

Haplogy of Reflexive Clitics in Czech

Alexandr Rosen

1. Clitics in Czech

As units exhibiting some properties of words and some properties of affixes, Slavic clitics (e.g. Franks & King, 2000) and clitics in general (e.g. Anderson, 1993; Zwicky, 1977) present a challenge to existing descriptive grammars and grammar formalisms. This is due mainly to their complex ordering properties, involving constraints of different types – phonological, morphological, syntactic, pragmatic and stylistic.

Czech clitics are no exception – in many aspects they are similar to clitics in other languages (inventory, position, order, climbing) but their behavior may be surprising in a ‘free-word-order’ language, i.e. language with a (constituent) order determined by information structure rather than by syntactic functions. Our focus will be on *special clitics* (henceforth simply *clitics*), whose word-order position is determined by constraints different from those determining the position of non-clitic words.¹

For space reasons and due to the focus of this paper we are not aiming at an exhaustive and precise coverage of all potential Czech clitics, but rather refer to previous work – Czech clitics have already attracted considerable attention (e.g. Fried, 1994; Avgustinova & Oliva,

¹ On the other hand, the position of *simple clitics* is the same as the position of non-clitic words of the same class, while the only difference is phonological (Zwicky, 1977).

1995; Toman, 1996; Rosen, 2001; Hana, 2007). They include short, mostly monosyllabic morphemes – auxiliaries, weak pronouns, short adverbs. Example (1) includes an auxiliary (*jsem*), a reflexive particle (*se*), a weak personal pronoun (*mu* ‘him_D’), and a pronoun which can occur both in a clitical and non-clitical position (*to* ‘it_A’).

- (1) Opravit₂ *jsem*₀ *se*₁ *mu*₂ *to*₂ včera snažil₁ marně.
 to-repair aux_{1sg} refl_A him_D it_A yesterday tried fruitlessly
 ‘I tried to repair_C it for him yesterday WITHOUT SUCCESS_R.’

All relevant clitics are given in italics for easier orientation. The subscripts co-index clitics with their syntactic governors, the finite auxiliary being labeled by 0. The subscripts increase with the degree of embedding of the governors. The subscript *A* stands for the accusative case, *D* for dative. Due to its information structure role as *Contrastive Theme* (marked by *C* and sans serif font), the embedded infinitive (*opravit* ‘to-repair’) precedes the clitics and *Rheme* (marked by *R* and UPPER CASE).²

There is a scale “clitic-hood”, depending on whether the morpheme can be used (i) only as a clitic, without a non-clitical variant, such as the forms of past tense and conditional auxiliary (e.g., *jsem*, *by*); (ii) only as a clitic, but there is a non-clitical variant of that morpheme, able to bear sentential stress (e.g., *se* vs. *sebe* ‘refl_A’ in (2)); or (iii) both as a clitic and a non-clitic (e.g., *jí* ‘her_D’). Forms that can only occur as clitics (e.g., *by*, *se*) are sometimes called *constant clitics*, while the promiscuous class (e.g., *jí*) is called *inconstant clitics*.

- (2) a. Marie *se* chválila v posudku.
 Marie refl_A praised in review
 ‘Marie praised herself in the review.’
 b. Marie chválila v posudku Petra / sebe / **se*.
 Marie praised in review Petr / refl_A / refl_A.
 ‘Marie praised PETR_R / HERSELF_R / *HERSELF_R in the review.’

² For more details about information structure, also in relation to Czech, see (e.g. Hajičová et al., 1998).

2. Ordering clitics

The rules governing the order of Czech clitics are remarkably different from those of full words. As Hana (2007, p. 67) observes, they are different at least in two important properties. (i) While the order of clausal constituents is “free” (i.e., determined by information structure), the order of clitics is quite rigid. Moreover, (ii) while full words tend to avoid discontinuous (non-projective) placement, except for a few marked cases and specific constructions, discontinuous order is rather common when clitics are involved (as in (1) above).

The most prominent constraints governing the placement of Czech clitics concern (i) the “second position” of a clitic cluster, (ii) the order within a clitic cluster, and (iii) the “single reflexive within a cluster” rule.

2.1 The second position

In a finite clause, clitics are in the “second position” (2P, Wackernagel position) as in (1) and (3-a). Any other position of the *clitic cluster* is ungrammatical (3-b)–(3-d).

(3)

a.

1P		2P				
Příští	sobotu	<i>bych</i>	<i>mu</i>	<i>to</i>	mohl	dát.
next	Saturday	would _{1sg}	him _D	it _A	could	give _{inf}

‘Next Saturday, I could give it to him.’

- b. *Příští sobotu mohl *bych mu to* dát.
- c. *Příští sobotu mohl dát *bych mu to*.
- d. **Bych mu to* příští sobotu mohl dát.

1P is usually the first clausal syntactic constituent, but sometimes 1P can be occupied by its part, as in (1) above, or by multiple clausal constituents.³ All such cases seem to be motivated by information structure. Moreover, in colloquial speech, 1P can be empty. In non-finite clause, the placement of clitics is less clear, but also rather rigid (e.g., if they follow the head, they follow it immediately).

2.2 Order within a cluster

Within a clitic cluster, clitics are ordered by their morphological properties (4), according to the pattern in (5).

- (4) a. Martin *by*₀ *se*₁ *jí*₂ *ho*₂ nakonec odhodlal₁ koupit₂.
 Martin would refl_A her_D him_A finally decided to-buy
 ‘Martin would find the courage to buy it for her in the end.’
 b. *Martin *se* *by* *jí* *ho* nakonec odhodlal koupit.
 c. *Martin *by* *jí* *se* *ho* nakonec odhodlal koupit.

- (5) **auxiliary** < reflexive < dative < accusative/genitive < *to* ‘it’

2.3 One reflexive only

Czech has two reflexive forms (either pronouns or particles) that have the status of constant clitics. Both have non-clitical counterparts, see (6).

- (6)

	full clitic
dative	<i>sobě si</i>
accusative	<i>sebe se</i>

There is one constraint that will be important throughout the rest

³ Such expressions occupying 1P are sometimes analyzed as a single constituent, defying the traditional criteria of constituency. However, the traditional dependency-based description of Czech syntax does not accommodate such a flexible view.

of this paper: no matter how many reflexive pronouns or particles are required by verbs or deverbal forms within a sentence, a cluster may include only one reflexive clitic.

- (7) a. Viděl *jsem se*, jak stojím na jeho místě.
 saw aux_{Isg} refl_A how stand_{Isg} at his place
 ‘I saw myself in his position.’
- b. *Představil *jsem si se*, jak stojím na jeho místě.
 saw aux_{Isg} refl_D refl_A how stand_{Isg} at his place
 ‘I imagined myself in his position.’
- c. Představil *jsem si* sebe, jak stojím na jeho místě.
 saw aux_{Isg} refl_D refl_A how stand_{Isg} at his place
 ‘I imagined myself in his position.’

However, a sentence can contain adjacent clusters, as in (8). The clitic *mu* is in the (singleton) cluster hosted by the infinitive *pomoci*, while *se* is in the (singleton) cluster hosted by the finite verb form *snažili*. Phonologically *mu* is an enclitic (i.e., following its host) while *se* is a proclitic (preceding its host), and there is a potential prosodic boundary between them.

- (8) [Pomoci₂ = *mu*₂] | *se*₁ = *snažili*₁ všichni.
 help_{inf} him_D refl_A tried all
 ‘Everybody tried to help him.’

3. Clitic climbing

As in other languages, Czech clitics governed by a syntactically embedded form may be positioned within a less embedded clitics cluster, even if there are no clitics dependent on the less embedded (or matrix) verb. While all clitics in (9-a) stay within the word-order domains of their syntactic governors (establishing embedded clitic clusters), those in (9-b) have “climbed” all the way to the second position of the whole sentence, forming a single clitic cluster. The resulting order within the cluster obeys the constraint presented in §2.2 above.

- (9) a. I Martin by_0 se_1 určitě snažil₁ [mu_2 pomoci₂ [ho_3 najít₃.]]
 even Martin would refl_A definitely tried him_D help_{inf} him_A find_{inf}
 ‘Even Martin would definitely try to help him to find it.’
- b. I Martin by_0 se_1 mu_2 ho_3 určitě snažil₁ pomoci₂ najít₃.
 even Martin would refl_A him_D him_A definitely tried help_{inf} find_{inf}
 ‘Even Martin would definitely try to help to help him to find it.’

Clitic climbing is not without constraints. One of them can be summed up as “climbing is monotonic” – a clitic cannot climb over another clitic. It may climb only half way to land in an intermediate cluster, but it can climb to a particular cluster only if all clitics with a less embedded governor had climbed to that or a higher cluster as well. None of the clitics in (10-a) have climbed, while in (10-b) both *mu* and *ho* have climbed one level up and in (10-c) both have climbed all the way to the highest cluster. On the other hand, in (10-d) the most deeply embedded clitic *ho* has climbed over *mu* at the intermediate position with an ungrammatical result.

- (10) a. Všichni jme_0 se_1 snažili₁ [mu_2 pomoci₂ [ho_3 najít₃.]]
 all aux_{ipl} refl_A tried him_D help_{inf} him_A find_{inf}
 ‘All of us tried to help him to find it.’
- b. Všichni jme_0 se_1 mu_2 snažili₁ [ho_3 pomoci₂ najít₃.]
 c. Všichni jme_0 se_1 mu_2 ho_3 snažili₁ pomoci₂ najít₃.
 d. *Všichni jme_0 se_1 ho_3 snažili₁ [mu_2 pomoci₂ najít₃.]

Interestingly, the order of syntactic governors is irrelevant. In (11-a) the two infinitives are in the first position, preceding the matrix clitic cluster and the rule that a more deeply embedded clitic cannot climb over a less deeply embedded clitic still holds (11-d).⁴

4 Note that the two adjacent clitics *mu ho* in can be analysed without additional data about the presence or absence of an intervening prosodic boundary either as an example similar to (10-a), where no clitic climbing takes place, or as a case of *ho* climbing into the intermediate cluster, hosted by *pomoci*.

- (11) a. [Pomoci₂ mu₂ ho₃ najít₃] jsme₀ se₁ snažili₁ všichni.
 help_{inf} him_D him_A find_{inf} aux_{1pl} refl_A tried all
 ‘It was all of us who tried to help him to find it.’
 b. Pomoci₂ ho₂ najít₃ jsme₀ se₁ mu₃ snažili₁ všichni.
 c. Pomoci₂ najít₃ jsme₀ se₁ mu₂ ho₃ snažili₁ všichni.
 d. *Pomoci₂ mu₂ najít₃ jsme₀ se₁ ho₃ snažili₁ všichni.

There are a few more facts about clitic climbing, such as its preference in a cluster of modal verbs and its impossibility from finite clauses and gerund phrases, or the constraint that reflexive clitics do not climb from object-controlled VPs (12).

- (12) a. Vláda občanům doporučila₁ se₂ pojistit₂.
 government citizens_D recommended refl_A insure_{inf}
 ‘The government recommended the citizens to get insurance.’
 b. *Vláda se₂ občanům doporučila₁ pojistit₂.
 government refl_A citizens_D recommended insure_{inf}

Another interesting constraint concerns a situation when two pronominal (non-reflexive) clitics meet in a single cluster. The order of such clitics corresponds to the degree of embedding of their governors. In (13) the first person pronoun *mi* precedes the third person pronoun *mu* with the only available interpretation of *mi* complementing the matrix verb *nepodařilo se* ‘failed’ and *mu* complementing the embedded infinitive *poslat* ‘send’.

- (13) Poslat₂ kurýrem se₁ mi₁ mu₂ ho₂ dnes nepodařilo₁.
 send_{inf} by-courier refl_A me_D him_D him_A today failed
 ‘I did not succeed in sending it to him by a courier today.’
 Avgustinova & Oliva (1995)

4. Haplogy

The term *haplogy* is usually applied to a strategy in morphology, resolving accidental repetition of identical morphemes. There are more attested strategies to deal with the phenomenon often seen as undesirable

or uncomfortable (Menn & MacWhinney, 1984): tolerance (14), avoidance (**manlily, likely*, sequences of interrogative and relative *which* or *who*), suppletion (15), portmanteau words (*apatheist* from *apathy* + *theist*) and deletion. As a strategy, haplology is identical with deletion.

- (14) Neshodne se se sestrou.
 agree_{neg} refl_A with sister
 ‘He/she doesn’t get along well with his/her sister.’
- (15) Goethe ist bekannter als Schriftsteller denn/*als als
 Naturwissenschaftler.
 Goethe is better-known as writer than/as as
 scientist
 ‘Goethe is better known as a writer than as a scientist.’
 Radford (1977,1979)

The term has also been used beyond morphology to cover cases of *syntactic* haplology (Neeleman & van de Koot, 2006). Thus the constraint presented above in §2.3, requiring that a clitic cluster should not include more than one reflexive marker, will be treated as a case of haplology, a situation where two or more reflexive markers, faced with the option of coexistence in a single clitic cluster as a result of climbing or multiple occurrence with a single governor, resort to the strategy of deletion instead.

4.1 Haplology of reflexive clitics

In (16) only a single reflexive may occur in the clitic cluster, representing both the reflexive particle complementing the inherently reflexive verb *stydět se* ‘be ashamed’ and the reflexive pronoun (homonymous with the particle) complementing the transitive verb *převléknout* ‘change dress’.

- (16) Děvče se₁₊₂ (*se) stydělo₁ převléknout₂.
 girl refl_A was ashamed change_{inf} dress
 ‘The girl was ashamed to change dress.’

Example (17) is different from (16) mainly because it includes reflexives in the dative case (both reflexive particles rather than pronouns). Because the matrix verb *netroufla* ‘didn’t dare’ occupies the first position and the infinitive *říct* ‘ask’ immediately follows the matrix clitic cluster, the two reflexives can either haplogologize within the matrix cluster (17-a), or the infinitive can retain its own cluster. Then the two occurrences of *si* must be prosodically separated.

- (17) a. *Netroufla₁ si₁₊₂ říct₂ o víc knedlíků.*
 dared_{neg} refl_D ask_{inf} for more dumplings
 ‘She didn’t dare to ask for more dumplings.’
- b. *Netroufla₁ si₁ | si₂ říct₂ o víc knedlíků.*
 dared_{neg} refl_D refl_D ask_{inf} for more dumplings
 ‘She didn’t dare to ask for more dumplings.’

Example (18) is different from (17) due to the presence of an auxiliary (required for past tense in 2nd person, not for 3rd person in (17)). The auxiliary *jsi* is usually pronounced (and sometimes written) in the same way as the reflexive *si*, but the auxiliary and the reflexive can still occur adjacently within the same clitic cluster. Together with (14) it shows that the constraint prohibiting multiple occurrence of *si* and *se* within a single cluster (or prosodical constituent) is restricted to reflexives and thus not phonological. On the other hand, the constraint is blind to the distinction between reflexive particle and reflexive pronoun (16).

- (18) *Netroufla₁ (j)si₀ si₁₊₂ říct₂ o víc knedlíků.*
 dared_{neg} AUX_{2sg} refl_D ask_{inf} for more dumplings
 ‘You didn’t dare to ask for more dumplings.’

4.2 Haplogy of unlikes?

Recall from §2.3 that Czech has two forms of reflexive clitics (dative and accusative) and that only one reflexive can occur in a clitic cluster. The logical question then is what happens when two different forms of reflexive are about to meet in a single cluster? Are these sentences simply impossible or can there be a “haplogy of unlikes”?

All the examples in (19) include a matrix verb *bát se* ‘be afraid’ and an embedded verb *vzít si* ‘put on’. The example (19-a) is the non-climbing version with both reflexive clitics next to their governors. In (19-b) the embedded clitic *si* occurs in the matrix cluster and the matrix clitic *se* disappears, while in (19-c) it is the embedded *si* that disappears. Both (19-b) and (less easily) (19-c) are acceptable to most speakers with the same meaning as (19-a). In line with the preceding considerations about clitic climbing and haplology we assume that a reflexive clitic may haplologize with another reflexive clitic even if they have different forms.

- (19) a. Jan se_1 $bál_1$ $vzít_2$ si_2 kravatu.
 Jan refl_A was-afraid take_{inf} refl_D tie
 ‘Jan was afraid to put on a tie.’
 b. Jan si_{2+1} $bál_1$ $vzít_2$ kravatu.
 Jan refl_D was-afraid take_{inf} tie
 c. ?Jan se_{1+2} $bál_1$ $vzít_2$ kravatu.
 Jan refl_A was-afraid take_{inf} tie

In (20) the reflexive clitics have switched their positions: *se* complements the embedded verb *usadit se* ‘sit down’ while *si* the matrix verb *troufnout si* ‘dare’. For the non-climbing version see (20-a). Similarly as in (19) above, the results after climbing are both acceptable, again with a preference for the embedded reflexive overriding the matrix one.

- (20) a. Troufla₁ si_1 usadit₂ se_2 v první řadě.
 dared refl_D to-sit refl_A in first row
 ‘She dared to sit in the first row.’
 b. ?Troufla₁ si_{1+2} usadit₂ v první řadě.
 dared refl_D to-sit refl_A in first row
 c. Troufla₁ se_{2+1} usadit₂ v první řadě.
 dared refl_A to-sit in first row

The assumption that the above examples involve haplology is not straightforward. They could be seen as similar to zeugma (grammatical

syllepsis – an irregular participation of a single item in two structures) or neutralisation (blurring of a distinction in the two forms). Our view that the phenomenon is best described as a case of haplology is supported by the fact that most assumptions involved in haplology are required anyway for the more standard “haplology of likes” and also by similar phenomena from other languages.

4.3 Haplology of unlikes in other languages

Polish, a language closely related to Czech, has just one form of reflexive clitic, namely *się*, corresponding to the Czech accusative form *se*. For the dative reflexive, Polish has only the full form *sobie* (6). Finding examples similar to Czech cases of “haplology of unlikes” is also harder because there are no inherently reflexive verbs with the dative full form *sobie*, but some similar examples can still be found. In (21) the embedded verb *przypomnieć* ‘remember’ should be complemented by a dative reflexive. Instead, the matrix clitic *się* seems to play a double role by complementing both the matrix verb *starać się* ‘try’ and the embedded verb *przypomnieć* (*sobie*). This is similar to the Czech example in (19-c).

- (21) Przywołuję te chwile usilnie, staram₁ się₁₊₂ przypomnieć₂.
 recall those moments hard try refl_A remember_{inf}
 ‘I’m trying hard to recall those moments, trying to remember.’

In (22) the same matrix verb *starać się* occurs both with (22-a) – (22-b) and without (22-c) the reflexive particle. As in (19-b) the accusative reflexive is replaced by the embedded dative reflexive.

- (22) a. Staralam₁ się₁ wytłumaczyć₂ swoją decyzję.
 tried refl_A explain my decision
 ‘I tried to explain my decision.’
 b. Staralam₁ się₁ sobie₂ wytłumaczyć₂ jego zachowanie.
 tried refl_A self_D explain his behavior.
 ‘I tried to explain to myself his behavior.’
 c. ?Staralam₁ sobie₁₊₂ wytłumaczyć₂ (jego zachowanie).
 tried self_D explain his behavior
 ‘I tried to explain to myself (his behavior).’

The pattern illustrated by (22-c) is not accepted by all native speakers of Polish, but the sequence *staralam sobie wytłumaczyć* returned 539 hits in Google (6 June 2013) and the query (23) in the IPI PAN corpus v.2 (250 million segments) yielded 8 results (Przepiórkowski et al., 2012).

(23) [base=“starać”] [orth=“sobie”] [pos=“inf”]

Breitbarth (2005) considers haplology of two adjacent finite auxiliaries split by a clause boundary in Early Modern German. However, an auxiliary can be dropped even if the two are not adjacent, suggesting that this could better be viewed as ellipsis.

(24) Als Swatoslaw seine Khinder versehen [_] / ist er in BVLGERN gezogen

(25) Wie die Bevestigung vnd di aussere Werck der Stadt beschaffen [_] / ist auß dem dieser Chronic beyfügtem Kupffer zu sehen.

As a last example of “haplology of unlikes” from foreign languages consider possessive markers in Romanian, which are deleted in the context of preceding definite determiners Neeleman & van de Koot (2006).

(26)	prietenul	(*al)	băiatului
	friend	POSS	boy
	<small>def, masc</small>	<small>sg, masc</small>	<small>def, dat, masc</small>
	the friend of the boy		

4.4 Haplology of unlikes in a corpus

The phenomenon calls for verification in a corpus. We used SYN, a corpus of written contemporary Czech, a part of the Czech National Corpus, including 1.3 billion words with automatically assigned lemmas and morphosyntactic tags.⁵ Using a list of inherently reflexive verbs taking infinitive we were looking for their finite forms in combination with an unexpected form of the reflexive clitic.

⁵ The data are based on the corpus version of May 2012, see Syn (2013).

Altogether there were about 2.7 million forms requiring *se* and 0.48 million those requiring *si*. Thus the ratio of *se*-reflexives : *si*-reflexives was 5.6 : 1.

The results are summarized in (27). In the queries we were searching for contiguous strings of inherent reflexives (IRs) and reflexive clitics (RCs), in any order (see the “Both IR” column). In a second round of searches the only the matrix verb was required to be an IR, while the embedded verb could be any verb taking a reflexive (the “Matrix IR” column). The results are of 4 types, depending on the direction of dependency between the two verbs taking *se/si* and the prevailing form of RC. The results were rectified by the relative frequency of *se/si* IRs.

(27)

Type	Matrix V	Embedded V	Resulting RC	Both IR	Matrix IR
A	<i>se</i>	<i>si</i>	<i>si</i>	52	901
B	<i>se</i>	<i>si</i>	<i>se</i>	64	–
C	<i>si</i>	<i>se</i>	<i>si</i>	34	281
D	<i>si</i>	<i>se</i>	<i>se</i>	45	56

The data do not attest any clear preference. While Stemberger (1981) claims that there is a preference for the deletion of the “outermost” morpheme in morphological haplology, the results show that there is no such parallel in syntactic haplology.

The results of the more general queries above were supplemented by a specific probe into a specific sequence combining the matrix verb *pokusit se* or *snažit se* ‘try’ with an embedded infinitive *zapamatovat si* ‘remember’. See (28) for the specification of the queries in the Corpus Query Language.

- (28) a. [lemma = “pokusit | snažit”] [word = “se | si”]
 [word = “zapamatovat”]
 b. [word = “se | si”] [lemma = “pokusit | snažit”]
 [word = “zapamatovat”]

The results in (29) for the specific matrix verb taking *se* and the embedded verb taking *si* show again a rather balanced outcome concerning the share of the prevailing RC, and a surprisingly high number of the cases where “haplology of unlikes” occurred (57), as compared with no haplology (60).

(29)

Type	Matrix V	Embedded V	Prevailing RC	No. of hits
A	<i>se</i>	<i>si</i>	<i>si</i>	36
B	<i>se</i>	<i>si</i>	<i>se</i>	21
Total for A+B				57
No haplology				60

4.5 Haplology of unlikes within a single clause?

Climbing is actually not necessary for two different reflexive clitics to meet within a single cluster. A Czech verb can undergo impersonal diathesis, resulting in a structure with an impersonal reflexive clitic *se*, as in (30).

- (30) a. Vy stále mluvíte.
 you always talk
 ‘You keep talking.’
 b. Vám *se* to mluví!
 you_D refl_A it complain
 ‘For you, it’s easy to talk (like this)!’

If the verb is inherently reflexive and taking *si*, the prevailing reflexive clitic is always the impersonal *se* (32).

- (31) a. Vy *si* stále stěžujete.
 you refl_D always complain
 ‘You keep complaining.’
 b. Vám *se*/**si* to stěžuje!
 you_D refl_A it complain
 ‘For you, it’s easy to complain!’

The impersonal *se* persists even when climbing takes place (32).

- (32) a. Na šéfa *se*/**si* špatně stěžuje.
on boss refl_A badly complain_{inf}
‘It is hard to complain about a boss.’
- b. Na šéfa *se*/**si* přestalo stěžovat.
on boss refl_A stopped complain_{inf}
‘People stopped complaining about the boss.’

It is not clear if the deletion of embedded reflexive clitic after impersonal diathesis can be seen as a case of “haplogy of unlikes”, especially in view of the fact that there is no alternative: *si* cannot occur instead or in addition to the impersonal *se* at all. We leave this phenomenon for further investigation.

5. A sketch of formalization in HPSG

The formalization is based on the proposal by Kupść (2000) for Polish, using clause union (aka argument composition/raising/amalgamation) as an independently motivated mechanism for solving mainly word-order phenomena involving any constituents Hinrichs & Nakazawa (1994). According to this solution, clitics may climb due to optional raising of arguments. Argument raising is a lexically specified option, restricting the class of syntactic categories from which arguments are raised and the class of “raising” verbs. A schematic lexical entry of a verb allowing for argument raising is shown in (33) (all irrelevant details are omitted). The entry specifies that some valency requirements of the verb’s infinitival argument (identified by $\boxed{1}$) may be left unsatisfied and treated as arguments of the verb.

(33) Argument raising

$$\left[\begin{array}{l} \text{HEAD } verb \\ \text{ARG-ST } \langle \dots, \left[\begin{array}{l} \text{HEAD } inf \\ \text{COMPS } \boxed{1} \end{array} \right] \rangle \oplus \boxed{1} \end{array} \right]$$

Argument raising is supplemented by an argument realization constraint to prevent RCs from surfacing more than once with a single host: multiple identical instances of RCs are resolved to a single form, while multiple different instances of RCs are resolved to an arbitrary form, unless one of the RCs is impersonal, then it persists. The crucial part of the constraint is presented in (34) as a relation **args2comps** between non-subject arguments [2] of a verb (a suffix of the verb's list of arguments – ARG-ST) and their realization as the verb's complements [2]. Technical details aside, the relation has access to any raised clitics in [2] and may select those appropriate for realization.

(34) Argument realization

$$word \rightarrow \left[\begin{array}{l} \text{SUBJ } [1] \\ \text{COMPS } [2] \\ \text{ARG-ST } [1] \oplus [2] \end{array} \right] \wedge \text{args2comps}([2], [2])$$

The linear position of a clitic cluster and the position of a clitic within the cluster is determined by constraints operating on morphosyntactic, phonological and prosodical properties. Details are beyond the scope of this paper (see, e.g., Kathol, 2000; Penn, 1999).

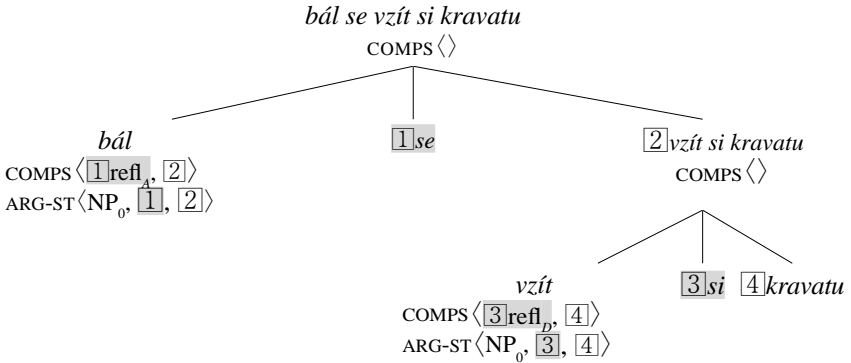
The order of arguments (within ARG-ST) is used to order clitics of same case in the cluster by their degree of embeddedness and to prevent clitics from skipping their less embedded neighbors.

The constraints above, together with other constraints of the grammar, license the three possible versions in (19), repeated as (35) for convenience.

- (35) a. Jan se_1 $bál_1$ $vzít_2$ si_2 kravatu.
 Jan refl_A was-afraid take_{inf} refl_D tie
 'Jan was afraid to put on a tie.'
- b. Jan si_{2+1} $bál_1$ $vzít_2$ kravatu.
 Jan refl_D was-afraid take_{inf} tie
- c. ?Jan se_{1+2} $bál_1$ $vzít_2$ kravatu.
 Jan refl_A was-afraid take_{inf} tie

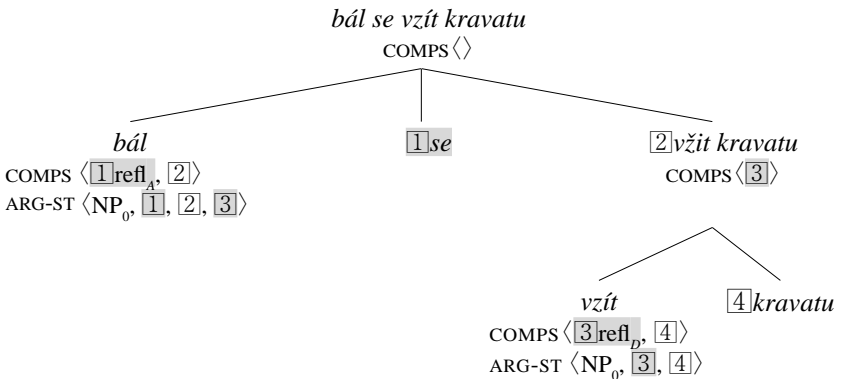
In (36) the embedded clitic does not climb ([1] in (33) is empty) and only the matrix clitic can be realized as a valency requirement of the matrix verb ([1]).

(36) No clitic climbing (35-a)



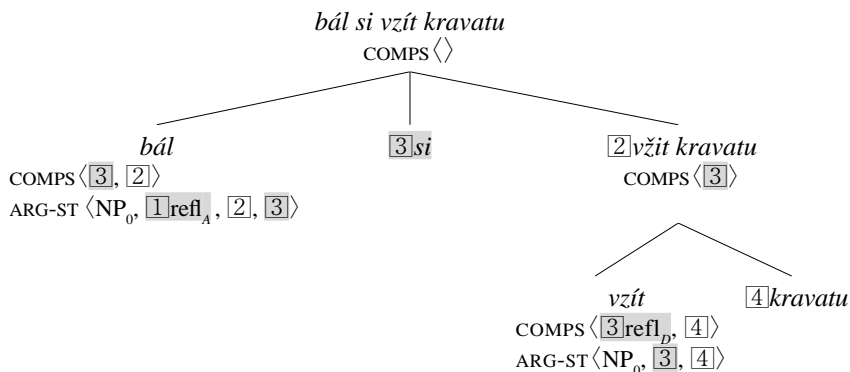
In (37) the embedded clitic does climb ([1] in (33) is non-empty, including exactly one item, namely [3]) and from the two reflexive clitics [1] and [3] the constraint (34) picks for realization the matrix clitic ([1]).

(37) Clitic climbing with haplogy of unlikes, result = matrix clitic (35-b)



In (38) the embedded clitic [3] climbs as in (37), but from the two reflexive clitics the embedded clitic ([3]) is picked for realization.

(38) Clitic climbing with haplogy of unlikes, result = embedded clitic (35-c)



6. Conclusion

Czech offers an interesting territory to investigate the phenomenon of syntactic haplogy in the case of clitics. With the strict clitic ordering rules, the “single reflexive per cluster” constraint, the clitic climbing option, a number of inherently reflexive verbs and two forms of reflexive clitics the phenomenon can be observed even in a corpus annotated by morphosyntactic tags. The results confirm the original hypothesis that haplogy of clitics often involves phonologically different items. In fact, in some lexical settings, the number of cases where two different forms haplogize equals the number of cases where the forms are kept apart.

“Haplogy of unlikes” is probably a relatively common phenomenon in more languages, too, calling for more detailed research with potential implications both for the development of competence grammars and for the study of linguistic performance.

7. Acknowledgments

To a large extent, this paper is based on common work with Jirka Hana. For a number of important comments I am also grateful to the audience of the workshop *Clitics and Beyond* at the University of Göttingen on May 3–5, 2012.

This work has been supported by the Grant Agency of the Czech Republic within the project No. P406/10/0434 and 13-27184S.

References

- Anderson, R. Stephen 1993. Wackernagel's revenge: Clitics, morphology, and the syntax of second position. *Language* 69: 68–98.
- Avgustinova, Tania and Karel Oliva 1995. The position of Sentential Clitics in the Czech Clause. In *Wiener Slavistischer Jahrbuch*, vol. 41: 21–42.
- Breitbarth, Anne 2005. *Live fast, die young – the short life of Early Modern German auxiliary ellipsis*. Ph.D. thesis, Universiteit Tilburg.
- Franks, Steven and Tracy Holloway King 2000. *A Handbook of Slavic Clitics*. New York: Oxford University Press.
- Fried, Mirjam 1994. Second-position clitics in Czech: Syntactic or phonological? *Lingua* 94: 155–175.
- Hajičová, Eva, Barbara H. Partee and Petr Sgall 1998. *Topic-focus articulation, tripartite structures, and semantic content* (Studies in Linguistics and Philosophy, vol. 71). Amsterdam: Kluwer Academic Publishers.
- Hana, Jirka 2007. *Czech Clitics in Higher Order Grammar*. Ph.D. thesis, The Ohio State University.
- Hinrichs, Erhard and Tsuneko Nakazawa 1994. Linearizing AUXs in German Verbal Complexes. In: John Nerbonne, Klaus Netter and Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, no. 46, 11–38.
- Kathol, Andreas 2000. *Linear Syntax*. Oxford: Oxford University Press.

- Kupść, Anna 2000. An HPSG grammar of Polish clitics. Ph.D. thesis, the Institute of Computer Science at the Polish Academy of Sciences, and the Université Paris 7.
- Menn, Lise and Brian MacWhinney 1984. The repeated morph constraint: Toward an explanation. *Language* 60(3): 519–541.
- Neeleman, Ad and Hans van de Koot 2006. Syntactic Haplology. In: Martin Everaert and Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, vol. 4, 685-710.
- Penn, Gerald 1999. Linearization and WH-extraction in HPSG: Evidence from Serbo-Croatian. In: Robert D. Borsley and Adam Przepiórkowski (eds.), *Slavic in HPSG*, Studies in constraint-based lexicalism, 149–182.
- Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski and Barbara Lewandowska-Tomaszczyk (eds.) 2012. *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.
- Rosen, Alexandr 2001. *A constraint-based approach to dependency syntax applied to some issues of Czech word order*. Ph.D. thesis, Charles University.
- Stemberger, Joseph P. 1981. Morphological haplology. *Language* 57(4): 791–817.
- Toman, J. 1996. A note on clitics and prosody. In Aaron L. Halpern and Arnold M. Zwicky (eds.), *Approaching second. Second position clitics and related phenomena*, CSLI Lecture Notes, no.61: 505–510.
- Zwicky, Arnold M. (1977). On clitics. Technical report. Ohio State University. Reproduced by the Indiana University Linguistics Club.