / *úmyslné balení$_{RN}$ těch léků je vzduchotěsné

(e) PNs cannot be modified by the adjectives denoting the properties of actual objects, RNs can:

(35) *Modré / *papírové balení$_{PN}$ těch léků ho překvapilo

a. Modré / papírové balení$_{RN}$ těch léků je vzduchotěsné

(f) PNs are not accessible for diminution, RNs are:

(36) Při psaní$_{PN}$ / *psaníčku$_{PN}$ toho dopisu zpíval

a. Psaní$_{RN}$ / psaníčko$_{RN}$ od milenky ho potěšilo

(g) PNs allow reflexiveness, RNs do not:

(37) Petrovo$_{A\text{-}1}$ holení$_{PN}$ se trvalo dlouho

a. Petrovo$_{Poss}$ *holení$_{RN}$ se zůstalo nepoužito

It is well known that the syntactic behavior of the reflexive *se* allows the interpretation that is ceases to be a clitic and it agglutinates as a postfix into the internal structure of a stem-derivate: (38) shows the contrast with the predicted clitic-behavior, (38a) shows the contrast with the behavior of a reflexive phrase with an overt case morphology, (38b) shows that *se* blocks the interpretation of a Possessive as an A-2:

(38) rychle se oholí / *rychle oholí se × *rychlé se oholení / rychlé oholení se

a. *oholení sebe Petra × oholení se Petra

b. Petrovo$_{A\text{-}1}$ seznámení se s Marií × Petrovo$_{A\text{-}1/A\text{-}2}$ seznámení s Marií

2.5

If we observe at least some of the data pertaining to the *root-derivate* syntax, we can see that some of them can be read both as a PN and as a RN (39), others only as a RN (17) and, some only as a PN (18) which is their idiosyncratic property :

(39) Východ$_{\text{RN}}$ z kina je ucpaný
Při východu$_{\text{PN}}$ z domu zahladil po sobě všechny stopy

(40) Sprcha$_{\text{RN}}$ je ucpaná
Při *sprše$_{\text{PN}}$ / sprchování$_{\text{PN}}$ zad ucítil chlad

(41) Ozývala se dlouhotrvající střelba$_{\text{PN}}$
*Střelba$_{\text{RN}}$ / střela $_{\text{RN}}$ vylétla z hlavně velkou rychlostí

Interestingly, root-derivates$_{\text{PN}}$ behave differently from stem-derivates$_{\text{PN}}$: root-derivates$_{\text{PN}}$ cannot be reflexive (42), they cannot take negation (43), their aspect is not deductible from their internal structure (44), manifestation of their internal argument is not obligatory (45), they cannot co-occur with VP adverbs (46), they require a SC with *jako* support (47l):

(42) hraní se / si na prezidenta × *hra se / si na prezidenta

(43) nečtení novin × *nečetba novin

(44) několik let se táhnoucí stavění / *dostavění stadionu × několik let se táhnoucí stavba / dostavba stadionu

(45) po *objevení došlo k poklesu úmrtnosti × po objevu došlo k poklesu úmrtnosti

(46) ?čtení té detektivy rychle je skoro hřích × *četba té detektivky rychle je skoro hřích

(47) rozhodl se pro odejití (jako) první × rozhodl se pro odchod jako první

An important property shared by root-derivates$_{\text{PN}}$ and stem-derivates, but different from genuine nouns (see (48) × (48a)) is the ability of their subjects to bind anaphoras:

(48) ?Petrovo procházení se ve své zahradě / ?Petrova procházka ve své zahradě
a. *Petrova lavička ve své zahradě

3 Conclusions

Neither the analysis of stem-derivates supposing a portion of a verbal structure containing little v, nor the analysis without little v can predict all the contrasts presented in this paper. One of the analytic options is the analysis according to H. Harley (2005) which cannot be properly tested yet. The presented data and their analyses can be relevant for a diachronic description already in this non-definite version. Hopefully, they can be a starting point for an analysis of the relation between neuter stem-derivates with a feature [mass] realized by inserting of the VI *-í-* , at which only the consonant endings are visible (M. Ziková, in press) (Sg: Nom, Gen, Dat, Acc, Loc: *kopání-*∅, Instr: *kopání-**m**;* Pl: Nom, Gen, Acc: *kopání-*∅, Dat*: kopání-**m**,* Loc: *kopání-**ch**,* Instr: *kopání-**mi***) and feminine stem-derivates with full visible adjective inflection (Nom:*kopan-**á***, Gen, Dat: *kopan-**é***, Acc: *kopan-**ou** ...*).

References

Abney, Steven Paul. 1987. *The English Noun Phrase in Its Sentential Aspect.* Ph.D dissertation. MIT, Cambridge, MA.

Ackema, Peter and Ad Neeleman. 2004. *Beyond Morphology.* Oxford: OUP.

Alexiadou, Artemis. 2005. Gerund Types, the Present Participle and Patterns of Derivation. Maienborn, C. & A. Wöllstein (eds.). *Event Arguments: Foundations and Applications*:139–152. Tübingen: Niemeyer.

Baker, Mark. 2005. *On Gerunds and the Theory of Categories.* Ms, Rutgers University.

Borer, Hagit. 1999-2001. *The Form, the Forming, and the Formation of Nominals.* Handout.

--. 2003. Exo-Skeletal vs. Endo-skeletal Explanations. Syntactic Projections and the Lexicon. Moore, J. & M. Polinsky (eds.). *The Nature of Explanation in Linguistic Theory*: 31–55. CSLI and University of Chicago Press.

Bresnan, Joan. 1997. Mixed Categories as Head Sharing Constructions. King, T. H. & M. Butt (eds.). *Proceedings of the LFG97 Conference.*

Caha, Pavel and Petr Karlík. 2005. Where Does Modality Come from? Hansen, B. & P. Karlík (eds.). *Modality in Slavonic Languages*:61–72. München: Otto Sagner.

Chomsky, Noam. 1970. Remarks on Nominalization. Jacobs, R. & P. Rosenbaum (eds.). *Reedings in English Transformational Grammar*: 184–221. Waltham, MA: Ginn & Co.

Engelhardt, Miriam and Helen Trugman. 1998. D as a Source of Adnominal Genitive in Russian. Bošković, Ž., S. Franks & W. Snyder (eds.). FASL 6. The Connecticut Meeting 1997:114–133. Ann Arbor: Michigan Slavic Publications.

Fu, Jingqi, Thomas Roeper and Hagit Borer. 2001. The VP within Process Nominals: Evidence from Adverbs and the VP Anaphor do-so. *Natural Language and Linguistic Theory* 19:549–582.

Grimshaw, Jane. 1990. *Argument Structure.* Cambridge, MA: MIT Press.

Harley, Heidi. 2005. *Syntactic Event Structure and Nominalizations.* Handout.

--. (in press) The Morphology of Nominalizations and the Syntax of vP. Giannakidou, A. & M. Rathert (eds.). *Quantification, Definiteness, and Nominalization.* Oxford: Oxford University Press).

Harley, Heidi and Rolf Noyer. 1998. Mixed Nominalizations, Short Verb Movement and Object Shift in English. Tamanji, P. N. & K. Kusumoto (eds.). *Proceedings of NELS* 28:143–157. Amherst: GLSA.

--. 2000. Formal vs. Encyclopedic Properties of Vocabulary: Evidence from nominalisations. Peeters, B. (ed.). *The Lexicon-Encyclopedia Interface*: 349–374. Amsterdam: Elsevier.

Jelínek, Milan. 1967: Jména dějová. Daneš, F., M. Dokulil et al. (eds.). *Tvoření slov v češtině* 2: 562–653. Praha: Academia.

Karlík, Petr. (in press) Několik drobných poznámek k dějovým jménům. *Sborník prací Filozofické fakulty brněnské univerzity* A 55.

Lees, Robert B. 1960. *The Grammar of English Nominalizations.* The Hague: Mouton & Co.

Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *U. Penn Working Papers in Linguistic* 4.2: 201–225.

Panevová, Jarmila. 2000. Poznámky k valenci podstatných jmen. Hladká, Z. & P. Karlík (eds.). *Čeština - univerzália a specifika* 2:173–180. Brno: Masarykova univerzita.

Scheer, Tobias. 2001. The Rhythmic Law in Czech: Vowel-final Prefixes. Zybatow, G., U. Junghans at al. (eds.): *Current Issues in Formal Slavic Linguistics*: 37–48. Frankfurt am Main: Lang.

Veselovská, Ludmila. 2001. K analýze českých deverbálních substantiv. Hladká, Z. & P. Karlík (eds.). *Čeština – univerzália a specifika* 3:11–28. Brno: Masarykova univerzita.

Ziková, Markéta. In press. Phonological cyclicity: case markers in the Czech nominal declension. *Proceedings of FDSL 6.5* (Nova Gorica).

Zucchi, Alessandro. 1993. *The Language of Propositions and Events*. Dordrecht: Kluwer.

1, 2, SE

Lucie Medová and Tarald Taraldsen

lmedova@Princeton.edu, knut.taraldsen@hum.uit.no

1 Introduction

The goal of this article is to contribute to the understanding of the morphosyntactic properties of the reflexive clitic (henceforth: SE) in Czech (and Slavic and Romance languages in general). In particular, we want to suggest a way of unifying its various uses, although we will focus mainly on the purely *reflexive* (1) and the *impersonal* SE (2), and we want to see how SE relates to the 1^{st} and 2^{nd} person clitics, with which morphology suggests that it forms a natural class[1].

(1) *reflexive* SE

Karel se učesal.
Karel$_{NOM.SG.M}$ SE$_{ACC}$ combed$_{M.SG}$
'Karel combed [his hair].'

(2) *impersonal* SE

Hrály se karty a zpívalo se.
played$_{F.PL}$ SE$_{ACC}$ cards$_{NOM.PL.F}$ and sung$_{N.SG}$ SE$_{ACC}$
'There was some cards playing and singing.' / 'People were playing cards and sang.'

We will begin by outlining a unified analysis of reflexive and impersonal SE. Then, we will discuss certain problems which this analysis seems to run up against. We entertain two possible ways to analyze SE, first as an analogue to antipassive morpheme and second as a spell out of an abstract SELF, a part of a special possessive-like construction. As the second type of analysis handles better the empirical evidence, we elaborate on its consequences. In the last section, we return to the connection we started with at the beginning, trying to link the morphology of the reflexive clitics with person and case.

2 The raising analysis of reflexives

We first present a somewhat simplified version of an analysis partially assimilating the derivation of sentences with *reflexive* SE to the one standardly assumed for (one type of) *impersonal*

[1] Abbreviations are the usual ones, case: NOM(inative), (GEN)itive, (DAT)ive, (ACC)usative, (LOC)ative, (INS)trumental, (OBL)ique, (ERG)ative, (ABS)olutive, gender M(asculine), MA=masculine animate, MI=masculine inanimate, F(eminine), N(euter), number: SG=singular, PL=plural; DO=direct object, IO=indirect object, IA=internal argument (subsuming both DO and IO objects), EA=external argument. The Czech examples (if not cited) are from Correct Czech (*spisovná čeština*), as required by the second author.

SE. As we imply that our discussion should extent both to Slavic and Romance reflexive clitics, we show examples both from Czech and Italian, the representatives of each group.

2.1 First step: Reviving Kayne (1986)

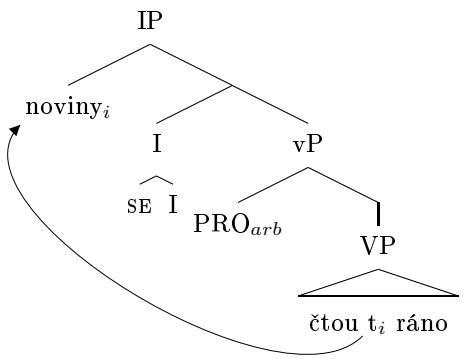
Routinely, a sentence with *impersonal* SE like (3)[2] is given a parse like (4)[3], where SE somehow licenses the occurrence of PRO_{arb} in Spec-vP. The arbitrary PRO gives rise to the (obligatory) human agent reading.

(3) *impersonal* SE

a. Noviny se čtou ráno (komiksy večer).
newspaper$_{NOM.PL}$ SE read$_{3.PL}$ morning comic strips evening
'People read newspaper in the morning, comic strips in the evening.'

b. I giornali si leggono la mattina.
the journals$_{NOM.PL}$ SE read$_{3.PL}$ the morning
'People read newspaper in the morning.'

(4)



A *reflexive* SE, on the other hand, is usually treated as an anaphoric object clitic (Burzio (1986), Dobrovie-Sorin (2005); for Czech Panevová (1999)). So, (5) on the reflexive reading would be analyzed as in (6), begging the question how the impersonal and the reflexive SE are related to each other[4]:

(5) *reflexive* SE

a. Chlapci se myjí každé ráno.
boys$_{NOM.PL}$ SE wash$_{3.PL}$ every morning

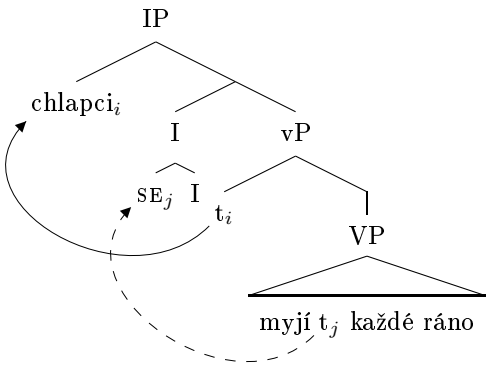
b. I ragazzi si lavano ogni mattina.
the boys$_{NOM.PL}$ SE wash$_{3.PL}$ every morning
'The boys bath themselves every morning.'

[2]The Czech version of (3), with the present tense and imperfective verb, implies somewhat deontic modality, cf. *One should read newspaper in the morning, comic strips in the evening* or habituality; in the text, we ignore this implication.

[3]To reduce clutter, we do not show V-raising to I (used agnostically to denote whatever functional head hosts a preverbal subject in its Spec). The (final) location of SE is given as I on the basis of Kayne (1986), Cinque (1988) and much other work. The choice of PRO_{arb} as the external argument (EA) accords with Kayne (1986). (A different choice will be made below.)

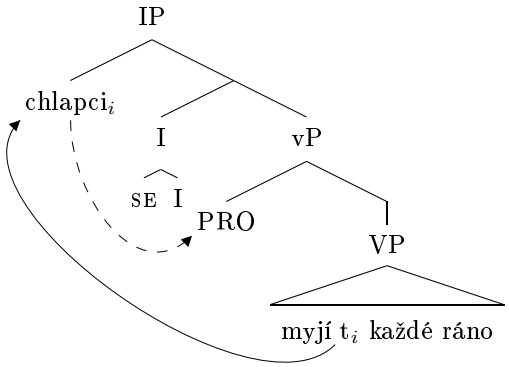
[4]However, there is a family of analyses represented by Grimshaw (1982) and Reinhart and Reuland (1993), for instance, which would see SE as the morphological reflex of a lexical operation converting a transitive argument structure to reflexive one.

(6)



Kayne (1986), however, points out the reflexive reading of (5) might actually come from a structure exactly like (4) except that PRO, the EA, doesn't have arbitrary reference, but is interpreted as controlled by the DP raised to Spec-IP:

(7)



On this account, impersonal and reflexive SE both license a PRO in the EA-position.

This analysis has the virtue of unifying the two SEs. But encounters a number of problems, two of which we will concentrate on below.

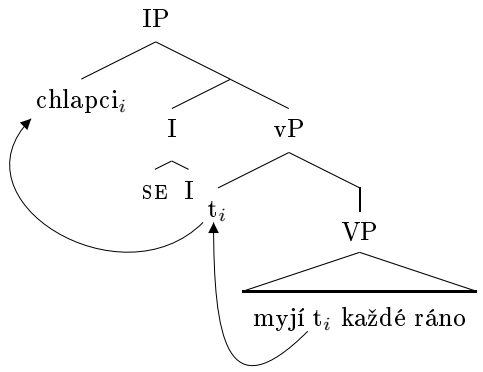
1. it seems to make it harder to understand why *reflexive* SE seems to group with the 1^{st} and 2^{nd} person clitics, if the latter are still treated as ordinary object clitics

2. by analyzing the subject of a reflexive sentence like (5) on a par with the subjects of unaccusatives, it predict – contrary to the facts – that reflexive verbs behave like unaccusatives with respect to auxiliary selection across Italian dialects[5]. Reflexives, however, select auxiliaries by and large like unergatives. Moreover, unlike unaccusatives, reflexives don't allow *ne*-extraction from a postverbal subject

Alboiu, Barrie, and Frigeni (2004) propose a way of updating Kayne's (1986) analysis which would eliminate the problems in 2. Adapting Hornstein's (1999) proposal that control is raising to a θ-position, they claim that the reflexive reading of (5) is the product of a derivation that raises the IA to Spec-vP before it is moved to Spec-IP, yielding (8) instead of (7):

[5]For Slavic, it is much harder to establish the unaccusative – unergative distinction, as the tests routinely used are unreliable. In our approach, true reflexives just must be unergative, as they 'pick up' the external θ-role.

(8)



On this view, the subject of (5) on its *reflexive* reading is an EA (the bearer of the agentive θ-role) as well as an IA, and reflexives are thus expected to group with transitives and unergatives rather than with unaccusatives.

On the face of it, however, such an approach runs straight against the θ-criterion: the single element *chlapci* in (8) accumulates two distinct θ-roles, one in the IA position and another one in the EA position. In the next subsection, we present an elaboration that cancels this violation.

2.2 Second step: θ-roles and Case

Medová (to appear) develops an analysis which shares the virtues of Alboiu, Barrie, and Frigeni (2004) account, and in addition accounts for the fact that a sentence like (9), with DAT SE, has a *reflexive* reading, but not an *impersonal* one; unlike (5) (repeated here as (10)), which is generally ambiguous between a reflexive reading and an impersonal one:

(9) *reflexive DAT* SI

a. Chlapci si myjí ruce každé ráno.
boys$_{NOM.PL}$ SE$_{DAT}$ wash$_{3.PL}$ hands$_{ACC.PL}$ every morning

b. I ragazzi si lavano le mani ogni mattina.
the boys$_{NOM.PL}$ SE$_{DAT}$ wash$_{3.PL}$ the hands every morning
1. REFL: 'The boys wash their hands every morning.'
2. *IMP: *'One washes the boys' hands every morning.'

(10) *reflexive ACC* SE

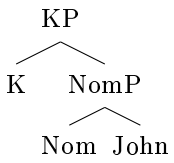
a. Chlapci se myjí každé ráno.
boys$_{NOM.PL}$ SE wash$_{3.PL}$ every morning

b. I ragazzi si lavano ogni mattina.
the boys$_{NOM.PL}$ SE wash$_{3.PL}$ every morning
1. REFL: 'The boys bath themselves every morning.'
2. IMP: 'One bathes the boys every morning.'

Her analysis incorporates Starke's proposal (Starke (2005); applied to Czech data by Caha (2006)) that a DP is in general embedded inside one or more *Case-layers* ordered in accordance with a universal hierarchy, for us, DAT > GEN > NOM. For example, the NOM subject of a run-of-the mill sentence like (11-a) might actually appear in Spec-vP as (11-b), with the NOM, the NomP-layer, embedded inside some oblique Case K:

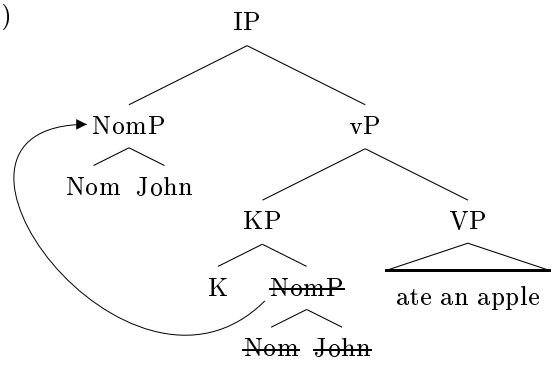
(11) a. John ate an apple.

b.



The NOM surfaces because the NomP is allowed to raise to Spec-IP from inside the larger KP, stranding K in Spec-vP[6]:

(12)



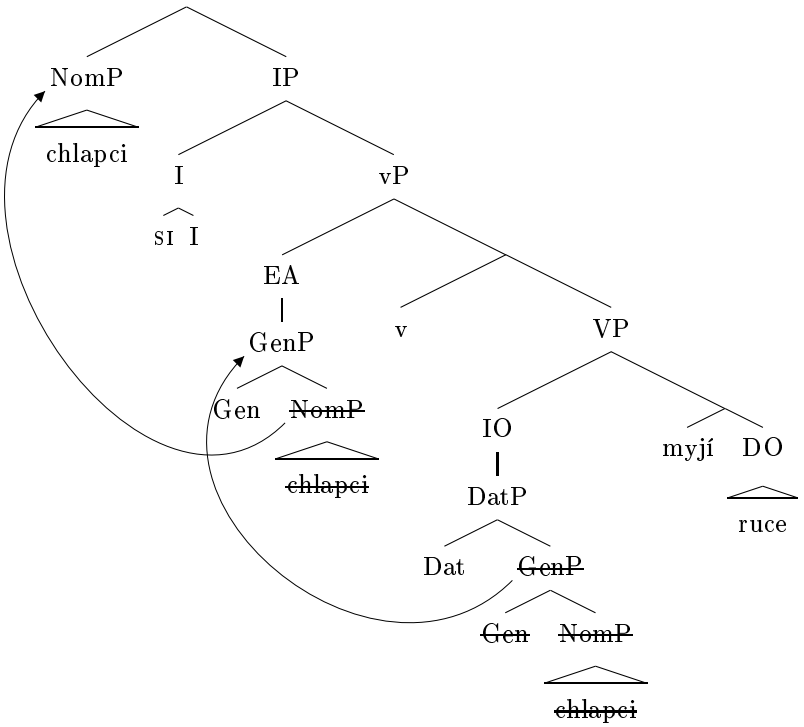
It is further assumed that the stranded K will be part of the substructure lexicalized by the verb root, under the conception of lexical insertion developed by Starke. That is, lexical insertion can target non-trivial subtrees. Like Starke, Medová (to appear) also assumes that only the sister of the highest head can subextract. This assumption is a crucial element in the explanation for (9).

Another crucial ingredient is the assumption that an EA, the element in Spec-vP, must be embedded inside a layer of oblique Case. For concreteness, we take the relevant K head to be GEN(itive). Similarly, Kayne (1993) and Mahajan (1994) have it that the EA argument is born as an OBL case, pointing toward the evidence from ergative languages. We return to the ergative connection in the section 3.1.

With this set of assumptions, each of which is independently motivated, Medová (to appear) predicts the lack of an impersonal reading with DAT SE. To reach the NOM position in Spec-IP, the DAT IA corresponding to the subject of (9) must be able to shed two layers of oblique Case, as depicted in (13).

[6]Whether or not there should be the trace (copy) of the raised NomP, depends on theoretical commitments irrelevant to our present concerns.

(13)



That is, the DP must raise in two steps: first to a GEN position, stranding the DatP-layer, and then, to the NOM-position, stranding the GenP-layer. But the only GEN-position between the DAT IA's initial position and the NOM-position is the EA-position, Spec-vP, and by moving to this position, the raised IA will pick up the agentive θ-role in addition to the one it received qua DAT IA, and a reflexive reading is produced[7].

Thus, to produce a reflexive reading, the DP has to have exactly two layers to shed: one in the place it has been born (DatP-layer in (13)) and another one in the EA-position, which, as we argued below, we take to be GenP[8].

Finally, concerning the θ-cirterion, notice that – strictly speaking – it is not *exactly* the same element that picks up the different θ-roles, as shown in (13). So, the θ-role in the IA position is associated with the DatP-layer while the EA θ-role with the GenP-layer. (We briefly discuss

[7]Another fact predicted by this is that a sentence like (i) has an impersonal reading, but not a reflexive one (for some DAT, see Medová (to appear) for discussion):

(i) Děti se vracej rodičům.
kids$_{NOM.PL}$ SE$_{ACC}$ return$_{3.PL}$ parents$_{DAT.PL}$
1. *REFL: 'The kids return (themselves) to their parents.'
2. IMP: 'The kids are being returned to their parents.'

The DAT in (i) c-commands the initial position of the IA raised to Spec-IP, so that Relativized Minimality would prevent a GenP to raise to Spec-vP from inside the IA, given that the GenP inside the non-reflexive DAT argument is closer to Spec-vP.

[8]If we take it seriously, we wind up with the prediction that only DAT (and animate, as discussed further) DPs should be able to reflexivize. It is a welcome result for a variety of Romance languages that have the animated DO introduced by a preposition *a* which is taken to be a DAT case marker, as in (i); inanimate nouns are not introduced by *a*. In Slavic, the ACC-to-GEN shift for MA nouns could be seen as an instantiation of an OBL case. See Medová (to appear) and section 4.2 for discussion.

(i) Vi a Antonio / (*a) un libro.
see$_{1.SG.PAST}$ *a* Antonio / a book
'I saw Antonio / a book.'

an alternative to this proposal in section 3.3.)

While perhaps unusual, the assumptions (each of them motivated, however) we made allow us to discuss the role of SE in the derivations of reflexives and impersonals, as, so far, the SE itself seemed rather superfluous. Also, we motivate our assumption that the Case layer associated with the EA is, indeed, GEN.

3 Internal Arguments have GEN inside

Under the assumptions so far, it is essential that the IA has a GEN inside: to derive a reflexive means to move the GenP (a GEN-Case layer) to the EA position as the first step of the reflexive derivation. Further, the NomP (a NOM-Case layer) moves from under the GenP to the NOM position, as depicted in (13).

But why should the layer associated with the IA position be GEN? And what exactly is SE needed for in the structure (13)?

There are two potential ways to substantiate the claim that the IA that derive reflexives contains in a way GenP (an OBL-Case layer, in any event) in it and that this GenP is in a particular way connected to an EA. The first option draws a parallel between the ergative languages and NOM/ACC languages via assimilating the antipassive morpheme of ergative languages to SE-morpheme of Slavic or Romance. The second option builds on the observation that GEN appears in possessive constructions and on a particular view of reflexivization going back to Helke (1971). We will discuss both options in turn, starting with the antipassive connection first.

3.1 Antipassive

Oblique case can be seen on a DO, if – for a reason – the DP does not appear in the DO position (which, essentially, translates as 'not being marked by a structural (ACC) case'). One particular instantiation of the OBL-marked DO is seen in antipassive constructions, characteristic for ergative languages[9]. So, in Chukchi (Chukotko-Kamchatkan), a transitive sentence has the EA in ERG case and the DO appears in ABS, as shown in (14-a). In antipassive, indicated by a special morpheme *-tko-* in Chukchi, the EA argument is in ABS case and the DO bears now DAT case, shown in (14-b). Notice, furthermore, that the ERG case marking of the EA is by an OBL case: a connection to our claim that EA are GEN.

(14) a. ətlə-e keyŋ-ən penrə-nen
father$_{ERG}$ bear$_{ABS}$ attack-3SG:3SG.AOR
'The father attacked the bear.'

b. ətlə-ən penrə-**tko**-ɣʔe kayŋ-etə
father$_{ABS}$ attack-AP-3SG.AOR bear$_{DAT}$
'The father rushed at the bear.'

(Kozinsky, Nedjalkov, and Polinskaja (1988)(2c,1c))

The DO thus starts out as an OBL case (DAT in (14-b)) and 'peels' to the structural DO-position – in parallel to the (oblique) DP in (13) 'peels' its way to the NOM(=structural case) position. Arguably, then, the DO has to remain OBL because the (structural) DO position is filled up with a different morpheme: *-tko-* in (14-b). This opens for an attractive analysis of

[9]But notice recent literature on antipassive constructions in NOM/ACC languages as well, for instance Postal (1977) on French, Ndayiragije (2006) on Kirundi (Bantu) and Say (2005) on Russian *sja*.

SE in (13): suppose, that in parallel to Chukchi's *-tko-*, the Czech SE 'blocks' the (structural) DO-position. The DP *Karel* in (13) starts out as DAT (an OBL case) and should move now to the DO-position, but it cannot as the position is filled by SE. Instead, it has to move to the EA position and further up to the NOM.

The idea of SE being a DO-position 'filler' is relatively well-supported. First, the SE itself is ACC-marked, if compared with the DAT-marked SI[10]. Second, SE indeed looks like an object clitic, as shown in (15) for Czech and Italian. Kayne (2000) argues that to complete the line *m-* (pronoun with the root *m*) for 1^{st} person, *t-* for 2^{nd} person, the *s-* type pronoun, not the *l*-type (as the definite articles) should be taken for 3^{rd} person. Plausibly, the same could be argued for Czech.

(15)

Czech		ACC	DAT
1^{st} person		mě	mi
2^{nd} person		tě	ti
reflexive		se	si
3^{rd} person	M&N	ho	mu
	F	ji	jí

Italian		
1^{st} person		mi
2^{nd} person		ti
reflexive		si
3^{rd} person	M&N	lo
	F	la

Third, when SE appears in the sentence, no other ACC-marked DP is possible; the other argument appears in (PP-introduced) OBL case (16).

(16) a. Karel se směje Ivoně.
Karel$_{NOM.SG}$ SE laugh$_{3.SG.PRES}$ Ivona$_{DAT.SG}$
'Karel laughs at Ivona.'

b. Karel se dotýká Ivony.
Karel$_{NOM.SG}$ SE touch$_{3.SG.PRES}$ Ivona$_{GEN.SG}$
'Karel touches Ivona.'

c. Karel se dívá na Ivonu.
Karel$_{NOM.SG}$ SE look$_{3.SG.PRES}$ on Ivona$_{ACC.SG}$
'Karel looks at Ivona.'

Fourth, SE is obligatorily introduced with certain prefixes (17). If we think that the prefixes transitivize the originally intransitive verb, then SE is, indeed, needed to fill in the DO position.

(17) Prželo. → Roz-pršelo *(se).
rain$_{3.SG.N}$ *roz*-rain$_{3.SG.N}$ SE
'It rained.' 'It started pouring down.'

However, this analysis of SE faces at least two problems. First, contrary to expectations, there are SE constructions with ACC case on another argument (18). How is the ACC assigned to *vodo* if SE, as argued above, fills the DO-position that assigns the structural ACC case to it? Second, if SE is a DO 'filler' in all of its uses, we are forced to posit a DO position for impersonal SE from intransitive verbs (19), making thus the unification of all the uses of SE rather impossible.

(18) Pilo se je vodo.
drun$_{SG.N}$ SE be$_{3.SG.PRES}$ water$_{ACC}$

[10] SE could be GEN as well: we suggest that for the analysis of impersonal SE presented in the section 3.3.

'People drunk water.'

(Slovene: Rivero and Sheppard (2003)(85c))

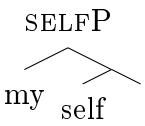
(19) Spalo se.
sleep$_{N.SG}$ SE
'People were sleeping. / Sleeping was going on.'

To summarize, the way we analyze antipassive substantiates the claim that each DO is essentially born as an OBL case and the role of SE in the derivation is to 'fill' the structural DO-position. While rather promising, this analysis still runs up against a number of problems, two of which were pointed above. Furthermore, it is not obvious why should the OBL case inside the DO should be GEN: as a matter of fact, the Czech examples in (16) show a number of different OBL case on the (DO) argument[11]. More pressingly with respect to the present inquiry, we just don't see how to extend this analysis to the DAT reflexive SI. Next, we turn to the other potential line of thinking that might suggest that an IA has a GenP inside itself: possessives and a particular view of SELF. After having introduced this idea, in the section 3.3, we will set each of the two approaches against the set of SE constructions (which we want to unify) and perhaps have a winner.

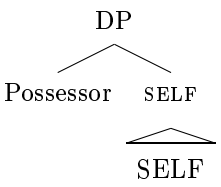
3.2 Possessor of SELF

We start by trying to solve the problem of what exactly is the SE needed for, if the derivation of reflexives must proceed as (13).

Our proposal is based on an idea going back to Helke (1971), who analyzes English SELF as a body-part nominal. Essentially, then, reflexives (20) are analyzed as a kind of a possessive construction with the head noun SELF and a pronoun.

(20) a. I hit myself.
b. 

Recently, Kayne (2005) revived Helke's analysis of reflexives extending it to Romance reflexive clitics. Adapting this proposal, we suggest that the DP which raises to the EA position in the derivation of reflexives originates as a possessor inside a complex DP, where the possessum is SELF which, according to us, spells out as SE, if the whole DP is in an ACC position, but as SI if it is in a DAT position. The structure is shown in (21).

(21) 

[11] On the other hand, in nominalizations, the ACC case routinely shifts into GEN (i).

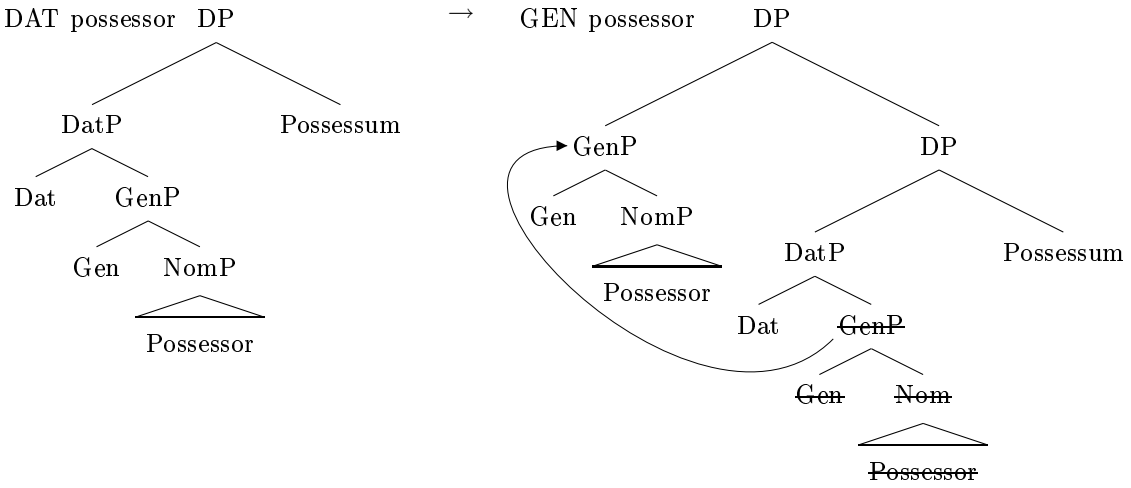
(i) překročit Alpy → překročení Alp
cross$_{INF}$ Alps$_{ACC.PL}$ crossing Alps$_{GEN.PL}$
'to cross the Alps' 'crossing of the Alps'

More precisely, we assume that the possessor DPs are in general introduced to the derivation as DatP. The motivation for this claim comes from a variety of languages. For instance, in French (22), the possessor is indeed introduced by the preposition *à*, a DAT case marker.

(22) Notre ami à tous.
our friend to all
'Our common friend.'

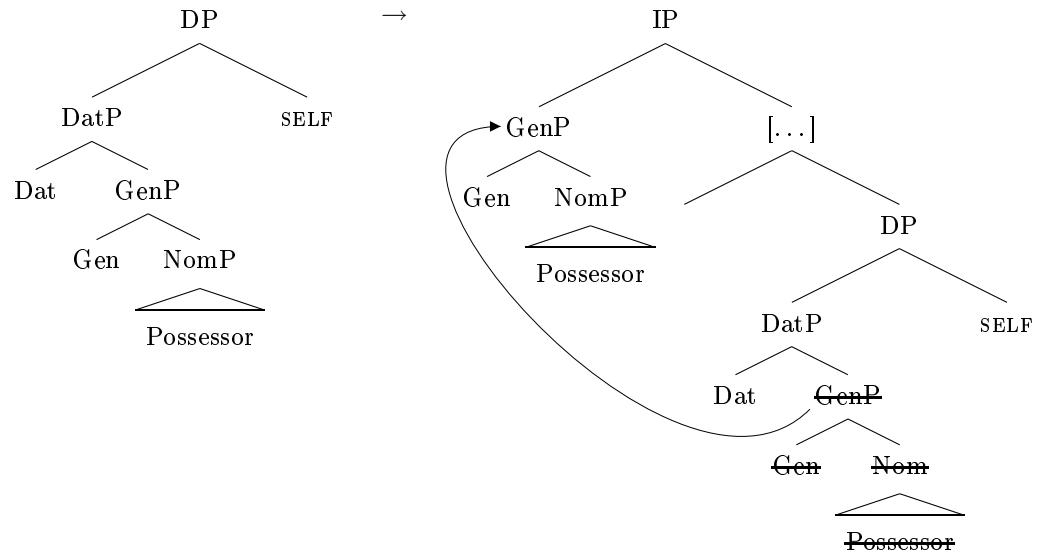
On the other hand, some possessors appear as GEN (*notre* in (22)). In our approach, as before, GenP emerges from under the DatP by a movement to a higher position in the complex DP, as shown in (23).

(23)



This higher position, we take it, is absent from the complex DPs headed by SELF, in other words, the PossessorSELF remains a DatP until the derivation reaches an EA position to which the GenP can raise from inside the possessive DatP, as shown schematically in (24)[12]. This way, the derivations we have, are consistent with the assumption that the highest head is always left behind under movement.

(24)



We will take it that SELF denotes something, which, by convention, can only be a part of an animate / human being. So, only animate DPs will combine with SELF and therefore only

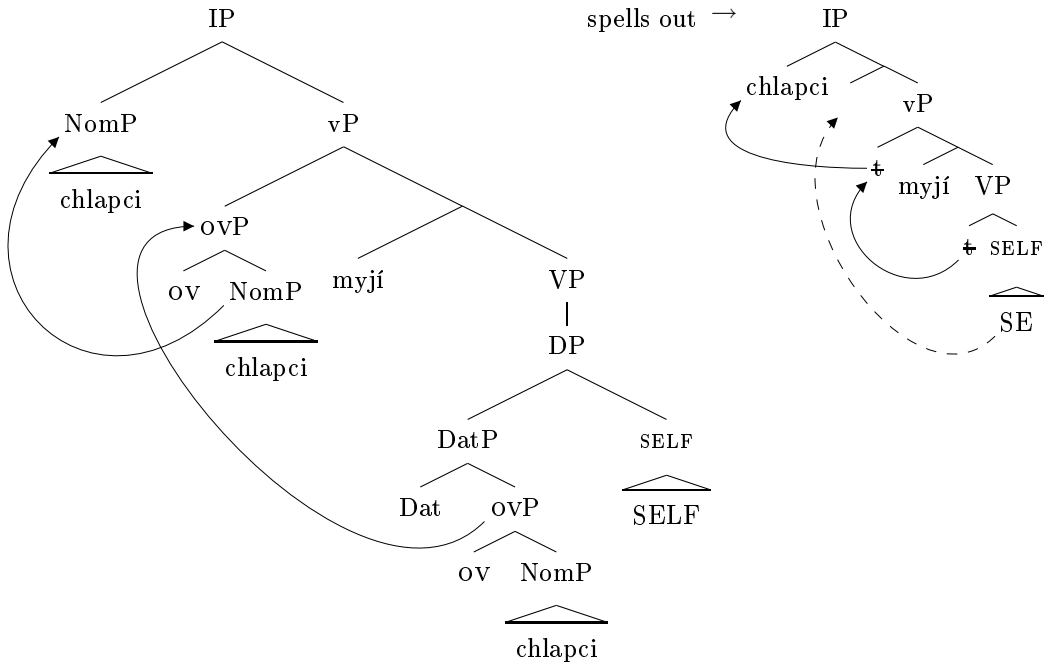
[12]In the section 4.2 we suggest an alternative.

animated DPs can reflexivize, as discussed in section 4.1.

Finally, the full derivation of reflexives (5), repeated here as (25), under these assumption would run as shown in (26).

(25) Chlapci se myjí každé ráno.
boys$_{NOM.PL}$ SE wash$_{3.PL}$ every morning
REFL: 'The boys bath themselves every morning.'

(26)



We leave it an open issue whether animate nouns are always introduced in the derivations as possessors of SELF, or, in fact, occur as possessors of SELF only when needed for the derivation of reflexives to converge. In the former case, we would have to say that the SELF part spells out independently (as SE) just in case the derivation splits the possessor from SELF (an option taken in section 3.3). Potentially, the facts that we are going to discuss in section 4.2 argue for this view.

At this point, however, we will compare the *antipassive* SE approach with the *Possessor of* SELF analysis. We want to see which analysis can explain more.

3.3 Open competition: Antipassive against Possessor of SELF analysis

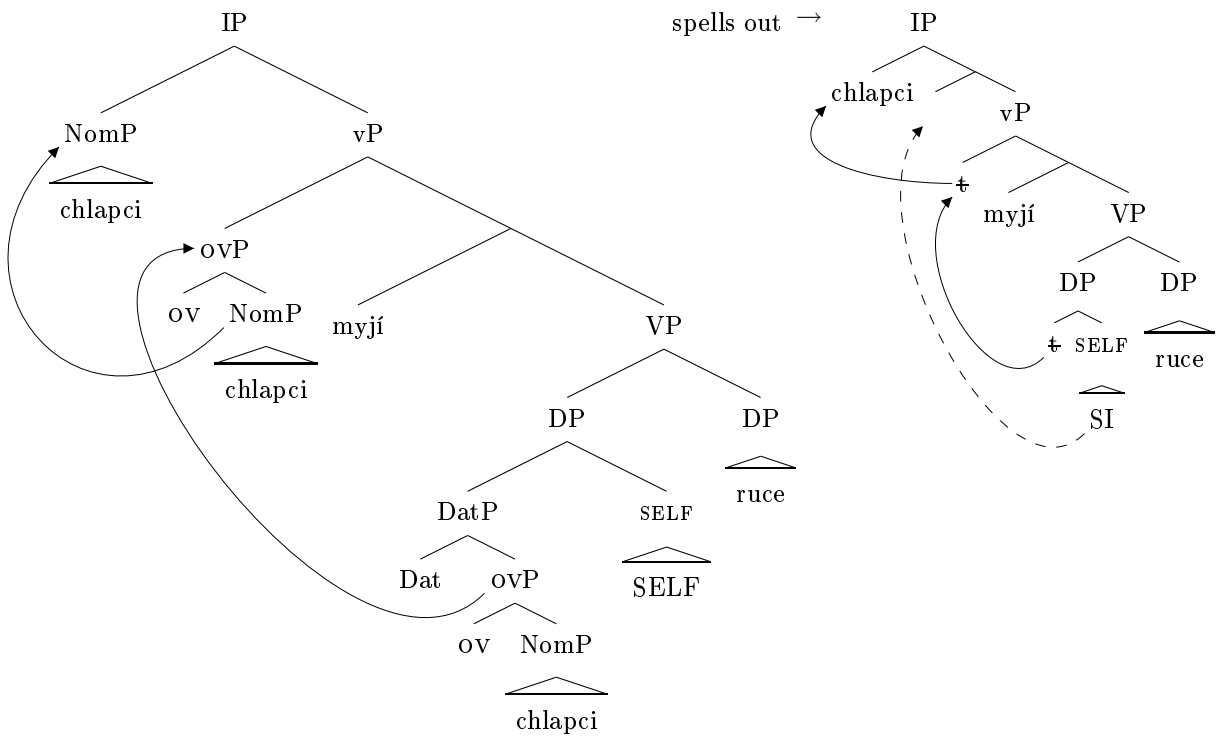
3.3.1 DAT reflexive

The antipassive analysis of SE links the morpheme SE to the DO position to 'block' it, so that no other DP can move into it. But as far as we see, there is no possibility to extend this view to DAT reflexive clitics SI. Moreover, from a comparative perspective, to the best of our knowledge, antipassives on indirect objects don't exist.

On the other hand, the Possessor of SELF analysis does provide a neat extension to DAT reflexives. The derivation of a DAT reflexive sentence (9), repeated here as (27), would proceed as (28) shows.

(27) Chlapci si myjí ruce každé ráno.
boys$_{NOM.PL}$ SE$_{DAT}$ wash$_{3.PL}$ hands$_{ACC.PL}$ every morning
REFL: 'The boys wash their hands every morning.'

(28)



Still, one might ask why exactly SELF is in the derivation. We have, as suggested above in section 3.2, two potential answers: first, SELF is in the derivation of reflexives simply because it is an inherent part of the structure of animate nouns. We take up this line in the section 4.2. On the other hand, SELF can be required for the θ-criterion, adapting essentially Kayne's (2005) approach, as an alternative to the approach we took in section 2.2[13]. So, if the complex DP splits up in the course of the derivation (as, indeed, happens with reflexives), then both parts must be spelled out.

3.3.2 Impersonals

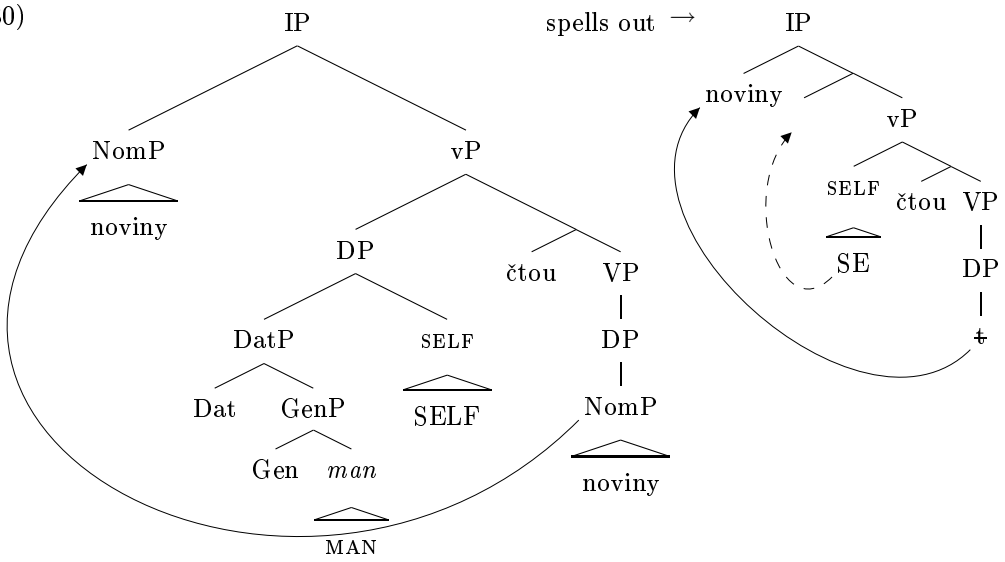
Under the Possessor of SELF story, impersonals are rather neatly derived. The example of impersonal SE is repeated below in (29).

(29) Noviny se čtou ráno (komiksy večer).
newspaper$_{NOM.PL}$ SE read$_{3.PL}$ morning comic strips evening
'People read newspaper in the morning, comic strips in the evening.'

On our approach we just cannot say that the obligatory human agent reading of impersonal SE comes from a PRO$_{arb}$, as is the standard assumption. Instead, we suggest, in line with the Possessor of SELF analysis, that the EA of an impersonal SE construction is, indeed, a complex DP, a Possessor of SELF of a particular kind: the possessor, in our view, is a silent human agent, labeled as MAN in (30). The complex DP in the EA position needs to be pronounced and the only way to do so is SE.

[13]We assumed that the θ-criterion is satisfied by the θ-roles being picked up by different Case-layers.

(30)



If we assume that the complex animate DP is born in the EA position, we need to make sure that the NomP layer of the MAN does not get to the NOM position: it would ban the DP from the DO *noviny* from getting there. But notice that the NomP layer of the EA argument MAN is buried down under the GenP layer, so, under Relativized Minimality, the NomP layer of the DP from the DO position is closer to the NOM position in IP.

Additionally, the movement of the whole complex DP from the EA position to the NOM position in IP should be disallowed as well. We would like to link the impossibility of such a movement to the fact that there are no NOM reflexives. The most striking example (or, rather, lack of such) come from Icelandic. A class of Icelandic verbs has oblique subjects (that bind anaphors and show other NOM-subject properties) and NOM objects: the expectation, then, is to see the oblique subject binding a NOM reflexive anaphor, the NOM counterpart to ACC *sig*. But such a form just does not exist, see Taraldsen (1994) Taraldsen (1995) for discussion.

Finally, we predict that the SE seen in impersonals must be GEN: we assume that the EA position is associated with GEN. However, there is a piece of evidence that SE, indeed, can be GEN: given that every instance of ACC case has to shift to GEN under nominalizations and given the fact that SE appears in nominalizations as it is (as discussed in Oliva (2001) and Hron (2005)), we might take it that SE is indeed both GEN and ACC form. (Morphologically, GEN and ACC pronouns have the same form in Czech.) Notice that if SE in impersonals is the GEN form, as predicted by our analysis, we might be a step closer to understanding the impersonals with ACC marked original DO, as shown in the Slovene example (18). We leave this issue for further research.

The antipassive analysis of SE does not extend to derive impersonals. As a matter of fact, it is not uncommon for a language to use the same morpheme to create antipassives and reflexives (Chukchi (Chukotko-Kamchatkan), Yidiɲ, Lardil (Australian), Greenlandic, Halkomelem (Salish), Medová (in progress) for discussion). However, non of the languages uses the reflexive/antipassive morpheme for an (equivalent to) impersonal. So, the fact that Slavic (and Romance) does might be because of the homonymy between the EA form (GEN) and the DO form (ACC) pointed to above.

To summarize, it seems that on the face of it the Possessor of SELF analysis can both derive the DAT reflexive SI constructions and the impersonals; the antipassive analysis cannot do neither.[14] In the next section, we thus discuss a potential extension for the winning analysis.

[14] Yet, the outcome is different for the *anticausative* use of SE exemplified in (i), that we did not discuss at all. In particular, our Possessor of SELF analysis does not lend itself in ant obvious way to derive such a

4 Extension of Possessor of SELF analysis

In the following section we will briefly discuss a set of other facts that show that animate nouns are different from inanimates in ways that might suggest that animate nouns always come with a SELF build in.

4.1 How much are animates different?

First, observe that quite generally, only animate nouns derive true reflexives. The examples in (31) are absolutely parallel with only one contrast: while (31-a) has an (masculine) animate noun, example (31-b) has a (masculine) inanimate; yet, the readings are diametrically opposite: true reflexive in (31-a) and impersonal in (31-b).[15]

(31) a. Karel se umyl.
Karel$_{NOM.MA.SG}$ SE wash$_{M.SG}$
REFL: 'Karel washed (himself).'
b. Talíř se umyl.
plate$_{NOM.MI.SG}$ SE wash$_{M.SG}$
IMP: 'The plate was washed.'

Another piece of evidence that animates are different from inanimates comes from morphological case marking of masculine nouns in Czech. (This evidence is, however, only partial as it concerns *only* masculine nouns.) Still, MA nouns show an extra piece of morphology in DAT/LOC.SG and NOM.PL nouns, as summarized in the table (32).

(32)

case	MI *plate*	MA *sir*
DAT/LOC.SG	talíř-i	pán-OV-i
NOM.PL	talíř-e	pán-OV-é

Third piece of evidence is based on Czech possessive adjectives. Generally, possessors appear in GEN case (33). For animate nouns, there is another option: the possessor can be expressed as a possessive adjective (34). Crucially, however, the -OV (for MA) or -IN (for F) suffixation

structure: if we assume that only animate nouns are Possessor of SELF (as discussed further in 4.1), there is no reason for anticausatives (with an inanimate noun in NOM) to be derived with SE. On the other hand, the antipassive analysis of SE could be extended to the impersonals along the following lines. SE in (ACC) reflexives 'blocks' the structural DO position. Now, if the structural position for EA is, as we argue, associated with an oblique case (GEN) (that is, it crucially is not the NOM position), then SE can 'block' this position. As a consequence, the DP from the DO position (or, more precisely, from the OBL case-layer) can move to the NOM position. That is, we predict that SE can 'block' both the (structural) EA position (associated with GEN case) and the (structural) DO position because SE is both ACC and GEN marked: the same outcome as above.

(i) (Karlovi) se rozbily brýle.
Karel$_{DAT.SG}$ SE broke$_{PL}$ glasses$_{NOM.PL}$
'The glasses broke (on Karel).'

[15] We acknowledge that there is a (marginal) impersonal reading of (31-a), but, importantly, there is no reflexive reading of (31-b). If we say that the distinction is purely pragmatic, we lose a possible explanation of the observed syntactic restrictions. However, even animate nouns can participate (more or less marginally) in impersonal constructions (as in (31-a) above) suggesting that animates can be coerced to inanimates. Conversely, in relevant (fairy-tale) contexts, inanimates can be conceived as animates.

that creates possessive adjectives is unavailable for inanimate nouns, as shown in (34-a) and (34-b)[16].

(33) a. tajemník fakulty
secretary$_{MA.SG}$ faculty$_{GEN.SG}$
'the secretary of the faculty'

b. noha stolu
leg$_{F.SG}$ table$_{GEN.SG}$
'the leg of the table'

c. prsten Jarmily Noskové
ring$_{MI.SG}$ Jarmila Nosková$_{GEN.SG}$
'Jarmila Nosková's ring'

d. motorka Karla Noska
motorcycle$_{F.SG}$ Karel Nosek$_{GEN.SG}$
'Karel Nosek's motorcycle'

(34) a. *fakult-in-Ø tajemník
faculty-IN-$_{MA.SG}$ secretary$_{MA.SG}$

b. *stol-ov-a noha
table-OV-$_{F.SG}$ leg$_{F.SG}$

(Veselovská (1998))

c. Jarmil-in-Ø prsten
Jarmila-IN-$_{MI.SG}$ ring$_{MI.SG}$
'Jarmila's ring'

d. Karl-ov-a motorka
Karel-OV-$_{NOM.F.SG}$ motorcycle$_{F.SG}$
'Karel's motorcycle'

There are at least two questions: first, why only animates have possessive adjectives and second, why only MA nouns have the OV in DAT/LOC.SG case endings. We will not make any specific proposals as to why it is that only animate nouns form possessive adjectives, but we will make a suggestion about the specific complex DAT/LOC forms of MA.SG. nouns, those shown in (32). In fact, one might analyze a form like *Michal-ov-i* as the DAT form of a complex DP, where Michal-ov is in fact a possessive adjective possessing SELF and the DAT/LOC *-i* marks the case of the complex DP headed by SELF. If this is correct, then, at least, animate SG. nouns must always be introduced as the possessors of SELF, since *Michal-ov-i* is the only

[16]We abstract away from the restrictions on possessive adjectives formation: only MA or F singular nouns can form possessive adjectives and only one item can appear in the possessive adjective, ruling out (i). Apparently, even inanimate could (marginally) have OV in the history of Czech, as shown in (ii).

(i) Jarmil-in-Ø (*Nosek-in-Ø) prsten
Jarmila-IN-$_{MI.SG}$ Nosek-IN-$_{MI.SG}$ ring$_{MI.SG}$
Intended:'Jarmila Nosková's ring'

(ii) k opravení ohně oltář-ov-a
to repair$_{DAT}$ flame$_{GEN}$ altar$_{POSS-GEN}$
'(about) the repair of the altar's flame'

(Gebauer (1929):159)

possible DAT/LOC form of *Michal*[17].

4.2 The animacy requirement

So far, we work with the assumption that animate nouns can combine with SELF in a kind of possessor structure. The analysis developed up to this point commits us to the view that the DatP Possessor of SELF does not shrink to a GenP internally to the DP in the derivation of a sentence with reflexive SE, since otherwise raising the GenP containing the non-pronominal part of the possessor to the EA-position would violate the requirement that the topmost Case-layer must be left behind under movement. But nothing precludes the possibility that a DatP Possessor of SELF may have its GenP raised to a DP-internal Gen-position in derivations in which the possessor does not raise to the EA-position, as discussed in (23). This in turn opens up for a solution to a problem that has nothing to do with reflexives.

In Czech and other Slavic languages, singular masculine animate nouns (MA) have GEN case as direct objects, as opposed to the masculine inanimate nouns (MI), in the literature referred to as the ACC-to-GEN shift, as shown in (35).

(35) Vidím Michal-a / hrad.
see$_{1.SG.PRES}$ Michal-$_{GEN.SG.MA}$ / castl$_{ACC.SG.MI}$
'I see Michal / a castl.'

Usually, this is regarded as a morphological quirk, an instance of syncretism between the GEN and the ACC. But significantly, GEN MA.SG direct objects group with oblique DPs (DPs bearing INS, LOC, DAT or 'real' GEN case) with respect to the obligatoriness of resumptive pronouns in *what*-relatives, as observed in Toman (1998). Whereas relativized NOM or ACC DPs do not require a resumptive pronoun, oblique DPs do, as shown in (36), and the 'Genitivus pro Accusativo' DPs do as well: compare (37-a) to non-obligatory resumptive pronoun *ho* with 'real' ACC of the MI nouns in (37-b).

(36) a. To je ten muž / ten nápoj, co jsem *(mu)
this is the man$_{NOM.MA}$ / the drink$_{NOM.MI}$ what be$_{1.SG.PRES}$ him$_{DAT.SG.M}$
včera propadla.
yesterday fall.for$_{F.SG}$
'This is the man / the drink that I felt for yesterday.'

b. To je ten muž / ten nůž, co sis *(ho) tak
this is the man$_{NOM.MA}$ / the knife$_{NOM.MI}$ what SE.be$_{2.SG.PRES}$ him$_{GEN.SG.M}$ so
obezřetně všimla.
cleverly noticed$_{F.SG}$
'This is the man / the knife that you so cleverly noticed.'

(37) a. To je ten chlap, co *(ho) viděli v tramvaji.
this is the guy$_{NOM.SG.MA}$ what him$_{GEN.SG.MA}$ see$_{pl}$ in street.car
'This is the guy they saw in the street car.'

b. To je ten nůž, co (ho) Petr našel na stole.
this is the knife$_{NOM.MI}$ what him$_{ACC.SG.MI}$ Petr$_{NOM}$ found on table
'This is the knife that Petr found on the table.'

(Toman (1998):(16,4c))

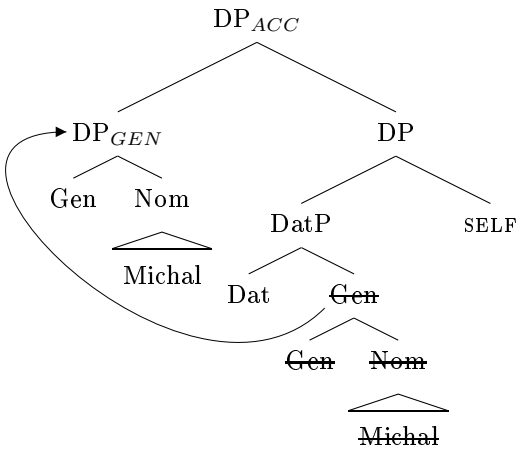
[17] We essentially play with the possibilities of this analysis, we have no answers as to why F nouns never show -IN- in their DAT/LOC case endings, nor do we know why a complex MA.SG. noun, like *Michal Valenta*, does not usually have the DAT/LOC form *Michal-ov-i Valent-ov-i*.

This suggests that the MA.SG.GEN seen on direct objects is in fact GEN case. But there are two objections that this assumption will run up against. First, there is the conceptual question why direct objects should have 'real' GEN case. Then, there is an empirical problem[18]: If a MA.SG. direct object is really a GenP, how come it can conjoin with something which is plainly not a GenP, as in (38)?:

(38) Viděla sem Michala a jeho kočku.
see$_{F.SG}$ be$_{1.SG.PRES}$ Michal$_{GEN.SG}$ and his cat$_{ACC.SG}$
'I saw Michal and his cat.'

Assumptions already made lead to a common answer to both questions. Suppose the GEN direct object *Michala* in (38) is a GenP Possessor of SELF, as discussed in (23) above. As before, the GenP *Michala* has been extracted from a possessor DatP to a DP-internal position (see (39) below), while the outer DP, headed by SELF, has ACC Case. Then, we have a partial answer to the question why the direct object has GEN case in (35), and the coordination problem dissolves as well since (38) in fact has a conjunction of two AccPs.

(39)



.

The solution is only partial, however, for at least two reasons. We still need to explain why a MA.SG. DP denoting a human being needs to combine with SELF to be a direct object, and, conversely, we need to understand why a MI.SG. DP cannot combine with SELF, when it is not a direct object, i.e. we need to account for those properties of GEN direct objects which make people postulate morphological GEN/ACC syncretism for MA.SG. nouns. In addition to this, we should also be able to say why only nouns denoting human beings combine with SELF this way. Well aware of these problems, we leave them for further research.

5 Person speculation

Given previous assumptions, the easiest way of aligning the reflexive clitics with their 1^{st} and 2^{nd} person counterparts seems to be by saying that 1^{st} and 2^{nd} person clitics always originate as possessor of SELF, extending the account of animate singular nouns sketched above. Then, we will have to add that when the complex DP doesn't split apart, the whole thing lexicalizes as *mě/me, tě/te* etc. In the derivation of a reflexive sentence, however, the complex DP does split up, since the possessor is raised to the EA-position, leaving SELF stranded. In that case, we, again, expect to see SELF spelling out as SE or SI, as is indeed the case in Czech and in

[18]Pointed to us repeatedly by Pavel Caha (p.c.), for which we thank him.

Slavic more generally. (Why this does happen in Romance, we leave an open question for now.)

Notice that this line of analysis won't extend to the nominative 1^{st} and 2^{nd} person pronouns originating in the EA-position, if they – like the MAN posited for impersonal SE – are embedded under a DatP-layer, and so cannot raise from the EA-position to the NOM position, since only the highest Case-layer can be left behind under movement. This might be a positive result, since the 1^{st} and 2^{nd} person nominative in fact seem morphologically unrelated to the corresponding non-nominative clitics. Yet, we could conceivably maintain the idea that all 1^{st} or 2^{nd} person pronouns are introduced as possessors of SELF by allowing 1^{st} and 2^{nd} person pronouns, but still no other possessor of SELF, to shrink to GenPs by moving to a DP-internal GEN position, a difference between $1^{st}/2^{nd}$ person forms and all other that might have something in common with the fact that some Romance varieties have the auxiliary *be* even with transitive and unergative verbs with 1^{st} and 2^{nd} person subjects only, assuming an analysis of Kayne (1993). A potential bonus of this approach might be that it would give a handle on the fact that there are no *impersonal* 1^{st} and 2^{nd} person subjects (40-a) comparable to (40-b), repeated here from (3-b).

(40) a. *I giornali mi/ti leggono la mattina.
the newspaper$_{NOM.PL}$ me/te read$_{3.PL.PRES}$ the morning
Intended: 'People read newspaper in the morning.'

b. I giornali si leggono la mattina.
the newspaper$_{NOM.PL}$ SE read$_{3.PL.PRES}$ the morning
'People read newspaper in the morning.'

References

Alboiu Gabriela, Michael Barrie, and Chiara Frigeni. 2004. *SE* and the Unaccusative-Unergative Paradox. Coene, M., G. de Cuyper & Y. D'Hulst (eds.). *Current Studies in Comparative Romance Linguistics. Antwerp Papers in Linguistic* 107:109–139.

Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel Publishing Company.

Caha, Pavel. 2006. *Case Shifting as Peeling and Some of the Consequences*. Unpublished, ms., CASTL Tromsø.

Cinque, Guglielmo. 1988. On SI Construction and the Theory of Arb. *Linguistic Inquiry* 19:521–581.

Dobrovie-Sorin, Carmen. 2005. The SE-anaphor and its role in argument realization. Everaert, M. & H. van Riemsdijk (eds.). *The Blackwell Companion to Syntax*, vol 4:118–179. Oxford: Blackwel.

Gebauer, Jan. 1929. *Historická mluvnice jazyka českého IV: Skladba. (Historical Grammar of Czech Language IV: Syntax.)*. F. Trávníček (ed.). Česká Akademie věd a umění: Praha.

Grimshaw, Jane. 1982. On the Lexical Representation of Romance Reflexive Clitics. Bresnan, J. (ed.). *The Mental Representation of Grammatical Relations*: 87–148. Cambridge: MIT Press.

Helke, Michael. 1971. *The Grammar of English Reflexives*. PhD dissertation, MIT.

Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30:69–96.

Hron, David. 2005. *On the derivation of reflexive nouns: the case of Czech*. MS thesis, Tel Aviv University.

Kayne, Richard S. 1986. *Participles, Agreement, Auxiliaries, Se/Si and PRO*. Unpublished, Handout to talk at Princeton University.

--. 1993. Toward a modular theory of auxiliary selection. *Studia Linguistica* 47:3–31, Reprinted in *Parameters and Universals*, Oxford University Press, 2000.

--. 2000. Person Morphemes and Reflexives in Italian, French, and Related Languages. In *Parameters and Universals*. Oxford University Press, Oxford Studies in Comparative Syntax.

--. 2005. *Pronouns and Their Antecedents*. New York: Oxford University Press: 105–135.

Kozinsky, Isaac, Vladimir P. Nedjalkov, and Maria S. Polinskaja. 1988. Antipassive in Chukchee: oblique object, object incorporation, zero object. Shibatani, M. (ed.). *Passive and Voice*: 651–706: Amsterdam and Philadephia: John Benjamins Publishing Company.

Mahajan, Anoop. 1994. The Ergativity Parameter: *have-be* Alternation, Word Order and Split Ergativity. Gonzalez, M. (ed.). *Proceedings of NELS 24*. GLSA. Amherst: University of Massachussetts.

Medová, Lucie. in progress. *Reflexive Clitics in Slavic and Romance*. Unpublished, Doctoral Dissertation, Princeton University.

--. to appear. Reflexive clitics, Movement and Layered Case. Marušič, F. (ed.). *Formal Description of Slavic Languages 6.5: Nova Gorica*.

Ndayiragije, Juvenal. 2006. The Ergativity Parameter: A view from Antipassive. Johns, A., D. Massam & J. Ndayiragije (eds.) *Ergativity. Emerging Issues*: 271–292. Dordrecht: Springer.

Oliva, Karel. 2001. Reflexe reflexivity reflexiv (*A reflexion on a reflexivity of reflexives.*). *Slovo a slovesnost* 62:200–207.

Panevová, Jarmila. 1999. Česká reciproční zájmena a slovesná valence (*Czech reciprocal pronouns and verbal valence*). *Slovo a slovesnost* 60:269–275.

Postal, Paul M. 1977. Antipassive in French. *Linguisticæ Investigationes* 1:333–374.

Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.

Rivero, Maria Luisa and Milena Moskojević Sheppard. 2003. Indefinite reflexive Clitics in Slavic. *Natural Language and Linguistics Theory* 21:89–115.

Say, Sergey. 2005. Antipassive *sja*-verbs in Russian. Dressler, W. U. (ed.). *Morphology and its Demarcations: Selected Papers from the 11th Morphology Meeting, Vienna, February 2004*:253–275. Amsterdam and Philadelphia: John Benjamins Publ.

Starke, Michal. 2005. *Nanosyntax. Class lectures*. Unpublished, University of Tromsø.

Taraldsen, Tarald. 1994. Reflexives, Pronouns and Subject/ Verb Agreement in Icelandic and Faroese. *Working Papers in Scandinavian Syntax 54*:43–58.

--. 1995. On Agreement and Nominative Objects in Icelandic. *Studies in Comparative Germanic Syntax*: 43–58. Dordrecht: Kluwer.

Toman, Jindřich. 1998. A Discussion of Resumptive Pronouns in Colloquial Czech. Bošković, Z., S. Franks & W. Snyder (eds.). *Formal Approaches to Slavic Linguistics. The Connecticut Meeting* 1997:303–318. Ann Arbor: Michigan Slavic Publ.

Veselovská, Ludmila. 1998. Possessive Movement in the Czech Nominal Phrase. *Journal of Slavic Linguistics* 6:255–300.

The Czech invariant demonstrative *to* is a Foc head

Radek Šimík

R.Simik@rug.nl

1 Introduction

This paper discusses the syntax and semantics of the Czech invariant demonstrative *to*. In order to avoid confusion with the standard demonstrative pronoun, which reflects the φ-features and case of the associated nominal (namely neuter/singular/nominative/accusative), I will refer to the invariant demonstrative using capitals (TO) in the body of the text, examples, as well as glosses, and with no implications for intonation.

TO appears in several seemingly unrelated syntactic contexts: it can be a subject of an NP/DP predicate; it can be attached to an NP in (both embedded and non-embedded) *wh*-questions; it is obligatorily attached to an internal NP head in appositive relative clauses; it can (must?) appear in DP appositions which are modified by an attributive adjective; it appears in exclamations; it can be attached to a projection of VP; in standard Czech and Moravian dialects it is a proximity marker adjoined to a demonstrative pronoun; in some exceptional constructions, it appears obligatorily. These contexts are exemplified below in the order just mentioned:[1]

- *subject of a DP predicate*

  (1) TO je moje učitelka.
  TO is my teacher
  'It is my teacher'

- *embedded and non-embedded wh-question*

  (2) a. [...] že se mohlo přímo uhádnout, kterou TO píseň pěli.
  that refl could directly guess which TO song sang-they
  '[...] that one could directly guess which song (it was that) they sang'

  b. Jaká TO síla zvedala ruce v sále [...]?
  what TO power raised hands in chamber
  'What power was it that was raising the hands in the chamber [...]?'

- *appositive relative with an internal head*

[1] The sentences from (2) to (6) are taken from the Czech National Corpus.

(3) [je] fejetonovým seriálem, kterýž TO název na něj nyní dodatečně připichuji
[is] feuilleton series which TO name on it now afterwards add
'It is a feuilleton series, which name I add to it now, afterwards'

- *apposition*

(4) Jednou jsem v Neapoli, v úžasném TO městě, na ulici zažil hádku
once aux-I in Naples in wonderful TO city on street experienced row
dvou chlápků
two chaps
'Once I in Naples, a beatiful city, I experienced a row between two chaps.'

- *exclamation*

(5) Jaká TO čeština, řekl by mnohý [...]
what TO Czech language said aux-he many
'What a Czech (language), many would say [...]'

- *attached to VP in wh-questions*

(6) Copak jsi mu TO tehdy radil?
what aux-you him TO then advise
'What was it that you gave him advice about then?'

- *proximity marker*

(7) ten-TO, ta-TO, to-TO; takový-TO, taková-TO, takové-TO
the/masc-TO fem-TO neut-TO such/masc-TO fem-TO neut-TO

- *obligatory expletive*

(8) Včelami se *(TO) tam jen hemžilo
bees-instr refl TO there only abound
'It abounded with bees there'

On the basis of syntactic and semantic arguments I will argue that TO in (1) through (6) is one and the same element. TO in (7) and (8) will not be taken into account for reasons of space and also because they are arguably different.

Following the insights of Kayne (1994), Rizzi (1997), and others, I will argue that TO is a focus-head Foc placed above IP. It will be shown that it can appear in analogous positions in both the verbal and nominal domains. There are two crucial observations leading to the proposed structure: (i) the XP with which TO is associated is obligatorily presupposed; (ii) in this XP, there must be a (properly bound) *wh-*/*focus*-variable.

The constructions dealt with in this article are related to (pseudo)cleft structures well-known from English. In fact, they are condensed ways of expressing the same semantics, which will be made explicit in the proposal. A brief but explicit analysis of English (pseudo)clefts based on movement (contras ellipsis) is included in section 5.3.

This paper is organized as follows. In section 2 I discuss some general properties of the constructions illustrated above. In section 3 I briefly review two existing accounts of some

aspects of TO in Czech, Polish, and Russian. Section 4 discusses the semantics of TO and section 5 deals with its syntax. Section 6 concludes the paper.

2 General characteristics

This section is devoted to setting the empirical scene. Most of the phenomena exemplified here will be subject to the proposed analysis in the following sections.

2.1 TO in copular constructions

TO can appear in all kinds of Czech predicative constructions involving the copula *být* 'be'.[2] The properties of this kind of TO are discussed in detail in Šimík (2006). The most notable property of TO is that it lacks φ-features and case. There are two pieces of evidence that show this quite clearly: (i) TO cannot trigger the agreement on the verb (copula) and (ii) nominative case has to be assigned to the predicate.

(9) TO jsou/*je moji kamarádi
TO are/is my friends
'Those are my friends'

(10) Co dělá maminka?
a. Maminka/ona/*pro* je učitelka/učitelkou.
mum/she/*pro* is teacher.nom/instr
b. Je TO učitelka/*učitelkou.
is TO teacher-nom/instr
'Mum/she is a teacher.'

Note that analogous NPs in the subject position, like *maminka* 'mum', *ona* 'she', or even an empty *pro*, allow for instrumental marking of the nominal predicate *učitelka* 'teacher'. In contrast, TO requires a nominative predicate: an instrumental is sharply ungrammatical. This state of affairs follows quite straightforwardly if we admit that φ-features are a necessary prerequisite for an NP to bear Case and that finiteness requires a relation with a nominative DP (Chomsky 1995). Because TO has no φ-features, as shown by the impossibility of agreement with the verb (9), it can bear no Case. Nonetheless, there must be an NP bearing a nominative Case, because the structure is finite. The only NP available is the predicate. Note that if the subject is not TO but rather a standard NP, as in (10a), an instrumental is available for the predicate.[3]

The reader should keep in mind that the invariance of TO (no agreement in case or φ-features with any potentially associated NP/DP/AP) is its defining property for all the occurrences of

[2] Geist and Blaszczak (2000) argue that the Polish and Russian equivalent of TO (*to* in Polish and *eto* in Russian) appears in specificational but not in predicational copular sentences. However, their claim is problematic for two reasons: firstly, they do not really show an ungrammatical sentence involving TO with a (nominative) NP predicate; secondly, they do not make it explicit how such a structure would be ruled out. The example (10b) below shows that TO is also licit with apparent nominative NP predicates in Czech. I leave this issue aside here for the moment.

[3] Geist and Blaszczak (2000) show that structures containing TO bar predicative instrumental NPs in Polish and Russian even if an overt subject is present—a potential nominative-bearer. This presents a problem for the assumption that the predicative NP is the only one which can be assigned nominative by T. I will come back to this problem in the analysis in section 5.3.

TO. I will not continue to emphasize this fact, but it is crucial throughout.

Another remarkable property of TO is that it is necessarily associated with an NP predicate (unlike other NPs). Observe the following example:

(11) Jaká je tvoje maminka?
'How is your mum?'
a. Moje maminka/ona/*pro* je úžasná (žena) / vždy nad věcí.
my mum/she/*pro* is wonderful woman / always above thing
b. Je TO úžasná *(žena) / vždy nad věcí.
is TO wonderful woman always above thing
'My mum/she is (a) wonderful (woman) / always stays calm.'

While 'normal' NPs can be subjects of an AP or PP predicate, this option is unavailable for TO. The explanation may well stem from the fact that the predicate is the only XP in the structure which can absorb case and therefore it must be an NP, and not e.g. an AP or PP. Examples like (10b) above then force us to admit that predicative NPs are (or at least can be) directly case-marked (by T).

2.2 TO as a second position clitic (?)

TO sometimes surfaces in the so-called Wackernagel position, i.e. immediately following the first constituent:

(12) Co TO Petr včera říkal?
what TO Petr yesterday said
'What did Peter say yesterday?'

When we include all the possible second-position clitics, we find out that TO occupies the right-peripheral position of the second position. This is typical of all Czech non-focused (argumental as well as adverbial) demonstratives (*t*-words), as shown below (for the sake of clarity I will use |x| when x appears in the second position):

(13) a. V kolik hodin |jste se mu ho TO| snažili předávat?
in how-many hours aux-you refl him.dat him.acc TO tried hand-over
'At what time did you try to hand it over to him?'

b. V kolik hodin |jste se mu ho tam tehdy| snažili předávat?
in how-many hours aux-you refl him.dat him.acc there then tried hand-over
'At what time did you try to hand it over to him there then?'

A question immediately arises: what happens if TO appears simultaneously with other *t*-words. Observe the following example:

(14) a. V kolik hodin |jste mu ho /TO/ tam tehdy /*TO/| chtěli
in how-many hours aux-you him.dat him.acc TO there then TO wanted
předat
give
'At what time did you want to give it to him there then?'

b. V kolik hodin |jste mu /TO/ ten (balík) /*TO/| chtěli předat
in how-many hours aux-you him.dat TO that package TO wanted give

'At what time did you want to give that (package) to him?'

Example (14) makes it clear that TO immediately follows verbal and pronominal clitics but has to precede any *t*-words (whether argumental or adverbial).

There is no space in this paper for any discussion of clitic-placement. I will tacitly assume that there is nothing like a Wackernagel position in the syntactic tree where all clitics move into. Rather, for every single clitic-like element there is a corresponding functional projection, which has a fixed place in the tree-hierarchy (in accord with Kayne 1994, whose LCA prohibits adjunctions of clitics to other clitics). The only position about which I will make explicit claims is the position of TO. It follows that all verbal and prenominal clitics have to precede this position and all *t*-words have to follow this position.

Let us also make clear here that the presence of TO (recall: taking the form of sg, neut, nominative/accusative demonstrative) signals no resumptive or clitic-doubling. The following examples show that as soon as any kind of *wh*-question is involved, TO may appear:

(15) a. Komu |jsi TO| ten obrázek nakonec dal?
whom.dat aux-you TO that picture in the end gave
'Whom did you give the picture in the end?'

b. Kdy |jsi mu TO| ten obrázek dal?
when aux-you him.dat TO that picture gave
'When did you give him that picture?'

(15a) shows that the presence of TO is independent of the presence of an sg/neut/accusative *wh*-phrase; (15b) shows that it is independent of the presence of a *wh*-argument.

2.3 TO in the nominal domain

Some properties of TO are observable in all the constructions from (2) through (5), i.e. (embedded and root) *wh*-questions, appositive relatives involving an internal head, nominal appositions, and exclamations. In this section, we focus on an adequate description of these properties. For independent reasons, not all of them are observable with all constructions (hence, not all possibilities can be exemplified).

The most remarkable property of all the constructions is the fixed PreAtt–TO–NP structure: the first element is a standard prenominal attribute (PreAtt), possibly in *wh*-form, the second one is TO, and the last one is the NP to which the attribute is related. The following examples show this for the case of questions, appositive relatives, appositions, and exclamations, respectively.

(16) a. Jaké/které/čí/kolik TO chleb-y/-ů
what kind of/which/whose/how many TO breads-nom/gen

b. *Od koho/koho/odkud TO chleby
from whom/of whom/from where TO breads

(17) a. chleba z Prahy, kterouž TO potravinu jsme koupili v pekárně
bread from Prague which TO food aux-we bought in bakery
b. *chleba z Prahy, odkud(-ž) TO potravinu jsme koupili v pekárně
bread from Prague from where TO food aux-we bought in bakery

(18) a. v Neapoli, (v) úžasném TO městě
in Naples in wonderful TO city
b. *v Neapoli, (v) dlouhé historie/bez starosty TO městě
in Naples in of long history/without mayor TO city

(19) a. (Babička,) jaká/šťastná TO žena
grandma what/happy TO woman
b. *(Babička,) z Náchoda/odkud/koho TO žena
grandma from Nachod/from where/of whom TO woman/wife

Note that the elements preceding TO above unexceptionally match the form of standard prenominal attributes, i.e. they are either agreeing adjectives or quantifying heads:

(20) a. *Pochutnali jsme si na chlebu [kmínovém/náchodském/dobrém]$_{\text{Att}}$
enjoyed aux-we refl on bread caraway/from Nachod/good
b. Pochutnali jsme si na [kmínovém/náchodském/dobrém/pěti]$_{\text{Att}}$
enjoyed aux-we refl on carraway/from Nachod/good/five
chleb-u/-ech
bread-loc/gen
'We enjoyed the caraway/(from) Nachod/good bread.'

(21) a. Pochutnali jsme si na chlebu [z Náchoda/bez kmínu/maminky]$_{\text{Att}}$
enjoyed aux-we refl on bread from Nachod/without carraway/by mum
b. *Pochutnali jsme si na [z Náchoda/bez kmínu/maminky]$_{\text{Att}}$ chlebu
enjoyed aux-we refl on from Nachod/without carraway/by mum bread
'We enjoyed the bread from Nachod/without carraway/from mum.'

All the other properties seem to follow, at least at the descriptive level, from the requirement of a rigid Att-TO-NP structure.

Thus, TO is ruled out from being in a relation with simplex *wh*-phrases, as the following examples show. Note that we disregard the occurrence of TO in the Wackernagel position.

(22) a. Kterou TO píseň |jste se ho| včera snažili naučit?
which TO song aux-you refl him yesterday tried teach
'Which song did you try to teach him yesterday?'

b. *Co TO |jste se ho| včera snažili naučit?
what TO aux-you refl him yesterday tried teach
'What did you try to teach him yesterday?'

c. *Kdy TO |jste se ho| snažili naučit tu píseň?
when TO aux-you refl him tried teach the song
'When did you try to teach him the song?'

The last significant property that I would like to mention here is the impossibility of left-branch extraction when TO is involved, which is otherwise a commonplace and fully grammatical structure in Czech (as (23a) shows).

(23) a. Kolik |jste| tam nechali ležet knížek?
how-many aux-you there left lie books

'How many books did you leave there lying?'

b. Kolik TO knížek |jste| tam nechali ležet?
how-many TO books aux-you there left lie
c. *Kolik TO |jste| tam nechali ležet knížek?
how-many TO aux-you there left lie books
d. *Kolik |jste| tam nechali ležet TO knížek?
how-many aux-you there left lie TO books

As expected, the same holds for appositive relatives with an internal head:

(24) a. *Borat, kterýž TO |jsem| viděl film už dvakrát
Borat which TO aux-I saw film already twice
b. *Borat, kterýž |jsem| viděl TO film už dvakrát
Borat which aux-I saw TO film already twice
'Borat, which film I have already seen twice'

A grammatical analogue to (23d) cannot be shown for appositive relatives of this kind, presumably because they obligatorily involve TO.

3 Previous accounts

The invariable demonstrative TO has received some attention in the descriptive as well as generative Slavic literature. The most analyzed structure involving TO is the copular construction in (1), repeated here:

(25) TO je moje učitelka. [=(1)]
TO is my teacher
'It is my teacher'

The first subsection is devoted to a short review of Geist and Blaszczak's (2000) account of the Polish and Russian equivalents of (25).

Concerning all the other constructions, only (2), repeated below, has been given any account.

(26) Jaká TO síla zvedala ruce v sále [...]? [=(2b)]
what TO power raised hands in chamber
'What power was it that was raising the hands in the chamber [...]?'

This TO is briefly analyzed in Veselovská (1995). I will review her account in the second subsection.

3.1 TO in copular constructions involves a hidden pseudocleft

Geist and Blaszczak (2000)[4] analyze Russian and Polish constructions like those in (27). They differ from the Czech one in that they can contain an explicit referential subject (*Ivan*). In Czech, this is possible only if there is a comma intonation between the subject and the rest of the structure. Otherwise, all the major characteristics (case restrictions, information structure properties, etc. are shared).

[4] A shorter version of this paper is Blaszczak and Geist (2001).

(27) a. (Ivan) to (jest) moj ojciec [Polish]
Ivan it is my father
b. (Ivan) eto moj otec [Russian]
Ivan it my father
c. (Ivan,) to je můj otec [Czech]
Ivan it is my father
'(Ivan,) it is my father.'

Geist and Blaszczak argue for a pseudocleft analysis along the following lines. *To/eto* is taken to be a conjunction-like element which conjoins two propositions where the first one contains a variable (it is an open proposition) and in the second one, which is syntactically identical, a referring expression substitutes for the variable (it is a closed proposition). Observe the following schema:

(28) $[_{\text{toP}}\ [_{\text{TP}}$ Ivan $x_i]$ to $[_{\text{TP}}$ (is) Ivan my father$_i]]$

Thus, what the constructions in (27) are explicitly compared to is the English pseudocleft, i.e. *What Ivan is is my father* (see e.g. den Dikken, Meinunger, and Wilder 1998). A similar analysis has been recently revived by Markman (2007), with some slight changes and an explicit claim that TO is a Top(ic) head.

Reviewing all the aspects of the proposed analysis would lead us far astray. Some critical comments on it will be put forth in the analysis below in section 5.3.

3.2 TO is definiteness agreement surfacing on the adjective

Veselovská (1995: Chapter 6) argues that the element TO is a morphological realization of "definiteness agreement", which surfaces attached to an adjectival specifier of a definite DP. She explicitly treats the cases exemplified in (2) and (7), exemplified below, as identical:

(29) a. jaký TO kamarád
what-kind-of TO friend
b. takový TO kamarád
such TO friend
'A friend like this'

This approach fails to capture the following facts.

Firstly, if TO is attached to a demonstrative pronoun, as in (29b), it yields a proximity interpretation, which is unexpected if it were simply a reflex of definiteness. Furthermore, even when TO is related to a *wh*-phrase, it is only partly true to say that TO involves definiteness. This will be shown in the section 4.

Secondly, if TO appears within a *wh*-phrase, as in (29a), it is not attached to the adjective, but rather to the NP with which it is associated. This is possible to see in cases where the adjectival and the NP do not form an (immediate) constituent. The first example involves left-branch extraction of the *wh*-word (30a), and the second concerns the presence of an echo-question marker, which takes the form of the standard Czech complementizer *že* (30b):

(30) a. *Jaké TO si-s včera koupila žvýkačky?
what-kind-of TO refl-you yesterday bought chewing-gums
b. Jaké /*TO/ že /TO/ žvýkačky si-s včera koupila?
what-kind-of TO that TO chewing-gums refl-you yesterday bought

'What kind of chewing gums did you buy yesterday?'

Veselovská (1995) wrongly predicts both examples in (30) to be grammatical. As we observed above (cf. (23)), left-branch extraction with TO is ungrammatical even if TO stays in situ along with the NP, so (30a) does not really show what I claim, that it is attached to the NP. (The impossibility of the left-branch extraction with TO will be discussed in 5.2.2.) The NP-attachment hypothesis, however, scores much better in the light of (30b).

Further evidence for the NP-attachment comes from Russian, which also makes use of TO. The Russian TO arguably involves the same structural properties and semantics as the Czech one, the only difference being its position: it appears right-attached to the NP:[5]

(31) Kakich sobak TO Maša chočit kupiť?
what-kind-of dogs TO Masha wants buy
'What kind of dogs does Masha want to buy?'

The following examples confirm the validity of the test in (31): the Russian TO is also ungrammatical with a simplex *wh*-phrase, showing that the presence of a full NP is crucial, see (32a); left-branch extraction is impossible for *wh*-phrases which involve TO but is possible otherwise, see (32b/c).

(32) a. *Čto TO Maša chočit kupiť?
what TO Masha want buy
'What does Masha want to buy?'

b. *Kakich TO Maša chočit kupiť sobak?
what-kind-of TO Masha wants buy dogs
c. Kakich Maša chočit kupiť sobak?
what-kind-of Masha wants buy dogs
'What kind of dogs does Masha want to buy?'

The analogy between the state of affairs in Czech and Russian is further supported by the fact that Russian allows TO also in the verbal domain. In such cases, it seems to be right-attached to the VP.

(33) a. Kakich /sobak/ Maša chočit kupiť /sobak/ TO
what-kind-of dogs Masha wants buy dogs TO
'What kind of dogs does Masha want to buy?'

b. Čto chočit Maša kupiť TO?
what wants Masha buy TO
'What does Masha want to buy?'

The grammaticality of (33b) shows that TO in such cases does not attach to NPs.

4 The semantics of TO

In this section I will try to formalize the intuition about the semantics of TO, which will give us the necessary background for the syntactic analysis in section 5. Considering the seeming optionality of TO, we have to ask how sentences with TO are different from those which lack

[5]I am grateful to Žeňa Markovskaja and Aysa Arylova for the data.

it. Let us first consider a VP-modifying TO, as described in section 2.2. Consider the following pair of questions and suppose they are asked by a mother who cares about the kindness of her son.

(34) a. Jaký dobrý skutek jsi dnes udělal?

b. Jaký dobrý skutek jsi TO dnes udělal?
what good deed aux-you TO today did
'What good deed did you do today?'

The question in (34a) is a straightforward one: the mother is inquiring of her son about a good deed which he did today, not actually knowing whether the son did anything good at all. If the mother asks (34b), however, we know that she knows more, namely, she knows that there is an event of her son doing a good deed during the day (even though she does not know what exactly he did). In more formal speech, the question containing TO involves an existential presupposition of the event that it relates to, 'doing a good deed' in the above case.

What happens when TO attaches to an NP? If TO in the nominal domain is the same element as in the verbal domain, the effect should be very similar. Observe the following example.

(35) a. Jaký dobrý skutek bys chtěl udělat?

b. Jaký TO dobrý skutek bys chtěl udělat?
what TO good deed aux-you want do
'What good deed would you like to do?'

Indeed, very much like in the example (34), the question in (35b), as opposed to (35a), involves an existential presupposition of a certain good deed (such that you would like to do it).

If TO causes its complement to be existentially presupposed, it strongly resembles the definite article. In fact, if nothing more were said about it, it would not differ from the way the definite article is standardly defined (e.g. Heim and Kratzer 1998). Approaching TO as a definite article, however, fails to capture the fact that it is ruled out in standard declaratives as well as *yes/no*-questions:

(36) *Dnes jsem /TO/ udělal dobrý /TO/ skutek.
today aux-I TO did good TO deed
'I did a good deed today'

(37) *Udělal jsi /TO/ dnes dobrý /TO/ skutek?
did aux-you TO today good TO deed
'Did you do a good deed today?'

The fact that TO cannot appear in these contexts is completely unexpected if we approach it purely in terms of definiteness. Thus, there seems to be another factor determining the distribution of TO. I propose that the phrase it operates on (VP/NP for the moment) obligatorily contains a *wh*-variable. The defining configuration of TO is thus the following:

(38) *wh*-operator ... TO [$_{XP}$ *wh*-variable]

where XP is interpreted as existentially presupposed

The generalization in (38) seems to spell-out our intuition about the semantics of TO correctly: if a speaker asks a question containing TO, he says that he knows about the existence

of something (an event, an individual) but there is a gap in the knowledge—expressed by the *wh*-word.

Let us now come back to the structures which contain TO but do not involve interrogative semantics, i.e. exclamations and appositions. Note that these structures seem to conform to the presupposition (definiteness) part of the story quite well: both exclamations and appositions presuppose the existence of the NP they relate to. For appositions this is a defining property, since they attach to an NP of which they give a closer description and as such they presuppose the existence of the NP which is modified. E.g. in *Naples, a beautiful city*, talking about Naples, being a city, entails the existence of a city and this existence is then taken for granted (is presupposed) in the apposition. The situation with exclamation seems to be very similar. If we pronounce a structure like *What a beautiful city!* it seems to involve an existential presupposition of a (particular) city.[6]

Note that some Czech exclamations with TO do not involve any *wh*-element:

(39) Šťastná TO žena!
happy TO woman

Such exclamations seem to indicate that the relation between exclamation and apposition is one of great resemblance. Note that these exclamations may easily be transformed into appositions, by simply adding the NP that we are referring to:

(40) a. Babička, šťastná TO žena!
grandma happy TO woman
b. Mark Twain, what a great humorist (he was)!

Examples of appositions and exclamations without *wh*-words thus indicate that the preliminary generalization made in (38) is not yet sufficient. Furthermore, even the VP-related TO can survive without a *wh*-element in a verbal domain, as soon as the structure involves a contrastive focus:

(41) a. *Včera |jsme se ho| TO snažili naučit tu píseň.

b. VČERA$_{CF}$ |jsme se ho| TO snažili naučit tu píseň [, ne před týdnem].
yesterday aux-we refl him TO tried teach the song not before week
'It was yesterday that we tried to teach him the song, not a week ago.'

The role of contrastive focus is to create a (finite and contextually determined) set of alternatives. Thus, in the example above, the contrastive focus (CF) *včera* 'yesterday' creates a set of alternative time slots of cardinality 2: 'yesterday' and 'a week ago'. The speaker then chooses to assert one of the (two) options and to negate the assertion (possibly implicitly) involving the other one.

Wh-expressions preceding TO in exclamations and appositive relatives arguably refer to a set of alternatives as well. In appositive relatives, it is (normally) a set of individuals, which gets quantified over by the head of the appositive relative. In exclamations it is mostly a set of degrees such that the degree chosen is the maximal one: when we say *What a beautiful woman!* then *what* refers to a set of measurable properties 'beauty' and the maximal degree of 'beauty' is chosen; hence the exclamation can be paraphrased by saying something like *It is a*

[6] In her dissertation, Miro (2002) argues that exclamatives do not assert but rather presuppose their propositional content. This is in accord with my argumentation.

most beautiful woman.[7] I suppose the semantics of exclamations and appositions without any *wh*-words may be sketched along similar lines, although perhaps with more difficulties.

This discussion forces us to supplement the generalization in (38) with the one in (42).

(42) *Foc*-operator TO [$_{\text{XP}}$ *Foc*-variable]

where XP is interpreted as existentially presupposed

Now, are there really two contexts where TO appears? The answer is negative:

It has been argued that (contrastively) focused (CF) phrases and *wh*-phrases in questions form a natural semantic class. Indeed, even *wh*-phrases, as CF, refer to a set of contextually determined alternatives. The only difference is that there is no positive assertion as to which of the alternatives is valid, and therefore questions denote a set of propositions such that the propositions vary with the denotation of the *wh*-phrase. In this, I am following one of the standard approaches to the semantics of questions, dating back to Hamblin (1958) and Karttunen (1977).[8]

There is also some syntactic evidence that the resemblance between (contrastive) focus and *wh*-phrases in questions is not accidental. Horvath (1986) for Hungarian and Bošković (1998) for Serbo-Croatian, among others, argue that if language has a position for focalized phrases, then *wh*-phrases move therein, suggesting that *wh*-phrases (in questions) and focus phrases belong syntactically to one family of expressions. Rizzi (1997) explicitly proposes that *wh*-phrases in questions move to SpecFocP.

With this in mind, let us move to the syntactic analysis, the core of this paper.

5 The syntax of TO

I propose the following syntactic analysis based on the observations in (38) and (42), where TO heads a Foc projection (in the sense of Rizzi 1997):

(43) [$_{\text{FocP}}$ *wh*/*CF*$_{\text{i}}$ Foc [$_{\text{XP}}$ ... *wh*/*CF*$_{\text{i}}$ ...]

where XP is presupposed and may be either an IP or NP

This analysis is in direct contradiction to previous proposals, which claim that TO is a topic head or its functional equivalent (see section 3.1). Top and Foc heads stand in a functional opposition. While Top makes its complement new information (comment) and its specifier is assigned a presuppositional interpretation, Foc does the exact opposite: it makes its complement a presupposition and its specifier is filled by a (contrastive) focus; see the following schema (Rizzi 1997: 285ff.).

(44) a. [$_{\text{TopP}}$ XP/topic - Top = [YP/comment]]

b. [$_{\text{FocP}}$ ZP/focus Foc [YP/presupposition]]

I assume that the presence of the Foc head and the presence of a *wh*-/*CF*-element within its complement are two sides of one coin.[9] The focalized phrase obligatorily moves to the specifier

[7] There are more contexts where *wh*-words obtain this so-called *maximalization* semantics, cf. Rullmann (1995) or Grosu (1998).

[8] For more references and alternative approaches I refer the reader to Hagstrom (2003).

[9] The Foc head may thus be seen as the syntactic source of Geurts and van der Sandt's (2004) principle BPR (Background Presupposition Rule), which says that focusing an XP always gives rise to a presupposition of the background, which holds of XP.

of Foc and functions there as an operator binding the variable within XP. The relation between the Foc head and a focalized element thus resembles the relation between an interrogative C and a *wh*-phrase, which, in Rizzi's (1997) framework comes as no surprise, since the interrogative (*wh*-marked) C actually *is* Foc. Thus, SpecFocP is a unified landing position for *wh*-phrases (in root/embedded) questions as well as for (contrastively) focused phrases.

I will further assume that some kind Comp-to-Spec filter is correct (Matushansky 2006, cf. Abels 2002). Such a filter prohibits any movement of a complement (XP) into the specifier of its own head (H); this restriction is schematized below (throughout my analysis the notation $<x_n>$ will be used for silent copies in a movement chain):

(45) Comp-to-Spec filter
*[$_{HP}$ XP_1 H [$_{XP}$ $<XP_1>$]]

Note that for our purposes the filter forces the focalized phrase to be a proper subpart of XP (in (43)). We will see that it is indeed the case that the whole XP cannot be focalized. Within our discussion, this filter in (45) has a "functional" explanation. Note that if the whole XP were focalized, then there would be nothing left in the complement to be existentially presupposed. It seems that Foc needs to "discharge" all its functions in order to obtain a grammatical result, a state of affairs reminding us, first, of another empirical phenomenon discussed here (section 2.1), i.e. finite T has to assign nominative, and more generally, of economy principles of the last resort kind: operations take place only if they have an effect on the outcome.

Last but not least, as suggested in (43), XP can be of verbal as well as nominal character. The present proposal may thus be seen as an explicit support for the parallelism between verbal and nominal functional domains (thus following e.g. Szabolcsi 1994, Bennis, Corver, and den Dikken 1998, and recently Hiraiwa 2005).

In the analysis below I first address TO in the verbal domain (section 5.1), then move to the nominal domain (section 5.2), and finally I provide an analysis of TO in the copular constructions (section 5.3). For reasons of space I leave out a detailed analysis of nominal appositions and exclamative constructions, but I believe their analysis is not much different, as already suggested in the previous section.

5.1 The verbal domain

In the verbal domain, the Foc head is placed above IP (alternatively TP/AgrP/FinP):

(46) ... [$_{FocP}$ *wh*/*CF*$_i$ TO/Foc [$_{IP/TP/AgrP/FinP}$ ... *wh*/*CF*$_i$ ...]

The following examples show that the choice of the focused constituent is relatively free: it may be a subject, object, and even an adverbial. However, it cannot be the whole IP, as the (d) examples below show.

(47) *wh*-phrases
a. Kdo TO včera nemohl najít boty? [subject]
who TO yesterday could-not find shoes
b. Co TO včera Petr nemohl najít? [object]
what TO yesterday Petr could-not find
c. Kdy TO včera Petr nemohl najít boty? [adverbial]
when TO yesterday Petr could-not find shoes
d. whIPs do not exist

(48) *focus*-phrases

a. PETR TO včera nemohl najít boty. [subject]
Petr TO yesterday could-not find shoes

b. BOTY TO včera Petr nemohl najít [object]
shoes TO yesterday Petr could-not find

c. VČERA TO Petr nemohl najít boty. [adverbial]
yesterday TO Petr could-not find shoes

d. *(*pro*) SPAL TO [, ne díval se na televizi] [clause/IP]
pro sleept-he TO not watched-he refl at TV
'He slept [, not watched TV].'

I take (48d) to be a violation of the Comp-to-Spec filter (45): the IP, being a complement of Foc, cannot move to Foc's specifier. The example below spells out both the grammatical and ungrammatical structures from the above.

(49) a. $[_{\text{FocP}}\ [_{\text{DP}}\ \text{shoes}]_1\ \text{Foc}\ [_{\text{IP}}\ \text{Peter could not find}\ <[_{\text{DP}}\text{shoes}]_1>\ \text{yesterday}]$ [≈(48b)]
b. *$[_{\text{FocP}}\ [_{\text{IP}}\ \textit{pro}\ \text{slept}]_1\ \text{Foc}\ <[\text{IP}]_1>]$ [≈(48d)]

Note that infinitive-VP focalization, which conforms to the Comp-to-Spec filter, does not lead to the same problems:[10]

(50) a. Najít boty TO Petr včera nemohl
find shoes TO Petr yesterday could-not
b. $[_{\text{FocP}}\ [_{\text{VP}}\ \text{find shoes}]_1\ \text{Foc}\ [_{\text{IP}}\ \text{Petr could not}\ <[\text{VP}]_1>\ \text{yesterday}]]$

A corresponding TP-topicalization is grammatical as well, which means that there is no general problem with preposing finite IPs in Czech. Because *Petr* is obligatorily read as contrastively focused in (51a), I assume that it moves to SpecFocP first and only after that the finite IP is topicalized (moved to SpecTopP).

(51) a. Knížku včera četl Petr.
book yesterday read Petr
b. $[_{\text{TopP}}\ [_{\text{IP}}\ <\text{Petr}_1>\ \text{read book yesterday}]_2\ \text{Top}\ [_{\text{FocP}}\ \text{Petr}_1\ \text{Foc}\ <[\text{IP}]_2>]]$

Note further that the structures in (47) and (48) are closely related to clefts and this relation is not accidental. Clefts realize the very same semantics: contrasted focus with a presupposed background.

(52) a. PETR TO byl, kdo včera nemohl najít boty.
'It was PETR who could not find the shoes yesterday.'

b. BOTY TO byly, co Petr včera nemohl najít.
'It was SHOES that Peter could not find yesterday.'

c. VČERA TO bylo, co Petr nemohl najít boty.
'It was YESTERDAY that Peter could not find the shoes.'

Cleft and pseudocleft structures do not sound very natural in Czech, the more condensed versions of them being the unmarked option. Therefore, I will not provide an explicit analysis

[10] The example (50) must be pronounced without comma intonation after *najít boty* 'find shoes' in order to get the desired (contrastive focus) reading. With a break, the VP is interpreted as topicalized (left-dislocated).

of (pseudo)clefts in Czech. However, see section 5.3 for a brief analysis of English clefts and pseufoclefts.

5.2 The nominal domain

In the nominal domain Foc is placed above NP (alternatively NumP):

(53) ... [$_{\text{FocP}}$ *wh/CF*$_{\text{i}}$ Foc [$_{\text{NP/NumP}}$ ... *wh/CF*$_{\text{i}}$ ...]

As was shown in 2.3, the choice of the focused constituent is much more restricted in the nominal domain than in the verbal domain. Namely, the only element which can move to SpecFocP in a nominal domain is one which also can be a prenominal attribute. Even though I am aware of the significance of this problem, I do not have an account of it at the moment.

In the following subsection I show how the Comp-to-Spec filter applies in the nominal domain. In 5.2.2 I provide an explanation for the fact that left-branch extraction cannot be carried out when TO is involved. Section 5.2.3 briefly deals with the postfix *-ž* which surfaces attached to the *wh*-word in appositive relatives.

5.2.1 Comp-to-Spec filter in the nominal domain

As was shown in 2.3, TO in *wh*-questions must be located between a prenominal attribute and a (full) NP, which means that using TO with simplex *wh*-phrases leads to ungrammaticality:

(54) a. Kterou TO píseň |jste se ho| včera snažili naučit? [=(22)]
which TO song aux-you refl him yesterday tried teach
'Which song did you try to teach him yesterday?'

b. *Co TO |jste se ho| včera snažili naučit?
what TO aux-you refl him yesterday tried teach
'What did you try to teach him yesterday?'

c. *Kdy TO |jste se ho| snažili naučit tu píseň?
when TO aux-you refl him tried teach the song
'When did you try to teach him the song?'

I argue that this is a consequence of the Comp-to-Spec filter, as defined in (45). The structures corresponding to the sentences in (54) are given in (55).

(55) a. [$_{\text{FocP}}$ [$_{\text{AdjP}}$ which] Foc [$_{\text{NP}}$ <[$_{\text{AdjP}}$which]> song]]
b.–c.*[$_{\text{FocP}}$ [$_{\text{whP}}$ what/when] Foc <[$_{\text{whP}}$what/when]>]

5.2.2 No left-branch extraction

Another observation which we made in 2.3 was that when TO is involved in the structure of the NP, then the whole NP has to move to the left periphery in WH questions:

(56) a. Kolik TO knížek |jste| tam nechali ležet? [=(23)]
how-many TO books aux-you there left lie
b. *Kolik TO |jste| tam nechali ležet knížek?
how-many TO aux-you there left lie books

c. *Kolik |jste| tam nechali ležet TO knížek?
how-many aux-you there left lie TO books
d. Kolik |jste| tam nechali ležet knížek?
how-many aux-you there left lie books
'How many books did you leave there lying?'

There are no doubt competing explanations of this state of affairs (ranging from prosody to semantics) and there is not enough space to discuss all of them here. Here, I opt for syntactic explanation. In the analysis so far, I have been following Rizzi (1997) in claiming that the landing position of *wh*-phrases in *wh*-questions is FocP. Consider the structure of a commonplace left-branch extraction in Czech, as exemplified above in (56d) (below, AmP refers to Amount Phrase, and FocV/FocN to a Foc in the verbal or nominal domain respectively):

(57) $[_{\text{FocVP}}\ [_{\text{AmP}}$ how many$]_1$ FocV $[_{\text{IP}}$ you left lying $[_{\text{NP}}$ <$[_{\text{AmP}}$how many$]_1$> books]]]

AmP lands in SpecFocP, receiving a standard interpretation of *wh*-phrases in questions. Now consider the hypothetical structure of the ungrammatical example in (56c):

(58) $[_{\text{FocVP}}$ [how many$]_1$ FocV $[_{\text{IP}}$ you left ... $[_{\text{FocNP}}$ [<[how many$]_1$> TO/FocN $[_{\text{NP}}$ <[how many$]_1$> books]]]]]

In (58) the AmP [how many] first moves to SpecFocNP after which it moves to SpecFocVP. Thus, if we are right at claiming that the verbal and nominal domain are identical as for their functional structure (FocV and FocN are just notational variants), the phrase [how many] occupies one functional position twice. I believe this is exactly the reason why (56c)/(58) is ungrammatical. I suppose there is a general restriction in syntax, which could be defined as follows:

(59) An XP cannot appear in SpecYP and SpecZP in one syntactic tree iff Y = Z.[11]

Although (56b) could be a potential example of this problem, it is ruled out for constituency reasons: it involves a movement of a Spec-Head complex, leaving the complement in situ. Thus, the only possibility for an NP with a Foc-layer is to move as a whole, as exemplified in (56a). The structure is given below:

(60) $[_{\text{FocVP}}$ $[_{\text{FocNP}}$ [how many$]_1$ Foc [NP <[how many$]_1$> books$]_2$ FocV $[_{\text{IP}}$ you left ... [<FocNP$]_2$>]]]]

In the grammatical structure above, first [how many] moves to SpecFocNP. Then the whole FocNP (or any other nominal structure that heads it, e.g. a DP) moves to SpecFocVP. Note that this structure does not yield the violation of (59) because in SpecFocNP and SpecFocVP there are two different phrases: $[_{\text{AmP}}$ how many] and $[_{\text{FocNP/DP}}$ how many TO books] respectively.

[11] An immediate objection to such a filter could be successive-cyclic movement (standardly taking place through intermediate SpeCPs). There are some suggestions in the literature that successive-cyclic movement in its strict form does not exist. For example Boeckx (2003) proposes a theory of chains where intermediate occurrences of a movement chain are strictly distinguished from the hierarchically top ones. I adopt such a chain theory as a background, without further discussion.

5.2.3 Appositive relatives and the postfix *-ž*

As discussed in section 2.3, appositive relatives with an internal head involve a postfix attached to the relative pronoun *který* 'which':[12]

(61) [je] fejetonovým seriálem, kterýž TO název na něj nyní dodatečně připichuji [=(3)]
[is] feuilleton series which TO name on it now afterwards add
'It is a feuilleton series, which name I add to it now, afterwards'

The same postfix appears in appositive relatives which relativize a category different from DP, as in (62a), or on a specialized Czech relative pronoun, which is composed of the (strong version of the) standard personal pronoun and the agglutinating postfix *-ž*, as in (62b).[13,14]

(62) a. Pavel pije denně, čemu-ž my jsme už odvykli
Pavel drinks daily which we aux-we already wean-away
'Pavel drinks daily, which we stopped doing already.'

b. ten muž, jeho-ž jsme včera potkali
the man him aux-we yesterday met
'the man who we met yesterday'

I suppose that the postfix *-ž* is a realization of a functional head placed very high in the hierarchy, presumably ForceP, to whose specifier some relative operators (arguably specific ones) may move. A similar movement of specific relative operators is proposed e.g. by Bianchi (2004). Consider the derivation of the relative clause in (61).[15]

(63) $[_{\text{ForceP}}$ which$_3$ *-ž* $[_{\text{TopP}}$ $[_{\text{FocNP}}$ <which$_{1/3}$> TO $[_{\text{NP}}$ <which$_1$> name$]]_2$ Top [IP ... <FocNP$_2$> ...]]]

An additional movement of the adjectival *wh*-word to a higher functional projection is attested also in questions, as we already noted in section 3.2. I repeat the relevant example below.

(64) Jaké že (TO) žvýkačky si-s včera koupila?
what that TO chewing-gums refl-you yesterday bought
'What (kind of) chewing gums did you buy yesterday?'

In (64) the *wh*-word *jaké* 'which/what' moves up, supposedly to SpecForceP as in the previous case, while the associated NP is left in SpecFocP. As the bracketed TO indicates, this movement is independent of its presence. Note that the Force head is realized by the standard Czech declarative complementizer *že* 'that', morphologically resembling the (relative) postfix *-ž*. Questions like (64) are interpreted roughly as echo-questions.

[12] I assume that appositive relatives are syntactically headed relatives (DP_{REL}) which are in appositional relation to an external head (DP_{EXT}); the appositional relation is then a kind of coordination: DP_{EXT} &: DP_{REL}. Thus, appositive relatives belong to a broader subclass of appositions. Cf. De Vries (2006).

[13] The only exception is a nominative relative pronoun. I refer interested readers to Šimík (2007) for more discussion.

[14] The postfixal *-ž* seems to appear in more contexts, e.g. in the temporal/conditional subordinator *až/když* 'when/if', comparative *než* 'than', or dialectal *tož* 'so'. I am not going to discuss these and do not want to make any claims about their relation to the postfix discussed in the body of the text.

[15] I assume that relative operators move to SpecTopP.

5.3 TO in copular constructions

The analysis above started off from discussing the clitic-like TO in the verbal and nominal domains. In this section I would like to accommodate this analysis to copular predicative constructions, like the ones below, which have been repeatedly discussed in the literature.

(65) a. (Maminka,) TO je učitelka [predication]
mum TO is teacher
'(Mum,) she is a teacher.'

b. (tamta žena,) TO je moje učitelka [specification]
that woman TO is my teacher
'(That woman,) she is my teacher.'

I would like to keep the view that TO is a Foc head and thus argue against the existing proposals of Geist and Blaszczak (2000) and Markman (2007), which suggest that TO (or its equivalents in Russian and Polish) is respectively a conjunction-like element and Top respectively.

I propose that in copular constructions of the kind in (65) TO is a Foc which heads a small clause (perhaps a defective IP?) containing (virtually any kind of) predication. The predicate moves to SpecFocP, thus obtaining the desired focus interpretation. Note also that everything contained in the small clause (apart from the predicate) is automatically interpreted as presupposed.

(66) $[_{\text{FocP}}$ teacher$_1$ TO/Foc $[_{\text{SC}}$ mum/*pro* <teacher$_1$>]

This small clause (more precisely FocP) is further selected by a predicative copula, which eventually enters into agreement with the predicate in SpecFocP.[16] This copula phrase (presumably an IP) is further selected by a Top head and FocP moves to SpecTopP:

(67) $[_{\text{TopP}}$ $[_{\text{FocP}}$ teacher TO $[_{\text{SC}}$ mum/*pro* <teacher$_1$>]]$_2$ Top $[_{\text{IP}}$ is <FocP$_2$>]]

I suppose that the subject of the small clause, being presupposed, is either *pro* or gets deleted. If an overt subject is present (as in Russian and Polish) then I take it to be base-generated above TopP, in a hanging topic position.[17]

The structure in (67) corresponds to the sentence in (68) where *učitelka* 'teacher' is indeed interpreted as a contrastive focus:[18]

(68) Učitelka TO je
teacher TO is

I assume that the unmarked order TO-copula-predicate, as in (65a), corresponds to a structure where *učitelka* 'teacher' moves to SpecIP, which is a standard case of raising from an infinitive

[16] Copulas seem to select CPs (FocP in our labeling) rather liberally, whether they do or do not contain something in their specifier. Consider run-of-the-mill sentences like *The reason is that we have no money* or *Peter is who we were looking for.*

[17] This assumption matches mainly the Czech data, where a potential overt subject must be divided by an intonational pause (comma intonation) from the rest, as indicated in (65) above. However, the Polish and Russian data fits this structure quite well, too, because the subject cannot enter into φ-feature agreement with the copula and is obligatorily assigned nominative (so it cannot appear in ECM construction, cf. Markman 2007), which is arguably the default case in Slavic.

[18] An alternative word order *Učitelka je TO* 'Teacher is it' is also possible. However, the semantics of the two is identical so I suppose prosody is involved here. What is decisive in the CF interpretation, is the predicate's position before TO.

structure (note that the predicate and the copula in I share φ-features, in any case).[19]

(69) $[_{\text{TopP}}\ [_{\text{FocP}}\ <\text{teacher}_{1/2}>\ \text{TO}\ [_{\text{SC}}\ \text{mum/}\textit{pro}\ <\text{teacher}_1>]]_3\ \text{Top}\ [_{\text{IP}}\ \text{teacher}_2\ \text{is}\ <\text{FocP}_3>]]$

The copula *je* 'is', being a clitic, is pronounced in the second position for independent reasons (see also footnote 17). (It can also be argued that it head-moves to Top; see the discussion of English pseudoclefts below and footnote 20.)

Note that with this analysis we can keep our elegant assumption from section 2.1 that nominal predicates in the instrumental, APs, and PPs are barred from predicative constructions with TO only because there is no other DP in the structure which could be assigned case.

Furthermore, the proposed structure can be assigned to a pseudocleft as well as a cleft in a straightforward way. Consider the following (a) examples with their structures in (b):

(70) a. What she (really) is is a teacher
b. $[_{\text{TopP}}\ [_{\text{FocP}}\ [\text{what}\ <\text{teacher}_2>]_1\ \text{Foc}\ [_{\text{IP}}\ \text{she is}\ <\text{what teacher}_1>]]_3\ \text{Top/is}\ [_{\text{IP}}\ \text{teacher}_2\ \text{I}\ <\text{FocP}_3>]]$

(71) a. It is a teacher what she (really) is
b. $[_{\text{TopP}}\ \text{it}\ [\text{Top/is}\ [_{\text{IP}}\ \text{teacher}_2\ \text{I}\ [_{\text{FocP}}\ [\text{what}\ <\text{teacher}_2>]_1\ \text{Foc}\ [_{\text{IP}}\ \text{she is}\ <\text{what teacher}_1>]]]]$

The derivation in (70b) starts out (ignoring the indefinite article) as *she is what teacher*; then *what teacher* moves to SpecFocP within the embedded clause and we get *what teacher*$_1$ *she is* t_1; then *teacher* moves to the SpecIP, where I is a copula selecting the embedded FocP (CP), i.e. *teacher is what she is*; then Top selects the whole structure and after I head-moves to Top,[20] the embedded FocP moves up to its specifier. We end up with *what she is is teacher.*

The derivation of a cleft sentence as in (71b) proceeds in exactly the same way up to the point where the embedded FocP should move. At this point, it stays in situ and the position (SpecTopP) is filled rather by an expletive *it*.[21]

Unfortunately, I have not enough space to fully compare my approach to copular constructions with the competing ones, briefly discussed earlier (section 3.1). I will just mention a few points where my analysis scores much better.

Firstly, TO can figure as a *wh*-phrase in questions, can be relativized, and even focalized:

(72) a. Co je foxteriér?
what is fox-terrier
'What is a fox-terrier?'
b. Tento pes, což je mimochodem foxteriér...
this dog which is by-the-way fox-terrier
'This dog, which is by the way a fox-terrier...'
c. Tady TO je foxteriér [, ne tam TO].
here TO is fox-terrier not there TO
'THIS (one) is fox-terrier [, not THAT (one)]'

[19] Interestingly, by moving to SpecIP the predicate somehow loses its contrastive interpretation. Currently, I have no idea why this should be the case.

[20] Den Dikken, Meinunger, and Wilder (1998) argue that the copula in pseudoclefts is actually a Top head.

[21] Note that this analysis of English pseudoclefts captures the paradoxical property that the focus (*a teacher*) binds the variable (*what*), and at the same time the subject of the pseudocleft (*she*) can bind an anaphor in the predicate, which property is often referred to as connectedness effects, cf. Heycock and Kroch (1999).

The question in (72a) can be asked in a context where the speaker failed to notice the object of reference in a contextually preceding sentence *To je foxteriér* 'This is a fox-terrier'. Under my approach, the question formation, focalization, and relativization of the above kind is no problem: the whole FocP headed by TO is assigned the features of an operator and can undergo *wh-/focus*-movement.[22] Thus, the small clause will move to SpecFocP in a question or in focalization, and to SpecTopP in a relative clause. Note that the relative operator bears the postfix *-ž*, suggesting that it moves further to SpecForceP, as discussed in 5.2.3.

Geist and Blaszczak (2000) and Markman (2007) both fail to capture these data: in the former approach, TO is a conjunction-like element and as such cannot bear operator features; in the latter approach, TO is a Top head whose complement is the predicate and the information structure properties of TO (if any at all) are expected to be those characteristic of a predicate and not those of a subject, as in my approach.

Secondly, my approach correctly captures the intuition that there is something definite about TO (recall that morphologically it is a definite-like demonstrative pronoun). As discussed above, TO renders its complement existentially presupposed. This comes quite naturally if TO is a Foc head. This aspect is problematic mainly for Geist and Blaszczak who have to make two unusual assumptions: a conjunction-like element bears definiteness features and this conjunction "agrees in definiteness" with its specifier (the first conjunct).

Thirdly, in my analysis there is at least one phase of the derivation where the predicate (being a focus) binds the focus-variable, which appears in the small clause (in fact, they are in movement relation). This seems to be necessary for independent reasons and both the competing approaches fail to capture this: in Markman the predicate originates in the complement of Top and is too low to bind the focus-variable, which is located in SpecTopP in his analysis. Geist and Blaszczak face exactly the same problem.

6 Conclusion

In this paper I have made a proposal about the categorical and semantic status of the invariant demonstrative-like element TO in Czech in a number of constructions, namely predications with the copula *být* 'be', *wh*-questions and structures involving a contrastive focus, appositions (including appositive relative clauses), and exclamations.

I have argued that TO is a realization of the Foc head, which is a low projection within the fine structured CP system, the "left periphery" in Rizzi's (1997) terms. Following the DP-CP parallelism hypothesis, I argued that TO may appear in the left periphery of verbal as well as nominal projections. The semantic import of TO is partly identical to that of the definite article: it causes its complement to be interpreted as existentially presupposed. The difference, however, is that TO also requires its complement to contain a variable which gets bound by a *wh-* or *focus*-operator, placed in the SpecFocP. The required operator-variable relation is created by a movement of a *wh-/focus*-phrase from the complement of TO to its specifier and can thus be captured by standard syntactic tools. I have showed that the choice of this variable is generally constrained by a version of the Comp-to-Spec filter (e.g. Abels 2002). Moreover, in the nominal domain there are more severe restrictions, a problem, which has to remain open in this paper.

The last part of the analysis deals with copular predicative constructions with TO. Under the present proposal they can be analyzed as involving a small clause with a standard subject-predicate structure, which is headed by TO. The analysis, based on movement of the predicate,

[22] Again, note that CPs/FocPs can be asked about: *What did he say? He said that it is true.*

is directly translatable to an account for English cleft and pseudocleft structures. Finally, I suggest that my analysis of the copular constructions is empirically and theoretically superior to the previous ones, which treat TO as a conjunction-like element (Geist and Blaszczak 2000) or a Top head (Markman 2007), the virtual inverse of my proposal.

References

Abels, Klaus. 2002. *Successive cyclicity, anti-locality, and adposition stranding.* PhD dissertation. University of Connecticut.

Bennis, Hans, Norbert Corver, and Marcel den Dikken. 1998. Predication in nominal phrases. *The Journal of Comparative Germanic Linguistics* 1:85–117.

Bianchi, Valentina. 2004. Resumptive relatives and LF chains. Rizzi, L. (ed.). *The structure of CP and IP: The cartography of syntactic structures, Volume 2:*76–114. Oxford: Oxford University Press.

Blaszczak, Joanna and Ljudmila Geist. 2001. Zur Rolle des Pronomens to/eto in spezifierenden Kopulakonstruktionen im Polnischen und Russischen. Zybatow, G., U. Junghanns et al. (eds.). *Current Issues in Formal Slavic Linguistics*: 247–257. Frankfurt: Peter Lang.

Boeckx, Cedric. 2003. *Islands and chains: Resumption as stranding.* Philadelphia, PA: John Benjamins Publishing Company.

Boškovic, Željko. 1998. *Wh-phrases and wh-movement in Slavic.* University of Connecticut, Storrs.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: The MIT Press.

den Dikken, Marcel, Andre Meinunger, and Chris Wilder. 1998. Pseudoclefts and ellipsis. *ZAS Papers in Linguistics* 10:21–70.

Geist, Ljudmila and Joanna Blaszczak. 2000. Kopulasätze mit den pronominalen Elementen to/eto im Polnischen und Russischen. *ZAS Papers in Linguistics* 16:115–139.

Geurts, Bart and Rob van der Sandt. 2004. Interpreting focus. *Theoretical Linguistics* 30:1–44.

Grosu, Alexander and Fred Landmann. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6:125–170.

Hagstrom, Paul. 2003. What questions mean. *Glot International* 7:188–201.

Hamblin, C. L. 1958. Questions. *The Australasian Journal of Philosophy* 36:159–168.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in generative grammar.* Oxford: Blackwell Publishers.

Heycock, Caroline and Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30:365–397.

Hiraiwa, Ken. 2005. *Dimensions of symmetry in syntax: Agreement and clausal architecture.* PhD dissertation. MIT, Cambridge, MA.

Horvath, Julia. 1986. *Focus in the theory of grammar and the syntax of Hungarian.* Dordrecht: Foris.

Karttunen, Lauri. 1977. The syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.

Kayne, Richard S. 1994. *The antisymmetry of syntax.* Cambridge, MA: The MIT Press.

Markman, Vita G. 2007. Pronominal copula contructions are what? Reduced specificational pseud-clefts! Talk presented at FASL 16, Stony Brook University.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37:69–109.

Miro, Elena Castroviejo. 2002. *Wh-exclamatives in Catalan*. PhD dissertation. Universitat de Barcelona.

Rizzi, Luigi. 1997. The fine structure of the left periphery. Haegeman, L. (ed.). *Elements of grammar: A handbook of generative syntax*: 281–337. Dordrecht: Kluwer Academic Publishers.

Rullmann, Hotze. 1995. *Maximality in the semantics of wh-constructions*. PhD dissertation. University of Massachusetts, Amherst.

Szabolcsi, Anna. 1994. The noun phrase. *Syntax and semantics 27: The syntactic structure of Hungarian*: 179–274. Academic Press.

Veselovská, Ludmila. 1995. *Phrasal movement and X^0 morphology: Word order parallels in Czech and English nominal and verbal projections*. PhD dissertation. Palacky University Olomouc.

de Vries, Mark. 2006. The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry* 37:229–270.

Wh-questions with conjoined wh-words

Hana Skrabalova

hanaskrabal@yahoo.fr

1 Two types of questions with conjoined wh-words

This paper examines two types of wh-questions containing two (or more) wh-words and where the last wh-item is introduced by a conjunction. In type 1 questions, all wh-words, including the one introduced by the conjunction, appear at the beginning of the interrogative clause, see (1). In type 2 questions, one or more wh-words appear at the beginning of the clause, and the wh-item introduced by the conjunction appears at the end of the clause, see (2). Through the paper, I will focus on questions with only two wh-words, but all what will be said also apply to questions with more then two wh-words.[1]

(1) a. Kdo a koho pozval? (type 1 questions)
who-NOM and who-ACC invited
„Who invited whom?“

b. Kdo koho a kdy pozval?
who-NOM who-ACC and when invited
"Who invited whom, and when?"

(2) a. Koho viděl a kde ? (type 2 questions)
who-ACC (he) saw and where
"Whom did he see, and where?"

b. Koho kde viděl a kdy?
who-ACC where (he) saw and when
"Whom did he see, where and when?"

Both question types exist in other Slavic and non-Slavic language, as show the following examples from Russian, French, and Hungarian:

(3) a. Kto i čto kupil? (Russian, Kazenin 2001)
who-NOM and what-ACC bought
"Who bought what?"

[1]The following abbreviations are used in the glosses: ACC = accusative case, ADJ = adjective, AUX = auxiliary verb ('to be'), CL = clitic, COND = conditional, DAT = dative case, FUT = future tense, GEN = genitive case, IMP = imperative, INSTR = instrumental case, NEG = negation, NOM = nominative case, PART = particle, POSS = possession, QUANT = quantifier, REFL = reflexive pronoun, SG = singular.

b. Kogda zakončilas' Perestrojka i čem?
when ended Perestroik and what-INSTR
"When did Perestroik end and how?"

(4) a. Quand et où a lieu cette conférence? (French, Skrabalova 2006b)
"When and where takes place this conference?"

b. Qui as-tu rencontré, et où ?
"Whom did you meet, and where?"

(5) a. Ki és mikor látta Marit? (Hungarian, Lipták 2001)
who-NOM and when saw-3SG Mary-ACC
"Who saw Mary, and when?"

b. Ki látta Marit és mikor?
who-NOM saw-3SG Mary-ACC and when
"Who saw Mary, and when?"

The aim of this paper is to describe the properties of Czech questions in (1) and (2) and to analyze their syntactic structure. I especially ask whether these questions involve clausal coordinations with ellipsis or phrasal coordinations. The paper is organized as follows. In the section 2, I show the syntactic, prosodic and semantic properties of wh-words in type 1 questions. In the section 3, I show the syntactic, prosodic and semantic properties of wh-words in type 2 questions. The section 4 summarizes the properties of the two question types. Finally, the section 5 proposes a syntactic analysis for each type of questions. It is argued that type 1 questions involve phrasal coordination, problematic though it can be from semantic point of view. Type 2 questions are argued to be clausal coordinations with one elliptical conjunct.

2 Properties of wh-words in type 1 questions

2.1 Syntactic functions of wh-words

Wh-words in type 1 questions can have either the same syntactic function as in (6), where both wh-words stay for the internal argument of the verb, or different syntactic functions as in (7)[2], where the first wh-word stands for the internal argument of the verb and the second one for a locative adjunct.

(6) Koho a čeho se bojíš?
who-GEN and what-GEN REFL fear-2SG
"Who and what do you fear?"

(7) Nechtěl říct, koho a kde viděl.
(he) NEG-wanted say who-ACC and where (he) saw
"He didn't want to say whom he had seen, and where."

Coordinations like in (6) can be considered as quite normal, since they obey the Wasow's generalization according to which conjoined items must have the same syntactic function and share the syntactic properties associated with this function (Pullum & Zwicky 1986). On the contrary, coordinations like in (7) are problematic with respect to this generalization (see section 5.2.2). Such coordinations are however completely natural in Czech (and other Slavic

[2]Wh-words *kdo* and *co* having different functions also bear different cases.

languages), where the conjoined wh-words can be both arguments and adjuncts. Czech is thus more permissive than languages like French, English or Dutch, where only adjuncts can be conjoined. The example (8) below shows a coordination of two arguments, a coordination of two adjuncts and a coordination of one argument and one adjunct. Note also that there are no superiority effects in type 1 questions, exactly as in multiples questions (Meyer 2004):

(8) a. Komu a co / Co a komu řekl?
who-DAT and what-ACC / what-ACC and who-DAT (he) said
"What did he say, and to whom?"

b. Kdy a jak / Jak a kdy skončila třicetiletá válka?
when and how / how and when finished 30-years war
"When and how did the 30-years war finish?"

c. Kdo a kdy / Kdy a kdo napsal tu knihu?
who-NOM and when / when and who-NOM wrote this-ACC book-ACC
"Who evaluated the students, and how?"

However, contrary to multiple questions, wh-words embedded within another phrase (NP, PP) cannot be conjoined with non-embedded ones, as shown in (9). This suggests that conjoined wh-words are not syntactically independent, as it will be shown in the following subsections.

(9) a. *Kdo a koho má rád dceru? (conjoined wh-words)
who-NOM and who-GEN likes well daughter-ACC
b. Kdo koho má rád dceru? (multiple wh-words)
who-NOM who-GEN likes well daughter-ACC
"Who likes whose daughter?"

2.2 Wh-words as prosodic and syntactic unit

Conjoined wh-words in type 1 questions form undoubtedly a prosodic unit. Indeed, it is not possible to separate by a pause the wh-word introduced by the conjunction from the first wh-word, as in (11), where the commas indicate the separation. All wh-words can nevertheless bear a heavy stress.

(10) a. Dej mi vědět, [kdo a koho] pozval.
give me know who-NOM and who-ACC (he) invited
"Let me know who invited whom."

b. Dej mi vědět, [kam a kdy] Petr šel.
give me know where and when Peter went
"Let me know where and when Peter went."

(11) a. *Dej mi vědět [kdo], [a koho], pozval.

b. *Dej mi vědět [kam], [a kdy], Petr šel.

But the conjoined wh-words in (10) form not only a prosodic unit, but also a syntactic unit. We can see indeed that neither second-position clitics nor the complementizer že ('that') may intervene between the conjoined wh-words in (12). In multiple questions, on the contrary, both

clitics and complementizer wh-words may appear between multiple wh-words, as shown in (13).

(12) a. Kdo (*mu) a co (mu) řekl? (conjoined wh-words)
who-NOM CL:he-DAT and what-ACC CL:he-DAT said
"Who said which thing to him?"

b. Kdo (*že) a co (že) řekl?
who-NOM that and what-ACC that said
"Who said which thing?"

(13) a. Kdo (mu) co (*mu) řekl? (multiple wh-words)
who-NOM CL:him what-ACC CL:him said
"Who said what to him?"

b. Kdo (že) co (*že) řekl?
who-NOM that what-ACC that said
"Who said what? "

The constrast between the conjoined words in (12) and the multiple wh-words in (13) suggests that conjoined wh-words are not independent constituents, but that they form a single constituent, which includes the conjunction.

2.3 2.4 Reading of type 1 questions

Type 1 questions receive a single-pair reading, contrary to multiple questions which receive a list-pair reading, compare (14) and (15). This means that type 1 questions do not presuppose the existence of a specific set of individuals from which the answer would be picked. In (14), the questions asks for the identification of the person who bought something and for the identification of the thing which has be bought by this person.

(14) a. Kdo a co koupil?
who-NOM and what-ACC bought
"Who bought what?"

b. Marie koupila auto.
Mary-NOM bought car-ACC
"Mary bought a car."

(15) a. Kdo co koupil?
who-NOM what-ACC bought
"Who bought what?"

b. Marie koupila auto, Petr motorku a Jan kolo.
Mary-NOM bought car-ACC, Peter-NOM motorcycle-ACC and John-NOM bike-ACC
"Mary bought a car, Peter (bought) a motorcycle, and John (bought) a bike."

Answers to type 1 questions are thus single propositions in which the constituents corresponding to the wh-words may be focused. The conjunction can also appear in the answer, backed up by a demonstrative particle *to*, as shown in (16).

(16) a. Kdo a komu koupil auto?
who-NOM and who-DAT bought car
"Who bought the car, and to whom?"

b. MARIE koupila kolo, a to PETROVI.
Mary-NOM bought bike-ACC and this Peter-DAT
"Mary bought the car to Peter."

Finally, all conjoined wh-words in type 1 questions are involved in a single event denoted by the verb of the interrogative clause. Therefore, it is impossible, in particular when wh-words are arguments, to use in these questions the conjunction *nebo* ('or'), as shown in (17a), or to paraphrase them by a sentential coordination, as shown in (17b). So, the data in (17) confirm that conjoined wh-words in type 1 questions form a single constituent.

(17) a. *Kdo nebo koho pozval?
who-NOM or who-ACC invited
b. *Kdo pozval a koho pozval?
who-NOM invited and who-ACC invited

3 Properties of wh-words in type 2 questions

3.1 Syntactic functions of wh-words

Conjoined wh-words in type 2 questions can also have either the same syntactic function or different syntactic functions, as shown in (18) and (19) respectively:

(18) Koho se bojíš, a čeho?
Who-GEN REFL fear-2SG and what-GEN
"Who do you fear, and what?"

(19) Chtěl jsem vědět, koho potkal a kde.
wanted AUX-1SG know who-ACC (he) met and where
"I wanted to know whom he had met and where."

Wh-words in type 2 questions can also be arguments or adjuncts, as shown in the example (20). However, argumental wh-words cannot appear in the clause-final position, except if they stand for an optional argument, as the word *komu* (to-whom) in (20c).

(20) a. Koho jsi viděl a kde? / *Kde jsi viděl a koho?
Who-ACC AUX-2SG saw and where / where AUX-2SG saw and who-ACC
"Whom did you see, and where?"

b. Kdy skončila třicetiletá válka a jak? / Jak skončila ... válka a kdy?
when finished 30-years war and how / how finished ... war and when
"When did the Thirty years war finish, and how?"

c. Co řekl a komu? / *Komu řekl a co?
what-ACC (he) said and who-DAT / who-DAT (he) said and what-ACC
"What did he say, and to whom?"

3.2 Prosodic properties of wh-words

The wh-item introduced by the conjunction can either form a prosodic unit with the interrogative clause, or form an independent prosodic unit. In the case the wh-word is not integrated into the clause (indicated by — in (21)), it will be emphasized:

(21) a. Kdy jsi potkal Janova bratra, — a KDE?
when AUX-2SG met Jean-POSS brother-ACC — and where
"When did you meet John's brother, and where?"

b. Chtěl bych vědět kdy skončila ta hrozná válka v Bosně, — a JAK?
wanted COND-1SG know when finished that terrible war in Bosnia — and how
" I would like to know when that terrible war in Bosnia finished, and how."

However, the sequence [conjunction wh-word] may not appear inside the interrogative clause, wheather it is prosodically integrated or not, as shown in (22). Since the sequence [conjunction wh-word] in Czech cannot move through the clause, it cannot thus function as an adjunct:

(22) a. *Kdy jsi potkal, a KDE, Janova nejmladšího bratra?
when AUX-2SG met and where Jean-POSS youngest-ACC brother-ACC
"When did you meet the John's youngest brother, and where?"

b. *Chtěl bych vědět kdy skončila, a JAK, ta hrozná válka v Bosně.
wanted COND-1SG know when finished and how that terrible war in Bosnia
"I would like to know when that terrible war in Bosnia finished, and how."

3.3 Reading of type 2 questions

Type 2 questions are interpreted as conjoined single questions. They can easily be paraphrased by sentential coordinations:

(23) a. Kdo přijde a kdy?
who-NOM come-FUT-3SG and when
"Who will come, and when"?

b. Kdo přijde a kdy přijde?
who-NOM come-FUT-3SG and when (he) come-FUT-3SG
"Who will come, and when will he come?"

The wh-words in type 2 questions receive thus a single-pair reading, as in type (1) questions. In the answer, the constituents corresponding to the wh-words can be focused and the conjunction can again appear in the answer. The sentences in (24) are possible answers to the questions in (23).

(24) a. Jan přijde dnes večer.
John come-FUT-3SG this evening
"John is coming this evening."

b. Přijde Jan a (to) dnes večer.
come-FUT-3SG John, and PART this evening
"John is coming, and this evening!"

Finally, in type 2 questions with more than two wh-words, the wh-words at the beginning of the clause can be either conjoined or adjacent, as we can see in (25a) and (26a). This means that the questions in (25a) and (26a) combine two different strategies of questioning. The question (25a) is a type 2 question which contains a type 1 question. The question (26a) is a type 2 question which contains a multiple question. Therefore, these questions do not have the same reading.

In the question (25a), the conjoined wh-words receive a single-pair reading which combines with a single reading of the conjoined final wh-word. The answer in (25b) identifies thus the two persons involved in the inviting and the times when this happened.

(25) a. Kdo a koho pozval, a kam?
who-NOM and who-ACC invited and where
"Who invited whom, and where?"

b. Jan pozval Marii do divadla.
John-NOM invited Mary-ACC to theatre-GEN
"John invited Mary to the theatre."

In the question (26a), the multiple wh-words receive a list-pair reading, which combines with a single reading of the conjoined wh-word. The answer in (26b) contains the list of pairs of inviting and invited persons. The wh-word introduced by the conjunction applies then to each pair of the list.

(26) a. Kdo koho pozval, a kam?
who-NOM who-ACC invited and where
"Who invited whom, and where?"

b. Jan pozval Marii do divadla, Petr Moniku do
John-NOM invited Mary-ACC to theatre-GEN, Peter-NOM Monika-ACC to
kina a Pavel Annu na večeři.
cinema-GEN and Paul-NOM Anna-ACC on dinner-ACC
"John invited Mary to the theatre, Peter (invited) Monika to the cinema, and Paul (invited) Anna to have dinner."

4 Summary of properties of conjoined wh-words

This section summarizes the properties of wh-words in type 1 and 2 questions:

Property	Type 1 questions *Wh₁ Conj Wh₂*	Type 2 questions *Wh₁…Conj Wh₂*
Wh_1 = argument	yes	yes
Wh_1 = adjunct	yes	yes
Wh_2 = argument	yes	no*
Wh_2 = adjunct	yes	yes
Adjacency of Wh_1 and [Conj Wh_2]	yes	no
[Conj Wh_2] prosodically autonomous	no	yes
Single-pair reading	yes	yes
Sentential reading	no**	yes

*unless (i) Wh_2 is an optional argument and (II) Wh_2 has the same function as Wh_1.

**unless Wh_2 has the same function as Wh_1.

5 Syntactic analysis

This section deals with the syntactic analysis of the two question type 1. Two issues in particular will be raised. The first issue is whether these questions involve clausal coordination with ellipsis or phrasal coordination. If they involve clausal coordination with ellipsis, the second issue is whether the elliptic conjunct contains deleted material or not.

5.1 Clausal coordination with ellipsis

According to Banréti (1992), both types of questions in Hungarian involve clausal coordination with deletion in one of the conjoined clauses. In type 1 questions, deletion occurs in the first clause, as shown in (27b). In type 2 questions, deletion occurs in the second clause, as shown in (28b).

(27) a. Kdo a kdy přišel?
who-NOM and when came
b. [[Kdo ~~přišel~~] a [kdy přišel]]
who-NOM came and when (he) came

(28) a. Kdo přišel a kdy?
who-NOM came and when
b. [[Kdo přišel] a [kdy ~~přišel~~]]
who-NOM came and when (he) came

The clausal analysis places the two question types on parallel grounds. We have seen, however, that these questions do not have the same properties. This conception must thus be wrong. It can be easily shown that type 1 and type 2 questions are different syntactic structures. The following subsections will show that clausal analysis must be rejected for type 1 questions, but can be maintained for type 2 questions.

5.1.1 Against clausal analysis of type 1 questions

The first evidence against clausal analysis of type 1 questions comes from their constituency. It has been shown in the examples (12) and (13) above that neither clitics nor complementizer, that is the elements that occur in the left periphery of the clause and intervene between multiple wh-words, may intervene between conjoined wh-words. This behaviour cannot be explained by the clausal analysis in (27) according to which the conjoined wh-words do not form a constituent.

The second argument comes from the questions where both wh-words are obligatory arguments of the verb. We have seen that type 2 questions where the final wh-word is an obligatory argument are ungrammatical. According to the clausal analysis, however, both question types have the same underlying structure and are thus semantically equivalent. This means that the clausal analysis cannot explain the difference in acceptability of examples like in (29).

(29) a. Kdo a koho uhodil?
who-NOM and who-ACC hit
"Who hit whom?"

b. *Kdo uhodil a koho?
who-NOM hit and who-ACC

Moreover, the underlying structure of the questions in (29) are ungrammatical since the conjoined clauses miss the obligatory arguments, see (30a). But even if all the arguments were present, we would have to admit that some are null, as in (30b). However, there are no other evidence in Czech for a null (referential) object pronoun and the null (referentially definite) subject pronoun cannot be questioned.

(30) a. *[[Kdo uhodil] a [koho uhodil]]?
who-NOM hit and who-ACC hit

b. [[Kdo uhodil pro_{obj}] a [koho pro_{suj} uhodil]]?
who-NOM hit (him) and who-ACC he hit

I conclude then that the clausal analysis cannot account for type 1 questions where conjoined wh-words have different syntactic functions[3].

5.1.2 Evidence for a clausal analysis of type 2 questions

Contrary to type 1 questions, type 2 questions are easily analyzed as clausal coordinations.

First, the clause containing the first wh-item can function as an independent interrogative clause. Second, the wh-item introduced by the conjunction cannot be an argument of the verb in the interrogative clause. That means the wh-word in this position is external to the interrogative clause.

Third, the wh-item introduced by the conjunction is always interpreted as a single question.

Finally, if type 2 questions involve two clauses, we predict that the first one can contain multiple or conjoined wh-words. This is what we have seen in (25) and (26) above. I conclude thus that type 2 questions involve clausal coordination:

(31) $[_{\text{CoordP}}\ [_{\text{CP1}}\ \text{Wh}_1\ [_{\text{TP}}\ldots]]\ \text{Conj}\ [_{\text{CP2}}\ \text{Wh}_2\ [_{\text{TP}}\ldots]]]$

[3] A clausal analysis would be plausible for type 1 questions with conjoined wh-words bearing the same function.

The questions arises however what is the syntactic status of the elliptic conjunct in type 2 questions. There are two possibilities. Either the elliptic conjunct is a clause with deleted material, or the elliptic conjunct is a clausal fragment (Ginzburg & Sag 2001). In the former case, the syntactic reconstruction with identity should always be possible. In the latter case, there would be no syntactic reconstruction, but the fragment would be interpreted as a clause. If we look now on the following examples, we see that the syntactic reconstruction with identity is possible only when the first wh-item is an adjunct, as in (32a). If the first wh-item is an argument, an NP or a pronoun must appear in the second clause, see (32b) and (32c).

(32) a. Kdy jsi potkal Jana a kde (jsi potkal Jana)?
when AUX-2SG met John-ACC and where AUX-2SG met John-ACC
"When did you meet John and where (did you meet John)?"

b. *Koho jsi potkal a kde (jsi potkal)?
who-ACC AUX-2SG met and where AUX-2SG met
(* "Who did you meet John and where did you meet? ")

c. Koho jsi potkal a kde (jsi ho potkal)?
who-ACC AUX-2SG met and where AUX-2SG CL:him met
"Where did you meet John and where did you meet him? "

I suggest thus that the elliptic conjunct would be better analyzed as a clausal fragment in the sense of Ginzburg and Sag (2001).

5.2 Phrasal coordination

5.2.1 Evidence for a phrasal analysis of conjoined wh-words in type 1 questions

There are three pieces of evidence for phrasal analysis of conjoined wh-words in type 1 questions. First, the conjoined wh-words are strictly adjacent and behave as a single constituent.

Second, the conjoined wh-words behave as a single prosodic unit.

Finally, the conjoined wh-words are all involved into a single event denoted by the verb of the interrogative clause. This strongly suggests that conjoined wh-words form a coordinated phrase (see Lipták 2001 for Hungarian), as in (33).

(33) $[_{CP}\ [_{CoordP}\ Wh_1\ Conj\ Wh_2]\ [_{TP}...]]$

We can however ask whether the morpheme *a* ('and') in type 1 question is really a conjunction, that means the head of the coordinate phrase (or Conjunction Phrase), as it has been proposed by Johannessen (1998). Penn (1999) claims that in type 1 questions in Serbo-Croatian, the morpheme *i* is not a conjunction, but a focus particle. Indeed, in Serbo-Croatian, the morpheme *i* which appears between the wh-words in (34a) also introduces the focused constituents, as we can see in (34b) and (34c), where *i* means 'also'. According to Penn, thus, type 1 questions are thus multiple questions where the wh-words are focused by the presence of the morpheme *i*.

(34) a. Ko i kome je kupio auto? (Penn 1999)
who-NOM and who-DAT AUX-3SG bought car
"Who bought the car for whom?"

b. Ivan je i danas sreo Mariju.
Ivan AUX-3SG also today met Mary-ACC

"Ivan also met Mary today (not only yesterday)."

c. Knjigu i Mariji odnesi.
book-ACC and/also Mary-DAT bring-IMP
" Bring the book to Mary."

I claim however that Penn's analysis is not a correct analysis, at least for Czech. In Czech, the morpheme *a* only appears in coordinate context, see (35a). It can never function as a focus particle, see (35b).

(35) a. Jan potkal Marii včera a dneska.
John met Mary yesterday and today
"John met Mary yesterday and today."

b. Jan potkal Marii *a / také dneska.
John met Mary and / also today
"John met Mary also today (not only yesterday)."

There is actually another morpheme in Czech, the morpheme *i*, which functions both as a conjunction (forcing a distributive reading) and as a focus particle (Skrabalova 2004), see (36a) and (36b). However, *i* cannot conjoin wh-words, see (36c).

(36) a. Jan potkal Marii včera i dneska.
John met Mary yesterday and today
"John met Mary both yesterday and today."

b. Jan potkal Marii i dneska.
John met Mary even/also today
"John met Mary even/also today (not only yesterday)."

c. Kdo a / *i komu koupil auto?
who-NOM and / and who-DAT bought car
"Who bought the car to whom?"

According to the data in (35) and (36), the morpheme *a* that appears between the wh-words in type 1 questions is a conjunction, and not a focus particle. Conjoined wh-words cannot be analyzed as multiple wh-words. This is also confirmed by the fact that multiple wh-words differ from conjoined wh-words by two other properties (see section 2): they do not form a single constituent, and they receive a list-pair reading.

5.2.2 How are the coordinations of wh-phrases licensed ?

It is generally assumed that conjuncts can neither be of different category nor have different syntactic functions (Williams 1981, Peterson 2004), as shown in (37a). Therefore, it should not be possible to conjoin wh-phrases in (37b), which is not correct.

(37) a. *I helped [$_{NP}$ Peter] and [$_{AdvP}$ quickly].

b. *Jan pozval [$_{CoorP}$ [$_{NP}$ Marii] a [$_{PP}$ do kina]].
John invited Mary-ACC and to cinema

There are however at least two kinds of coordinations where conjuncts may have different functions. The first kind of unlike coordinations involves emphasized constituents. It has been noted (Lipták (2001) among others) that unlike constituents can be conjoined provided they are emphasized:

(38) a. John read [$_{NP}$ a book] and [$_{AdvP}$ quickly]!

b. John met [$_{CoorP}$ [$_{NP}$ MARY] and [$_{PP}$ IN HER HOUSE]].

If we consider the coordinations in (39b) and (40b) in Czech, we see indeed that these coordinations are felicitous because the conjuncts are emphasized, or focused. The focusing here is natural if these sentences are used as answers to the questions in (39a) and (40a) respectively. The focused conjoined constituents in the answer correspond to the conjoined wh-words in the question. This is not surprising since wh-words are focus elements.

(39) a. Chtěl bych vědět, koho a kam chce Jan pozvat.
wanted COND-1SG know who-ACC and where want-3SG John invite
"I would like to know whom John wants to invite, and where."

b. Myslím, že chce pozvat MARII a DO KINA.
think-1SG that want-3SG invite Mary-ACC and to cinema
"I think that he wants to invite Mary to the movie."

(40) a. Nevíš, kam a proč Petr jel?
NEG-know-2SG where and why Peter went
"Do you know where Peter went and why?"

b. Nejsem si jistý, ale myslím, že jel DO LONDÝNA a NA
NEG-am REFL sure but think-1SG that (he) went to London and for
NĚJAKÝ KONGRES.
some congress
"I'm not sure, but I guess he went to London for a congress."

The coordinations of focused unlike phrases like in the examples (39b) and (40b) suggest thus that wh-words may be conjoined precisely because they are focus elements.

The second kind of unlike coordinations involves quantifiers. The data below show that universal and negative quantifiers bearing different functions may indeed be conjoined:

(41) a. Kdykoli a kamkoli jdu, vždycky ho potkám.
when-QUANT and where-QUANT go-1SG, always him meet-1SG
"Whenever and wherever I go, I always come across him."

b. Kdykoli a kdekoli ho potkám, vždycky je dobře naladěn.
when-QUANT and where-QUANT him meet-1SG always is well mood-ADJ
"Whenever and wherever I meet him, he is always in a good mood."

(42) a. Pavel je pořád zalezlý doma. Ten nikdy a nikam nechce jít.
Paul is always hidden home he never and no-where NEG-wants go
"Paul always hides home. He never wants to go anywhere."

b. Pavel má něco za lubem, ale nechce nic a
Paul has something in mind, but NEG-want-3SG nothing-ACC and
nikomu říct.
nobody-DAT say
"Paul has something in mind, but he does not want to say anything to anybody."

If we assume that wh-words are quantified expressions (see Beghelli & Stowell 1997), the examples of quantifier coordinations in (41) and (42) suggest that wh-words may be conjoined precisely because they are quantified expressions.

5.3 Remaining issues

The analysis proposed in the previous sections leaves at least two remaining issues. The first issue concerns the fact that the coordinate wh-phrases have no equivalent *in situ*, contrary to other constructions involving extraction. The same problem arises however with the constructions such as partial VP fronting in German (Haider 1990 among others) or CP topicalization in English (Bresnan 1972). This issue is thus not specific to conjoined wh-word phrases. The second issue is why wh-coordination is not possible to the same extent in other languages which allow coordination of focused or quantified unlike constituents. This question suggests that there are other parameters licensing the wh-coordinations. Unfortunately, the answer to that question goes beyond the limits of this paper.

6 Conclusion

In this paper, I examined two types of questions with conjoined wh-words. In type 1 questions, all conjoined wh-words appear at the beginning of the interrogative clause. In type 2 questions, one or more wh-words appear at the beginning of the interrogative clause, while the wh-word introduced by the conjunction is clause-final. I showed that type 1 and 2 questions have different syntactic, semantic, and prosodic properties. In particular, conjoined wh-words in type 1 questions form a single constituent and type 1 questions cannot be interpreted as coordinations of two (or more) single questions. Consequently, I argued that these two question types involve different syntactic structures. The differences between type 1 and type 2 questions can indeed be explained if we analyze type 1 questions as involving clause internal coordination of wh-phrases, and type 2 questions as involving coordination of clauses, one of them being elliptic. The coordination of wh-words bearing different syntactic functions seems problematic, but it is not an isolated phenomenon. Focused constituents and quantifiers bearing different functions can also be conjoined. Wh-coordination is thus another counter-example to the generalization that conjuncts must bear the same syntactic function.

References

Abeillé, Anne. 2003. A Lexicalist and Construction-based Approach to Coordinations. Müller, S. (ed.). *Proceedings of the HSPG03 Conference*. Stanford: CSLI publications.

--. 2005, Les syntagmes conjoints et leurs functions syntaxiques. *Langages* 160:42–66.

Banréti, Zoltan. 1992. A mellénrendelés [La coordination]. *Strukturális magyar nyelvtan I. Mondattan*: 715–796. Budapest: Akadémiai Kiadó.

Beghelli, Filippo and Tim Stowell. 1997. Distributivity and Negation: the Syntax of *each* and *every*. Szabolcsi, A. (ed.). *Ways of Scope Taking*:71–109. Dordrecht: Kluwer.

Bresnan, Joan. 1972. *The theory of complementation in English syntax*. PhD Dissertation. MIT, Cambridge, Massachussets.

Dik, Simon. 1997. *The Theory of Functional Grammar*. Berlin: Mouton de Gruyter.

Ginzburg, Jonathan and Ivan Sag. 2000. *Interrogative Investigations. The Form, Meaning, and Use of English Interrogatives*. Stanford: CSLI Publications.

Haider, Hubert. 1990. Topicalization and othes puzzles in German syntax. Grewendorf, G. & W. Sternefeld (eds.). *Scrambling and Barriers*: 93–112. Amsterdam: John Benjamins.

Johannessen, Jane B. 1998. *Coordination*. Oxford: Oxford University Press.

Kazenin, Konstantin. 2001. On coordination of Wh-phrases in Russian. Talk given at *The Fourth European Conference on Formal Description of Slavic Languages*. Potsdam, Nov., 28–30.

Lipták, Aniko. 2001. *On the Syntax of Wh-words in Hungarian*. Utrecht: LOT.

Meyer, Roland. 2004. Superiority Effects in Russian, Polish, and Czech. *Cahiers linguistiques d'Ottawa* 32:44–65. University of Ottawa.

Penn, Gerald. 1999. Linearization and WH-Extraction in HPSG: Evidence from Serbo-Croatian. Borsley, R. & A. Przepiorkowski (eds.). *Slavic in HPSG*. Stanford: CSLI Publications.

Peterson, Peter. 2004. Coordination: consequences of a lexical-functional account. *Natural Language and Linguistic Theory* 22:643–679.

Pullum, Geoffrey and Arnold Zwicky. 1986. Phonological resolution of syntactic features conflicts. *Language* 62:751–773.

Rudin, Catherine. 1988. On Multiple questions and Multiple Wh-fronting. *Natural Language and Linguistic Theory* 6:445–501.

Skrabalova, Hana. 2004. *Recherches sur la syntaxe and la sémantique des expressions coordonnées, avec application particulière à la coordination nominale en tchèque*. PhD dissertation. University Paris 3-Sorbonne nouvelle.

--. 2006a. Parataxe and coordination des interrogatifs en tchèque. *Faits de langue* 28:232–241. M.-A. Morel et al. (eds.). *Coordination et subordination: typologie et modélisations*. Paris: Ophrys.

--. 2006b. Two types of wh-questions with conjoined wh-words in Czech. Talk given at the conference *Coordination and ellipsis*. University Paris 7 & LLF (UMR7-CNRS) June 23.

Wilder, Chris. 1997. Some Properties of Ellipsis in Coordination. Alexiadou, A. & T. Hall (eds.). *Studies on Universal Grammar and Typological Variation*: 59–107. Amsterdam: John Benjamins.

Williams, Edwin. 1981. Transformationless Grammar. *Linguistic Inquiry* 17:645–653.

Reflexive verbal forms in Czech from the Romance perspective

Andrea Volencová

A.Volencova@seznam.cz

Reflexive verbal forms with a clitic *se/si* and an indefinite or general meaning are present in all Romance and Slavic languages. They are part of a wider class of reflexive constructions. Representative examples of reflexive verbal forms in Czech and Italian are given in (1) and (2) respectively[1].

(1) a. Tady se často jedí špagety.
here SE often eat.pres.3pl spaghetti.pl
'Here spaghettis are often eaten.'

b. Pracuje se tu příliš mnoho.
work.pres.3sg SE here too much
'One works too much here.'

c. Z Francie se často přijíždí pozdě.
from France SE often arrive.pres.3sg late
'One often arrives late from France.'

(2) a. Qui, gli spaghetti si mangiano spesso.
here spaghetti.pl SE eat.pres.3pl often
'Here spaghettis are often eaten.' (Cinque 1988: 554)

a' Qui, si mangia spesso gli spaghetti.
here SE eat.pres.3sg often spaghetti.pl
'Here one often eats spaghettis.' (Cinque 1988: 554)

b. Si lavora sempre troppo.
SE work.pres.3sg always too much
'One always works too much.' (Cinque 1988: 522)

c. Spesso si arriva in ritardo.
often SE arrive.pres.3sg late

[1]The following abbreviations are used in the glosses: acc = accusative case, instr = instrumental case, sg = singular, pl = plural, msg = masculine singular, nsg = neuter singular, mpl = masculine plural, 3sg = 3rd person singular, 3pl = 3rd person plural, pres = present tense, past = past tense, fut = future tense, inf = infinitive, expl = expletive, SE = reflexive clitic *se/si*

'Often one arrives late.' (Cinque 1988: 522)

d. Non si è mai contenti.
not SE be.pres.3sg never satisfied.mpl
'One is never satisfied.' (Cinque 1988: 522)

e. Si è spesso trattati male.
SE be.pres.3sg often treated badly
'One is often ill-treated.' (Cinque 1988: 522)

As can be seen in (1) and (2), the range of reflexive verbal forms varies cross-linguistically and it is more restricted in Czech than in Italian.

This paper will focus on reflexive verbal forms in French, Italian, Spanish, Romanian and Czech. I will compare reflexive verbal forms of those languages in order to trace their similarities and differences. Then, I will attempt to show what reflexive verbal forms and their description in Romance can tell us about reflexive verbal forms in Czech.

The paper is organized as follows. Section 1 delimits reflexive verbal forms as opposed to other existent reflexive constructions in given languages. Section 2 summarizes cross-linguistic variation of reflexive verbal forms in Romance and Czech. Section 3 mentions some generative analysis proposed for reflexive verbal forms in Romance. Section 4 accounts for Czech reflexive verbal forms from the Romance perspective. The findings are summarized in section 5.

1 What does a reflexive verbal form mean?

Reflexive verbal forms are a subclass of syntactically productive constructions with a reflexive clitic SE[2] which serves various functions in Romance and Slavic. Examples of reflexive constructions which will be discussed in this paper are given in (3) for Czech and in (4) for French.

(3) a. Jan se myje.
Jan SE wash.pres.3sg
'Jan washes (himself).'

b. Váza se rozbila.
vase SE break.past.3sg
'The vase broke.'

c. Tento oblek se pere snadno.
this suit SE wash.pres.3sg easily
'This suit washes easily.'

(4) a. Jean se lave.
'Jean washes (himself).'

b. Le vase s'est cassé.
'The vase broke.'

[2] SE in capitals refers uniformly to different instances of a reflexive clitic in Romance and Czech: *si* in Italian and *se* in French, Spanish, Romanian and Czech.

c. Ce costume se lave facilement.
'This suit washes easily.'

Examples (a) instantiate reflexive verbs expressing coreference between their agent and patient. Sentences in (b) are based on unaccusative verbs and have an inchoative meaning with nonagentive interpretation. Sentences in (c) represent the middle construction which expresses a characteristic of the syntactic subject.

Furthermore, I will have nothing to say here about idiosyncratic reflexive forms. Unlike the reflexive constructions mentionned above, these forms are not productive and vary inexplicably in different languages. Czech and French examples are given in (5) and (6) respectively.

(5) a. Jana omdlela.
Jana faint.past.3sg
'Jana fainted.'

b. Jana se usmála.
Jana SE smile.past.3sg
'Jana smiled.'

(6) a. Jeanne s' est évanouie.
Jeanne SE be.pres.3sg fainted
'Jeanne fainted.'

b. Jeanne a souri.
Jeanne have.pres.3sg smiled
'Jeanne smiled.'

What are then characteristic properties of reflexive verbal forms given in (1) and (2) as compared with the other reflexive constructions in (3) and (4)?

First, an agent is syntactically present in reflexive verbal forms (1) and (2). Therefore, reflexive verbal forms are compatible with agent-oriented adverbs like „deliberately", as shown in (7) for Czech and (8) for French.

(7) Tyto byty se záměrně prodávají obydlené.
these apartment.pl SE deliberately sell.pres.3pl occupied
'These apartments are deliberately sold occupied.'

(8) Ces appartements se vendent délibérément occupés.
these apartment.pl SE sell.pres.3pl deliberately occupied
'These apartments are sold deliberately occupied.'

On the contrary, in middles and unaccusative-inchoative constructions, no agent takes part in the event. Middles predicate a specific property of the syntactic subject irrespective of any agent. Inchoatives express an event with an unspecified cause. In both cases, the agent is not present syntactically. Hence, agent-oriented adverbs are not permitted. Examples from Czech are given in (9) and the same is true in French (10) and Romance in general.

(9) a. Tento oblek se (*záměrně) pere snadno.
this suit.sg SE (*deliberately) wash.pres.3sg easily
'This suit (deliberately) washes easily.'

b. Váza se (*záměrně) rozbila.
vase SE (*deliberately) break.past.3sg
'The vase broke (*deliberately).'

(10) a. Ce costume se lave (*délibérément) facilement.
'This suit washes (*deliberately) easily.'

b. Le vase s'est cassé (*délibérément).
'The vase broke (*deliberately).'

Second, the agent in reflexive verbal forms is either indefinite or generic and distinct from the patient, as shown by the Spanish examples in (11) and (12).

(11) a. Se cantan canciones de protesta en las fiestas.
SE sing.pres.3pl song.pl of protest at the party.pl
'Protest songs are sung at the parties.' (Campos 1989: 2)

(12) a. Se canta canciones de protesta en las fiestas.
SE sing.pres.3sg song.pl of protest at the party.pl
'One sings protest songs at the parties.' (Campos 1989: 2)

b. Se habla espanol aquí.
SE speak.pres.3sg Spanish here
'One speaks Spanish here.' (Campos 1989: 1)

c. Se es feliz en ocasiones.
SE be.pres.3sg happy sometimes
'One is sometimes happy.'

As can be seen in (11) and (12), reflexive verbal forms consist of two subsets of constructions. (11) is an example of the reflexive passive and (12) of the impersonal SE-construction. The reflexive passive form occurs with transitive verbs. It is analogous to the periphrastic passive in that the verb's internal argument moves to the subject position and the verb agrees with it. The impersonal SE-construction is possible with all kinds of verbs. The internal argument, if there is one, stays in its base-position and is assigned accusative case there. The verb has the default form of the 3rd person singular. In the reflexive passive, the agent is indefinite; in the impersonal SE-construction, it has a generic interpretation.

On the contrary, in constructions with reflexive verbs (3a) and (4a) the agent is definite and nondistinct from the patient.

To summarize, the fundamental properties of reflexive verbal forms in Romance and Czech are their productivity, syntactic presence of an agent and its indefinite or generic interpretation. The next section provides comparison of reflexive verbal forms and their syntactic properties in Romance and Czech.

2 The range of reflexive verbal forms in Romance and Czech

2.1 Italian and Spanish

Italian and Spanish manifest the widest range of reflexive verbal forms among Romance languages. The two languages allow transitive, unergative and unaccusative verbs, copulas and even passive verbs to occur in a reflexive verbal form. Examples from Italian were given in (2) above, those from Spanish follow in (13) below.

(13) a. Los apartamentos se alquilan.
the apartment.pl SE rent.pres.3pl
'The apartments are rented.' (Contreras 1973: 86)

a' Se alquila (los) apartamentos.
SE rent.pres.3sg (the) apartment.pl
'One rents the apartments.' (Contreras 1973: 84)

b. Se trabaja mucho aquí.
SE work.pres.3sg a lot here
'One works a lot here.' (Campos 1989: 2)

c. Los domingos se va al parque a pasear.
on Sundays SE go.pres.3sg to the park to stroll
'On sunday one goes to the park for a stroll.' (Campos 1989: 11)

d. Se es humano o no se lo es.
SE be.pres.3sg human.msg or not SE it.acc.sg be.pres.3sg
'One is human or one is not.' (Contreras 1973: 83)

e. Se es explotado por los poderosos.
SE be.pres.3sg exploited.msg by the powerful
'One is exploited by the powerful.' (Campos 1989: 3; cited from Otero 1985)

As can be seen in the example (13 a/a'), Spanish as well as Italian (2 a/a') allows for two different reflexive verbal forms with transitive verbs. The forms differ in the agreement of the verb. In the (a) examples, the verb agrees in number and person with its internal argument that has moved to the subject position. In the (a') examples, the verb has a default agreement in the third person singular and the internal argument stays in its base-position where it is assigned accusative case.

As Campos (1989) extensively argued for, while the (a) sentences are interpreted as passive, the (a') sentences have an active meaning. His test for determining the passive or active meaning of the sentences is shown in (14).

(14) ¿Qué hacen en las fiestas en tu país?
'What do they (people in general) do at parties in your country?'
(Campos 1989: 2)

a. #Se cantan canciones de protesta en las fiestas.

'Protest songs are sung at the parties.'

a' Se canta canciones de protesta en las fiestas.
'One sings protest songs at the parties.'

While (a') can be used as an aswer to an active question, (a) is not acceptable in such a context. On the other hand, (a) is required as an answer to a passive question.

Furthermore, Campos observes that reflexive verbal forms with unergative verbs are also ambiguous between impersonal and what he calls "passive impersonal interpretation" in the present tense. That is, they can be used in an answer to both active and passive questions, as shown in (15).

(15) ¿Qué hacen los sábados en tu país?
'What do people do on Saturdays in your country?' (Campos 1989: 5)
¿Qué se hace los sábados en tu país?
'What is done on Saturdays in your country?'

a/a' En mi país se trabaja los sábados.
'In my country one works (people work) on Saturdays.'

In the preterite, however, only the passive meaning ("passive impersonal interpretation" in his terms) survives and a reflexive verbal form is not an acceptable answer to an active question.

(16) ¿Qué hicieron ayer? (Campos 1989: 6)

a. #Se trabajó todo el día ayer.
SE work.past.3sg all day yesterday
'It was worked the whole day yesterday.'

On the contrary, as he further argues, unaccusative verbs seem not to be ambiguous. In Spanish, they appear only as an active construction in the present tense, but they do not occur in the preterite where only the passive interpretation is supposed to survive.

(17) a. ?*El aňo pasado se fue al parque a pasear.
last year SE go.past.3sg to the park for a stroll
'Last year, one (people) went to the park for a stroll.' (= it was gone) (Campos 1989: 11)

a' Los domingos se va al parque a pasear.
on Sundays SE go.pres.3sg to the park for a stroll
'On Sunday, one (people) goes to the park for a stroll.'

2.2 Romanian

Romanian reflexive verbal forms are more restricted than in Italian and Spanish. Romanian allows them with transitive, unergative and unaccusative verbs but not with copulas and passive verbs as shown in (18).

(18) a. S-au recitat poezii de Eminescu.
SE-have.3pl recited poem.pl by Eminescu

'Poems by Eminescu were recited.' (Dobrovie-Sorin 1998: 405)

b. Se munceşte.
SE work.pres.3sg
'It is worked.' (Dobrovie-Sorin 1998: 405)

c. Ieri s-a ajuns tîrziu acasa.

'Yesterday SE arrived home late' (Cinque 1988: 571)

d. *Nu se este niciodată mulţumit.
not SE be.pres.3sg never satisfied
'One is never satisfied.' (Dobrovie-Sorin 1998: 405)

e. *Adesea se este trădat de prieteni falşi.
frequently SE be.pres.3sg betrayed by friend.pl false
'One is frequently betrayed by false friends.' (Dobrovie-Sorin 1998: 405)

As is shown by Dobrovie-Sorin (1998), unlike in Italian and Spanish, only a passive interpretation is available for reflexive verbal forms in Romanian. I will come back to this observation in more detail in section 4.2.

2.3 French

French demonstrates the narrowest range of reflexive verbal forms among Romance languages. Only transitive verbs but not others can appear in reflexive verbal forms.

(19) a. Ce roman se lira bientôt à Moscou.
this novel SE read.fut.3sg soon in Moscow
'This novel will be read soon in Moscow.' (Zribi-Hertz 1982: 349)

a' Il s'est déjà lu beaucoup de romans dans cette bibliothèque.
it.expl SE be.pres.3sg already read a lot of novel.pl in this library

'A lot of books has been already read in this library.'

b. *Il s'est couru sur ce stade récemment.
it.expl SE be.pres.3sg run at this stadium recently
'One has recently run at this stadium.'

c. *Dans cette grotte il s'entre une fois par an.
into this cave it.expl SE enter.pres.3sg once a year
'One enters this cave once a year.'

d. *Il ne s'est jamais heureux.
it.expl not SE be.pres.3sg never happy
'One is never happy.'

e. *Il s'est souvent exploité par les puissants de ce monde.
it.expl SE be.pres.3sg often exploited by the powerful.pl of this world
'One is often exploited by the powerful of this world.'

As shown in (19a) and (19a'), two types of construction are possible for reflexive verbal forms with a transitive verb in French. In the former, the verb's internal argument moves to the subject position. In the latter, the internal argument stays in its base-position and the overt expletive pronoun *il* occurs as syntactic subject. However, the impersonal construction (19a') is not available for other kinds of verbs in French.

French is the only non-pro-drop Romance language and it seems that independent reasons may be involved in constraining impersonal reflexive verbal forms in it. Belletti (1982) and Cinque (1988) argued that these factors involve Case constraints, while Dobrovie-Sorin (1998) and Guéron (1996) argued that interpretive constraints account for the differences.

2.4 Czech

In comparison to Romance languages, Czech reflexive verbal forms are very similar to those found in Romanian. That is, Czech accepts reflexive verbal forms with transitive, unergative and unaccusative verbs, repeated here as (20 a-c). On the contrary, copulas and passive verbs are not allowed:

(20) a. Tady se často jedí špagety.
here SE often eat.pres.3pl spaghetti.pl
'Here spaghettis are often eaten.'

b. Pracuje se tu příliš mnoho.
work.pres.3sg SE here too much
'One works too much here.'

c. Z Francie se často přijíždí pozdě.
from France SE often arrive.pres.3sg late
'One often arrives late from France.'

d. *Není se nikdy šťasten(-o).
not-be.pres.3sg SE never happy.msg (nsg)
'One is never happy.'

e. *Je se využíván(-o) mocnými tohoto světa.
be.pres.3sg SE exploited.msg (nsg) by the powerful of this world
'One is often exploited by the powerful of this world.'

A more detailed analysis of Czech reflexive verbal forms will be given in section 4.2.

2.5 Summary

The following table summarizes the range of reflexive verbal forms found in Romance and Czech.

The languages divide into three groups. Italian and Spanish show the widest range of reflexive verbal forms which are available for all kinds of verbs. Romanian and Czech are more restricted

Table 1: Table 1

	Italian, Spanish	Romanian, Czech	French
Transitive verbs	√	√	√
Unergative verbs	√	√	*
Unaccusative verbs	√	√	*
Copulas	√	*	*
Passive	√	*	*

and do not allow for reflexive verbal forms with copulas and passive verbs. Finally, the most limited range of reflexive verbal forms is attested in French, which allows only transitive verbs to occur in them.

The question which now arises is how to account for this cross-linguistic variation in a simple and principled way. To do that, I will first review a few generative proposals designed to explain cross-linguistic differences of this kind.

3 Generative analysis of Romance

Reflexive verbal forms in Romance have received a lot of attention in the generative literature: Cinque 1988, Dobrovie-Sorin 1998, Lyons 1995, Raposo-Uriagereka 1996, Reinhart – Siloni 2005, Wehrli 1986 and Zribi-Hertz 1982 to mention only some of them.

Cinque's analysis (1988) of SE-constructions in Italian and Romance counts among the most influential ones and it has given rise to a lot of discussion. This section reviews his analysis briefly and shows how it works for Romance languages.

Cinque observes that Italian *si* can appear in all tensed and even in some untensed clauses. Nonetheless, he shows that its distribution in untensed contexts is restricted only to transitive and unergative verbs in raising constructions (21). In control clauses (22), it is uniformly excluded.

(21) a. Sembra non essersi ancora scoperto il vero colpevole.
'It seems one not to have yet discovered the true culprit.' (Cinque 1988: 524)

b. Sembra non essersi lavorato a sufficienza.
'It seems one not to have worked sufficiently.'

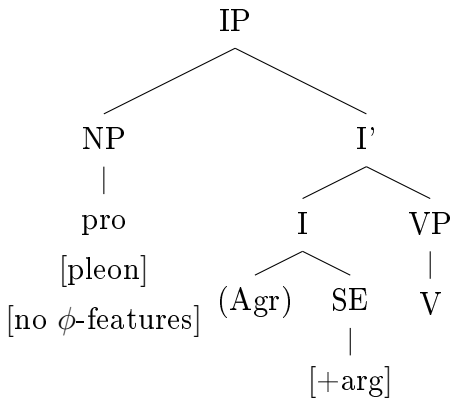
(22) *Sarebbe meglio scoprirsi il colpevole.
'It would be better one to discover the culprit.' (Cinque 1988: 522)

On the basis of the data, Cinque concludes that two kinds of SE are available in Italian: an argument [+arg] SE and a nonargument [-arg] SE. While [+arg] SE is licensed by verbs with an external argument, [-arg] SE is licensed by a finite inflection. Therefore, while a [+arg] clitic can appear only with transitive and unergative verbs, a [-arg] one is limited to finite contexts:

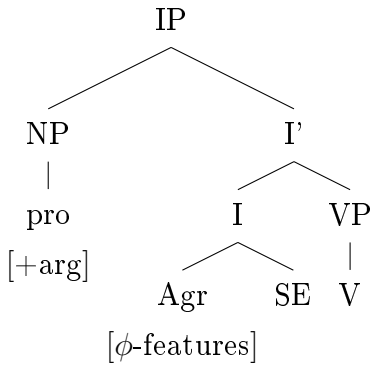
> "In finite contexts [-arg] *si* is possible (with all verb types) because it can 'amalgamate' with personal Agr. [+arg] *si*, on the other hand, is possible only with those verb types that assign an external θ-role." (Cinque 1988: 531)

Cinque's structure configurations for [+arg] and [-arg] SE are given in (23a) and (23b) respectively.

(23) a.



b.



The argument clitic (23a) has to be associated with an external θ-role at every level of representation. It is coindexed with a *pro* in [NP, IP] which is merely a pleonastic pronoun with no φ-features. On the contrary, the nonargument clitic (23b) combines with a finite Agr as a syntactic marker for unspecified (generic) Person, while an argument *pro* discharges the verb's θ-role.

Since the [+arg] clitic is licensed by verbs assigning an external θ-role and the [-arg] clitic by finite inflection, it follows that the clitic is ambiguous between [+arg] and [-arg] in finite clauses with a transitive or an unergative verb. Cinque takes advantage of this ambiguity in the case of Italian reflexive forms with transitive verbs. As has been shown above, transitive verbs allow either for an agreeing or a non-agreeing reflexive form, repeated in (24 a) and (24 a') respectively.

(24) a. Qui, gli spaghetti si mangiano spesso.
here spaghetti.pl SE eat.pres.3pl often
'Here spaghettis are often eaten.'

a' Qui, si mangia spesso gli spaghetti.
here SE eat.pres.3sg often spaghetti.pl
'Here one often eats spaghettis.'

Cinque claims that it is possible to explain the existence of these two constructions by the fact that Italian possesses both [+arg] and [-arg] SE.

In his view, the reflexive passive construction (24a) contains the [+arg] clitic which is associated with the verb's external θ-role. The internal argument is not assigned accusative, so it

has to move to the position of the syntactic subject to receive Case. The verb then agrees with the derived syntactic subject.

On the other hand, the impersonal SE-construction (24a') contains the [-arg] clitic and the verb's external θ-role is assigned to an argument *pro*. Consequently, according to Burzio's generalization, the internal argument of the verb can be assigned accusative case. The verb is in the default form of the 3rd person singular because its subject *pro* is specified for generic person by the nonargument [-arg] clitic.

If finite verbs with an external argument accept both [+arg] and [-arg] clitics, then the ambiguous interpretation of transitive verbs as well as unergative ones in Italian and Spanish is exactly what one expects. Hence, the differences in the interpretation noted by Campos (1989) which were summed up in section 2.1. can be accounted for in Cinque's terms.

As Cinque shows, it is tempting to set the given [+arg] / [-arg] distinction as a parameter and thus to account for cross-linguistic variation. He argues on the example of Romance languages that Italian, Spanish and Portuguese possess both kinds of SE, while French and Romanian have only [+arg] SE. Therefore, whereas in the former languages SE occurs in all finite and some nonfinite contexts, in French and Romanian its use is more restricted.

Cinque's parameter-based analysis of Romance has been further developed by Dobrovie-Sorin (1998) for Romanian. She confirms that the nature of Romanian SE can be best captured by setting the parameter as [+arg]. Then Romanian reflexive forms with transitive and unergative verbs are then accounted for. What remains to be solved are reflexive verbal forms with unaccusative verbs (18c) which are also available in Romanian. Dobrovie-Sorin proposes that these forms are re-analysed as unergatives in Romanian. I will return to her proposal in section 4.2. below.

It has been shown in section 2 that the range of reflexive verbal forms is parallel in Czech and Romanian. Could then such a parameter proposed by Cinque (1988) and Dobrovie-Sorin (1998) account in a simple way for reflexive verbal forms in Czech as well?

4 Generative analysis of Czech

There exist a few generative accounts for reflexive verbal forms in Czech as well, for example Rivero 2002 and Růžička 1992. The next section summarizes Růžička's analysis of Slavic, including Czech.

4.1 [-arg] analysis of Czech

Following Cinque's analysis (1988) of Romance languages, Růžička (1992) attempts to deal with reflexive verbal forms in Slavic. First, he observes that Slavic does not allow reflexive verbal forms to occur in untensed contexts. Therefore, he proposes to set the Cinque's parameter as [-arg] for Slavic.

Second, he distinguishes two kinds of reflexive verbal forms, namely a reflexive passive with transitive verbs (25a) and an impersonal SE-construction with other kinds of verbs (25b).

(25) a. Okno / malý chlapec se myje mýdlem.
window / little boy SE wash.pres.3sg soap.instr
'The window / The little boy is (being) washed with soap.' (adapted from Růžička 1992: 153)

b. Tancovalo se až do rána.
danse.past.3sg.n SE until to morning
'One danced into the morning.'

For Růžička, a reflexive passive is merely a contextual interpretation of certain reflexive verbs. His conclusion is based on the observation that the reflexive passive is formed only with verbs that exist also as reflexive verbs. Therefore, the reflexive passive is a contextually determined interpretation allowed by the inherent meaning of these reflexive verbs, and it is not structurally different from them:

> "Indeed, hardly a reflexive form in a passive configuration is found to occur which cannot be a lexically reflexive verb 'otherwise'. (...) Clearly, the pertinent verbs are susceptible and accessible to passive interpretation by virtue of their inherent meaning. (...) I conclude that reflexive passive is a contextual interpretation of certain classes of lexically reflexive verbs." (Růžička 1992: 153 – 155)

It means that example (25a) can be interpreted either as passive or as containing a reflexive verb. If there is an inanimate noun in the subject position, the sentence is usually interpreted as having a passive meaning. On the contrary, if the subject is animate, the sentence allows for both interpretations. Hence, it can be either the little boy who washes himself or someone else who washes him. Because of their ambiguous meaning such sentences should be avoided, as noted by Czech grammars (cf. Karlík et al. 1995).

On the other hand, Růžička claims that unlike the reflexive passive, impersonal SE-constructions are structurally specific and have the structure (23b) proposed by Cinque. However, due to the absence of the [+arg] clitic in Slavic, they are uniformly excluded in all nonfinite contexts, namely in control infinitives (26a) and raising structures (26b).

(26) a. *Bylo by lépe jít se domů teď.
'It would be better (for one) to go home now.' (Růžička 1992: 143)

b. *Otázky struktury věty se zdají se projednávat na zasedání komise.
'Questions of the structure of the sentence seem to be dealt with on the committee's meeting.' (Růžička 1992: 144)

In sum, Růžička argues that the reflexive passive in Czech is a context-dependent interpretation of reflexive verbs, whereas the impersonal SE-construction is a specific structure based on the [-arg] value of the reflexive clitic.

However, this conclusion does not explain why the range of reflexive verbal forms in Slavic varies in different languages nor notably why reflexive verbal forms with copulas and passive verbs are not available in Czech, as noted above.

4.2 [+arg] analysis of Czech

As has been shown in section 2, Czech reflexive verbal forms are parallel to those found in Romanian. The two languages allow for reflexive verbal forms with transitive, unergative and unaccusative verbs, while excluding copulas and passive verbs from them.

I mentioned in the previous section that an analysis of Romanian data based on Cinque's [±arg] parameter has been proposed by Dobrovie-Sorin (1998). She claims, according to Cinque, that the parameter is set as [+arg] for Romanian. It follows that Romanian has only the reflexive passive and no impersonal SE-constructions at all. Given the correspondence of Romanian and

Czech data, would it be possible to apply this analysis to Czech, as well? To put it differently, does it mean that Czech, which is similar to Romanian, has no impersonal SE-constructions, as opposed to Růžička's account?

The assumption that Czech SE is [+arg] would account correctly for Czech reflexive verbal forms with transitive and unergative verbs, repeated here in (27).

(27) a. Tady se často jedí špagety.
here SE often eat.pres.3pl spaghetti.pl
'Here spaghettis are often eaten.'

b. Pracuje se tu příliš mnoho.
work.pres.3sg SE here too much
'It is worked too much here.'

Moreover, such an analysis explains why only agreeing constructions with transitive verbs are available in Czech.

(28) a. Ty byty se prodávají od ledna.
the flat.pl SE sell.pres.3pl since January
'The flats have been selling since January.'

a' *Ty byty se prodává od ledna.
the flat.pl SE sell.pres.3sg since January
'One has been selling the flats since January.'

If Czech is supposed to have a [+arg] clitic, its reflexive verbal forms should be interpreted as passive. This prediction is borne out as shown in the following examples by Campos' test.

(29) a. #Co (lidi) prodávají od ledna?
'What have they (people) been selling since January?'

b. Co se tu prodává (je tu prodáváno) od ledna?
'What has been selling since January?'

Prodávají se ty byty.
sell.pres.3pl SE the flat.pl
'The flats have been selling.'

(30) a. #Jak mnoho tady (lidi) pracují?
'How much do they (people) work here?'
Pracuje se tu příliš mnoho.
'It is worked too much here.'

Although differences are subtle, reflexive forms seem to be more acceptable as an answer to a passive rather than an active question. The same holds for Romanian, as pointed out by Dobrovie-Sorin (1998):

(31) a. Se cîntă / doarme / munceşte / mănîncă.
SE sing.pres.3sg / sleep.pres.3sg / work.pres.3sg / eat.pres.3sg
'It is sung / slept / worked / eaten.' (Dobrovie-Sorin 1998: 405)

b. Si canta / dorme / lavora / mangia.
SE sing.pres.3sg / sleep.pres.3sg / work.pres.3sg / eat.pres.3sg
'One sings / sleeps / works / eats.' (Dobrovie-Sorin 1998: 406)

While the Italian examples in (31b) are interpreted as active and can be potentially ambiguous between active and passive meaning, as shown by Campos (1989) for the Spanish examples, Romanian reflexive forms are unambiguously passive, as indicated by the English gloss.

Moroever, the same meaning survives in the past tense, more evidence that reflexive forms in Czech are passive rather than impersonal.

(32) a. Tady se často jedly špagety.
here SE often eat.past.3pl spaghetti.pl
'Here spaghettis were often eaten.'

b. Pracovalo se tu příliš mnoho.
work.past.3sg.n SE here too much
'It was worked too much here.'

The [+arg] status of SE in Romanian and Czech explains its incapacity to act as a syntactic controller and binder. As shown in (33) below, [+arg] SE with an expletive *pro* in the subject position cannot count as a syntactic controller of the embedded subject PRO in Romanian and Czech. It thus contrasts with Spanish (34) and Italian as well.

(33) a. *S-a promis a respecta dispoziţiile.
SE-has promised to obey the instruction.pl
(Dobrovie-Sorin 1998: 427)

b. *Přislíbilo se poslouchat pokyny.
promise.past.3sg.n SE obey.inf instruction.pl
'It was promised to obey the instrutions.'

(34) Se espera llegar al final del camino.
SE hopes arrive.inf at the end of the road
'One hopes to arrive at the end of the road.' (Dobrovie-Sorin 1998: 426)

Similarly, the ability of SE to function as an anaphor binder in Czech seems to be questionable, at least for plural anaphors:

(35) a. *Mluvilo se tam jeden o druhém.
speak.past.3sg.n SE there one about another
'It was talked there about each other.' (Růžička 1992: 152)

b. ?*Mluvilo se tam jen o sobě.
speak.past.3sg.n SE there only about oneself
'It was talked there only about oneself.' (Růžička 1992: 152)

However, two problems remain under the [+arg] analysis of Czech reflexive verbal forms. First, the proposed analysis predicts the existence of raising constructions with reflexive verbal forms of transitive and unergative verbs. But in Czech this expectation is not fulfilled, as has been shown in (26).

Nevertheless, Růžička argues that raising constructions with reflexive verbal forms in Czech

may be ruled out for independent reasons. The Czech raising verb *zdát se* is itself a reflexive verb which combines with an infinitival clause from which an element is raised to the matrix clause. As already pointed out by Šmilauer (1947), Czech does not allow structures where one reflexive verb or form is dependent on another one. This is true not only for reflexive forms (26b) dependent on the raising verb *zdát se* but also for reflexive verbs as in (36).

(36) *Lidé se zdáli navzájem se setkávat.
people SE seem.past.3pl each other SE meet.inf
'People seemed each other to meet.' (Růžička 1992: 144)

Given the ban on doubled SE in Czech clauses and the non-existence of non-reflexive raising verbs in Czech, it is difficult to test the possibility of nonfinite reflexive verbal forms with raising verbs in Czech. Their existence should be in principle allowed by the [+arg] parameter but they might be ruled out for independent reasons.

Second, the proposed analysis in terms of [+arg] SE does not explain the existence of the reflexive verbal forms with unaccusative verbs attested in Romanian as well as in Czech, repeated here.

(37) a. Ieri s-a ajuns tîrziu acasa. (Cinque 1988: 571)

b. Včera se domů přijelo pozdě.
'Yesterday SE arrived home late."

Since unaccusative verbs have no external argument, reflexive verbal form with a [+arg] clitic, which has to be associated with the external θ-role, should not be available for them. Dobrovie-Sorin (1998) proposes that Romanian reflexive verbal forms with unaccusative verbs re-analyze as unergatives. This could be supported by Růžička's (1992) observation that reflexive verbal forms with unaccusative verbs are not licensed if the form does not already exist with unergative verbs:

> "In other words, if complementless structures containing an external argument are not available, no structures without an external argument are licensed." (Růžička 1992: 141)

A more detailed analysis of the nature of refexive verbal structures and behavior of reflexive clitics in Romanian and Czech will be needed to decide if this account is on the right track. In any case, reflexive verbal forms with unaccusatives seem to be the only remaining problem for the proposed [+arg] analysis.

5 Summary

In this paper, I have reviewed a sample of reflexive verbal forms in Romance languages and Czech. The inspected languages split up into two different groups as shown in Table 2.

The first group including Italian and Spanish allows for a wide range of reflexive verbal forms with all kinds of verbs. In all reflexive verbal forms, an active interpretation with a general agent (meaning people in general) is available under certain conditions. As for transitive verbs, the meaning of their reflexive forms depends on the verb's form. If the verb has a default 3rd person singular agreement and its internal argument stays in its base-position, the construction has an active meaning. If a verb agrees with its internal argument moved to the position of the

Table 2: Table 2

	Italian, Spanish, (Portuguese)	***Romanian, Czech, (French)***
Reflexive passive	√	√
Impersonal SE-construction	√	*

syntactic subject, the construction is interpreted as passive. As for unergative verbs, they have an ambiguous interpretation in the present tense and only a passive meaning in the past tense.

Properties of Italian and Spanish can be captured by Cinque's assumption that they have both [+arg] and [-arg] reflexive clitics. While the [+arg] clitic has to be associated with an external argument, the [-arg] clitic is limited to finite contexts. Whereas [+arg] SE accounts for reflexive passives, [-arg] SE accounts for the impersonal SE-construction. Since both clitics overlap in finite constructions with transitive and unergative verbs, exactly these reflexive verbal forms show structural and interpretive ambiguity.

The second group includes Romanian, Czech and French. Romanian and Czech have a narrower range of reflexive verbal forms, excluding those with copulas and passive verbs. French is even more restricted. As the only non-pro-drop language, it appears to be subject to other language-specific constrains. These three languages do not have impersonal SE-constructions, and all their reflexive verbal forms have a passive meaning.

It seems that what the languages in this group have in common is the lack of a [-arg] clitic as argued by Cinque (1988) and Dobrovie-Sorin (1998) for French and Romanian. It has been shown in section 4.2. that this account can be extented to Czech, as well. If these languages are supposed to have only [+arg] SE, passive interpretation of their reflexive verbal forms as well as a bundle of other properties they have in common can be accounted for. Minor problems that remain unsolved, especially the status of reflexive verbal forms with unaccusative verbs in Czech, will be a focus for future research.

References

Belletti, Adriana. 1982. "Morphological" Passive and Pro-Drop: The Impersonal Construction in Italian. *Journal of Linguistic Research* 2:1–34.

Campos, Héctor. 1989. Impersonal Passive „se“ in Spanish. *Linguisticae Investigationes* XIII:1–21.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, Mass.: The MIT Press.

Cinque, Guglielmo. 1988. On *Si* Constructions and the Theory of Arb. *Linguistic Inquiry* 19:521–581.

Contreras, Heles. 1973. Grammaticality Versus Acceptability: The Spanish Se Case. *Linguistic Inquiry* 4:83–88.

Dobrovie-Sorin, Carmen. 1998. Impersonal *se* Constructions in Romance and the Passivization of Unergatives. *Linguistic Inquiry* 29:399–437.

Faltz, Leonard M. 1985. *Reflexivization – A Study in Universal Syntax.* New York: Garland Publishing.

Fried, Mirjam. In press. Agent back-grounding as a functional domain: reflexivization and passivization in Czech and Russian. Solstad, T. & B. Lyngfelt (eds.). *Agent demotion.* Amsterdam & Philadelphia: John Benjamins.

Geniušiene, Emma. 1987. *The Typology of Reflexives.* Berlin: Mouton de Gruyter.

Grimshaw, Jane. 1992. *Argument Structure.* Cambridge, Mass.: The MIT Press.

Guéron, Jacqueline. 1996. Qu'est-ce qu'une phrase impersonnelle: remarques sur le rôle du clitique se dans les langues romanes. *Recherches linguistiques de Vincennes* 25:53–82.

Halle, Morris and Alec Marantz. 1994. Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics* 21:275–288.

Havránek, Bohuslav. 1928. *Genera verbi v slovanských jazycích.* Praha: Kr. česká spol. nauk.

Karlík, Petr et al. 1995. *Příruční mluvnice češtiny.* Praha: Nakladatelství Lidové noviny.

Keyser, Samuel Jay and Thomas Roeper. 1984. On the Middle and Ergative Constructions in English. *Linguistic Inquiry* 15:381–416.

Kopečný, František. 1954. Passivum, reflexivní forma slovesná a reflexivní sloveso. *Studie a práce lingvistické* 1:224–247. Praha: Nakladatelství ČSAV.

Kosta, Petr and Jens Frasek. 2004. Neakuzativita (ergativita) vs. Neergativita v češtině, polštině a jiných slovanských jazycích na rozhraní morfologie a syntaxe. Hladká, Z. & P. Karlík (eds.). *Čeština – univerzália a specifika* 5:172–194. Praha: Nakladatelství Lidové noviny.

Lidz, Jeffrey. 2001. The Argument Structure of Verbal Reflexives. *Natural Language and Linguistic Theory* 19:311–353.

Lyons, Christopher. 1995. Voice, Aspect, and Arbitrary Arguments. Smith, Ch. & M. Maiden (eds.). *Linguistic Theory and the Romance Languages. Current Issues in Linguistic Theory* 122:77–114. Amsterdam / Philadelphia: John Benjamin Publishing Company.

Marantz, Alec. 1984. *On the nature of grammatical relations.* Cambridge, Mass.: MIT Press.

--. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. Dimitriadis, A. et al. (eds.). *Proceedings of the 21st Penn Linguistics Colloquium. UPenn Working Papers in Linguistics*: 201–225. Philadelphia.

Oliva, Karel. 2001. Reflexe reflexivity reflexiv. *Slovo a slovesnost* 62:200–207.

Petr, Jan et al. 1987. *Mluvnice češtiny 3.* Praha: Academia.

Raposo, Eduardo and Juan Uriagereka. 1996. Indefinite Se. *Natural Language and Linguistic Theory* 14:749–810.

Reinhart, Tanya and Tal Siloni. 2004. Against the Unaccusative Analysis of Reflexives. Alexiadou, A. et al. (eds.). *The unaccusativity puzzle: Studies on the syntax-lexicon interface*: 159–180. Oxford: Oxford University Press.

--. 2005. The Lexicon-Syntax Parameter: reflexivization and other arity operations. *Linguistic Inquiry* 36:389–436.

Rivero, María-Luisa. 2002. On Impersonal Reflexives in Romance and Slavic and Semantic Variation. Camps, J. & C. R. Wiltshire (eds.). *Romance Syntax, Semantics and L2 Acquisition*: 169–195. Amsterdam: Benjamins.

Rizzi, Luigi. 1986. Null Objects in Italian and the Theory of pro. *Linguistic Inquiry* 17:501–557.

Růžička, Rudolf. 1992. Slavic and Italian impersonal constructions with reflexive clitics. Zimmermann, I. & A. Strigin (eds.). *Fügungspotenzen: Zum 60. Geburtstag von Manfred Bierwisch*: 133–161. Berlin: Akademie Verlag.

Šmilauer, Vladimír. 1947. *Novočeská skladba. První část nové mluvnice moderního českého jazyka.* Praha: Knižnice KPČJ.

Veselovská, Ludmila and Petr Karlík. 2004. Analytic Passives in Czech. *Zeitschrift für Slawistik* 49:163–236.

Wehrli, Eric. 1986. On some properties of French clitic se. Borer, H. (ed.). *The Syntax of Pronominal Clitic. Syntax and Semantics 19*: 269-284. New York: Academic Press.

Zribi-Hertz, Anne. 1982. La construction „se-moyen“ du français et son statut dans le triangle: moyen – passif – réfléchi. *Lingvisticae Investigationes* VI:345–401.

Why Czech case markers sometimes get lost

Markéta Ziková

zikova@phil.muni.cz

1 Introduction

The aim of this paper is twofold: (i) to elaborate a plausible explanation of the misbehavior of a Czech declensional paradigm, (ii) to provide independent morphological evidence for empty Nuclei and lexically floating segments, the phonological objects introduced in the standard Government Phonology and developed in the CVCV model of Scheer (2004).

Among Czech nominal declensional paradigms, the *dělání* paradigm is peculiar - it shows only consonant-initial markers. I claim that its pecularity follows not from its morphology, this paradigm has the same morphological structure as other paradigms, but from the phonology. More specifically, from the phonological representations of the case markers and nominal stems belonging to this paradigm.

I assume that short vowels at the beginning of case suffixes are lexically floating segments which, in order to be spelled out, need to associate to the stem-final Nucleus. They can do it only when this Nucleus is empty, otherwise they remain floating, i.e. phonetically unrealized. Because only consonant-final, but not vowel-final stems are followed by empty Nuclei, only the former, but not the latter display vowel-initial markers. I argue that the vowel *í* which appears in the *dělání* paradigm is a stem building suffix instead of part of case markers. I claim that it is associated with the stem-final Nucleus hence prevents initial vowels of the case suffixes from surfacing.

2 When vowel–initial markers are absent

Czech nominal declensions are traditionally organised into fourteen paradigms, each of which is conceived of as a set of case markers which combine with gender-specified stems. Among them the neuter paradigm *dělání* "making" is outstanding because it shows only consonant-initial markers. Its surface forms are shown in the following table.

(1) Paradigm *dělání*

	Singular	Plural
Nom	dělání	dělání
Acc	dělání	dělání
Gen	dělání	dělání
Dat	dělání	dělání-m
Loc	dělání	dělání-ch
Ins	dělání-m	dělání-mi

In all other paradigms, case markers which contain consonants are always compound: the consonant is preceded by a vowel, either short or long. Furthermore, each of all other paradigms shows no less than three different markers consisting of single vowels. By way of illustration, in table (2), I show the structure of two other neuter paradigms: *město* "city" and *moře* "sea".

(2) Neuter paradigms

	město "city"		moře "sea"	
	Singular	Plural	Singular	Plural
Nom/Acc	měst-o	měst-a	moř-e	moř-e
Gen	měst-a	měst-∅	moř-e	moř-í
Dat	měst-u	měst-ům	moř-i	moř-ím
Loc	měst-u/ě	měst-ech	moř-i	moř-ích
Ins	měst-em	měst-y	moř-em	moř-i

Comparing these two tables this question arises: Why are initial vowels absent from case markers in the *dělání* paradigm? In the following section, I argue that the vowel *í* which appears across all paradigm cells in (1) is not a case marker itself, but a stem-building suffix.

3 Part of the stem? Part of the case marker?

In Czech reference grammars, the *í* which appears in the *dělání* paradigm is traditionally assumed to be part of the case markers. In that case, the paradigm has the structure as described in (3).

(3) Paradigm *dělání*: *í* is part of the case markers

	Singular	Plural
Nom	dělán-í	dělán-í
Acc	dělán-í	dělán-í
Gen	dělán-í	dělán-í
Dat	dělán-í	dělán-ím
Loc	dělán-í	dělán-ích
Ins	dělán-ím	dělán-ími

There exist several arguments against this traditional analysis, two of which are related to syncretism. I assume that case syncretism (i.e. homophony of case markers) is not accidental, but arises from the sharing of the morphosyntactic features among the syncretised cases.

A systematic approach to case syncretism was first introduced in Jakobson (1936) and then formalised within many frameworks (for an overview see Matthew Baerman's annotated syncretism bibliography available at http://www.surrey.ac.uk/LIS/MB/Bibliography). In this paper, I adopt a Distributed Morphology view that case syncretism follows from the lexical underspecification of given case markers (for details see Caha & Ziková (2006) where we propose an analysis of Czech nominal declension couched within the framework of Distributed Morphology).

If the *-í* were a case marker, the *dělání* paradigm would show massive syncretism which is unprecedented in other paradigms. In this case twelve paradigm slots (i.e. six cases times two numbers) receive only four phonologically different markers: *-í*, *-ím*, *-ích*, *-ími*. Also, the syncretism predicted by (3) is highly suspicious: table (3) assumes that the locative and the nominative singular are syncretic as well as the genitive and the nominative plural. Table (4) illustrates why these types of syncretism are suspicious: they are not attested anywhere else in the declensions, neither nominal nor adjectival.[1]

(4) Other declensions

	"castle" [Masc]	"machine" [Masc]	"woman" [Fem]	"city" [Neu]	"sea" [Neu]	"young, adj." [Fem]
NomSg	hrad-∅	stroj-∅	žen-a	měst-o	moř-e	mlad-á
LocSg	hrad-u/ě	stroj-i	žen-ě	měst-u/ě	moř-i	mlad-é
NomPl	hrad-y	stroj-e	žen-y	měst-a	moř-e	mlad-é
GenPl	hrad-ů	stroj-ů	žen-∅	měst-∅	moř-í	mlad-ých

Now if the vowel *í* is not part of the case markers, it must belong to the stem. Furthermore, I claim that it is not only part of the stem, but also a stem-building suffix. As shown in (5), it either serves as a PP- or V-nominalizer, or derives collective nouns.

(5) a. pří - moř - í "seaside" (PP-nominalization)
by - √sea - noun
b. děl - á - n - í "making" (V-nominalization)
√make - th - part - noun
c. lan - ov - í "ropes" (collective)
√rope - adj - noun

Examples shown in (5) indicate what is the contribution of this suffix to the structure of the stem. First, the merger of it always produces stems of the neuter gender (there are no masculine or feminine stems formed by this suffix); compare e.g. *loket* "elbow, masc." and *před-lokt-í* "forearm, neu.", *zem* "ground, fem." and *pod-zem-í* "under-ground, neu.", or *kámen* "stone, masc." and *kamen-í* "stones, neu.".[2]

[1] This holds true not only for Czech, but for Slavic languages in general; see e.g. Müller (2004).

[2] Of course, there exists a handful of feminine and masculine stems ending in a high front long vowel, e.g. *rallye* [rƐli:] "rally, fem." or *mufti* "Mufti, masc.". However, all of them are loanwords and the final [i:] is not a suffix, but an inseparable part of the stem. As a consequence of this, none of these stems show case forms observable in the *dělání* paradigm: masculine animate stems of this type feature an adjectival declension

Second, neuter stems derived by this suffix are either count (PP-nominalizations), or mass (V-nominalizations and collectives). If we adopt a Distributed Morphology view that affixes are just phonological forms which realize particular portions of the syntactic structure, the stem-building *í* would be lexically specified to realize that part of the syntactic structure which is a projection having its own gender and number features.

Does the fact that the *-í* is a stem-building suffix mean that the *dělání* paradigm has the structure in (6)? Certainly not because (6) produces the same strange types of the syncretism as mentioned under (3) above. In both cases, eight out of twelve paradigm cells are occupied by the suffixes of the same form. The only difference lies in that in (3) they are positive, but in (6) they are zero markers.

(6) Paradigm *dělání*: zero markers

	Singular	Plural
Nom	dělání-∅	dělání-∅
Acc	dělání-∅	dělání-∅
Gen	dělání-∅	dělání-∅
Dat	dělání-∅	dělání-m
Loc	dělání-∅	dělání-ch
Ins	dělání-m	dělání-mi

4 Underlying structure is regular

What is my proposal how to solve the syncretism problem described above? I claim that the *dělání* paradigm has the same morphology as all other paradigms. Its underlying structure is in (7): all paradigm cells are occupied by vowel-initial markers as paradigm cells in all other nominal paradigms normally are.[3]

What makes this paradigm special is the fact that its stems are vowel-final while all other paradigms are built on consonant-final stems: it is derived by the same phonology as all other paradigms, but the phonology prevents the initial vowels of the case markers from surfacing.

(7) Paradigm *dělání*: vowel-initial markers are underlyingly present

(compare *muftího* "Mufti, GenSg" and *severního* "northern, adj., GenSg", *muftímu* "Mufti, DatSg" and *severnímu* "northern, adj., DatSg", and so on), stems of all other genders, including neuter stems like *karí* "curry" or *čatný* "chutney", do not display case markers at all. Indeclinable loanwords are further discussed in section 7.

[3] Also in comparison with the structure of other paradigms, zero markers may be assumed either for the nominative and accusative singular or for the genitive plural. I leave this possibility open.

	Singular	Plural
Nom	dělání-V	dělání-V
Acc	dělání-V	dělání-V
Gen	dělání-V	dělání-V
Dat	dělání-V	dělání-m
Loc	dělání-V	dělání-ch
Ins	dělání-m	dělání-mi

Having no positive (i.e. phonetic) evidence what the vowels in (7) look like, I use the symbol V for them. Of course, one can object that the structure in (7) raises the same syncretism problems as structures in (3) or (6). That would be true if the symbol V represents the same vowel. But provided that we accept a hypothesis that the *dělání* paradigm is not special, at least the locative and the nominative singular marker and the genitive and the nominative plural marker as well should be represented by phonologically different vowels.

5 CVCV

In what follows, I argue that the best way to get the surface structures of the case forms shown in (1) from their hypothetical underlying structures in (7) is by using the theory of CVCV phonology, abbreviated hereafter as CVCV.

In CVCV (Lowenstamm 1996, Scheer 2004), phonological structure is represented on two separate levels whose items are connected through association lines. The syllabic level consists of a strict sequence of non-branching Onsets (i.e. consonantal constituents, C) and non-branching Nuclei (i.e. vocalic constituents, V), hence CVCV. The segmental level is built from phonological expressions which are considered to have a hierarchical structure as well. In this respect, two things are important: (i) CV units are the minimal building blocks (the existence of C implies the existence of V, and *vice versa*), (ii) only segments linked to syllabic constituents are phonetically realized. Furthermore, CVCV assumes that the syllable structure is recorded in the lexicon, and then projected into the derivation. This is a phonological version of the syntactic Projection Principle.[4]

Two main consequences follow from these assumptions:

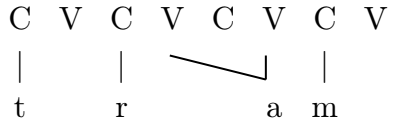
1. The parts of all consonant clusters are separated by empty Nuclei.

2. In the lexicon, the syllable structure of all morphemes starts with an Onset (empty or full) and ends in a Nucleus (empty or full).

To illustrate these consequences, under (8), I give the lexical representation of the root √trám "beam". It consists of eight constituents, three of which are empty. It contains two empty Nuclei, a final empty Nucleus and an empty Nucleus separating two morpheme-internal consonants, and one empty Onset which is enclosed within a long vowel, i.e. within two Nuclei linked simultaneously to one segment.

[4]The phonological Projection Principle was originally formulated within the Standard Government Phonology; see e.g. Kaye et al. (1990).

Representation of the root √trám

(8)


Scheme (8) indicates that mapping between the syllabic structure and the segmental level need not be one to one: there exist syllabic constituents without segments and segments linked to multiple constituents. In the following section, I explore an other case of such non-isomorphism, namely the case where segments that lack any syllabic support.

6 Stem–final Nuclei

Having outlined the core principles of CVCV, let us turn our attention to the phonological structure of case markers. From what has been said it follows that there are two types of them: zero markers and positive markers, all of them beginning with vowels. In this section, I submit two independent arguments that marker-initial vowels are lexically specified to associate to the stem-final Nucleus. The first argument is based on an $e \sim \emptyset$ alternation; the second argument is grounded in the fact that in the *dělání* paradigm, no marker-initial vowels surface.

6.1 Vowel–zero alternations

Czech (as other Slavic languages) features vowel-zero alternations. In CVCV, vowels alternating with zero are lexically floating segments. (It should be noted that in Czech, only a mid front vowel alternates.) The assumption that Nuclei which host vowel-zero alternations are already present in the lexical representation follows from the Projection Principle mentioned above. By way of illustration, I show the lexical representation of the root √kotøl "boiler", which contains an alternation site between *t* and *l*. (This alternation site is marked by ø.)

Representation of the root √kotøl

(9)


Outside CVCV, V $\sim$ ø alternations are examined in terms of epenthesis (the alternating vowel is not lexically present) or deletion (the vowel disappears during the derivation). If alternating vowels are either inserted or deleted by rule; their distribution should be predictable.

However, what is predictable is the distribution of alternants, but not the distribution of the alternation sites themselves. This can be illustrated by three roots: √kotøl "boiler", √trotl "prune", and √hotel "hotel". If we adopt an epenthetic scenario, the root √trotl should behave in the same way as the root √kotøl because in the underlying structure they both end in a *tl* cluster. In fact, they do not behave alike: in the context of a zero marker, the root √kotøl shows the vowel *e*, but the root √trotl does not; compare *kotel-*Ø "boiler, NomSg" and *trotl-*Ø "prune, NomSg". From this it follows that information about epenthesis must be somehow encoded in the lexical representation. Adopting the deletion scenario raises the same problem. Even though the roots √kotøl and √trotl will differ lexically, additional information about the alternating vowel is still needed to capture the difference between the root √kotøl

whose *e* undergoes deletion and the root √hotel whose *e* does not; compare *kotel-∅* "boiler, NomSg" and *kotl-e* "boiler, NomPl" with *hotel-∅* "hotel, NomSg" and *hotel-y* "hotel, NomPl". This behaviour pleads in favour of the analysis proposed by CVCV: alternating vowels are lexically floating segments whose phonetic realization depends on Government.

Government is a regressive relation that holds between the syllabic constituents: Nuclei govern either other Nuclei, or their own Onsets. What is important is that only those Nuclei which are not governed display Government. In case that there are two full Nuclei in a row, the second one always governs the closest constituent, i.e. its own Onset. From this it follows that full Nuclei are never governed, hence always govern: whenever a Nucleus with a lexically floating vowel is followed by a full Nucleus, it is governed. As for morpheme-final empty Nuclei and their effect on vowel-zero alternations, they are governed (due to the morphology), therefore do not govern. In that case, an association line between the floating segment and its Nucleus is created.

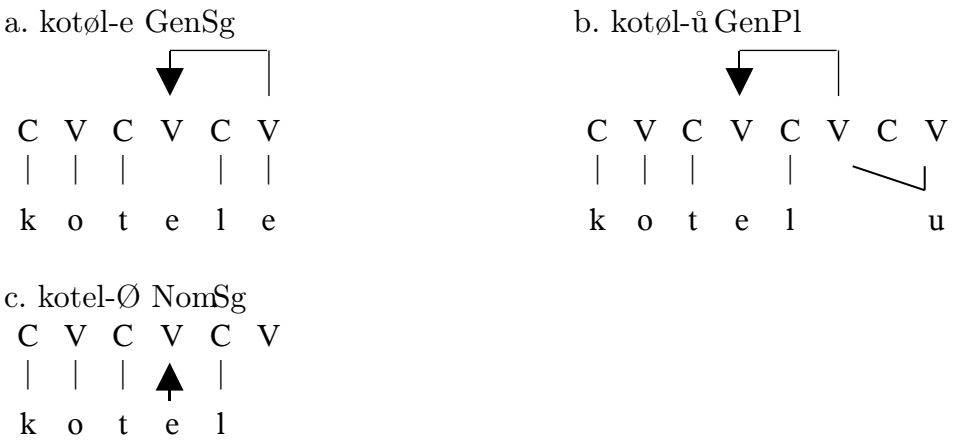
In table (10), I show the structure of the case forms derived from the root √kotøl whose lexical structure has been introduced in (9). We can see that the merger of positive markers always produces a zero alternant within the root, but the merger of zero markers causes vocalisation. And all other positive and zero markers, not only those merging the root √kotøl, behave alike.

(10) Paradigm of the root √kotøl

	Singular	Plural
Nom	kotel-∅	kotl-e
Acc	kotel-∅	kotl-e
Gen	kotl-e	kotl-ů
Dat	kotl-i	kotl-ům
Loc	kotl-i	kotl-ech
Ins	kotl-em	kotl-y

From what has been said about Government and its effect on vowel-zero alternations it inevitably follows that no empty Nucleus intervenes between the root and the positive marker. The assumption that marker-initial vowels must belong to the stem-final Nucleus is illustrated in (11).

(11) Derivation of e ~ ø alternants: √kotøl "boiler"



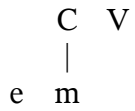
In (11a,b) the Nucleus following the alternation site is occupied by the vowel of the case marker, hence it servers as a governor. As a consequence the Nucleus with the floating *e* is governed and thereby prevented from being spelled out. In (11c), I show the nominative singular form derived by the zero marker. In this case, the floating vowel associates with its Nucleus because it is not governed.

The picture outlined in (11) raises the question about the lexical representation of marker-initial vowels: If they associate with the empty Nucleus of the preceding morpheme, how should they be represented lexically?

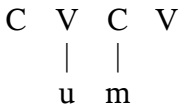
In case of markers which begin with short vowels, we have no choice but to let these vowels lexically float. Marker-initial short vowels must be lexically floating segments that lack any syllabic support. In order to be pronounced, they need to associate with an empty Nucleus. On the other hand, marker-initial long vowels are lexically associated to a Nucleus. In order to be pronounced as long, they must spread to their left. The lexical difference between long and short marker-initial vowels is illustrated in (12): in (12a), I show the structure of the instrumental singular marker *-em*; in (12b) I outline the structure of the genitive plural marker *-ům*. Their merger with the root √kotøl "boiler" is illustrated in (13).

(12) Marker initial short vs. long vowels

a. -em [InsSg]

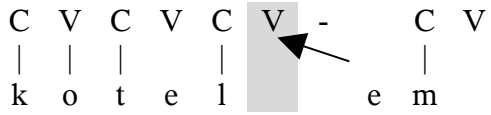


b. -ům [DatPl]

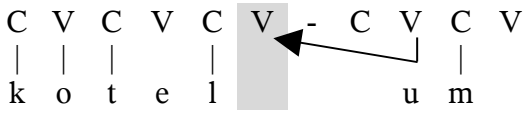


(13) Derivation of the case forms

a. kotl-em
√kotøl-[InsSg]



b. kotl-ům
√kotøl -[DatPl]



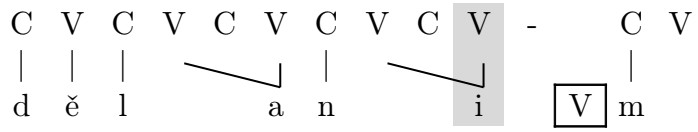
To sum up, on the one hand, positive case markers differ in their syllabic structure (their initial vowels either have their own constituent, or they are without any syllabic support), on the other hand, they must associate with the preceding empty Nucleus in common.

6.2 When lexically floating segments are not realized

In the *dělání* paradigm, marker-initial vowels cannot be pronounced because the stem is vowel-final, it ends in the suffix *-í*. No empty Nucleus it is available that could receive case-marking segment. Hence affixes containing only vowels are not pronounced at all. Com-

pound suffixes with both lexically floating and associated melody can realize only the latter. This is illustrated in (14) where the derivation of the surface structure of the instrumental singular form *dělání-Vm* is shown: the lexically floating V fails to be pronounced while the lexically associated *m* normally surfaces.

(14) Derivation of the surface forms



From what has been said it follows that those case markers which are underlyingly present in the *dělání* paradigm should start with short vowels. Why? There is no empty Nucleus after the stem, so therefore long marker-initial vowels could not branch and should be realized as short. However, in this paradigm no short vowels surface.

Of course, one can ask whether the proposal that all markers in the *dělání* paradigm begin with floating vowels is plausible if in the dative plural, other neuter paradigms show long vowels, as we have seen in table (2). In fact, in Czech nominal declensions, only a minority of case markers are gender-bound (see Caha & Ziková (2006) for details). From this perspective a dative plural allomorph which begins with the short vowel, *-em*, and which appears in the feminine paradigm *kost* "bone" could be assumed to merge with neuter stems derived by the suffix *-í* as well. Moreover, the fact that the instrumental plural marker in the *dělání* paradigm is *-mi* indicates that this paradigm is indeed a mixture of neuter and feminine markers, since the *-mi* appears in all feminine paradigms as well.

7 The rodeo story

The proposed model of the derivation of case forms predicts that vowel-final stems should be unable to realize those case markers which begin with short vowels. However, there exist neuter loanwords like *duo*, *rodeo*, *video*, or *stereo* which do display the vowel of vowel-initial markers even though their stem ends in the vowel *o*. Table (15) shows that they take on the same markers which appear in the main neuter paradigm *město*, with one exception in the genitive plural: *město* takes ∅, while *rodeo* takes *-í*.

(15) Rodeo and město

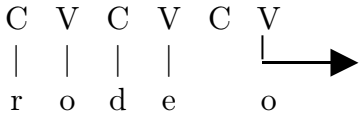
	město "city"		rodeo "rodeo"	
	Singular	Plural	Singular	Plural
Nom/Acc	měst-o	měst-a	rode-o	rode-a
Gen	měst-a	měst-∅	rode-a	rode-í
Dat	měst-u	měst-ům	rode-u	rode-ům
Loc	měst-u	měst-ech	rode-u	rode-ech
Ins	měst-em	měst-y	rode-em	rode-y

Is there any explanation for why just these *o*-final stems behave the same way as those stems which are consonant-final and why in the genitive plural they take a pronounced allomorph

instead of a zero? I claim that *o*-final neuter stems can display the same case markers as consonant-final stems of the *město* paradigm because the stem-final vowel can be easily re-interpreted as the nominative singular marker -*o*. As a consequence, a final empty Nucleus is created. Therefore all floating vowels of the markers can be normally spelled out.

Reinterpretation of the stem final vowel

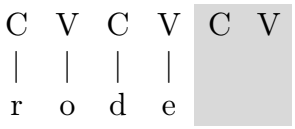
(16)



This analysis can also answer the question why do we get *měst-∅*, but not **rode-∅*? This is because unlike the former, the latter bears a final empty CV unit and a word may not end in an empty CV sequence.

GenPl *rode

(17)



To sum up, there exists a critical difference between neuter stems of the *dělání* type and neuter stems of the *rodeo* type. They both end with vowels; however this vowel either is (*rodeo*) or is not (*dělání*) re-interpreted as a case marker.

In this connection, the question arises as to what decides whether root vowels are interpreted as case markers. If we explore vowel-final stems which display vowel-initial markers, we can see that vast majority of them are those whose final vowels serve as the nominative singular markers in the default paradigms for feminine and neuter genders, i.e. those which end either in *a* (the default feminine marker) or in *o* (the default neuter marker). (As for masculines, in the nominative singular, the default merge is with a zero marker.) The following table illustrates the fact that in case of *o*- and *a*-final stems the final vowel is re-interpreted as a case marker, but in case that the stem ends in an other vowel it is not.

(18) Stem–final vowels

a, o	other vowels
duo → du-o	emu → *em-u
idea → ide-a	brandy → *brand-y

Finally, to avoid deriving wrong case forms like the dative plural form **emu-m*, or the locative plural form **brandy-ch*, where lexically associated parts of the case markers realize, we must conclude that to be indeclinable is a lexical property of particular stems. In other words, those stems whose final vowels are not re-interpreted as case markers and which do not belong to the *dělání* paradigm must be lexically specified as indeclinable, i.e. unable to combine with any case markers.

8 Conclusion

In this paper, I have brought together three independent pieces of evidence that initial vowels of case markers in a Czech nominal declension are either lexically floating segments (short vowels), or are lexically associated with Nuclei (long vowels). What they have in common is that they are lexically specified to search for an empty Nucleus to their left and associate with it.

References

Caha, Pavel and Markéta Ziková. 2006. Princip synkretismu aneb Augiášův chlév české deklinace. *Lingustica ONLINE* 1/2.

Kaye, Jonathan et al. 1990. Constituent structure and government in phonology. *Phonology Yearbook* 7:193–231.

Lowenstamm, Jean. 1996. CV as the only syllable type. Durand, J. & B. Laks (eds.). *Current trends in Phonology. Models and Methods*: 419–441. Salford, Manchester: ESRI.

Müller, Gereon. 2004. On Decomposing Inflection Class Features: Syncretism in Russian Noun Inflection. Müller, G., L. Gunkel & G. Zifonun (eds.). *Explorations in Nominal Inflection*: 189–227. Berlin: Mouton de Gruyter.

Scheer, Tobias. 2004. *A Lateral Theory of Phonology. Vol.1: What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.