ARTIFICIAL HOMONYMY

JAROSLAVA HLAVÁČOVÁ

Institute of the Formal and Applied Linguistics, Charles University, Prague, Czech Republic

Abstract: The paper presents a discussion of homonymy of Czech nouns with different or varying genders. The lemmas with this type of homonymy are treated in the new release of the MorfFlex dictionary as separate. We show that the separation of paradigms according to the gender is not only superfluous, but also clumsy, because it forces a choice when making one is not necessary. That is why we call this type of hononymy "artificial".

Keywords: homonymy, polysemy, gender variation, dictionary

1   BASIC CONCEPTS

There are several definitions of homonymy. For the purpose of this paper, we will use the following one:

Homonyms are words with the same spelling but accidentally different meanings.

The definition concerns only one part of homonymy, namely the homography (identical written forms). There is also the homophony (identical pronunciation), but in this paper, only homography will be dealt with under the term homonymy.

There are two terms in this definition that might be a source of misunderstanding.

As for the "words", there are two basic ways how to capture them: as individual wordforms or as lemmas. Artificial homonymy relates to the homonymy of lemmas.

The more difficult term in the definition is the one of "meaning". No unambiguous, simple definition of the "meaning" exists. That is the reason why we will use the term meaning in accordance with "common sense". We consider two lemmas homonymous if their meanings are not connected by any means. In other words, if their spelling is the same only by chance. The example is the lemma *kolej,* which has two independent meanings in Czech: 1. a housing facility for students (college, dormitory), 2. a track or rail.

On the other hand, if a word is used in a figurative meaning, for instance as a metaphor, we consider it "only" polysemous, not homonymous, though we are

aware that the distinction between the two terms is fuzzy.[1] Thus, the lemma *ušák* (1. a hare with big ears, 2. a chair with "ears" resembling a hare, or a pot with big handles resembling ears of a hare) is polysemous, not homonymous, because all the meanings relate to the ears. Another example of the polysemy is using the same proper name for a personal name, as well as for the name of his or her company or firm (e.g. *Albert*). We will discuss individual types of such polysemy later.

If two words with the same spelling belong to different parts of speech, we always consider them homonymous, without regard to their meanings. In other words, in our interpretation, words of different parts of speech have always different meanings.

Such "inter-POS" homonymy is widespread in English, where many words (lemmas) can be used as a verb, a noun, and an adjective. An example is the lemma *house*. The following examples are from the British National Corpus (BNC).[2]
The guard was still in the <u>house</u> (noun).
The practice in medieval times was to <u>house</u> (verb) *all of the grain crops in the barn.*
The <u>house</u> (adjective) *door was locked.*

In the rest of the paper, we will cope with the Czech language only.

1.1 Homonymy in Czech

With its rich morphology, the homonymy is very common in Czech, but not so much in our sense – among the lemmas.[3] Contrary to English, there are only several lemmas that can be used as different parts of speech similarly to the above English example with the *house*. It does not mean that homonymy does not exist. Lemmas with the same spelling and different meanings do exist in Czech. However, we want to show that there is a large set of homonymous lemmas where the homonymy (in the sense presented above) is "artificial". In other words, it is not necessary to call it homonymy as we are convinced that there are not two, but only one word with polysemous property.

There are two basic types of artificial homonymy.

The first one is the homonymy of nouns with different genders, the second one is the homonymy among adverbs, particles, conjunctions, possibly also interjections and prepositions.

The latter type is highly dependent on the definitions of the parts of speech included in the list. This is the reason why we lay aside this type of homonymy. This

---

[1] The simple distinction between homonymy and polysemy is given by dictionaries – polysemous words usually have a single headword, while homonymous ones are divided into more headwords.
[2] Data cited herein have been extracted from the British National Corpus, distributed by the University of Oxford on behalf of the BNC Consortium. All rights in the texts cited are reserved. The examples are from texts A03, A79 and A0N.
[3] The Czech language is especially rich in the so called morphological homonymy – homonymy among word forms. See the impressive treatise in Petkevič [1].

problem is very complex and its scope extends beyond the possibilities of this paper. Our recent paper deals only with the homonymy of nouns.

2   MOTIVATION – NOUNS WITH VARYING GENDER

The basis of our study is the Czech morphological dictionary MorfFlex CZ ([2], [3]), examples were taken from the corpora of the SYN[4] [4] and Aranea[5] [5] series. For its latest edition, version 2.0, several principles were applied to make the content of MorfFlex consistent (see [2]). One of those principles is the "Principle of unique paradigm" saying that there are no two identical paradigms (sets of lemma-tag pairs) in the dictionary. It means that every paradigm has only one lemma, even if it has more meanings. This rule was adopted for the reason of simplicity. The lemma *kolej* presented above, is a typical example.

Another basic principle – "Principle of morphological differentiation" – implies that nouns with different genders are different.

The Czech language has three genders – masculine, feminine, and neuter. The masculine gender may be animate or inanimate. These two subgenders have partially different inflections. It was probably the reason why they are usually considered two separate genders in the field of NLP. Thus, in most of the Czech morphological tagsets, there are codes for 4 genders: masculine animate, masculine inanimate, feminine, and neuter.

The great majority of Czech nouns have a single gender within their paradigm. However, there are nouns with varying genders.

There are two basic ways how to describe that situation morphologically. The way adopted by the authors of the new version of the MorfFlex (see above) was the division of the paradigm with the varying gender into more paradigms, each having all the wordforms of a single gender. In such way, the paradigms became separated, each represented by its own lemma. As both lemmas have the same spelling, they become homonymous. Technically, in the morphological dictionary, they are distinguished by means of a numerical index added to the lemma. See the example of the word *kredenc* in Tab. 1.

In our view, this solution is superfluous and the resulting homonyms are artificial. We suggest another solution – rejection of the part of the Principle of morphological differentiation concerning the noun gender. There is no reason why wordforms within one paradigm should have only one gender. It is even in contradiction with the reality.

Let us illustrate both approaches on an example with the lemma *kredenc*. In the present version of the dictionary, we have *kredenc-1* with the masculine inanimate

---

[4] Accessible at: http://www.korpus.cz.
[5] Accessible at: http://unesco.uniba.sk/aranea/.

gender and *kredenc-2* with the feminine gender. The meaning of the both is the same – a cupboard.

If we admitted both genders in the same paradigm, there would be no need to have two lemmas. The set of wordform-tag pairs will be the union of the pairs from both paradigms in the former approach (see Tab. 1).

There are several more types of lemmas divided under that principle. They will be discussed in the following sections.

In the field of NLP, the only thing that should not be violated is the Golden rule of Morphology ([2], [6], [7]) saying that every combination of a lemma and a morphological tag must not be represented by more than one wordform. If the two wordforms with the same lemma differ in their gender, even if the rest of the morphological features is identical, their tags do differ, which is a sufficient condition for meeting the Golden rule requirement. However, there are some issues that have to be mentioned and resolved.

2.1 Lemma of a varying gender paradigm

If we merge the paradigms of artificial homonyms with a varying gender, a question may arise what will the gender of its lemma be? The spelling of the lemma is unique, but within a merged paradigm, it can be assigned two genders, depending (only) on the context.

		kredenc			
kredenc-1	*kredence*	NNIP1-----A----	*kredence*	NNFP1-----A----	*kredenc-2*
	kredenců	NNIP2-----A----	*kredencí*	NNFP2-----A----	
	kredencům	NNIP3-----A----	*kredencím*	NNFP3-----A----	
	kredencum	NNIP3-----A---6			
	kredence	NNIP4-----A----	*kredence*	NNFP4-----A----	
	kredence	NNIP5-----A----	*kredence*	NNFP5-----A----	
	kredencích	NNIP6-----A----	*kredencích*	NNFP6-----A----	
	kredenci	NNIP7-----A----	*kredencemi*	NNFP7-----A----	
	kredencema	NNIP7-----A---6	*kredencema*	NNFP7-----A---6	
	kredenc	NNIS1-----A----	*kredenc*	NNFS1-----A----	
	kredence	NNIS2-----A----	*kredence*	NNFS2-----A----	
	kredenci	NNIS3-----A----	*kredenci*	NNFS3-----A----	
	kredenc	NNIS4-----A----	*kredenc*	NNFS4-----A----	
	kredenci	NNIS5-----A----	*kredenci*	NNFS5-----A----	
	kredenci	NNIS6-----A----	*kredenci*	NNFS6-----A----	
	kredencem	NNIS7-----A----	*kredencí*	NNFS7-----A----	

Tab. 1. Merged paradigms of the lemma *kredenc-1* and *kredenc-2*

In the sentence (1), the adjective wordform *malou* is described by the lemma *malý* (small) and the tag AAFS4----1A---- (the 3rd position F says that it is feminine), while in the sentence (2) the lemma is the same, *malý*, but the tag differs in the code I for the gender (masculine inanimate) at the position 3: AAIS4----1A----. From the forms of the adjective, the gender of the noun *kredenc* is deduced. The noun *kredenc* can have the same lemma in both examples, but its tags will differ in gender. Both sentences mean the same: 'We have a small cupboard.'

(1) *Máme malou kredenc.* (feminine)
(2) *Máme malý kredenc.* (masculine inanimate)

The answer is simple. There is no need to assign any gender to the lemma. The morphological tag is not part of the lemma. The lemma is a wordform in nominative (usually singular, but there are also pluralia tantum – see later). Its written form can be described with two tags, which differ in gender, but the lemma itself is unique.

2.2 Gender of undistinguishable wordforms

Another problem could be assigning a gender to a wordform with an undistinguishable gender in the given context. An example is the sentence (3), where it is not clear which gender is the right one:

(3) *Máme kredenc.* ('We have a cupboard.')

It differs from the previous examples (1) and (2) by the missing adjective – there is no clue how to decide about the gender of the wordform *kredenc*. Thus, in the present setting of the dictionary, with two lemmas for *kredenc*, it is necessary to choose one of them arbitrarily. If we reject the artificial homonymy, the lemma assignment is easy. However, the necessity of a choice will not disappear. We still have to choose between the two genders, more precisely – between the tags with different genders, because the lemma is now unique – *kredenc*.

In fact, the necessity of choice is the same. Evidently it is not important, the gender can be assigned randomly in such cases. Sometimes, an objection appears, that it is necessary that the gender should be the same throughout a single text. With sofisticated advanced automatic tools this is achievable, or will probably be soon. However, very often even human authors are not consistent throughout a single text. That is why automatic tools need not be consistent, either. It follows that the selection of an appropriate gender for such an occurrence can really be arbitrary.

There are several ways how to decide on a unique solution for each word in a context without any clue for its gender. The gender selection might be random, or according to a criterion. The simplest solution would be an ordering according to a gender preference, the same for all the lemmas. The most natural would probably be this one: M, I, F, N. The rule for the gender assignment would be: If the gender of a wordform cannot be decided from the context, pick one which is the leftmost in the above list. According to that simple rule, the gender of *kredenc* from the example (3) would be I (masculine inanimate).

3 GENDER COMBINATIONS AMONG NOUNS

Let us have a look at possible combinations of genders in the set of all artificial homonyms from the morphological dictionary MorfFLex CZ 2.0.

The following sections will have the names according to codes of genders: M masculine animate, I masculine inanimate, F feminine, N neuter.

3.1 MN

There is 15 lemmas of that kind in MorfFlex, two of them being problematic, possibly wrong. The majority of them (9) are lemmas ending wigh –*e* or –*ě*. They are old words denoting mainly members of nobility (*hrabě* – 'earl', *markrabě* – 'margrave' etc.). They have very unusual morphology for masculine gender in Czech. All these words have the paradigm typical for the neuter gender. However, according to contexts, both genders are plausible. In corpora, we can find expressions or sentences such as:

hrabě hohenembský (masculine) 'Earl of Hohenemb';

V 18. století ji vlastnili *hrabata* z Bubnu (masculine) 'In the 18th century, it was owned by the earls of Buben';

S nimi spříznění *hrabata* Stadničtí ho drželi až do počátku druhé světové války. (masculine) 'earls of Stadnicty, related to them, owned it until the outbreak of World War II';

V návštěvních knihách pak čteme další jména, svědčící o tom, že zámek navštěvovala knížata, *hrabata* (neuter) 'In guest books, we read further names, which shows that the castle was visited by counts and earls';

Hrabata Desfours-Walderodové na Dřínově byla podle vypravování typická *hrabata*, jak je známe z anekdot a divadelních frašek. (neuter) 'The earls of Desfours-Walderod on Dřínov were reputed to be typical earls, as we know them from anecdotes and theatrical farces'.

There are two more words with a typical neuter ending: *pako* ('nitwit') and *libero* ('libero'). The former one is colloquial, appearing often in a nonstandard context, the latter one belongs to sport slang. Their gender really varies but we can see the tendency to use neuter gender in singular and masculine gender in plural (Tobě *podobní paka* mě opravdu nepřekvapí. (masculine animate) 'Nitwits like you won't really surprise me.')

Finally, there is the word *cockney*. In the neuter gender, it means a dialect, in the masculine gender, a man speaking in that dialect. This particular gender selection may be subject to discussion but in any case, *cockney* is one word and as such should be represented with a single lemma. The same is true for the whole group.

3.2 IN

All the words in this group are loanwords. Some of them originally come from ancient Latin or Greek (*ostrakon* – 'ostracon' – a piece of ancient pottery) where it has the neuter gender, but after becoming part of the Czech vocabulary, people

started to decline them according to their formal ending, which resembles the Czech masculine gender. Thus, in texts, both declensions appear.

Some of the words are loanwords with unstable declension (*blues*, *interview*).

There is only one word in this group where the separation is reasonable. *House-1* is the Czech word meaning a gosling (young goose) and *house-2* is a loanword with the English meaning house, but with the Czech declension. Here, strict separation of genders is justified, as the identical spelling is only incidental. They are true homonyms.

3.3 IM

The combination of animate and inanimate masculine declension contains 571 nouns. There are four main semantic groups:

A: inanimate lemmas used in figurative meaning for masculine animate persons. For instance *truhlík* ('box')*, věchýtek* ('bundle of straw')*, hajzl* ('toilet' – vulgar)*, klenot* ('jewel');

B: animate lemmas used in figurative meaning for inanimate things: *špaček* ('starling'), *hlemýžď* ('snail'), *ušák* ('rabbit' or 'hare'), *žralok* ('shark');

C: a tool or a person doing the same thing as the tool: *kompilátor* ('compilator'), *konstruktor* ('constructor'), *komunik*átor ('comunicator'), *dělič* ('divider'), *držák* ('holder');

D: other nouns used in both genders with the same or very similar meaning: *tenor* ('tenor' as a man or as a voice), *exot* ('freak'), člen ('member'), *solitér* (something or someone appearing uniquely).

The groups A and B contain lemmas that appear often in an expressive metaphorical meaning. It is the reason why we cannot proclaim their meaning independent. They are polysemous, but not homonymous.

The group C contains nouns denoting either men, or tools/means of an activity (*constructor* is someone or something that constructs, *držák* ('holder') is someone or something that holds, etc.). There are contexts in which it is even not possible to guess the correct gender. In such case, however, the gender should probably be consistent throughout a single text. Thus, the precedence rule would not be appropriate here.

The group D contains words with very similar meanings, the gender of which often cannot be distinguished even in some contexts. For instance in the sentence: *Tenor* se většinou ... vůbec neprosadí. ('The tenor usually doesn't succeed at all.') it is not clear, if the *tenor* is a singer (animate), or his voice (inanimate).

There is one word in the group IM that is really homonymous: *rys*. Its two meanings are not connected (animate 1. lynx, inanimate 2. feature), the same spelling is accidental. In this case, the two paradigms with two lemmas are reasonable. In the inanimate gender, the word *rys* has more meanings, but according to the Principle of morphological differentitaion, there is only one lemma with that gender, regardless of more meanings.

3.4 FI

Nouns having both genders, feminine and masculine inanimate, are typical words with varying gender. One of them – *kredenc* – has been already discussed. Many Czech names of geographical objects, villages and towns, belong to this group. Very often only inhabitants of the place know the correct gender, as it is usually a matter of dialect or tradition. A famous example is the Moravian town of *Olomouc,* but there are many other (*Bubeneč, Černíč, Dobrovíz, Radom*). There are even different villages with the same name but different genders, according to the local tradition, but it is not reasonable to have two lemmas for them, as the usage varies in those cases, too.

This type of gender variation comprises also nouns that appear only in plural – the so called pluralia tantum. It is usually impossible to guess their gender from the lemma. There are only several grammatical cases, from which it is possible to deduce their gender. However, they are very often not unique. In other words, their gender varies, too.

Compare the paradigms of the lemma *varhany* ('organ' – musical instrument) in Tab. 2. In the two left columns, there are wordforms and their cases. The last two columns contain their frequency in the corpus Araneum Bohemicum IV Maximum of the Aranea series. The black lines mark wordforms that are identical for both genders in the given case. White lines contain only feminine wordforms, the grey lines contain only masculine gender.

Wordform	case	gender F	gender I
varhany	1 (nom)	16 874	1 276
varhan	2 (gen)	10 089	3 360
varhanů	3 (dat)	---	62
varhanám	3 (dat)	556	---
varhanům	3 (dat)	---	170
varhany	4 (acc)	9 988	1 170
varhanech	6 (loc)	---	108
varhanách	6 (loc) non-stand.	1 000	31
varhanami	7 (instr)	1 567	---
varhanama	7 (instr) non-stand.	5	28
varhany	7 (instr)	---	416

Tab. 2. Gender distribution across all wordforms of *varhany* according to the annotation in the corpus Araneum Bohemicum IV Maximum

The black lines are unfathomable. There is no reason why to assign different gender to those wordforms, as there cannot be a single clue for their distinction in

any context. On the other hand, that division causes no problem. It is only strange and cannot be explained.

The more natural solution could be assigning a single gender to the whole paradigm. Where there are different forms for a particular case (dative, locative, instrumental), the wordforms could be considered as variants. Nevertheless, preserving the current state with different genders is also reasonable. The only change should be merging all the wordforms under a single lemma of *varhany*. Having two lemmas, *varhany-1* and *varhany-2,* does not make any sense. The same applies to all pluralia tantum from this group, including, again, the proper names (*Lažánky, Sudety*).

Generally, in the case of pluralia tantum with the gender varying between F and I, the gender is not important at all. It has no influence on any type of agreement. That is why it is not necessary to assign two genders to them. It is reasonable to choose one according to etymology, dialect, or any other clue, or even randomly, and to proclaim the forms resembling the other gender inflectional variants.

Another solution would be selection of a single gender for those wordforms that do not differ. Wordforms with a "visible" gender can keep the different gender. In any case, it is not necessary to create a different lemma for them.

3.5 FM

There are 206 lemmas having this combination of genders. As the masculine animate gender is involved, it is clear that feminine "homonyms" will also denominate living creatures, persons, or animals. There are 78 lemmas ending with –í type (*strojvedoucí* – 'train driver'). These nouns follow the soft adjective declension where the gender manifests itself only in several combinations of case and number.

Another large group (121) are nouns ending with –*a*. These nouns are usually semantically related as they are sort of expressive nicknames for persons, both men and women: *pápěrka* ('weakling'), *sirota* ('orphan'), *trouba* ('simpleton', but also 'an oven'), etc. Many of them are derived from general words and their usage for a person denomination is a sort of a metaphor, similar to the group A in the section IM.

The rest are loanwords with indefinite gender: *hippie*, *pair*, (*super*)*star*, *sfinx*, *okapi*.

There are two more old words, namely *choť* ('spouse') and *sršeň* ('hornet'). The former one can refer to both a man or a woman, while the latter one appears ambivalently – as a scientific name it is feminine, but it is commonly used as masculine.

3.6 FN

This group contains mainly loanwords. Examples: *panorama, scifi, promile*. However, the most interesting (and problematic) words are three old Czech words,

namely *oko* ('eye'), *ucho* ('ear'), and *dítě* ('child') and their derivatives (for instance *biodítě* 'bio-child'). They have an unusual morphology, because their varying gender has a system; they are neuter in singular and feminine in plural.

The solution adopted in the new MorfFlex is the division of the paradigm according to the grammatical gender, which corresponds with the grammatical number. Thus, we have the lemma *dítě-1* having only neuter wordforms in singular, and lemma *dítě-2* with only feminine wordforms in plural.

Oko and *ucho* are even more complicated. They have the regular declension in neuter for plural, too, but only for the figurative meanings. When speaking about an organ of a vision or hearing, the gender changes to feminine in plural. In this case, the results are lemmas *oko-1, ucho-1* with the regular declension and the both numbers, and *oko-2, ucho-2* that have only plural feminine wordforms in their paradigms.

All those irregular words could be captured simply within a single paradigm with a single lemma and a varying gender.

3.7 FIN

There is one word that appears in texts in three genders. It is the loanword *image* that became quite popular but as its ending, and the disagreement of its written and pronounced form, does not correspond with any Czech pattern, people use it in all three genders.

3.8 FMN

There are two words in this group: *budižkničemu* ('good-for-nothing') and *rukojmí* ('hostage'). Their belonging to more genders follows from the fact that their endings are not typical for any nominal gender. The gender in a particular context can be derived from the agreement rules. Whenever it is not possible, the gender M seems to be the most appropriate. In any case, all the paradigms can (should) be merged into a single one, as the meanings are the same for all the genders.

3.9 Gender combinations among foreign proper nouns

Foreign proper names appear very often in the language data. If they denote persons, they can get the gender according to the sex of the person – either masculine animate or feminine. If he or she has a company with the same name, we usually assign it the masculine inanimate or neuter gender. Sometimes, names of persons are the same as geographical names, with different genders. Thus, the same proper name can have all genders. In such cases, the gender is often subspecified with the code X saying that any gender is possible. For a particular noun, every possible combination of genders can appear in corpora, including the subspecified X.

This situation is visible in MorfFlex CZ 2.0. In fact, the genders were assigned to foreign proper names according to findings in the data, namely the new PDT-C

corpus [3], that was manually annotated. On the one hand, it is nice that the dictionary is in agreement with the corpus, on the other hand, this solution is not general. We are convinced that a generalization should be made so that foreign names may fit in any context in which they appear in the future. Thus, a single lemma and a maximally subspecified morphological tag would be the most general solution. For names that can undergo Czech inflection, the paradigm may include all the appropriate wordforms with their tags, no matter which gender the wordforms will acquire.

Moreover, adding foreign proper names to the lexicon is a neverending task. Guessers should be used instead of increasing the dictionary with this type of words.

4   CONCLUSION

We discussed polysemous nouns and their treatment in the morphological dictionary MorfFlex CZ 2.0. If they have the same gender, they are now treated as a single lemma with a unique paradigm. Where the genders differ or, they are treated as homonyms. It follows that there are two (or even more) paradigms, each of them with a unique gender, represented by several lemmas distinguished by means of numerical indexes. We call these homonyms artificial, because we are convinced that they are not homonyms at all. Meanings of the great majority of them are interconnected, if not even the same (esp. for lemmas with varying gender).

We presented an overview of possible gender combinations, using the dictionary MorfFlex and Czech corpora, to show that dividing such lemmas according to their gender is not necessary. We suggest all the wordforms of possibly more meanings merge into a single paradigm with a unique lemma. It will make morphological annotation simpler. Also, maintenance of the lexicon will become easier, especially with respect to foreign words, as there will be no need to add and to number new lemmas if they appear with a different gender of a foreign word in future data.

ACKNOWLEDGEMENTS

References

[1]   Petkevič, V. (2016). Morfologická homonymie v současné češtině. NLN, Praha, 588 p.
[2]   Mikulová, M., Hajič, J., Hana J., Hanová, H., Hlaváčová, J., Jeřábek, E., Štěpánková, B., Vidová Hladká, B., and Zeman, D. (2020). Manual for Morphological Annotation, Revision for the Prague Dependency Treebank – Consolidated 2020 release (technical report).

[3]   Hajič, J., Hlaváčová, J., Mikulová, M. et al. (2020). MorfFlex CZ 2.0, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-3186.

[4]   Hnátková, M., Křen, M., Procházka, P., and Skoumalová, H. (2014). The SYN-series corpora of written Czech. In Proceedings of the Ninth International Conference on Language Resources and Evaluation, pages160–164. Reykjavik: ELRA.

[5]   Benko, V. (2014). Aranea: Yet Another Family of (Comparable) Web Corpora. In P. Sojka, A. Horák, I. Kopeček and K. Pala (eds.), Text, Speech and Dialogue. 17th International Conference, TSD 2014, Brno, Czech Republic, Proceedings. LNCS 8655. Springer International Publishing Switzerland, pages 257–264.

[6]   Hlaváčová, J. (2017). Golden rule of morphology and variants of wordforms. Jazykovedný časopis, Bratislava, Slovak Academic Press. DOI 10.1515/jazcas-2017-0024.

[7]   Hlaváčová, J. (2009). Formalizace systému české morfologie s ohledem na automatické zpracování českých textů. Ph.D. thesis, FF UK, Praha.