Athens Institute for Education and Research
ATINER

ATINER's Conference Paper Series
LNG2013-0512

Usage of some Non-finite Constructions in Czech and Russian

Natalia Klyueva
PhD Student
Charles University in Prague
Faculty of Mathematics and Physics
Czech Republic

An Introduction to
ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. The papers published in the series have not been refereed and are published as they were submitted by the author. The series serves two purposes. First, we want to disseminate the information as fast as possible. Second, by doing so, the authors can receive comments useful to revise their papers before they are considered for publication in one of ATINER's books, following our standard procedures of a blind review.

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research

This paper should be cited as follows:

Usage of some Non-finite Constructions in Czech and Russian

Natalia Klyueva
PhD Student
Charles University in Prague
Faculty of Mathematics and Physics
Czech Republic

Abstract

Czech and Russian are closely related Slavic languages that share many morphological and syntactic features. This paper describes one of the discrepancies between the languages that occur in non-finite clauses, namely in participial clauses and transgressives. We will concentrate mainly on transgressives because they demonstrate greater discrepancies.

The usage of Russian transgressives(gerunds) is very similar to that of English, they are typical mostly of a written language rather than spoken, whereas in Czech they are nowadays archaic in any genres and generally not used at all.

This discrepancy can pose some challenge for language learners as well as for translators or machine translation systems. We will present some examples of Russian participial/gerund clauses and the respective translations into Czech from a parallel corpus of news and belletristic texts.

Participles (verb forms ending in -ing in English) are indefinite forms of a verb that might have several functions in a sentence, here we will concentrate on their adjectival function in a non-finite clauses.

Transgressive constructions are clauses with non-finite verbs - transgressives (also gerunds, or adverbial participles) that express an action done simultaneously with/or right after the main verb.

Gerund clauses are rather common in Russian as well as in English, whereas modern Czech avoids it and uses other descriptive constructions – such as relative clauses, coordination of two verbs, or just dropping the verb.

Keywords:

Corresponding Author:

Introduction

The aim of my research is comparing Czech and Russian on different language levels. One of the syntactic differences between the languages is in the usage of non-finite clauses that include transgressives and participles. Those constructions exist in both languages and have almost identical form of surface morphemes for verbs, but in Russian language they are used rather frequently, whereas in Czech only occasionally. Following are the examples of the clauses in both languages1:

Transgressive phrase (clause, construction) - underlined:

(1cz) *Vaněk **jda_tra** se Švejkem, **řekl:***

(1ru) *Ванек, **идя_tra** рядом со Швейком, **сказал***

(1en) *Vanek, **going_tra** behind Svejk, **said**...*

Participles:

(2cz) *Stala se z něho vyrovnaná **osobnost**, **žijící_par** plným životem.*

(2ru) *Получилась несгибаемая **личность**, **живущая_par** полноценной жизнью.*

(2en) *He became a stable **personality**, **living_par** a full life.*

In this paper we make a comparative analysis on a material of a parallel corpus. In order to detect and calculate the discrepancies we explored two parallel Czech-Russian corpora from two domains - the one containing mainly news articles and commentaries in modern language and another consisting of literature, both classic and contemporary. Each word in the corpora has a morphological tag with information on the part of speech. The annotation provides us a part-of-speech tag that indicates a gerund, so the gerundive clause is rather easy to detect. It is more harder with participles, as they do not have their specific marker when used in non-finite constructions. They can be easily confused with adjectives or even nouns that were derived from a participial form and are not a part of non-finite clauses. So the statistics for the gerunds can be rather reliable, and as for the participial clauses we have analyzed only a small sample of examples.

Parallel text analysis showed that gerund clauses almost do not occur in modern Czech texts, either in news or in literature, though older Czech texts (such as "The Good Soldier Švejk") contain plenty of gerunds. In Russian they are quite frequent especially in the official language of the news. When a Russian text contains a gerund/participial clause, translators may sometimes opt to use a similar construction in Czech, but the general tendency is to avoid it. In the paper I also suggest a possible reason why Czech language is not so inclined to such constructions as compared with other Slavic languages.

Transgressives

---

[1]We will mark a transgressive with a tag **_tra,** participle – with a tag **_par** and a finite verb in a main clause - **_fin**. The generally used example subscript with information on other linguistic features is not necessary in our study.

Many researchers have studied the transgressive construction (in Russian – deeprichastnyj oborot) both from synchronic and diachronic point of view. Especially interesting is the history of development of such constructions from functioning as finite verbs, through being the second verb in a sentence up to the present state of adverbial modifications (Dombrowski, 2006; Barnet, 1962; Zeman et al., 2005).

In English tradition this form of a verb is called gerund, and sometimes we will use it as a synonym for a transgressive. Though, the phenomena in English and Slavic languages are different, rather the term 'gerund' is the closest translation equivalent of the term 'transgressive' for this case.

In Czech they are considered to be archaic, whereas in some other Balto-Slavic languages, like Russian, Polish, Lithuanian etc., they are used rather frequently especially in the official style. Czech transgressive formation is more morphologically complex, as the form agrees with the actor of a main clause in number and in gender. Polish and Russian gerunds are not that complex, they have only one form for all numbers and genders, so there are 3 forms for Czech gerund and only one in Russian for either past or present tense.

The system of Czech transgressives is more complicated than that in Russian also due to the fact, that there are two pattern paradigms of gerund declension (*a/ouc/ouce vs. e/-íc/-íce)*. This might be the reason why most native speakers avoid transgressives, and several researchers have already proposed this hypothesis.

We will make a comparative analysis of gerunds in Czech and Russian within two genres. Firstly, we will analyze InterCorp (Čermák & Rosen, 2012) a parallel corpus of belletristic texts. Those texts are either direct translation of Russian books from Czech into Russian or otherwise.

Secondly, we will have a look at news and commentary parallel texts downloaded from the web site Project Syndicate (www.project-syndicate.org/). Unlike the belletristic texts, they are not directly translated from Czech into Russian. The original language of the short articles is mainly English, and the texts are translated from English into Czech and from English into Russian independently and then "parallelized".

The method to find sentences with transgressive constructions is rather obvious in this case. The Russian tagger (Sharoff et al., 2008) marks a transgressive with a "*Vmg*...." sequence: *Vm* - main verb, *g* -gerund. A tag for a Czech gerund is "*Ve...*". After dots there may go several possibilities depending on whether it is perfective or progressive, active or medial. For Czech the categories of gender and number are also presented. We are not interested in these detailed features though. The table below presents the number of sentences with gerunds in the two genres. The last line shows number of sentences, in which the gerund constructions were used both in Czech and Russian. As the figures were derived automatically, the small number of cases can be calculated wrongly because the tagger prediction accuracy is not 100%.

Table 1. *Statistics on the Number of Transgressives*

Language/genre	literature	news
Czech	732	167
Russian	10460	6547
Transgressives in both	571	92

After the analysis of a small sample of gerunds in the corpora, we concluded, that, in news texts , when a gerund is used in English, it is translated as a gerund in Russian, and it can be translated in many different ways in Czech depending on the sense of the gerund phrase. As for the belletristic texts, when the gerund construction is translated from Russian into Czech, the Czech equivalent might be also gerund, or it might be some descriptive construction. We are mostly interested in finding the typical translation equivalents of Russian gerunds in Czech, so we have collected samples of those translations.

1. Dependent clauses in Czech:

(3cz) *V březnu **učinil_fin** Mušaraf nejsmělější krok , **když odvolal_fin** z funkce hlavního soudce ...*

(3ru) *В марте Мушарраф **сделал** смелый шаг , **отстранив_tra** от должности главного судью*

(3en) *Last March , Musharraf **took** his boldest step , (cz **when he removed**)/(ru **removing**) the Chief Justice...*

2. Coordinated clauses in Czech:

(4cz) *Mozart se **vzdálil_fin** a **ponechal_fin** Nicholase o samotě.*

(4ru) *Моцарт **удалился_fin, оставив_tra** Николаса наедине .*

(4en) *Mozart **went away** (cz: **and left**)/(ru: **leaving**) Nicholas alone*

3. Other means (noun phrases):

(5cz) *To si **získá** podporu mezi mnoha členskými státy OSN , **mimo jiné_NP** těmi , které mají rozsáhlejší nespokojené etnické menšiny .*

(5ru) *Это **найдет** поддержку у многих стран - членов ООН , **включая_tra** государства , имеющие крупные недовольные национальные меньшинства .*

(5en) *That will **attract** support among many UN members , (cz: **above all**)/(ru: **including**) those that have major dissatisfied ethnic minorities .*

4. Sometimes only <u>one verb</u> (either main or a gerund) is left in a Czech equivalent sentence:

(6cz) *Vyzývavě **se usmál_fin**.*

(6ru) *Задорно **улыбнувшись_tra**, вор сказал...*

(6en) ru: *Cheerfully **smiling**, the thief said...*cz:*The thief cheerfully **smiled**:*

5. A sentence is divided into two:

(7ru) *Мимо, враждебно (2)**косясь_tra** на сидящего, (1)**ковыляла_fin** старая дама, бог весть какими судьбами (3)**занесенная_par** сюда в этот поздний час.*

*lit.: Around, **looking_tra** askance at him, **hobbled_fin** an old woman, by God know what destiny **being brought_par** here.*

(7cz) *Kolem něho se (1)**belhala_fin** stará paní a (2)**vrhala_fin** na něj nepřátelské pohledy. Bůhví jaký vítr ji sem (3)**zanesl_fin** v tuto pozdní hodinu.*
*lit.: Around him **hobbled_fin** an old woman and **looked_fin** askance at him. God knows what **brought_fin** her here.*

The translator got a bit embarrassed by so many non-finite clauses in a Russian sentence, and decided not only to change the ordering of verbs, but also to divide a sentence into two smaller ones.

6. Some examples of Russian gerund are actually <u>lexicalized</u>: (ru) *Судя* по (according to) , the same can be found in Czech: *takříkajíc (so speaking), nehledě na (not looking at) , etc.*

7. And, finally, some examples with <u>the same translation- as a gerund</u>:

(8cz) *"Nikdy, " **odsekla_fin** jsem, **obouvajíc_tra** si kozačky.*

(8ru) *Никогда, - **отрезала_fin** я, **натягивая_tra** сапоги.*

(8en) *Never – I told, **putting_tra** on boots.*

A remark is due on the ordering of elements in a transgressive clause. Generally, when translating from one language into another, generally, the translator chooses the same order of clauses, as in the examples above, but sometimes the order is reversed when it is more appropriate to do so:

(9cz) *Mohl se třeba **posadit** na bradla a **kouřit**. (He could **seat** on bars and **smoke**)*

(9ru) *Он мог **курить**, **сидя_tra** на гимнастических брусьях. (He could **smoke, sitting** on bars)*

As far as we consider translation of gerund phrases between Czech and Russian, there is more than one way to translate those types of constructions dependent on their sense.

Transgressives in contemporary Czech have a flavor of an old bookish style. At some point in the past they were common, but as their paradigm was rather complicated, speakers tended to avoid using them until they eventually became archaic elements. There might be and additional reason why gerund phrases are used more often in Russian than in Czech. Czech is a pro-drop language, so in both main and a non-finite clause the subject pronoun is left out. In Russian there is a strong preference to use the pronoun in both cases, and so the variant with the gerund phrase might be more economical than that with a relative clause. The same explanation of a wide-spread usage of gerunds may be applied to English language as well :

(10cz) *Když uviděl Tomáše, zavrtěl ocasem.* - no pronouns for the subject

(10ru) *Увидев Томаша, **он** вильнул хвостом. Когда **он** увидел Томаша, **он** завертел хвостом.*

(10en) *Having seen Tomas, **he** waved his tail.   When **he** saw Tomas, **he** waved his tail.*

Non-finite clauses with participles

Participles (verb forms ending in -ing in English) are indefinite forms of a verb that might have various functions in a sentence. Again, the notion of

'participial clause' is rather ambiguous for different researchers. For example, in (Mala, 2006) the participial clauses are what we call gerund clauses in this work. More often they participate in a formation of different verb tenses. Here we do not use the notion of the participle in its common sense, but as a verbal form participating in dependent non-finite constructions, called in Russian "prichastnyj oborot".

For instance, let us have a look at a Czech passive participle "*inspirován*" - "inspired". The following three sentences (11) are the examples of a participial non-finite clause, and examples (12) are the cases when participials just form the past tense.

(11cz) *Profesor, **inspirován** článkem, přednášel o nových problémech.*

(11ru) *Профессор, **вдохновленный** статьей, читал лекцию о новых проблемах.*

(11en) *Professor, **inspired** by an article, lectured on new problems.*

(12cz) *Profesor byl **inspirován** článkem.*

(12ru) *Профессор был **вдохновлен** статьей.*

(12en) *Professor was **inspired** by an article.*

In contrary to transgressive phrases, we can not detect participial phrases in both Czech and Russian text with the help of tags because participles are marked with the same tag – verbal adjective - Vmpp in Russian tag or AG...-----A---- in Czech, and they can indicate a participle in the function of a common attribute, e.g. *вдохновленный профессор - inspired professor.* Moreover, many of them after the process of a lexicalization became nouns, e.g. *vedoucí(cz), ведущий(ru) (leading person).* So we could not provide the statistics of such constructions like in the previous case. Nevertheless, we made an analysis of a small sample of Russian/Czech sentences that contain a participial tag occurring straight after the dot, which can be the indicator (often, but not always) of a participial non-finite phrase. In this case, the Czech translation equivalents of Russian adjectival participial phrase can be:

 1. Dependent constructions, ex. relative clauses:

(13cz) *Známý bankéř, **který_REL** si vydržuje smečku body-guardů...*-(en) *famous banker, who keeps ...*

(13ru) *Известный банкир, **содержащий_par** отряд боевиков..*

(en) *famous banker, keeping*

 2. Czech adjectival phrase in postposition without a comma:

(14cz) *Walzer .. posvítil na přistavený **žebřík vedoucí_par dolů.** -*

(en) *a leadder coming down*

(14ru) *Вальзер .. осветил приставную **лесенку, ведущую_par вниз.***

(en) *a leadder, coming down*

 3. Czech participial phrase in postposition, with a comma – the same in both languages:

(15cz) *Možná že si ruští mafiáni, **pátrající_par** po Libereji, přečetli celý text listu...*

(15ru) *Возможно, русские мафиози, **разыскивающие_par** Либерею, прочли полный текст письма...*

(15en) lit. *Possibly the Russian mafias, **searching_par** for Liberia, had read the letter...*

The difference here is only in punctuation mark - which is obligatory in Russian and non-obligatory in Czech depending on whether or not the participle is a simple attribute(14) or some more complex and independent structure(15cz). Thus, the reason for the differences in usage of those phrases in the two languages is only formal. Maybe in Russian the punctuation mark shows that the construction is more independent of the governing noun than the Czech one.

Applications

The differences described in this paper can pose some challenge in the real world - for language learners, translators and machine translation systems. The first one - language learners are affected least.

Russian learners of Czech language will probably at some time see the paradigm of Czech transgressive, which is more complex than Russian. But they will not encounter many of them in the real world, only while reading some old Czech texts. As for the Czech learners of Russian, they will not find gerund or participial phrases in colloquial speech either, but they will have to recognize it in a written speech, especially in news and formal texts.

Translators from Czech into Russian generally do not change the structure of a Czech text. And on the contrary, when translating Russian non-finite gerund or participle phrases into Czech, they tend to use some other descriptive measures described above, like relative or coordinated clauses.

Probably, those non-finite phrases pose the biggest challenge for Russian-Czech Machine Translation systems, which have no intuition of a human translator and could not always opt for the best translation of a Russian transgressive or participial phrase.

Conclusion

In this paper we have described two non-finite constructions - transgressive and adjective participial phrases in Czech and Russian. Whereas for participial phrases the discrepancies are more formal - in punctuation or ordering, for transgressives they are not so trivial. In Czech they are archaic, in Russian they are not, but they occur mostly in written texts. The commonly known reason for the fact, that transgressives had become archaic is their complex paradigm. We have suggested one more hypothesis. One of the functions of transgressives can be economy of references, as there is no need to use a pronoun referencing to the subject of a main clause. In pro-drop Czech there is no necessity to use it anyway - and in finite constructions as well. On the contrary, in a "good-styled" Russian text it is better to use the reference, so Russian chooses more economical ways to present the sense - transgressive constructions.

Bibliography

Barnet, V. (1962) 'Ke genezi souvětí'. *Otázky slovanské syntaxe : sborník brněnské syntaktické konference*, 17.-21.IV.1961. Vyd. 1. Praha: Státní pedagogické nakladatelství, 1962. [In Czech]

Čermák, F. & A. Rosen (2012) 'The case of InterCorp, a multilingual parallel corpus.' *International Journal of Corpus Linguistics*, 13(3):411–427, 2012.

Dombrowski, A. (2006). 'Towards a Typology of Participial Development: Evidence from Romance and Slavic.' in LSO *Working Papers in Linguistics 6* (Proceedings of WIGL2006).

Malá, M. (2006) 'Some Remarks on Participial Adverbial Clauses from the Point of View of Functional Sentence Perspective' . In: *Acta Universitatis Carolinae - Philologica 2*, 2005, Prague Studies in English, roč. 24, č. 1, 2006, pp 45-58.

Sharoff, S., M. Kopotev, T. Erjavec, A. Feldman & D. Divjak (2008) 'Designing and evaluating Russian tagsets', In *Proc. LREC 2008*, Marrakech, May, 2008.

Zeman, D., J. Hana, H. Hanová, J. Hajič, E. Jeřábek & B. Vidová Hladká (2005) 'A Manual for Morphological Annotation, 2nd edition'. In: *ÚFAL Technical Report*, 27, MFF UK, Prague, Czech Republic, 2005.