

ONDŘEJ BLÁHA

Czech Language Institute of the Czech Academy of Sciences, Prague

ORCID: 0000-0002-4397-2083

e-mail: o.blaha@ujc.cas.cz

## PREPOSITIONAL CASES IN CZECH: TOWARD A TYPOLOGICAL CHANGE

### Introduction

The modification of the meaning of the case by a preposition is rightly considered a structurally significant and typologically relevant grammatical feature (Sgall 1958: 77; Skalička 2004 [1960]: 767; Anusiewicz 1978: 139; Comrie 1989 [1981]: 56; Lotko 1997: 64; Uličný 2003; Uličný 2021: 134; Osolsobě 2017). There are differences in the use of simple and prepositional cases between the Slavic languages, as well as between dialects of the same language<sup>1</sup>, which are often interpreted as the result of a tension between the synthetic and analytic expression of the relations between the nouns in a sentence. While e.g. Czech appears from this perspective as a language expressing the relations between nouns mainly synthetically (i.e. by means of endings), Bulgarian and Macedonian can serve as examples of Slavic languages which, in the spirit of the analytic or isolation language type, have completely transferred the expression of the so-called deep case from endings to prepositions, respectively to the word order: e.g. Bulgarian sentence *В случай на превод от карта във валута, различна от лева, за картодържателя на картата в различна от лева валута могат да възникнат разходи*, ‘In the event of a transfer from a card in a currency other than Bulgarian leva, the cardholder may incur costs in a currency other than Bulgarian leva’, cf. with its Czech equivalent *V případě převodu z karty v jiné měně než v bulharských leva mohou držitelé karty vzniknout náklady v jiné měně než v bulharských leva*. In the Czech equivalent, the prepositional cases are often used, but the inflection of nouns is preserved

---

<sup>1</sup> In some traditional dialects of Czech, prepositional cases are replaced by cases without prepositions in certain context: *boudě* ‘in the shed’ (< *v boudě*), *pátek* ‘on Friday’ (< *v pátek*), *doktorovi* ‘to the doctor’ (< *k doktorovi*), *máslem* ‘with butter’ (< *s máslem*). See *Český jazykový atlas*, vol. 5, map 323 and 324. On-line: <https://cja.ujc.cas.cz>. Quoted April 14th, 2022.

here (except for the name of the currency, which is characteristically preserved in the uninflected form). The locative, which is strictly prepositional in modern Slavic languages, occurs several times in the sentence. However, in the earlier stages of the development of Slavic languages, including Czech, the locative without preposition also occurred, e.g. Old Czech *Pavel dal jest Ploskovicích zem'u, Vlach dal jest Dolás zem'u Bogu* 'Pavel gave the land in Ploskovice, Vlach gave the land in Dolany to God' in a marginalia from the beginning of the 13th century, inscribed in the founding charter of the chapter in Litoměřice.

A cursory comparison of Old Czech texts with the modern Czech gives the impression that Czech has gradually reduced the use of simple cases and has moved towards prepositional expressions of the case meanings. Cf. e.g. these examples from Old Czech *Chronicle of Dalimil* (finished before 1314), to which I am adding a 'translation' into present-day Czech: *Však dievek bojem utkati nesmniechu vs. Nesměli se utkat s dívkami v boji* 'They were not allowed to fight the girls in combat'; *Tuto sě Pražan skryte vs. Tady se skryjte před Pražany* 'Hide from the Pragueers here'; *Až Bieliny řeky dojeđu vs. Až přijeli k řece Bílině* 'When they came to the river Bílina'). Close links between the preposition and the noun can be observed as early as in the 14th and 15th centuries, when the preposition – temporarily – expressed the meaning of the accusative in phrase idioms quite "analytically" and without the addition of an ending (*Jmieti někoho za pán, za biskup vs. Mít někoho za pána, za biskupa* 'To have someone as lord, as bishop', *Před cisař vs. Před cisaře* 'In front of the emperor'). In this case, it was one phase of the development of the category of animacy (Komárek 2012: 168).

However, as I previously suggested in my frequency survey into Old and Middle Czech (Bláha 2016: 94), there are no reasons to believe that in the last seven centuries of Czech development there has been any massive proliferation of prepositional cases that would even remotely resemble the developmental trends e.g. in Old and Middle Bulgarian. From the 14th century to the present day, about a quarter of all noun cases have been modified by prepositions in Czech, and this value has fluctuated. The impression of significant changes in the use of prepositional cases arises from the relatively few 'textual dominants' in older Czech texts: cases where the case form is markedly different from that in present-day Czech.

This is not to say, however, that the frequency of prepositional cases in Czech has not gradually increased over the centuries. The essence of this phenomenon was given by M. Jelínek (2002: 79), who stated that in contemporary written Czech the choice of prepositional cases instead of simple cases increases the explicitness in expressing the relations in the sentence, and thus, on a more general level, increases the precision of expression. The demands for greater precision of expression in written language are also the cause of the striking proliferation of improper prepositions (on which see below) in 20th century Czech. A contrast to the changes in written Czech may be the relatively lower frequency of prepositions and prepositional cases in spoken Czech (Šonková 2008: 33), i.e. in spoken language, which, unlike written language, is situationally and contextually bound, and thus in which a higher degree of imprecision and vagueness of expression is possible.

The described developmental tendencies in the written language, i.e. the refinement of expression by means of prepositional cases, could not take place if there were no intra-linguistic, structural, or typological prerequisites for them. At least since Middle Czech (16th to 18th centuries) and certainly in present-day Czech, quite diverse analytization tendencies are evident, an overview of which is given most recently by O. Uličný (2021: 160). The aim of this article is to illustrate empirically, on the basis of surveys into the Czech National Corpus, the dynamics of prepositional cases (and also of individual prepositions) in Czech of the last thirty years, thus complementing some of the observations made in the works of the aforementioned M. Jelínek (2002) and O. Uličný (2021), as well as those of M. Hrdlička (2014-2015) and M. Čechová (2016).

## 1. Preliminaries

Although the spread of prepositional cases is undoubtedly slow and difficult to track even over large periods of time, it is clear that today's times are bringing unprecedented changes in the way we deal with language. The causes of these changes are the advent of new media, increased time pressure in the production of texts intended for the public, a certain casualness in the stylization of texts in general, the weakening of the feedback control of the written text, etc. All of this contributes to the accentuation of the developmental tendencies that are evident in written texts of a public nature and, in general, also contributes to a more rapid stabilisation of some of the results of these tendencies.

In particular, the journalistic style, in which developmentally progressive language is used (there is a direct contact with a more sophisticated form of contemporary spoken language), plays a significant role in the aforementioned stabilisation of the results of developmental tendencies. At the same time, the authors of journalistic texts strive for representativeness and, last but not least, for reasonable accuracy of the message. As a result, journalistic style provides a suitable field for the pursuit of a truly living and representative language.

My survey into the developmental tendencies of prepositional cases in Czech is based on the analysis of journalistic texts. From the latest, ninth version of the SYN corpus I created five temporally defined subcorpora (corresponding to the time span of the SYN2000, SYN2005, SYN2010, SYN2015 and SYN2020 corpora)<sup>2</sup>. These subcorpora contain only texts from national daily newspapers *Hospodářské noviny*, *Lidové noviny*, *Mladá fronta DNES* a *Právo*. Thus, I do not observe the form of language of leisure time journalism (e.g. in lifestyle magazines, which are usually also included in the journalistic parts of the corpora), and I work only with comparable types of texts, i.e. with texts that report on past, ongoing or expected events

---

<sup>2</sup> The SYN2000 corpus contains almost exclusively texts from 1990 to 1999, the SYN2005 corpus from 2000 to 2004, the SYN2010 corpus from 2005 to 2009, the SYN2015 corpus from 2010 to 2014 and the SYN2020 corpus from 2015 to 2019. Description of corpora: *Korpus SYN2000* [online]. Quoted 2022, 26th April. <<https://wiki.korpus.cz/doku.php/cnk:syn2000>>.

and that provide follow-up comments on these events. The considerable size of all five subcorpora (the numbers of tokens in these subcorpora are 41,174,874, 18,255,494, 20,877,916, 10,143,482 and 10,195,199)<sup>3</sup>, as well as the uniform way of tagging them, makes it almost certain that the data will be statistically significant<sup>4</sup> – I am dealing mainly with grammatical phenomena with high frequency.

The first important finding from the corpus survey is that the frequency of nouns in journalistic texts has been steadily decreasing over the last three decades. In the SYN2000 subcorpus, nouns had a frequency of 34.12%, in subsequent time layers this value decreased to 34% (SYN2005), 33.37% (SYN2010), 32.69% (SYN2015) and in the most recent text layer (SYN2020) the frequency of nouns was 32.59%<sup>5</sup>. Thus, there was an overall decrease in the frequency of nouns by 1.53%. In this respect, the written journalistic texts are somewhat closer to the commonly spoken language, in which nouns have a frequency of 13.3% (Šonková 2008: 33), i.e. much lower than in the written texts.

Nouns, which make up approximately one third of all tokens in journalistic texts, usually express the case only by the ending, i.e. without determination by a preposition. When a preposition does occur, it can be of two kinds: first, a proper preposition (or essential, *primární předložka* in Czech), second, an improper preposition (or accidental, *sekundární předložka*). The first group includes a total of nineteen prepositions that are not derived from any other word and most of which are shared with other Slavic languages (for definition of proper prepositions see *Mluvnice češtiny*, i.e. Komárek – Kořenský et al., 1986: 199-201). These prepositions are: *bez(e)* ‘without’<sup>6</sup>, *do* ‘(in)to’, *k(e/u)* ‘to, toward(s)’, *na* ‘on’, *nad(e)* ‘above, over’, *o* ‘about, of, for’, *ob* ‘every other/second’, *od* ‘from’, *po* ‘after, along, for’, *pod(e)* ‘under, below’, *pro* ‘for’, *před(e)* ‘before’, *přes(e)* ‘across, over’, *při* ‘at, near’, *s(e)* ‘with’, *u* ‘at, by, next to’, *v(e)* ‘in’, *za* ‘behind’ and *z(e)* ‘from, of’<sup>7</sup>. As the proper prepositions cannot be considered prepositions like *mezi* ‘between’ (< noun *mez* ‘limit, balk’), *mimo* ‘outside, beyond’ (closely related to verb *minout* ‘to pass by’) nebo *proti* ‘against, opposite’ (< preposition *pro* ‘for’), as well as e.g. composed prepositions *zpod* ‘from under’ or *zpoza* ‘from behind’.

As improper prepositions are also defined, in addition to those just mentioned, many dozens of other forms, e.g. *během* ‘during, within’, *díky* ‘thanks to’, *kolem* ‘around, along’, *kromě* ‘except (for)’, *kvůli* ‘because of, due to’, *podle* ‘according to’, *uprostřed* ‘in the middle’ or *včetně* ‘including’. Improper prepositions are a very dynamic phenomenon and have been spreading rapidly in Czech since the first half of the 20th century. They most often govern the genitive case.

<sup>3</sup> Corpus query: [tag="][ABCDEFGHIJNPRTV].\*"].

<sup>4</sup> Statistical significance of the results of the corpus surveys is verified by  $\chi^2$  test: *Calc: Corpus Calculator* [online]. Quoted 2022, 26th April. <<https://www.korpus.cz/calc/>>.

<sup>5</sup> The numbers of tokens from nouns in subcorpora were as follows: 14,048,115 (SYN2000), 6,206,008 (SYN2005), 6,967,150 (SYN2010), 3,315,404 (SYN2015) and 3,322,791 (SYN2020).

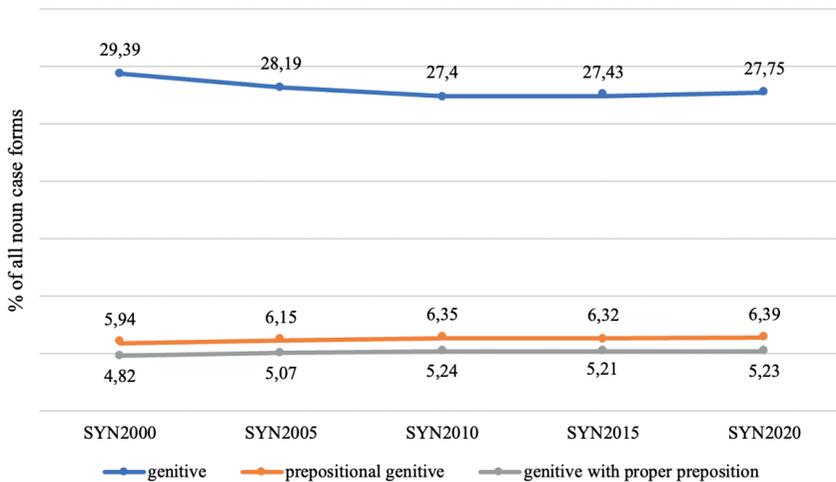
<sup>6</sup> In brackets I list the vowel that is added to the preposition in certain verbal contexts.

<sup>7</sup> The historical arguments for this definition of a total of nineteen proper prepositions can be supported by detailed data on their etymology (Kopečný 1973).

On the occasion of the explanations of prepositional cases, I also track the frequency of both proper and improper prepositions in the five subcorpora. By comparing the values of “instances per million (words)” in the SYN2000 and SYN2020 subcorpora, I conclude which prepositions have been increasing and which have been decreasing in the last decades. I notice only those prepositions that had a value of “instances per million (words)” of at least 100 in the SYN2020 subcorpus. This is to avoid misleading information: the frequency of some rather rare prepositions may increase by leaps and bounds, e.g. as a result of their occurrence in the idiolect of some journalist.

## 2. Prepositional genitive

As the following graph shows, the frequency of the genitive in journalistic texts has fallen by 1.64% over last three decades<sup>8</sup>. The reason for this is probably the long-term tendency to suppress “peripheral” cases, i.e. those with more marked surface semantics (genitive, dative, instrumental), in favour of “syntactic” or “structural” cases, i.e. of nominative and accusative (Uličný 2003; Čechová 2016; Uličný 2021: 160). The retreat of the genitive, especially from the syntactic function of the object, has been documented in Czech since the 17th century (Hausenblas 1958: 198), and this tendency is probably even more noticeable today.



The frequency of the prepositional genitive (with both proper prepositions and improper prepositions)<sup>9</sup> increased by 0.45% (the percentage given in the graph

<sup>8</sup> Corpus query: [tag="N...2.\*"].

<sup>9</sup> I use the query given in the monograph by V. Cvrček et al. (2020: 42). This formulation of the query allows the result to capture cases where there are developing members (attributes) between the

refers to the total number of nouns, see note 5). In direct proportion, the representation of the genitive with proper prepositions (by 0.41%), which accounts for the largest share of the total number of prepositional genitives, also increased.

Among the proper prepositions that govern genitive case, the greatest frequency increase is seen in the preposition *do* ‘(in)to’ (+0.4%)<sup>10</sup>. The preposition *od* is closely followed by prepositions *od* ‘from’ (+0.2%) and *u* ‘at, by, next to’ (+0.1%). These are prepositions with stronger semantics: the increase in their frequency is undoubtedly motivated by the need to express more precisely, cf. *obraz Josefa Lady* ‘a picture painted by a painter Josef Lada’, but also (theoretically) ‘a painting depicting the painter Josef Lada, his portrait’. Prepositional case solves this ambiguity: *obraz od Josefa Lady* ‘a picture painted by a painter Josef Lada’. As the analyzed subcorpora show, some proper prepositions slightly recede: *z* ‘from, of’ (−0.3%), *bez* ‘without’ (−0.05%) and *za* ‘behind’ (−0.05%)<sup>11</sup>.

Turning to improper prepositions, we see that the preposition *během* ‘during, within’ has become more frequent (+0.08%), followed by prepositions *včetně* ‘including’ (+0.04%), *vedle* ‘beside, next to’ (+0.02%) and *kromě* ‘except (for)’ (0.01%). On the contrary, two improper prepositions recede: the frequency of the preposition *podle* ‘according to’ decreases very significantly (−0.26%) and also *kolem* ‘around, along’ recedes (−0.02%).

### 3. Prepositional dative

The frequency of the dative case, as the journalistic texts from 1990 to 2019 show, is increasing very slightly (by 0.13%)<sup>12</sup>, but it is still a case that has a very low frequency in Czech.

The frequency of the prepositional dative<sup>13</sup> in texts is also increasing very slightly (+0.1%), but this increase is probably due to improper prepositions *kvůli* and *díky*. The frequency of both of these prepositions in Czech has increased significantly: *kvůli* ‘because of, due to’ (+0.3%), *díky* ‘thanks to’ (+0.13%). The other two improper prepositions, which govern the dative case, tend to weaken in frequency: *proti* ‘opposite’ (−0.11%) and *vůči* ‘towards, against’ (−0.01%).

It is worth noting that the only proper preposition that governs the dative also recedes, namely *k* ‘to, toward(s)’ (−0.37%). The preposition *k* receded in historical

---

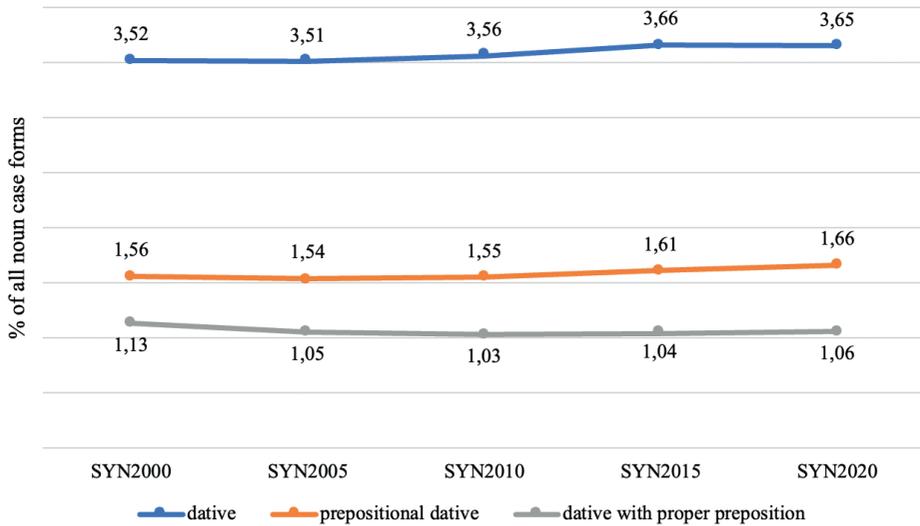
preposition and inflected noun. Corpus query: ([tag="R...2.\*"] within ([tag="R...2.\*"][tag="P[149EJ].\*"])? [tag="R...2.\*"]\*[tag="N...2.\*"]+([tag!="...2.\*"]</s>) within <s/>)).

<sup>10</sup> The numbers in brackets, given for each prepositions mentioned here and further, reflect the difference in frequency when comparing data from the subcorpus SYN2000 and SYN2020. The basis for calculating the percentages here is the sum of all tokens from the prepositions in the subcorpora SYN2000 (4,746,677 tokens of prepositions) and SYN2020 (1,193,106 tokens). Corpus query: [tag="R.\*"].

<sup>11</sup> However, the frequency of the preposition *za* is increasing when it governs instrumental case (see below in section 6).

<sup>12</sup> Corpus query: [tag="N...3.\*"].

<sup>13</sup> Corpus query: ([tag="R...3.\*"] within ([tag="R...3.\*"][tag="P[149EJ].\*"])? [tag="R...3.\*"]\*[tag="N...3.\*"]+([tag!="...3.\*"]</s>) within <s/>)).



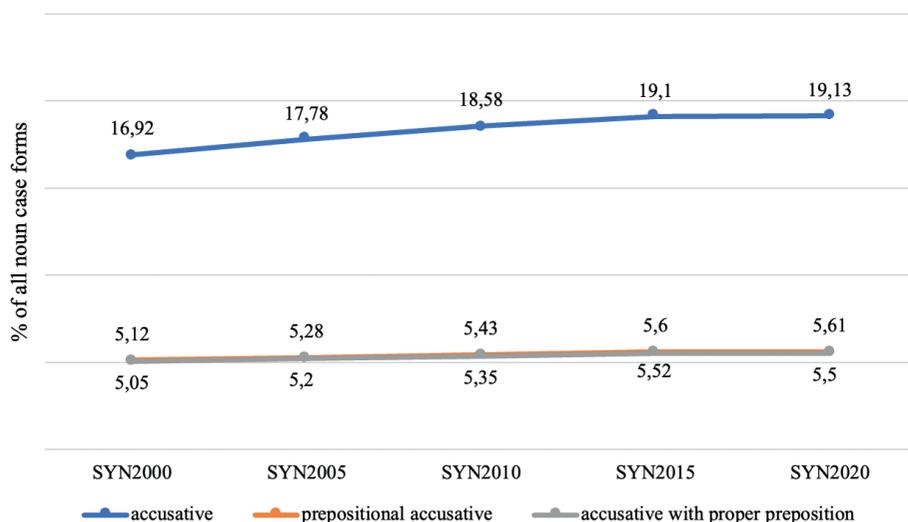
times in several Slavic languages – in Polish since the mid-16th century (Kopečný 1973: 100), today in written Polish the preposition *k* does not exist, cf. *Poszedł do lekarza* and Czech *Šel k doktorovi* ‘He went to see the doctor’. The preposition *k* is also gradually replaced by the preposition *do* in Ukrainian (*Він пішов до лікаря*) and Belarusian (*Ён пайшоў да доктара*). In Macedonian, which is a strongly “prepositional” language, the proper preposition *k* is also completely absent (*Toj otiđe na lekap*). The current decline of the preposition *k* in Czech is, in my opinion, an exact parallel to earlier processes in other Slavic languages. In this context, the aforementioned increase in the frequency of the preposition *do* in Czech is also striking<sup>14</sup>.

#### 4. Prepositional accusative

The accusative is one of two structural cases, i.e. cases with a strong syntactic function and relatively weak surface semantics. It is not surprising that the frequency of the accusative in journalistic texts increases most significantly among all the studied cases (+2.21%, comparing SYN2000 and SYN2020)<sup>15</sup>. The accusative has long been used as an almost exclusive object case in Czech (Hausenblas 1958: 45), cf. *Nemám žádné peníze<sup>acc</sup>* and Polish *Nie tam żadnych pieniędzy<sup>gen</sup>* or Ukrainian *У мене немає грошей<sup>gen</sup>*. In contemporary Czech, there is also a marked tendency for the accusative to replace some persistent genitive cases in the syntactic function of the object (*Užili si léta<sup>gen</sup>* > *Užili si léto<sup>acc</sup>* ‘They enjoyed the summer’), locative cases (*Diskutovali o té možnosti<sup>loc</sup>* > *Diskutovali tu možnost<sup>acc</sup>* ‘They discuss the possibility’) etc.

<sup>14</sup> The preposition *k* occurs little or not at all in Czech dialects in Silesia (Ostrava and its surroundings), where is common *do doktora* ‘to the doctor’, cf. Standard Czech *k doktorovi*.

<sup>15</sup> Corpus query: [tag="N...4.\*"].



The fact that the accusative is used as a “signal of syntactic object” is reflected in the fact that the accusative is not as often prepositional as the dative or instrumental (see below). It also does not occur as often as the genitive with improper prepositions, so that the increase of prepositional accusative (+0.49%) almost coincides with the increase of accusative governed by proper preposition (+0.45%)<sup>16</sup>.

The accusative occurs mainly with proper prepositions in the material studied. The frequency of these prepositions governing the accusative is increasing: *na* ‘on’ (+0.51%), *pro* ‘for’ (+0.14%), *přes* ‘across, over’ (+0.06%) and *o* ‘about, of, for’ (+0.05%). On the contrary, the frequency of the preposition *v* ‘in’ decreases significantly when it governs the accusative (−0.34%), but not when it governs the locative (see below).

Although the accusative is mainly associated with proper prepositions typically and frequently, it can be concluded from our material that frequency of these improper prepositions is slightly increasing: *mezi* ‘between’ (+0.04%) and *mimo* ‘outside, beyond’ (+0.03%) governing the accusative.

## 5. Locative (always prepositional)

For the sake of completeness, let us also present data for the locative case, which in Czech, as in other Slavic languages, is strictly prepositional. Frequency of locative case<sup>17</sup> slightly increases in Czech (SYN2000: 10.74% – SYN2005:

<sup>16</sup> Corpus query: ([tag="R...4.\*"] within ([tag="R...4.\*"][tag="P[^149EJ].\*"])?[tag="R[^R]...4.\*"]\*[tag="N...4.\*"]+([tag!="...4.\*"]</s>) within <s/>)).

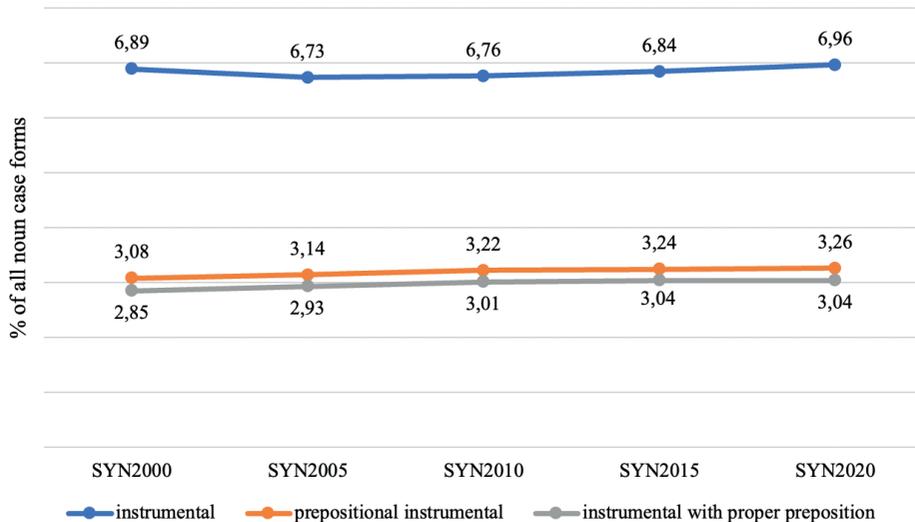
<sup>17</sup> Corpus query: ([tag="R...6.\*"] within ([tag="R...6.\*"][tag="P[^149EJ].\*"])?[tag="R[^R]...6.\*"]\*[tag="N...6.\*"]+([tag!="...6.\*"]</s>) within <s/>)).

10.82% – SYN2010: 11.09% – SYN2015: 11.06% – SYN2020: 11.16%), or the whole period under review (30 years) is an increase of 0.42%. Czech locative case is governed by proper prepositions only. The frequency of two prepositions is increasing: *na* ‘on’ (+0.39%), *v* ‘in’ (+0.02%). The other three proper prepositions governing the locative are decreasing in frequency: *o* ‘about, of, for’ (–0.74%), *při* ‘at, near’ (–0.26%) and *po* ‘after, along, for’ (–0.11%).

## 6. Prepositional instrumental

Although the frequency of the instrumental in Czech, as shown by comparing the subcorpora at the edges of the timeline, increased only very slightly over the whole period (by 0.07%)<sup>18</sup>, the overall frequency of this case seems rather constant. Some reduction in the frequency of the instrumental in contemporary Czech is probably due to the retreat of this case from the verbonominal predicate (Uličný 2021: 159), where the instrumental is replaced by the nominative, e.g. *Autor byl pedagogem*<sup>ins</sup> > *Autor byl pedagog*<sup>nom</sup> ‘The author was a teacher’. In this case, Czech is developing in a different direction than other languages in the Slavic north, cf. *Autor byl pedagogiem*<sup>ins</sup> (\**byl pedagog*<sup>nom</sup>) and Ukrainian *Автомор був педагогом*<sup>ins</sup>. Parallels to the results of the tendencies observed in Czech can be found more in the Slavic South, cf. *Autor je bio pedagog*<sup>nom</sup>.

The indicated frequency decline of the instrumental is, as the graph shows, offset to some extent by an increase in the frequency of prepositional instrumental<sup>19</sup>.



<sup>18</sup> Corpus query: [tag="N...7.\*"].

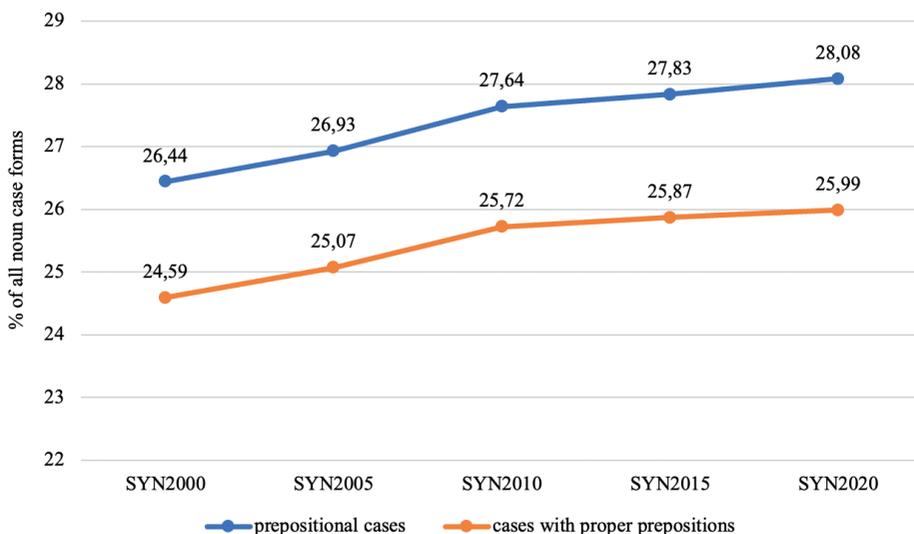
<sup>19</sup> Corpus query: (([tag="R...7.\*"] within ([tag="R...7.\*"] [tag="P[^149EJ].\*"])? [tag="^[^R]...7.\*"]\* [tag="N...7.\*"] + ([tag!="...7.\*"] </s>) within <s/>)).

Unlike the dative, but in agreement with the genitive, the instrumental in Czech, when prepositional, appears mainly with proper prepositions. A frequency increase has been noted for these prepositions: *s* ‘with’ (+0.19%) and *za* ‘behind’ (+0.07%). On the contrary, the frequency of the preposition *nad* ‘above, over’ (−0.16%), as well as the frequency of preposition *před* ‘before’ decreases (−0.06%).

Of the improper prepositions governing the instrumental case, only the preposition *mezi* ‘between’ has a more markable frequency. However, as the analysed sub-corpora show, the frequency of the preposition *mezi* somewhat decreases (−0.08%).

## 7. Prepositional cases and analytization tendencies

To sum up, the frequency of prepositional cases in Czech, as shown by journalistic texts from the last three decades, is increasing (+1.64%). The growth was more pronounced between 1990 and 2009, then it slowed down somewhat. The largest share of the total number of prepositional cases is accounted by cases with proper prepositions, which increased by 1.4% over the period under review.



It is problematic to explain this increase in the frequency of prepositional cases as a phenomenon that clearly indicates a typological change of Czech. At the very least, in my opinion, it cannot be said that the degree of syntheticity in contemporary Czech is significantly weakening. This impression, which can be obtained simply by observing the language, can be supported empirically. A suitable tool is the typological indicator, designed by I.-I. Popescu (Popescu et al. 2008). This indicator expresses the degree of syntheticity of certain language based on the relationship between the length of the text sample and the value of the *h*-point. The *h*-point

is defined as a place in the frequency distribution of a text where the rank of a given word equals its frequency<sup>20</sup>. I.-I. Popescu's typological indicator  $a$  can be calculated using the following equation, where  $N$  is the total length of the text (number of tokens).

$$a = \frac{N}{h^2}$$

Popescu's typological the indicator has the following values for each subcorpus analysed: SYN2000 – 8,27; SYN2005 – 8,69; SYN2010 – 8,78; SYN2015 – 8,99; SYN2020 – 9,04. It is true that the higher the value of Popescu's typological indicator, the higher the syntheticity degree of certain language. Thus, the observed values show (perhaps surprisingly) that the syntheticity degree of Czech is somewhat increasing. This fact can perhaps be interpreted as a consequence of the above-mentioned striving for ever more precise, explicit expression in journalistic texts<sup>21</sup>. This may also be manifested, for example, by more frequent and consistent inclusion of various specific data, such as names of persons, places, etc. Perhaps consistent with this interpretation is the relatively steadily increasing value of the type-token ratio<sup>22</sup>, which is one of the basic indicators of lexical richness of texts: SYN2000 – 0,014; SYN2005 – 0,023; SYN2010 – 0,021; SYN2015 – 0,032; SYN2020 – 0,031.

## Conclusion

It can be stated that the frequency of prepositional cases in Czech journalistic texts has increased over the last three decades. However, a clear typological interpretation of this increase (in the sense of “analytic tendencies in Czech are getting stronger”) is problematic, because the increase in the frequency of prepositional cases is undoubtedly a reflection of the efforts of the authors of journalistic texts to express themselves at least “on the surface” of the texts as accurately and explicitly as possible. This tendency seems to be evidenced by the increasing value of Popescu's typological indicator and type-token ratio, measured on the basis of frequency lists from all five subcorpora under study. Thus, the increase in the frequency of prepositional cases is a stylistic rather than a typological phenomenon (the overall typological “tuning” of Czech, in which certain analysing tendencies are evident, allows this rather stylistically motivated process to occur). The fact that

<sup>20</sup> Values of  $h$ -point, found out on the basis of frequency list of types: 2 231 (SYN2000, total number of types: 583 254), 1 449 (SYN2005, total number of types: 421 286), 1 542 (SYN2010, total number of types: 445 920), 1 062 (SYN2015, total number of types: 320 507), again 1 062 (SYN2020, total number of types: 317 924). As throughout this article, I consider only word types, including abbreviations and foreign words, but not punctuation and unrecognized tokens. Query: [tag="\*[ABCDFIJNPRTV].\*"].

<sup>21</sup> At least in terms of the chosen lexemes and syntactic devices (semantics or possible interpretations of journalistic expression cannot be dealt with here).

<sup>22</sup> Type-token ratio: number of types / number of tokens. A more detailed description: Cvrček – Čech – Kubát (2020). Quoted May 5th, 2022. On-line: <https://korpus.cz/quitau/>.

e.g. in Slovenian (and in Upper and Lower Sorbian), i.e. in languages in which nouns are inflected in a relatively rich and archaic way, the instrumental is strictly prepositional, probably reflects the old results of contact with German, cf. Slovenian *Se je tja peljal z avtobusom* ‘He travelled there by bus’ and German *Er fuhr mit dem Bus dorthin*. The slight increase in prepositional cases in contemporary Czech may also be the result of the revival of old dispositions in this language, which were determined by many centuries of contact with German, or Czech-German bilingualism. This revival of old dispositions is also happening because there has been no anti-German purism among Czechs for more than half a century, and most Czechs do not even know German, i.e. unlike their ancestors, they cannot recognise Czech-German parallels in their language (and possibly avoid them).

In line with the assumption, now clearly determined by typological factors (preference for “structural” cases), the frequency of the accusative, both simple and prepositional, is increasing in Czech (+2.21%). The frequency of the locative (+0.42%) and dative (+0.13%) is increasing to a much lesser extent. The frequency of the instrumental appears to be stable on the material of journalistic texts from 1990 to 2019 (+0.07%), while the frequency of the genitive, again in line with the typological assumption (suppression of “peripheral” cases with more pronounced surface semantics), is decreasing.

Regarding overall changes in the frequency of proper prepositions, these prepositions are mostly significant in this respect: *za* (governing the instrumental), *přes* (accusative), *od* (genitive), *u* (genitive) and *na* (accusative). It is worth noting that all of these prepositions are syllabic (unlike prepositions such as *s*, *z*, *v*) and thus are not, especially in spoken language, variant according to the phoneme with which the first word that follows them begins, cf. *v lese* [vlese] ‘in the forest’, *v Praze* [fpraze] ‘in Prague’, *ve vodě* [vevod’e] ‘in the water’. In present-day Czech, the preference for syllabic prepositions also increases regularity (invariance) of language and therefore its economy.

Finally, the improper prepositions whose frequency increases the most include the following: *kvůli* (governing the dative), *díky* (dative), *mezi* (accusative), *během* (genitive) and *mimo* (accusative).

## Bibliography

- Anusiewicz J., 1978, *Konstrukcje analityczne we współczesnym języku polskim*, Wrocław.
- Bláha O., 2016, *Poznámky k morfologickému vývoji češtiny*, Olomouc.
- Comrie B., 1989 [1981], *Language universals and linguistic typology: syntax and morphology*, Chicago.
- Cvrček V., Laubeová Z., Lukeš D., Poukarová P., Řehořková A., Zasina A.J., 2020, *Registry v češtině*, Praha.
- Cvrček V., Čech R., Kubát M., 2020, *QuitaUp – a tool for quantitative stylometric analysis*, Ostrava, <https://korpus.cz/quitaup/>.
- Čechová M., 2016, *Slovesné vazby a předložkové konstrukce v pohybu*, in: *Přednášky z 59. běhu Letní školy slovanských studií*, Praha, s. 29-38.
- Hausenblas K., 1958, *Vývoj předmětového genitivu v češtině*, Praha.

- Hrdlička M., 2014-2015, *O čem vypovídají české předložky?*, “Český jazyk a literatura” 65, s. 14-20.
- Jelínek M., 2020, *Pády prosté ustupují pádům předložkovým*, in: *Čeština – univerzália a specifika*, t. 4, Z. Hladká, P. Karlík (eds), Praha, s. 77-84.
- Komárek M., 2012, *Dějiny českého jazyka*, Brno.
- Komárek M., Kořenský J. et al., 1986, *Mluvnice češtiny*, t. 2, *Tvarosloví*, Praha.
- Kopečný F., 1973, *Etymologický slovník slovanských jazyků: slova gramatická a zájmena*, t. 1, *Předložky, koncové partikule*, Praha.
- Lotko E., 1997, *Synchronní konfrontace češtiny a polštiny*, Olomouc.
- Osolsobě K., 2017, *Deflektivizace*, w: *Nový encyklopedický slovník češtiny*, P. Karlík, M. Nekula, J. Pleskalová (eds), Praha, <https://www.czechency.org/slovník/DEFLEKTIVIZACE>.
- Popescu I.-I. et al., 2009, *Word frequency studies*, Berlin.
- Sgall P., 1958, *Vývoj flexe v indoevropských jazycích, zejména v češtině a v angličtině*, Praha.
- Skalička V., 2004-2006, *Souborné dílo*, F. Čermák, J. Čermák, C. Poeta (eds), Praha.
- Šonková J., 2008, *Morfologie mluvené češtiny: frekvenční analýza*, Praha.
- Uličný O., 2003, *K deflektivizačním tendencím ve slovanských jazycích*, in: *Česká slavistika 2003: české přednášky pro 13. mezinárodní sjezd slavistů v Lublani*, I. Pospíšil, M. Zelenka (eds.), Praha, s. 155-163.
- Uličný O., 2021, *Lingvobohemistické analýzy*, Praha.

ONDŘEJ BLÁHA

### Prepositional cases in Czech: toward a typological change

#### Abstract

This article examines the increase in frequency of use of the prepositional noun cases in Czech over the last three decades (1990-2019). The author has based his research on the Czech National Corpus where he created five temporally defined subcorpora containing only journalistic texts from the major Czech daily newspapers. Over the period under study, the share of prepositional cases in the total noun tokens has increased from 26.44% to 28.08% (by 1.64%). The share of cases with proper prepositions has increased the most, while the number of cases with improper prepositions has increased less. The author has identified developmental trends for specific frequented prepositions. According to the author, the increase in the frequency of prepositional cases in Czech is a phenomenon rooted in the “typological dispositions” of the Czech language. However, he considers the stylistic factor (the desire for more explicit, precise expression) to be a factor of more importance, contributing to the increase in the frequency of prepositional cases. On the other hand, some changes in the frequency of use of specific cases, most notably the observed greater frequency of the accusative (+2.21%), is a purely typologically motivated developmental change.

**Keywords:** Czech, language change, typology, journalistic style, noun, prepositional cases.