

Lexical Semantics and Grammatical Aspects of a Czech Text from the Perspective of a Rule-Based Machine Translation for the Czech-Slovak Language Pair

Martin Diweg-Pukanec
Katedra slovanských filológií
Filozofická fakulta Univerzity Konštantína Filozofa v Nitre

Lexical Semantics and Grammatical Aspects of a Czech Text from the Perspective of a Rule-Based Machine Translation for the Czech-Slovak Language Pair

Litikon, 2024, Vol. 9, No. 2, pp. 139-147

ORCID ID: <https://orcid.org/0000-0003-4175-1727>

This paper describes the development of a shallow-transfer rule-based machine translation system from Czech to Slovak on the Apertium platform, mechanically built by combining the lexical, semantic, and grammatical data contained in the newspaper article of the Lidovky.cz: Schwarzenberg: Na Ježíška věřím pořád, kšeftování s Vánoci mě šťve. A lightweight manual post-processing is carried out in order to add multi-words that are needed. The rule-based machine translation system called Apertium-Apersky (the latter name means from a Slavic language to Slovak on the Apertium platform) is evaluated and outperforms Google Translate system when it comes to this text concerning one of the greatest Christian holidays and centuries-old Christian moral values in aristocratic families.

Keywords: Apertium, Apertium-Apersky, Czech language, machine translation, Slavic languages, Slovak language

1. Introduction

The objective of this paper is to describe the development of a shallow-transfer rule-based machine translation system from Czech to Slovak. Czech and Slovak are closely related West Slavic group languages suitable for this type of machine translation approach. Each of them has its own standard variety, roofing the vernacular dialects within political borders today. The languages are spoken predominantly in what are now two independent countries, the Czech Republic and the Slovak Republic, in Central Europe. The Czech and Slovak linguistic area thus spreads from the borders of Germany in the west to those of Ukraine in the east, and from the borders of Poland in the north to the borders of Austria and Hungary in the south (cf. Nábělková, 2016, p. 140).

In the 10th century, Czech and Slovak were still the same language, which then began to diverge from each other, but the living and written Czech word started to influence the Slovak language as early as the 14th-15th centuries. The 16th century was a turning point in the development of Czech in both Czech and Slovak lands. In the Czech lands, the Czech language was standardised by translation of the Bible. In the Slovak land, Czech was being forgotten and priority was given to Latin. In the 17th and 18th centuries, the Slovaks were already represented by the pre-standard Slovak, although Latin continued functioning as a representative of aristocracy and science and Czech became the cultic language of Slovak evangelicals. The end of the 18th and 19th centuries present the evidence of the rise of standard Slovak, even though in 1849 the Austrian government attempted to banish standard Slovak from schools and offices and to introduce the so-called Old Slovak, i.e. Czech-Slovak of the poet J. Kollár (cf. Krajčovič, 1980, pp. 59, 78, 99, 174, 205).

The codification of the Slovak language in the 18th and 19th centuries signified a definitive linguistic separation of the two languages, in spite of the existence of several Czech-Slovak republics in the course of the 20th century. During the Czechoslovak era, Slovak was even “purified” of bohemisms. This is a proof that standard languages behave a bit differently from natural languages. And standard languages are not always used in a standard way. Czech holds a unique position among languages of Europe by its internal stratification, called diglossia or code-switching. Instead of having one underlying code, it has two, Standard or Literary Czech, and Common Czech, the predominant spoken form of the Czech lands. Although other languages, e.g. French or German, exhibit multiple codes, only in Czech do the two codes share such a central position in the language. The increasing use of Common Czech in mass-media and other areas (cf. Eckert, 1989, p. 84) makes any machine translation system from the Czech language(s) more complicated.

The Czech language (or Czech languages) and Slovak language are, however, still more or less mutually intelligible. The languages are typologically similar. They are medium inflected languages with relatively free word order, sharing verb tenses, present, past, future, three persons, as well as two grammatical numbers, singular and plural (dual being mostly lost), three grammatical genders, and seven cases, nominative, genitive, dative, accusative, vocative, local, instrumental (vocative case being in standard Slovak mostly lost). In addition, languages draw a distinction between animate and inanimate masculine nouns and adjectives. This immensely reduces the complexity of transfer between not only these Slavic languages (cf. Ruth – O'Regan, 2011, p. 69).

2. Apertium-Apersky

Apertium-Apersky is a shallow-transfer rule-based machine translation system from Slavic languages to Slovak on the Apertium platform. The name *Apersky* is intended to sound Slavic because of the Slavic suffix *-sky* which includes *SK* as an Apertium designation (and an international abbreviation) for Slovak, *Aper-* refers to the platform Apertium as well as to the designer of the Apertium-Apersky system (Latin *aper* means ‘(wild) boar’, in Slovak *diviak*, *kanec*). Unlike in the usual Apertium systems, in this system a lot of words, including some verbs, are treated as nouns, which is a great reduction of work for a designer.

The Apertium platform is one of the open-source machine translation systems which originated within the project Open-Source Machine Translation for the Languages of Spain,

founded by the central Spanish government. It is basically a shallow-transfer machine translation system, initially designed for the translation between related language pairs. In 2010, Apertium could translate between the pairs Spanish-Galician, Spanish-Catalan, Catalan-Occitan, Catalan-French, and was supposed to be used to build translators between other related language pairs, such as Danish-Swedish or Czech-Slovak (cf. Forcada et al., 2010, pp. 1-2). However, no developed Czech-Slovak machine translation system on the Apertium platform has been designed until today. It is therefore a challenge to apply the success of projects based on the Apertium platform for closely-related languages in the Romance family to this Slavic language pair.

The design of the Apertium platform makes it possible to produce machine translation systems that are fast (translating tens of thousands of words per second on ordinary desktop computers) and that achieve results that are, in spite of the errors, reasonably intelligible and easily correctable. In the case of related languages such as the ones involved in the project (Spanish, Galician, Catalan) or some others (e.g. Italian, Sardinian; Tyers et al., 2017), a mechanical word-for-word translation (with a fixed equivalent) would produce errors that, in most cases, can be solved with a quite rudimentary analysis (a morphological analysis followed by a superficial, local and partial syntactic analysis) and with an appropriate treatment of lexical ambiguities (mainly due to homography). The design of the Apertium system follows this approach with very interesting results (Forcada et al., 2010, p. 5).

The Apertium architecture uses finite-state transducers for lexical processing (thanks to four modules: morphological analyser, lexical transfer module, morphological generator and post-generator), hidden Markov models for part-of-speech tagging and finite-state-based chunking for structural transfer. Finite-state transducers for lexical processing are a type of finite-state automata, which may be used as one-pass morphological analysers and generators. The part-of-speech tagger is based on first-order hidden Markov models (cf. Rabiner, 1989), that is, on statistical data. The structural transfer system is made of three modules, the first one of which (called *chunker*) can be used in isolation in order to run a shallow-transfer system (which is the transfer system used so far for related language pairs such as Spanish-Catalan or Spanish-Galician). When the system is used for less related language pairs and, therefore, an advanced transfer becomes necessary, the three transfer modules are executed (cf. Forcada et al., 2010, pp. 5, 17, 56, 73).

3. Development

The modules Apertium-Apersky and all the typical Apertium systems consist of are the following (cf. Forcada et al., 2010, pp. 6-9):

- The *de-formatter*, which separates the text to be translated from the format information (RTF, HTML, etc.).
- The *morphological analyser*, which tokenizes the text in surface forms (lexical units as they appear in texts) and delivers, for each surface form, one or more lexical forms consisting of lemma (the base form commonly used in classic dictionary entries), the lexical category (noun, verb, etc.) and morphological inflection information (number, gender, person, tense, etc.). Tokenization of a text in surface forms is not straightforward due to the existence, on the one hand, of contractions (in English: *didn't*, *can't*) and, on the other hand, of lexical units made of more than one word (in English: *in front of*, *taken into account*). Lexical units made of

more than one word (multi-words) are treated as single lexical forms and processed specifically according to its type.

- The *part-of-speech tagger* chooses, using a statistical model (hidden Markov model), one of the analyses of an ambiguous word according to its context. A sizeable fraction of surface forms are ambiguous, that is, they can be analysed into more than one lemma, more than one part-of-speech or have more than one inflection analysis, and are therefore an important source of translation errors when the wrong equivalent is chosen.

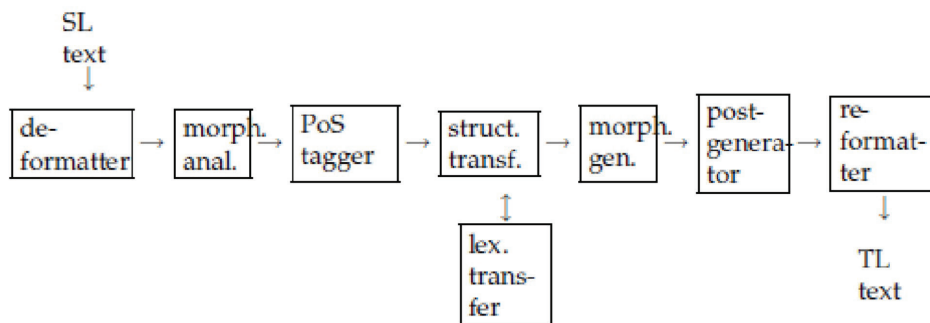
- The *lexical transfer module* that uses a bilingual dictionary and is called the structural transfer module, reads each lexical form of the surface form and delivers the corresponding target language lexical form. The dictionary contains a single equivalent for each source language lexical form; that is, no word-sense disambiguation is performed. Multi-words are translated as a single unit.

- The *structural transfer module*, which detects and processes patterns of words (chunks or phrases) that need special processing due to grammatical divergences between the two languages (gender and number changes, word re-orderings, changes in prepositions, etc.).

- The *morphological generator* that, from a lexical form in the target language, generates a suitably inflected surface form.

- The *post-generator* that performs some orthographic operations in the target language such as contractions and apostrophations.

- The *re-formatter*, which restores the original format information into the translated text.



Ilust. 1: The modular architecture of a typical Apertium machine translation platform.

The Apertium-Apersky system for the Czech-Slovak language pair is based on the Apertium machine translation platform. The platform was originally aimed at the Romance languages of the Iberian Peninsula, but has also been adapted for other language pairs, and in particular languages from the Celtic group, e.g. Welsh (cf. Tyers – Donnelly, 2009), and the Germanic group (cf. Tyers – Nordfalk, 2009). The whole platform, both programs and data, are licensed under the Free Software Foundation's General Public Licence (GPL) and all the software and data for the supported language pairs (and the other pairs being worked on) are available for download from the project website. The development of the Czech-Slovak pair owes a lot to previous work on other language pairs, including Slavic ones.

Basic Apertium linguistic data for both Slovak (<https://github.com/apertium/apertium-slk>) and Czech can be found (<https://github.com/apertium/apertium-ces>) on the related website. On the website, there is also an Apertium translation pair for Czech and Slovak (<https://github.com/apertium/apertium-ces-slk>), however simple it is. None of the materials are at an advanced stage of development and Czech materials obviously ignore the existence of Common Czech. Moreover, the Apertium-Apersky system for the Czech-Slovak language pair is simplified in that way that a lot of words, including some verbs, are treated as nouns. The above mentioned linguistic data must therefore be manually checked and revised, and the system is mostly mechanically built. A helpful set of initial transfer rules is located in the incubator (which houses works in early development) for the Polish-Slovak language pair (<http://svn.code.sf.net/p/apertium/svn/incubator/apertium-pl-sk/>), since Polish is typologically similar to Czech and Slovak languages. A more advanced set of rules can be found for the Polish-Czech language pair (<https://github.com/apertium/apertium-pol-ces>). The rules for Slavic language pairs such as Bulgarian-Macedonian (cf. Rangelov, 2011) or Serbian/Croatian-Macedonian (cf. Peradin – Tyers, 2012) are of no use, since both Bulgarian and Macedonian are not grammatically complex.

4. Evaluation

The *ad hoc* evaluation of the quality of translation has been based on the newspaper article of the Lidovky.cz Schwarzenberg: *Na Ježíška věřím pořád, kšeftování s Vánoci mě štve* (Plavcová, 2014), and the comparison of the Apertium-Apersky system for the Czech-Slovak language pair with the Google Translate system for the same language pair. The article, its part, was supposed to be about Christmas and centuries-old Christian moral values in aristocratic families. The main part of the newspaper article (which is actually an interview) on Christmas contains the first paragraph of the article, the first question and the first answer, plus we have included the third question about manners/morals and the third answer for research, for a total of 247 words, manually added to the Apertium-Apersky system where necessary. Thus, the lexical coverage of the system under development has not been evaluated.

The quantitative evaluation has not been performed, but the qualitative evaluation has been done to determine where the system can be improved. As Czech and Slovak are closely related languages sharing the same history, manual evaluation has shown that lexical and semantic errors are rare. The Google Translate system for the Czech-Slovak language pair translated Czech *krucipísek* 'God damn it' > Slovak *krucipánik* (the word in Slovak does not exist), three Czech adjectives *český* 'Czech' were two times surprisingly translated into Slovak as *slovenský* 'Slovak', and Czech adverb *poblíž* 'near' was translated into Slovak as noun *blízkosti* 'proximity', the correct form could have been *v blízkosti* 'in proximity'. The third question, which is interesting from the point of view of ethics: *Pocitujete i o svátcích úbytek dobrých mravů?* 'Do you feel a decline of good manners/morals even during the holidays?', was translated correctly by Google Translate, but, for example, translator DeepL translated it with a clear semantic shift: *Pocitujete stratu dobrých mravov aj počas sviatkov?* 'Do you feel a loss of good manners/morals even during the holidays?', apparently because this context is more common in Slovak. Semantically, however, it is a mistranslation, which is particularly socially sensitive with moral words.

Since the Apertium-Apersky system for the Czech-Slovak language pair was mechanically built by combining the lexical, semantic, and grammatical data contained in the article, it translated not only the sentence but also all three of the above-mentioned words correctly: Czech *krucipísek* > Slovak *krucinál*, 3x Czech *český* > Slovak *český*, Czech *poblíž* > Slovak *blízko*. However, it may be rightfully supposed that the Apertium-Apersky system would have translated the sentence and the words correctly anyway. Both systems translated correctly Czech multi-word *(po) mši svaté* (noun-adjective) '(after) Holy Mass' > Slovak *(po) svätej omši* (adjective-noun), but a lightweight manual post-processing was carried out in order to add multi-words that were needed in the Apertium-Apersky system. Both systems translated incorrectly the Czech pronoun *je* 'them' > Slovak verb *je* 'is' because of homonymy and higher statistical position of the Czech form of the verb *být* 'to be', the correct translation would have been Slovak *ich*.

As Czech and Slovak languages are typologically similar, sharing relatively free word order, verb tenses, three persons, two grammatical numbers, three grammatical genders, seven cases and so on, manual evaluation has shown that grammatical errors are rare, too. The main problem of the Google Translate system is the declension of the proper nouns. These are treated by the Apertium-Apersky system correctly. The proper nouns *(z rodu) Fürstenbergů'* (of) the Fürstenbergs (family)' (genitive plural, anthroponym) and *(Maxmilián) z Hohenbergu* '(Maximilian) of Hohenberg' (nobiliary particle + toponym in genitive singular) are translated into Slovak as *(z rodu) Fürstenberg* (nominative singular) and *(Maxmilián) z Hohenberg* (*z* = genitive preposition, *Hohenberg* = nominative singular) respectively by the Google Translate system but *(z rodu) Fürstenbergovcov* and *(Maxmilián) z Hohenbergu* respectively by the Apertium-Apersky system, which is correct. The latter, rule-based system offers a very good possibility to treat this kind of problem as is shown below:

M. Animate: <e lm="Maximilán"><i>Maximilán</i></e><par n="chlap__n"/></e> (the type *chlap* at least in singular)

M. Inanimate: <e lm="Hohenberg"><i>Hohenberg</i></e><par n="rod__n"/></e> (the type *rod* being preferred to the type *dub* because of genitive singular)

M. Animate Familial: <e lm="Hohenberg"><i>Hohenberg</i></e><par n="Novák__n"/></e>

<e lm="Fürstenberg"><i>Fürstenberg</i></e><par n="Novák__n"/></e>

<e lm="Schwarzenberg"><i>Schwarzenberg</i></e><par n="Novák__n"/></e>

<pardef n="Novák__n">

<e><p><l></l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="nom"/></r></p></e>

<e><p><l>a</l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="gen"/></r></p></e>

<e><p><l>ovi</l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="dat"/></r></p></e>

<e><p><l>a</l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="acc"/></r></p></e>

<e><p><l></l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="voc"/></r></p></e>

<e><p><l>ovi</l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="loc"/></r></p></e>

<e><p><l>om</l><r><s n="n"/><s n="ma"/><s n="sg"/><s n="ins"/></r></p></e>

<e><p><l>ovci</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="nom"/></r></p></e>

<e><p><l>ovcov</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="gen"/></r></p></e>

<e><p><l>ovcom</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="dat"/></r></p></e>

<e><p><l>ovcov</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="acc"/></r></p></e>

```
<e><p><l>ovci</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="voc"/></r></p></e>
<e><p><l>ovcoch</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="loc"/></r></p></e>
<e><p><l>ovcami</l><r><s n="n"/><s n="ma"/><s n="pl"/><s n="ins"/></r></p></e>
</pardef>
```

The biggest translation problem of both systems, however, is how to treat Czech adjectives of the type *jarní* (soft-ending adjectives). The adjective case homonymy rate is too large, cf. the form *jarní*:

Singular	Plural			
	M. Inanimate	Feminine +Neuter	Masculine+Neuter+Feminine	
Nominative	<i>jarní</i>	<i>jarní</i>	<i>jarní</i>	<i>jarní</i>
Genitive			<i>jarní</i>	
Dative			<i>jarní</i>	
Accusative		<i>jarní</i>	<i>jarní</i>	<i>jarní</i>
Vocative	<i>jarní</i>	<i>jarní</i>	<i>jarní</i>	<i>jarní</i>
Local			<i>jarní</i>	
Instrumental			<i>jarní</i>	

That is why the translations of the adjectives of this type are often incorrect: Czech *knížecí vyprávění* ‘prince’s story (telling)’ (nominative singular neuter) > both incorrect Slovak *kniežací rozprávanie* (Google Translate) / *kniežací rozprávanie* (Apertium-Apersky, *kniežací* = nominative singular masculine), the correct form being *kniežacie rozprávanie* or, even better, *rozprávanie kniežaťa*; Czech (*včetně*) *nejznámější vánoční písně (světa)* ‘(including) the best-known Christmas song (in the world)’ (*nejznámější, vánoční, písně* = genitive singular feminine) > incorrect Slovak (*vrátane*) *najznámejšie vianočné piesne (sveta)* (Google Translate, *najznámejšie, vianočné* = nominative plural feminine) / (*vrátane*) *najznámejší vianočný piesne (sveta)* (Apertium-Apersky, *najznámejší, vianočný* = nominative singular masculine) before multi-word treatment / (*vrátane*) *najznámejší vianočné piesne (sveta)* (Apertium-Apersky, *najznámejší* = nominative singular masculine, *vianočné* = nominative plural feminine) after multi-word treatment, the correct form being (*vrátane*) *najznámejšej vianočnej piesne (sveta)*.

In addition, the other type of Czech basic adjectives, the type *mladý* (hard-ending adjectives) sometimes presents grammatical problems of the same kind: Czech *rakouské a české Vánoce* ‘Austrian and Czech Christmas’ (nominative plural) > both incorrect Slovak *rakúskej a českej Vianoce* (Google Translate, *rakúskej, českej* = genitive singular feminine, *Vianoce* = nominative plural feminine) / *rakúskych a českých Vianoce* (Apertium-Apersky, *rakúskych, českých* = accusative plural masculine animate, *Vianoce* = nominative plural feminine), the correct form being *rakúske a české Vianoce*. Further transfer rules will be needed to solve this problem properly.

Except for these two systematic grammatical errors, the former being treated correctly by the Apertium-Apersky system, the latter being treated chaotically by both systems, there are some more or less random errors concerning the verbs in the translations: a typical Common Czech (*já*)... *vyrost* ‘I grew up’ (Standard or Literary Czech (*já*) *jsem vyrostl*) > both incorrect

Slovak (*ja*)... *vyrástol* (Google Translate) / (*ja*)... *vyrástol* (Apertium-Apersky), the correct form being (*ja*)... *som vyrástol*; Czech (*já*) *se*... *zeptal* 'I asked' > incorrect Slovak (*ja*) *sa*... *spýtal* (Google Translate) / correct Slovak (*ja*) *som sa*... *spýtal* (Apertium-Apersky), because the reflexive *se* after a pronoun in Czech makes it easy to treat the problem in the Apertium-Apersky system; Czech *Vánoce*... *se liší* (3rd person plural) > correct Slovak *Vianoce*... *sa líšia* (Google Translate, 3rd person plural) / incorrect Slovak *Vianoce*... *sa líši* (Apertium-Apersky, 3rd person singular). The latter ambiguity exists in Czech but does not exist in Slovak. It is not easy to be solved for the Apertium-Apersky because both Czech and Slovak are pro-drop languages, which means that the subject pronoun can be omitted since it can be usually inferred from the verb form.

5. Conclusion

The first translator between free word order languages developed for the Apertium platform was developed and presented in 2011. It was a translator from Czech to Polish and its results were far from impressive (Ruth – O'Regan, 2011, p. 75). However, the results of the *ad hoc* test of a rule-based machine translation system for the Czech-Slovak language pair seem to be much more promising. This is partly due to the fact that the system does not ignore the existence of Common Czech and that Czech and Slovak are more closely related languages to each other than Czech as a source language is related to Polish.

A significant part of the above-mentioned translation problems is typical for a system under development. There remains a lot of work to do on improving both the quality of transfer rules of the system and their quantity, since the number of disambiguation rules is too small. But in terms of future work, the essential task, of course, is to increase coverage. Nevertheless, the rule-based machine translation system called *Apertium-Apersky* qualitatively outperforms Google Translate system for the Czech-Slovak language pair, which is a very good beginning for the development of a shallow-transfer rule-based machine translation system from Slavic languages to Slovak.

REFERENCES

- ECKERT, Eva. 1989. Standard and Common Czech: Attitudes and Usage. *Bohemia*, 1989, Vol. 30, No. 1, pp. 84 – 93. ISSN 0523-8587.
- FORCADA, Mikel L., et al. 2010. Documentation of the Open-Source Shallow-Transfer Machine Translation Platform Apertium [online]. Alicante: Departament de Llenguatges i Sistemes Informàtics Universitat d'Alacant, 2010 [cit. 31.07.2024]. Available at: <https://aclanthology.org/www.mt-archive.info/Forcada-2010.pdf>.
- KRAJČOVIČ, Rudolf. 1980. *Svedectvo dejín o slovenčine*. Martin: Matica slovenská, 1980.
- NÁBĚLKOVÁ, Mira. 2016. The Czech-Slovak Communicative and Dialect Continuum: With and Without a Border. In: KAMUSELLA, Tomasz – NOMACHI, Motoki – GIBSON, Catherine (eds.). *The Palgrave Handbook of Slavic Languages, Identities and Borders*. London: Palgrave Macmillan, 2016, pp. 140 – 184. ISBN 978-1-137-34838-8.
- PERADIN, Hrvoje – TYERS, Francis. 2012. A rule-based machine translation system from Serbo-Croatian to Macedonian. In: ESPAÑA-BONET, Cristina – RANTA, Aarne (eds.). *Proceedings of the Third International Workshop on Free/Open-Source Rule-Based Machine Translation* [online]. Gothenburg: Department of Computer Science and Engineering Chalmers University of Technology and University of Gothenburg, 2012, pp. 55 – 64 [cit. 31.07.2024]. Available at: <https://aclanthology.org/volumes/2012.freeopmt-1/>.

- PLAVCOVÁ, Alena. 2014. Schwarzenberg: Na Ježíška věřím pořád, kšeftování s Vánoci mě štve [online]. *Lidovky.cz*, 2014 [cit. 31.07.2024]. ISSN 1213-1385. Available at: https://www.lidovky.cz/domov/schwarzenberg-na-jeziska-verim-porad-kseftovani-s-vanoci-me-stve.A141218_172510_In_kultura_mmu.
- RABINER, Lawrence, R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 1989, Vol. 77, No. 2, pp. 257 – 286. ISSN 0018-9219.
- RANGELOV, Tihomir. 2011. Rule-based machine translation between Bulgarian and Macedonian. In: SÁNCHEZ-MARTINEZ, Felipe – PÉREZ-ORTIZ, Juan Antonio (eds.). *Proceedings of the Second International Workshop on Free/Open-Source Rule-Based Machine Translation* [online]. Barcelona: Universitat Oberta de Catalunya, 2011, pp. 53 – 59 [cit. 31.07.2024]. Available at: <https://aclanthology.org/volumes/2011.freeopmt-1/>.
- RUTH, Joanna – O'REGAN, Jimmy. 2011. Shallow-transfer rule-based machine translation from Czech to Polish. In: SÁNCHEZ-MARTINEZ, Felipe – PÉREZ-ORTIZ, Juan Antonio (eds.). *Proceedings of the Second International Workshop on Free/Open-Source Rule-Based Machine Translation* [online]. Barcelona: Universitat Oberta de Catalunya, 2011, pp. 69 – 76 [cit. 31.07.2024]. Available at: <https://aclanthology.org/volumes/2011.freeopmt-1/>.
- TYERS, Francis M. – DONNELLY, Kevin. 2009. *apertium-cy* – a collaboratively-developed free RBMT system for Welsh to English. *The Prague Bulletin of Mathematical Linguistics*, 2009, No. 91, pp. 57 – 66. ISSN 0032-6585. ISBN 978-80-904175-1-9.
- TYERS, Francis M. – NORDFALK, Jacob. 2009. Shallow-transfer rule-based machine translation for Swedish to Danish. In: SÁNCHEZ-MARTINEZ, Felipe – PÉREZ-ORTIZ, Juan Antonio – TYERS, Francis M. (eds.). *Proceedings of the First International Workshop on Free/Open-Source Rule-Based Machine Translation* [online]. Barcelona: Universitat Oberta de Catalunya, 2009, pp. 27 – 33 [cit. 31.07.2024]. Available at: <https://aclanthology.org/2009.freeopmt-1.0/>.
- TYERS, Francis M., et al. 2017. Rule-Based Machine Translation for the Italian-Sardinian Language Pair. *The Prague Bulletin of Mathematical Linguistics*, 2017, No. 108, pp. 221 – 232. ISSN 0032-6585.

Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V04-00670 *Inverted Morality: Reversed Semantics in Old Church Slavonic Moral Words (OXYMORAL)*.

.....
prof. Mgr. Martin Diweg-Pukanec, PhD.
Katedra slovanských filológií
Filozofická fakulta Univerzity Konštantína Filozofa v Nitre
Štefánikova 67
949 01 Nitra
Slovenská republika
mdpukanec@ukf.sk