

**FORMAL PROPERTIES OF
PRESENT (and Past) CONDITIONAL IN CZECH**

Vladimír Petkevič

**Institute of Theoretical and Computational Linguistics
Faculty of Arts
Charles University**

Contents:

Formal properties of conditional mood in Czech:

- **present conditional**
- **past conditional**

Textual base: SYN2010 corpus of contemporary Czech comprising 100 million word forms

Objectives of the study

Theoretical (descriptive) objective:

Conditional structures are described in a more detailed way

Practical objectives:

- automatic **morphological disambiguation** can be improved
- automatic **syntactic analysis** can be improved
- automatic identification of the
 - **preterite tense**
 - **present conditional**
 - **past conditional**

1. FORMAL PROPERTIES OF THE PRESENT CONDITIONAL

The present conditional is formed by:

A. the conditional particle *by* (бы), in the variety of forms for 1st / 2nd / 3rd person singular / plural
+
a verbal past participle of the Slavic I-type

- (1) *Podívejme se ještě, zda bychom nenašli nějaké další stopy.*
(Let us see whether we would not find some other marks.)

B. the conditional conjunction *aby* (*чтобы*) and *kdyby* (*если бы*) in the variety of forms for 1st / 2nd / 3rd person singular / plural
+
a verbal past participle of the Slavic I-type

- (2) *Byl bych velice nerad, kdyby to věděl Scotland Yard.*
(*I would be very unhappy if the Scotland Yard knew about it.*)
- (3) *A domů to pak stihneš včas, abys to umyl.*
(*And you will make it home, so that you washed it.*)

1.1. Structures with the conditional morpheme *by*

The conditional morpheme *by* has the following variety of forms:

- bych/bysem* – 1st pers. SG (*I would*)
- bys* – 2nd pers. SG (*thou would*)
- by* – 3rd pers. SG (*he / she / it would*)
- bychom/bysme* – 1st pers. PL (*we would*)
- byste* – 2nd pers. PL (*you would*)
- by* – 3rd pers. PL (*they would*)

The **past participle** expresses:

either active voice:

(4) *Mnoho lidí **by** rádo vědělo, jak to dělá.*
(*Many people **would like to know**, how it is done.*)

or

it has a form of the auxiliary verb *být* (быть), and with a form of a **passive participle** it expresses passive voice:

**(5) Slova ožívají, jako *by byla* napsána
neviditelným inkoustem.**

**(Words revive as *if they were written in an
invisible ink.*)**

1.1.1. Principal syntactic facts

Statement 1. Both the conditional morpheme *by* and a corresponding **past participle** belong to the same clause.

Statement 2. The conditional morpheme *by* cannot stand independently in a clause, there must be a corresponding **past participle** form there as well. If not, the clause is a defective one.

Statement 3 – agreement. The conditional morpheme *by* and the corresponding **past participle** coforming the conditional mood agree in number.

Agreement – example:

(6) *V tom případě **bychom**(PL) **si raději měli**(PL) pohovořit.*

*(In that case we **should** better **have** a chat.)*

Statement 3a – agreement

The conditional form *byste* (2nd person plural) and the corresponding **past participle** need not agree in number.

If they do (the **past participle** is in **plural**), the conditional construction refers to a plural subject.

If they do not (the **past participle** is in **singular!**), the conditional construction has the meaning of a polite form of address referring to a singular entity.

(7) **Nechtěl** (SG) **byste** (PL) *přijít k nám a zeptat se ho sám?*

(*Would you not like to come to us and ask him yourself?*)

1.1.2. WORD ORDER

(i) either the conditional *by* precedes its **PP**,

(ii) or the conditional *by* follows its **PP**.

Important to consider:

(i) *by* is a clitic taking up a specific position in a clitic cluster

(ii) the clitic cluster takes up a specific – 2nd (Wackernagel's) – syntactic position in a clause:

(8a) *V zoufalství by už se vzdala všech pokusů o záchranu.*

(8b) *V zoufalství by se už vzdala všech pokusů o záchranu.*

(In despair she would already give up all attempts at a rescue.)

cluster: by se už

The word order in the cluster is syntactically fixed, the conditional *by* always preceding:

- reflexives *se / si*
- DAT and ACC pronominal clitics
- the other clitics.

The ordering of the enclitics in the cluster is as follows:

- i) conjunction **-li** (**если**)
- ii) short adverbs (**уж**, **уже**) (but they can be almost anywhere in the cluster!)
- iii) conditional **by** in all of its forms
- iv) reflexive particle **se / si** (**-ся**)
- v) clitic pronominal form in the dative case (**mi**, **мне**; **ti**, **тебе**; **mu**, **ему**,...)
- vi) clitic pronominal form in the ACC case (**mě**, **меня**; **tě**, **тебя**; **ho**, **его**; **ji**, **её**...)
- vii) other pronominal and adverbial clitics such as **to**, **это**; **rak**, **потом**

...

Conditional BY: 1st word order variant:

BY1. Conditional *by* precedes its **past participle**

(unmarked case)

Example:

**(9) *Ty bys chtěl odejít z Utahu?*
(*You would like to go from Utah?*)**

The distance between the conditional **by** and the **past participle**:

0. **by** is immediately followed by its **past participle** (immediate adjacency):

141929 occurrences in the SYN2010 corpus

1. **by** – 1 word in between – pastpart: **114949** occ.
2. **by** – 2 words in between – pastpart: **59239** occ.
3. **by** – 3 words in between – pastpart: **31954** occ.
4. **by** – 4 words in between – pastpart.: **15461** occ.
5. **by** – 5 words in between – pastpart.: **7282** occ.
6. **by** – 6 words in between – pastpart.: **3375** occ.
7. **by** – 7 words in between – pastpart.: **1709** occ.

Except for

distance(n=0) : distance(n=1) (141929 : 114949)

we see that:

distance(n) \approx 2 x distance(n+1)

Thus, the number of occurrences roughly decreases by the power of 2 – the regularity of the decrease is an astounding phenomenon!

Example – distance 1:

(10) *Ale raději **by** si **vyřízl** jazyk.*
(*But he **would** better **cut** out his tongue.*)

A clause can also be formed by two non-contiguous parts since an embedded clause can separate these parts:

(11) *... **jako** **by** **poté**, co na mě svalila všechny své hříchy, **potřebovala** něco na oplátku.*
(*... **as if**, after she laid all her sins on me, **she would** **need** something in return.*)

Conditional BY: 2nd word order variant:

BY2. Conditional *by* follows its **past participle**

(marked case)

Example:

(12) *Mohl by se jmenovat ‚Policejní zpravodaj‘.*
(It could be called the ‘Police bulletin’.)

In this case the distance is very short:

- either the conditional **by** immediately follows its participle
- or only the following elements can stand in between:
 - conjunction **-li** (**если**)
 - short adverbs/particles (**už**, **уже**; **prý**, **мол**)

A typical example:

(13) *Řekl-li bych, že na to do smrti nezapomenu...*
(*Would I say that I won't forget it till my death...*)

Conjunction **ABY-KDYBY**

ABY-KDYBY. Structures with the conditional expressed by conjunctions *aby* (*чтобы*) and *kdyby* (*если бы*)

Conjunction *aby* / *kdyby*:

- always precedes its **past participle**
- must be followed by a **past participle** in the same clause (**no infinitival constructions as in, e.g., Russian are possible!**)

The distance between the conjunction *aby* (*чтобы*) and the **past participle**:

0. *aby* is immediately followed by its **past participle** (immediate adjacency):

82237 occurrences in the SYN2010 corpus

1. *aby* – 1 word in between – pastpart: **81163** occ.
2. *aby* – 2 words in between – pastpart: **39884** occ.
3. *aby* – 3 words in between – pastpart: **22581** occ.
4. *aby* – 4 words in between – pastpart.: **9895** occ.
5. *aby* – 5 words in between – pastpart.: **4430** occ.
6. *aby* – 6 words in between – pastpart.: **1898** occ.
7. *aby* – 7 words in between – pastpart.: **845** occ.

Except for $n=0$:

$\text{distance}(0) \approx \text{distance}(1)$ (**82237** \approx **81163**)

we see that again:

$\text{distance}(n) \approx 2 \times \text{distance}(n+1)$

Distance for $n=0$ and $n=1$:

by: 141929 : 114949

aby: 82237 \approx 81163

by – 1 word in between – pastpart: **114949** occ.

aby – 1 word in between – pastpart: **81163** occ.

Conjunction **KDYBY**

The distance between the conjunction *kdyby* (*если бы*) and the **past participle**:

kdyby: 0-word distance: **18675** occ. in SYN2010

kdyby: 1-word distance: **17641** occ. in SYN2010

With bigger distances the decrease in frequencies of *kdyby* structures is also very similar to the *by* and *aby* ones:

$$\text{distance}(n) \approx 2 \times \text{distance}(n+1)$$

Why is there a difference between the *by* and *aby/kdyby* case for distance(0) and distance(1)?:

by: 141929 : 114949

aby: 82237 \approx 81163

Attempt at an explanation:

The conjunction *aby* – unlike the *by* morpheme – almost always introduces a dependent clause

The elements that take up the first position in the *by* structures take up – quite often – the second position in the *aby* clauses (i.e. they immediately follow the introductory conjunction *aby*) and seem to make for the balance of:

aby_distance(0) ≈ aby_distance(1)

- (14a) *Ted' přidám trochu mléka, aby nápoj byl chutný.*
(*I will add some milk, so that the drink were tasty.*)
- (14b) *Nápoj by byl chutný.*
(*The drink would be tasty...*)

2. FORMAL PROPERTIES OF THE PAST CONDITIONAL (the **by** case only)

The past conditional differs from the present conditional:

1 or 2 additional occurrences of a **past participle** of the auxiliary verb *být / bývat* (**быть**) are used.

The past conditional has a meaning of an **unreal condition**:

(15) *Ten chlap **by byl** potřeboval pár skřítků.*
(*That chap **would have needed** several elfs.*)

Here the past conditional is expressed by three word forms:

- conditional *by*
- past participle *byl* of the auxiliary verb **být/bývat**
- past participle *potřeboval* of the autosemantic verb

2.1. Agreement in past conditional

Statement 3c – agreement. The conditional *by* and the corresponding **past participles** coforming the past conditional agree in number except for *byste* (2nd person plural).

If the form *byste* is used, the corresponding **past participles** all agree in number but they need not agree with *byste* in number.

If they do, the past conditional refers to a PL subject.

If they do not (**all past participles** are in SG), the conditional construction expresses a polite form of address referring to a singular entity.

(16) Mohli jsme se tam sejít všichni čtyři, a konečně *byste byla poznala* jednu moji manželku.

(We could have met there all four, and at last you could *would have got to know* one of my wives.)

2.2. Word order in past conditional

The most frequent word order configurations in past conditional:

by PPbýt PPverb: 1122 occ.

by 1word PPbýt PPverb: 1107 occ.

Typical example:

(17) *Kde by byla měla ta holka vystoupit?*
(*Where should that girl have got off?*)

(17a) *Pochopitelně bych ho byl nepoznal.*
(*Of course, I would not have recognized him.*)

More words standing in between → sharp decrease:

<i>by</i>	w	PPbýt	PPverb:	1107 occ.
<i>by</i>	w w	PPbýt	PPverb:	314 occ.
<i>by</i>	w w w	PPbýt	PPverb:	112 occ.
<i>by</i>	w w w w	PPbýt	PPverb:	29 occ.

Word form(s) standing between **PPbýt** and **PPVerb**:

the decrease in frequency is as follows:

<i>by</i>	PPbýt	w	PPverb : 414 occurrences
<i>by</i>	PPbýt	w w	PPverb : 166 occurrences
<i>by</i>	PPbýt	w w w	PPverb : 105 occurrences
<i>by</i>	PPbýt	w w w w	PPverb : 38 occurrences

(18) **Pak** *by* *byl* *tyhle věci neudělal?*

(*Then he would these things not have done.*)

All 3 word forms coforming the **past conditional** are separated by 1 word at least:

<i>by</i>	w	PPbýt	w	PPverb: 313 occ.
<i>by</i>	w	PPbýt	w w	PPverb: 124 occ.
<i>by</i>	w w	PPbýt	w	PPverb: 66 occ.
<i>by</i>	w w	PPbýt	w w	PPverb: 22 occ.

(19) *Jinak by tam byli ještě čekali.*

(*Otherwise, they would still have waited there.*)

Another relatively frequent configuration of the past conditional:

PPbýt *by* PPverb

(20) ... a *byl by dostal navíc ještě nějakou pěknou maličkost.*

(... and, moreover, he *would have got a nice little something.*)

Note on negation. The past conditional structure can be negated; the negation is expressed as follows:

- either **PPverb** is negated
- or **PPbýt** is negated

PPverb is negated:

(21a) *Byla **by** nedokázala Edith s Polly vysvětlit, že se jí chce smát.*

*(She **would not have managed** to explain Edith with Polly, that she wants to laugh.)*

PPbýt is negated:

(21b) *Nemohl výpověď zapisovat, to **by** nebyl stačil ani zkušený stenograf.*

*(He could not make notes of the testimony, no experienced stenographer **would have coped** with it.)*

THANK YOU FOR YOUR ATTENTION!

Past conditional structures can be formed not only by the conditional morpheme *by*, but also by conjunctions *kdyby* and *aby*. The structures with *kdyby* are fully grammatical but not frequent. For instance, the configuration:

kdyby *pastpart_of_aux_být* *pastpart_of_autosem_verb*

has 528 instances in the SYN2010 corpus, e.g.:

(17) Ovlivnilo by výši majetku, *kdyby* byl spáchal sebevraždu?

E. Would the amount of property be influenced if he had committed suicide?

and the configuration:

kdyby *1_word* *pastpart_of_aux_být* *pastpart_of_autosem_verb*

has 423 instances in the SYN2010 corpus, e.g.:

(18) *Kdybych* to byla tušila!

E. If only I had suspected it!

Unlike the conjunction *kdyby*, past conditional structures introduced by the conjunction *aby* are much rarer (150 occurrences):

(19) Teď bych si přál, *abys* byl přišel dříve, Gandalfe.

E. lit. Now I would wish, so that you had come sooner, Gandalf.

E. Now I wish you had come sooner, Gandalf.