

FORMAL PROPERTIES OF PRESENT CONDITIONAL IN CZECH¹

Abstract. The text deals the main formal properties of the present conditional in Czech based on the extensive SYN2010 corpus of contemporary Czech. The present conditional is formed analytically by (a) the free morpheme *by* or the conjunctions *aby* / *kdyby* in all their gender and number variations, and (b) a past participle of a verb. Formal properties such as agreement and word order are studied.

1. Introduction

One of the type of problems to be solved by automatic morphological disambiguation and mainly by automatic parsing is an identification of multi-word units having a specific function. One of such multi-word types of structures is a form of conditional mood in Czech where the conditional mood is almost always formed analytically, i.e. by more than one word form. Two kinds of conditional mood are to be distinguished in Czech:

- present conditional;
- past conditional.

In this paper, formal properties of much more frequent present conditional is studied on the textual material of the SYN2010 corpus of contemporary Czech comprising 100 million word forms.

2. Formal properties of the present conditional

The present conditional is formed by the following variants:

(a) by the conditional particle *by* (E. *would*), in the 1st / 2nd / 3rd person singular / plural, and by a verbal past participle of the Slavic I-type:

¹ This study was written within the Programme for the Development of Fields of Study at Charles University, No. *P11 Czech national corpus, sub-programme Czech national corpus*.

(1) *Podívejme se ještě, zda bychom nenašli nějaké další stopy.*

E. lit. *Let us see whether **we would not find** some other marks.*

(b) by a conditional conjunction *aby* or *kdyby* (E. *in order to* or *if*, respectively), in the 1st / 2nd / 3rd person singular / plural, and by a verbal past participle of the Slavic I-type:

(2) *Byl bych velice nerad, **kdyby** to věděl Scotland Yard.*

E. lit. *I would be very unhappy **if** the Scotland Yard **knew** it.*

(3) *A domů to pak stihneš včas, **abys** to umyl.*

E. lit. *And you will make it home, **so that you could wash** it.*

The type (a) and (b)-*kdyby* have a true conditional meaning, whereas (b)-*aby* has only a form of conditional, its meaning being that of purpose, rather than condition. All the three types display a similar formal behaviour, their specific features are discussed below.

2.1. Structures with the conditional morpheme by

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

bych/bysem – 1st person singular (E. *I would*)

bys – 2nd person singular (E. *thou would*)

by – 3rd person singular (E. *he/she/it would*)

bychom/bysme – 1st person plural (E. *we would*)

byste – 2nd person plural (E. *you would*)

by – 3rd person plural (E. *they would*)

The past participle expresses either the active voice exemplified by sentence (4):

(4) *Mnoho lidí **by** rádo vědělo, jak to dělá.*

E. *Many people **would** like to **know**, how it is done.*

or it has a form of the auxiliary verb *být* (E. *to be*), and with a form of a passive participle it expresses the passive voice:

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

E. lit. *George's words revive as if **they were written** with an invisible ink.*

2.1.1. Principal syntactic facts

For space reasons we shall study only the active present conditional. The following syntactic facts are of crucial importance:

Statement 1. Both the conditional morpheme *by* and a corresponding past participle belong to the same clause. (This is a self-evident fact.)

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. A notable exception to this rule is specified in Statement 3a below.

In example sentence (6)

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

E. *In that case we should better have a chat.*

the conditional morpheme *bychom* and the corresponding past participle *měli* are both in plural.

Statement 3a – agreement. The conditional form *byste* (2nd person plural) and the corresponding past participle need not agree in number. If they do, the conditional construction refers to a plural subject; if they do not, i.e. in case the past participle is in singular, the conditional construction has the meaning of a polite form of address referring to a singular entity.

This anomaly is exemplified by:

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

E. lit. *Would you not want to come to us and ask him yourself?*

2.1.2. The word order

In a Czech clause there are two word order variations of the mutual position of the conditional *by* and its past participle:

- (i) either the conditional *by* precedes its past participle,
- (ii) or the conditional *by* follows its past participle.

The key syntactic property of the conditional morpheme *by* concerning the prosody and accent in a clause is the fact that the morpheme is always a clitic taking up a specific position in a clitic cluster which, in its turn, takes up the 2nd (Wackernagel's) syntactic position in a clause. The following example elucidates this fact:

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

E. *In despair she would give up all attempts at the rescue.*

Here *v zoufalství* (E. *in despair*) is an adverbial of manner taking up the first position in sentence (8), the conditional morpheme *by* and the reflexive particle *se* coform the enclitic cluster on the 2nd syntactic position prosodically leaning on the accented noun *zoufalství* (E. *despair*). The word order within the cluster is syntactically fixed, the conditional always preceding reflexives *se / si*, dative and accusative pronominal clitics and the other clitics. Generally, the ordering of the enclitics in the cluster is as follows:

1. conjunction *-li* (E. *if*)
2. short adverbs (*už*, E. *already*; *prý*, E. *reportedly*)
3. conditional *by* in all of its forms (E. *would*)
4. reflexive particle *se/si* (E. rough analogy: *-self*)
5. clitic pronominal form in the dative case (*mi*, E. *to me*; *ti*, E. *to you*; *mu*, E. *to him*,...)
6. clitic pronominal form in the accusative case (*mě*, E. *me*; *tě*, E. *you*; *ho*, E. *him*; *ji*, E. *her*...)
7. other pronominal and adverbial clitics such as *to* (E. *it*), *pak* (E. *then*) etc.

2.1.2.1. The unmarked case – the conditional morpheme precedes its past participle

This word order type is exemplified by sentence

(9) *Ty bys chtěl odejít z Utahu?*

E. lit. *You would want to go from Utah?*

E. *You would like to leave Utah*

Here the conditional morpheme *bys* immediately precedes its past participle *chtěl* (E. *want*). Such structures in which the past participle immediately follows the conditional morpheme is the most frequent one. Let us look more closely at the distance between the two elements with respect to their frequency:

0. *by* is immediately followed by its past participle (immediate adjacency): 141929 occurrences in the SYN2010 corpus

1. *by* – 1 word in between – past part.: 114949 occurrences

2. <i>by</i> – 2 words in between – past part.:	59239 occurrences
3. <i>by</i> – 3 words in between – past part.:	31954 occurrences
4. <i>by</i> – 4 words in between – past part.:	15461 occurrences
5. <i>by</i> – 5 words in between – past part.:	7282 occurrences
6. <i>by</i> – 6 words in between – past part.:	3375 occurrences
7. <i>by</i> – 7 words in between – past part.:	1709 occurrences

We see that the number of occurrences roughly decreases by the power of 2 – the regularity of the decrease is an astounding phenomenon! Specially, if one word is in between, it is predominantly the reflexive *se/si* or a clitic pronoun or adverb as in:

(10) *Ale raději by si vyřízl jazyk.*

E. lit. *But he would better cut out his tongue.*

The conditional *by* and its participle always belong to the same clause (Statement 1). However, the clause can 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.*

E. lit. ... *as if, after she laid all her sins on me, she would need something in return.*

Here we have two clauses, (a) and (b):

(a) *jako by poté ... potřebovala něco na oplátku*

E. *as if ... she would need something in return*

(b) *co na mě svalila všechny své hříchy*

E. lit. *after she laid all her sins on me*

The conditional *by* is in the first part of the (a) clause, the corresponding past participle is in the second one, the (b) clause being embedded in the (a) clause. Such structures are, however, extremely rare.

2.1.2.2. *The marked case – the conditional morpheme follows its past participle*

This word order type is exemplified by sentence

(12) *Mohl by se jmenovat ‚Policejní zpravodaj‘.*

E. *It could be called the ‘Police bulletin’.*

Here the variety of possible distances between the conditional *by* and the preceding participle radically differ from the reversed word order. The distance is very short: either the conditional *by* immediately follows its participle, or only the following elements can stand in between:

- conjunction *-li* (E. *if*)
- one instance of short adverbs/particles (*už*, E. *already*; *prý*, E. *reportedly*)

A typical example follows:

(13) **Řekl-li bych**, že na to do smrti nezapomenu...

E. lit. **Would I say** that I won't forget it till my death...

2.2. Structures with the conditional expressed by conjunctions *aby*, *kdyby*

This type of conditional, where conjunctions *aby* (E. *in order to*) and *kdyby* (E. *if*) are used, has a similar behaviour as the *by* type in case a past participle **follows** the conjunction. The following table shows the distance between the conjunction *aby* and the corresponding past participle:

0. *aby* is immediately followed by a past participle (immediate adjacency): 82237 occurrences in the SYN2010 corpus

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

Unlike the *by* case, the number of adjacent configurations and those where there is one word between the conjunction and the corresponding past participle is roughly the same. It is partly due to the fact that the conjunction *aby* – unlike the *by* morpheme – takes up almost always the first position in a dependent clause (the conjunction introduces the clause) and the elements that take up the first position in the *by* structures take up – in a vast majority of cases – the second

position in the *aby* clauses immediately after the introductory conjunction *aby*. In other cases, the decrease in occurrences resembles the *by* situation. The behaviour of the structures with the conjunction *kdyby* is similar to the *aby*-structures since the number of occurrences of the 0-word and 1-word distance is almost the same:

kdyby 0-word distance: 18675 occurrences in SYN2010

kdyby 1-word distance: 17641 occurrences in SYN2010

With longer distances the relative decrease in number in *kdyby*-structures also roughly corresponds to the decrease in *aby*-structures.

The reverse order, where a past participle is followed by the conjunction *aby* / *kdyby*, is entirely ungrammatical in Czech: in this respect, the *by* type considerably differs from the *aby* / *kdyby* one.

3. Conclusion

We have shown some of the formal properties of the present conditional in Czech such as agreement and specific features of the word order variation.