Година 7 (2021), брой 1 (13)
Year 7 (2021), Issue 1 (13)

LINGUISTIC TYPOLOGY IN KINSHIP TERMINOLOGY OF WESTERN AND SOUTHERN SLAVIC LANGUAGES

Jana Demelová
Masaryk University (Czechia)

Яна Демелова
Масариков университет (Чехия)
demelova@phil.muni.cz

Abstrakt: Tento článok sa zameriava na posledné zmeny v príbuzenskej terminológii západoslovanských a južnoslovanských jazykov. Sústredí sa na inovácie, ktoré sa v západoslovanských jazykoch presadili, a využíva korpusové dáta z tých južnoslovanských, aby situáciu porovnal. Cieľom je preskúmať situáciu ohľadom príbuzenskej terminológie a určiť, či je vývoj v týchto oblastiach obdobný, prípadne akým smerom sa oblasť príbuzenskej terminológie môže vyvíjať v budúcnosti. Článok je rozdelený na dve hlavné časti – prvá je venovaná teoretickým predpokladom a približuje aj tému jazykovej typológie, či metodológiu, zameriavajúcu sa na kvantitatívnu analýzu korpusových dát. Druhá časť predstavuje dáta z národných korpusov skúmaných jazykov a diskutuje zistené výsledky. V práci bolo potvrdené, že poľština nie je v tejto oblasti najviac analytickým jazykom – toto miesto zabrala slovenčina. Južnoslovanské jazyky sa v otázke veľmi líšia, avšak bulharčina a čiastočne slovinčina sa ukázali byť tiež otvorenými analytickým termínom.

Kľúčové slová: príbuzenská terminológia, jazyková typológia, analytické jazyky, syntetické jazyky, slovanské jazyky.

Abstract: This paper discusses recent changes in the kinship terminology of West and South Slavic languages. It focuses on innovations that have become prominent in West Slavic languages and uses corpus data from South Slavic languages to compare the situation in the two geographic areas. The aim of this paper is to examine if the situation with respect to kinship term innovation is the same in South Slavic languages and in what directions the development of kinship terminology in these languages might continue in the future. The paper is divided into two main parts, the first being theoretical assumptions, where theoretical background of linguistic typology is introduced, with remarks about the general categorization of the studied languages and linguistic typology in lexicology. The methodology employed is also discussed in this part, specifically the quantitative study of the corpus data referenced above. The second part brings together data from national corpora of studied languages and discusses the results of the related comparison. Hence, Polish has been rejected as the most analytical language in this area – this place has been taken over by Slovak. Southern Slavic languages differ a lot, however Bulgarian and to some extent Slovenian are open to analytical terms, too.

Keywords: kinship terminology, linguistic typology, analytic languages, synthetic languages, Slavic languages.

Theoretical assumptions

Linguistic typology focuses on the classification of languages based on their phonology, morphonology, morphology, lexicology and syntax. The difference between genetic classification and linguistic typology lies in both aims and outcomes. Linguistic typology tries to explain language reality from a totally different point of view – from a structural one, focusing on similarities in language structure, not in shared origin of the languages in question. This is particularly useful for better understanding the relations between genetically related languages and for finding similarities between non-related languages.

Languages are typically divided into two main groups (that can then be divided into subgroups) – analytic and synthetic. In the first group we can define a subgroup of isolating languages (such as Chinese or English). Basically, this group uses specific parts of speech (e. g. prepositions, particles, and others) to express grammatical relationships in sentences. On the other hand, synthetic languages change the form of words to express the above-mentioned relationships. Synthetic languages can be divided into two (three) subtypes: agglutinative (Turkish, Japenese), fusional (Russian, Czech), and eventually polysynthetic (Inuit). The difference between agglutinative and fusional languages lies mainly in the number of grammatical features one affix accommodates. Agglutinative languages add separate affixes for every feature (grammatical number, case…), but fusional languages use one affix holding all the features. Polysynthetic languages can be viewed as an extreme case of synthetic languages, or as a totally different category, that creates a line analytic – synthetic – polysynthetic. There is a number of features that can be associated with one of the groups, and that can be found listed in the literature.[1]

However, as this paper is dealing with lexicology, we need to state the main differences in this area, too. Linguistic typology in lexicology can be seen mainly in production of new terms, composition, or comparison of adjectives. As was noted by Vladimír Skalička, linguistic typology is not an area that deals with vocabulary in its entirety. For example, semantic differences must be left for other areas of examination.[2]

[1] For more about linguistic typology and different views on it please see Dolník 2009 or Skalička 2004.
[2] Skalička 2004: 910.

Despite the above, we can identify some features in lexicology that could help us define a language by means of linguistic typology. Analytic languages are more willing to accept words and terms from different languages, and synthetic languages are more prone to create their own names using derivation.[3] Two-word terms are typical of analytic languages, composition of polysynthetic, and introflection and using of affixes are typical of synthetic ones.[4]

West Slavic languages are fusional languages (and therefore synthetic), however, the situation among South Slavic languages is more complicated, including both fusional and agglutinative languages. Even despite this rather simplified distinction, we need to stress again that the level of fusion and agglutination might be different for each language. Even fusional languages might have some agglutinative characteristics (some more, some less), and even the lexical stock of any of these languages might have characteristics pertaining to different types. Even though we cannot define a language as solely analytic or synthetic, fusional, or agglutinative (or polysynthetic), we can trace some tendencies, that might be interesting.

In this paper, we will focus on West Slavic (Czech, Polish, Slovak) and South Slavic languages (Bulgarian, Croatian, Macedonian, Slovenian, Serbian). This will allow us to focus on terms in both subdivisions closely enough and to compare them.

It is generally accepted that "standard Czech is the most fusional of Slavic languages and that Polish is in many areas (mainly syntactics and lexicology) branching away from fusional type."[5] As is widely noted in literature, Polish tends to create two-word expressions in many situations in which Czech or Slovak would use one-word term.[6] As Jiří Damborský concludes, these two-word expressions usually are stylistically marked as more official.[7] Polish also tends to create composites and appositional expressions where Czech and Slovak behave strictly fusionally, and also uses analytic comparison. However, we cannot overlook an important fact, i.e. that both synthetic and analytic production of terms is productive in Polish.[8]

While most of Slavic languages on the Balkan peninsula are undoubtedly sticking to the typical Slavic fusional type of language, Bulgarian (and Macedonian) have moved away from this significantly. The absence of declination (with some minor exceptions), the presence of definite

[3] Lotko 1981: 64.
[4] Skalička 2004: 517.
[5] „[...] spisovná čeština je ze všech slovanských jazyků nejflexívnější a že polština se v mnoha směrech, zejména však v syntaktické a lexikální rovině, dost zřetelně odchyluje od flexívního typu." (Lotko 1986: 37).
[6] See Damborský 1977: 55.
[7] Damborský 1977: 59.
[8] Damborský 1977: 59.

postposed article and several other features lead Vladimír Skalička to assign Bulgarian to the agglutinative type.[9] This is confirmed in more detail in the subsequent work of Skalička.[10]

The selection of kinship terminology for this study is not accidental. Kinship terminology is one of the most stable ones, and therefore only major influences can cause a change in it. It allows us to see how ancient terms, originating several thousand years ago, can be changed, based on external or internal factors.

Methodology and corpus data

Based on West Slavic diversions from the standard lexicon, we defined three areas of kinship terminology where changes have occurred in the last decades. Firstly, we have the area of naming grandparents, where two-word naming seems to gain a solid position in Slovak. The second area of change is naming stepparents and stepchildren, that seems to be happening in all West Slavic languages. The third and last area is naming godparents and godchildren, with traces of two-word terms in Slovak and Polish.

After collecting all possible terms for each language, we verified the occurrences for each of the terms in national corpora. To gain as much data as possible, and ensure that it is comparable, we used the Sketch Engine application. To observe more natural language use, a wider stylistic variety and a greater number of texts incorporated, we decided to use a web corpus for each of the languages. The most recent and extensive web corpora were used.[11] Using a quantitative method supports our outcomes, as empirical data will be provided.[12] Comparing all the West and South Slavic languages will provide further benefits due to the number of cases compared (multi-case study).[13]

All the terms were then categorized into four groups: analytic, non-analytic, secondary, and other. The first group refers to the above analytic features of lexicology (therefore newly-created), the second one is the original Indo-European (and therefore fusional) vocabulary, the third group contains secondary derivative terms (from either of the first two groups) and the fourth group contains all that cannot be – for various reasons – included in either of the above groups.

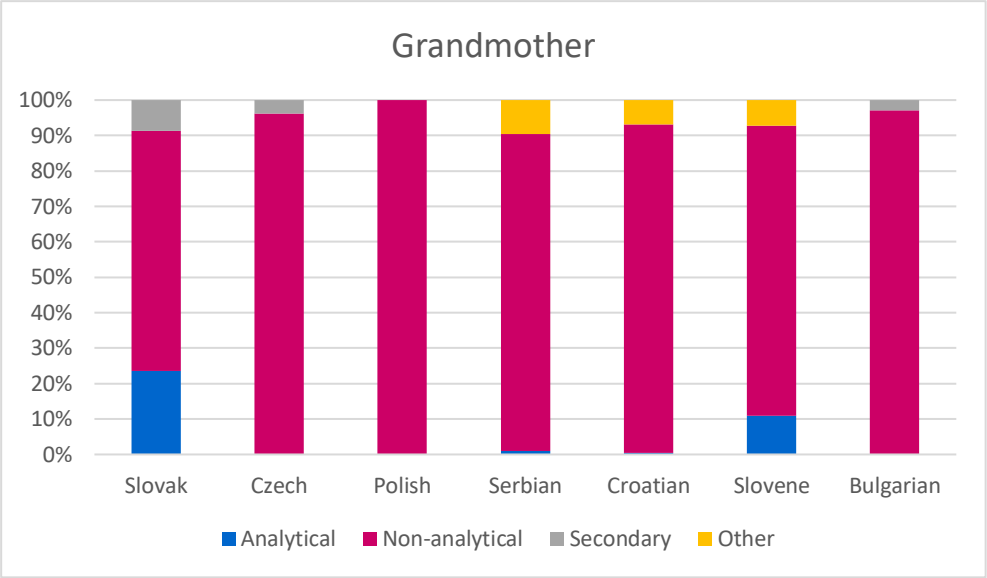
[9] Skalička 1972: 27.
[10] Skalička 1974: 5-14.
[11] The full list of corpora used can be found in bibliography.
[12] Itkonen 1979: 350.
[13] Beekes 1995: 108.

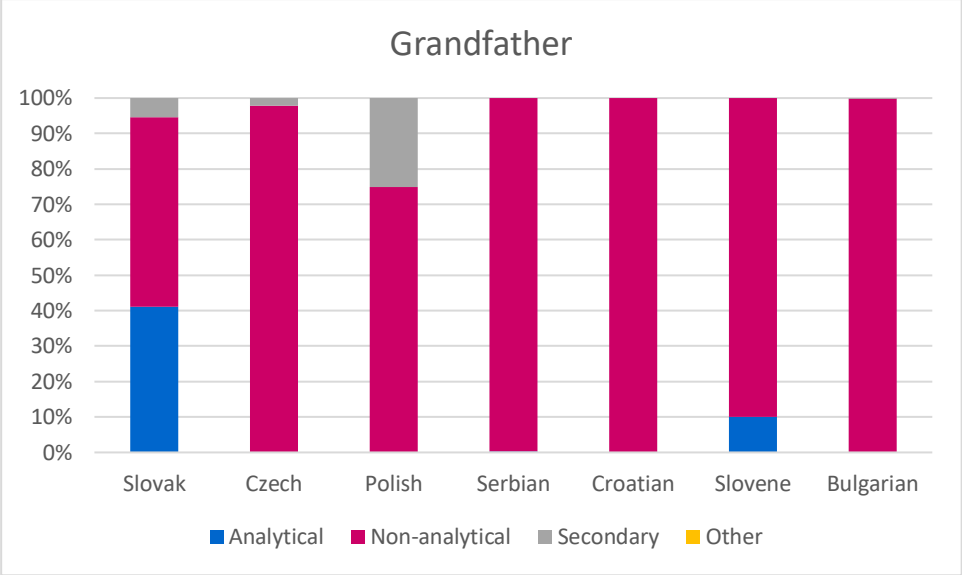
Comparison

Grandparent terminology

The first category, naming grandparents, is particularly interesting, as terms such as *starý otec* or *stará matka* (and variants, in blue) are stable only in Slovak (over 20% in case of grandmother and over 40% in case of grandfather) and Slovenian (approximately 10% in both cases). A very small incidence can be attested in Serbian (interestingly, not Croatian). Serbian, Croatian, and Slovenian also show an incidence of special terms for grandmother (*nana, nona, oma*), which may also be a loan from another language and therefore are analytic terms (in yellow). One possible explanation can be that of external motivation, as both Slovenian and Slovak might be influenced by German or Hungarian lexicons (*Großmutter, Großvater* in German, *nagymama*, *nagypapa* in Hungarian), what would be like terms *nana, nona, oma*. This, however, does not explain why other Slavic languages, also in contact with German and Hungarian, such as Slovak and Slovenian, were not influenced into using such terminology.

Grey also shows traces of secondary terms (in this case derived from a different original term or shortened form of analytic term – e.g. *starucha, stařenka* or *старица*, and *starký, stareček, staruszek* or *старик*), but these are quite rare (except of Polish for 'grandfather') and their origin unsure. If these are traces of different original terms (distinguishing between maternal and paternal parents), these would need to be included in the non-analytic term group. On the other hand, if these are just shortened versions of analytic terms, these would need to be included in the first category. There is also a third option, i.e. that two-word analytic terms were created out of these terms, trying to give them a more official form. The third option seems not too probable, as we cannot see them in Slovene, even though we can find two-word analytic terms in this language.



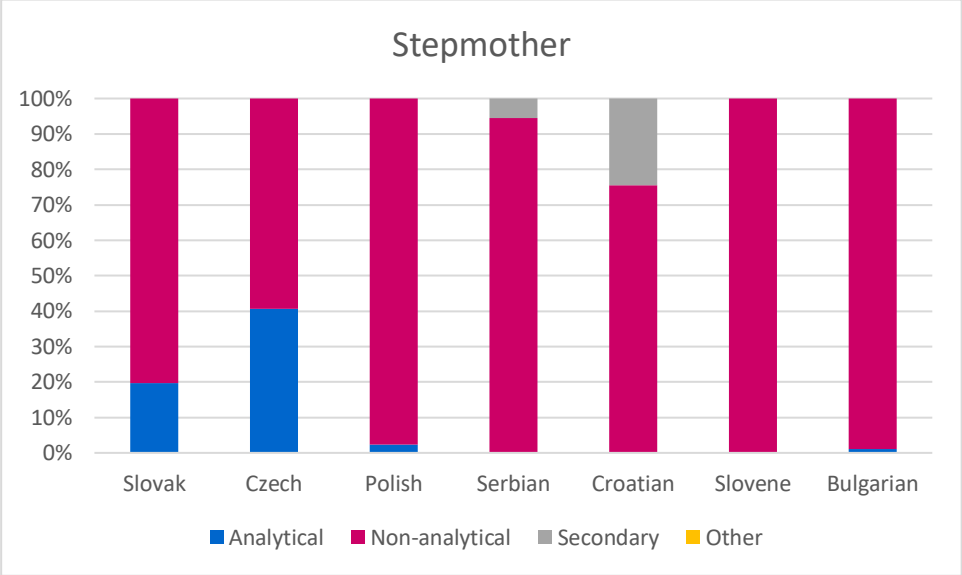


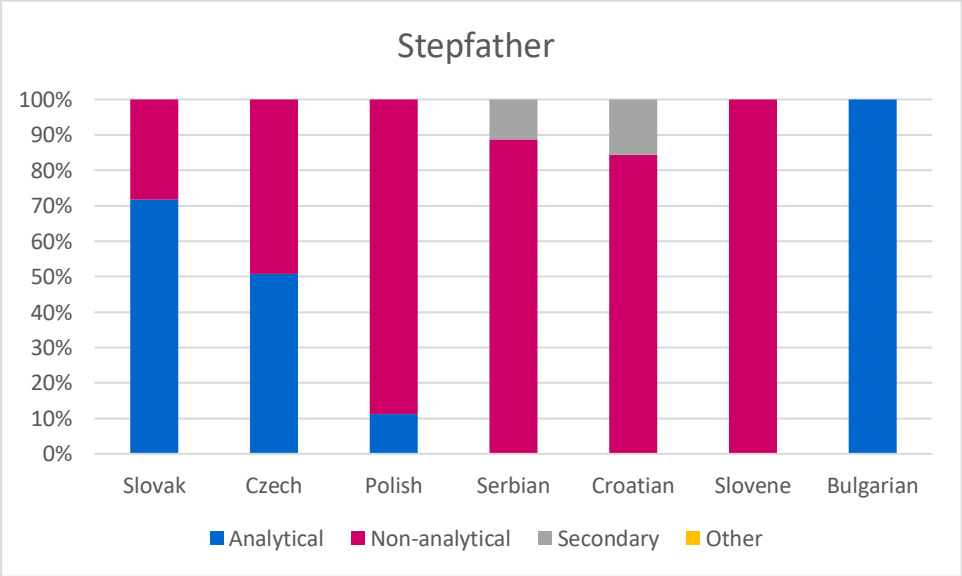
Stepparent terminology

In this category, the situation changes significantly. All three West Slavic languages possess analytic terms, as well as Bulgarian. Again, Slovak shows higher incidence of the analytic term for male kin, reaching almost 70% (compared to less than 20% for females). Usage in Czech is more balanced, with 50% for male and 40% for female relatives. Surprisingly, Polish has the lowest incidence of analytic terms, with only stepfather analytic names exceeding 10%.

All West Slavic languages, therefore, show a higher incidence of analytic terms for male relatives, what is further confirmed by data on Bulgarian, which does not use a non-analytic term for stepfather at all and uses analytic terms for females very rarely. Even though some explanations of this situation might arise from the higher incidence of terms for stepmother, a full explanation would require further study on the issue.

Let us conclude by looking at the terminology in Serbian and Croatian, which have developed a special term distinction between stepparent in relation to an adopted child, compared to stepparent as a new partner of one of the parents (with terms *pomajka* and *poočim*, here in the category of secondary terms). These cannot be assigned to analytic terms, but need to be distinguished from the typical Slavic terms for stepparents.



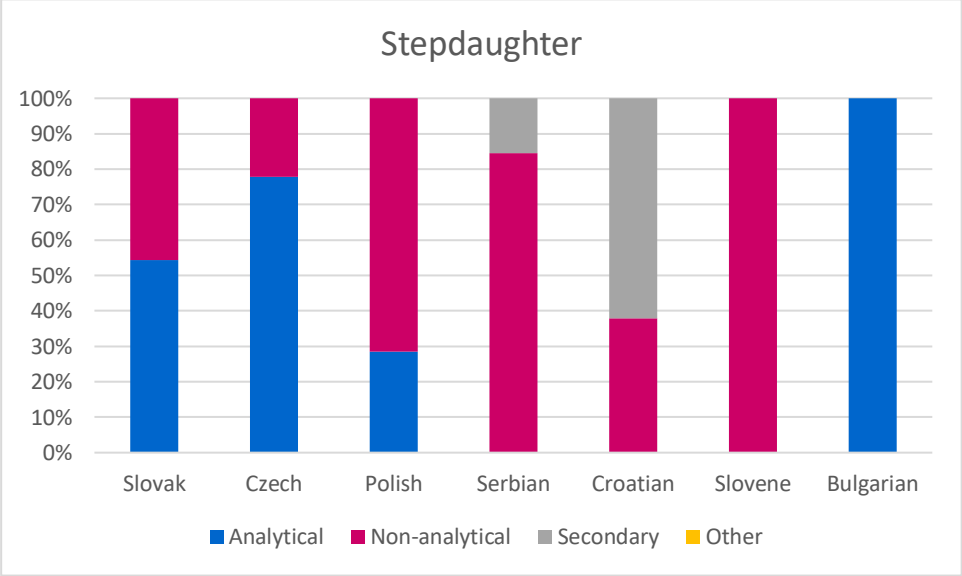


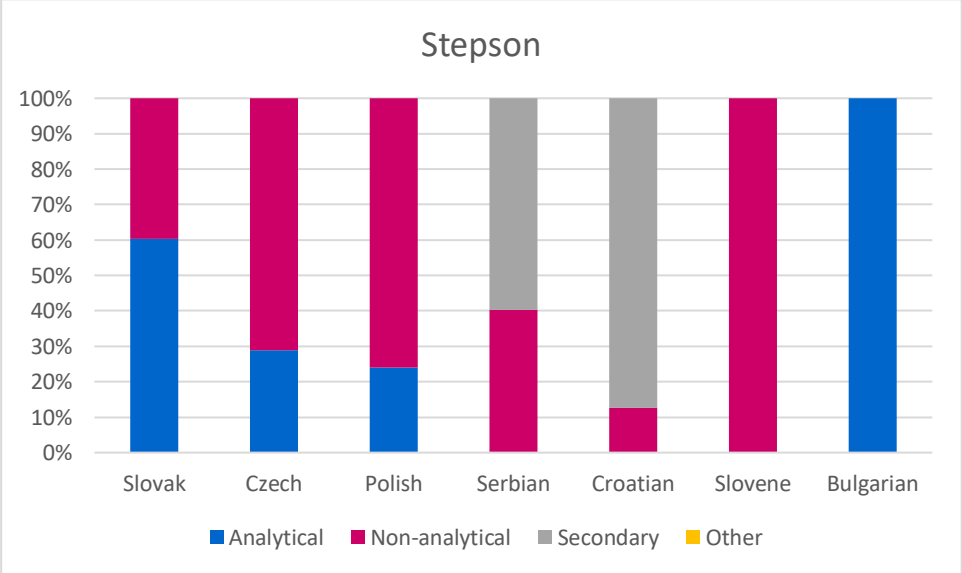
Stepchild terminology

The situation is very similar with respect to stepchild terminology. We can find traces of analytic terms in all the West Slavic languages and Bulgarian, but results are more convincing this time. Slovak has an incidence of analytic terms in this category well above 50% (for both stepdaughter and stepson), while Czech shows an incidence of above 70% for stepdaughter and almost 30% for stepson. Polish, again, seems to be the least prone to analytic terms of the West

Slavic languages, with an incidence of below 30% for both terms. Bulgarian, for its part, shows fully analytic terms for both stepdaughter and stepson. Analytic terms might be even more frequent in West Slavic languages, as non-analytic terms are homonymous and therefore raw corpus data might include other meanings, too.

Here again, Croatian and Serbian developed special terms for adopted children (*posinak*, *pokćerka*) compared to those for children of a partner. We will not count them as being part of either category here, as in the previous case.



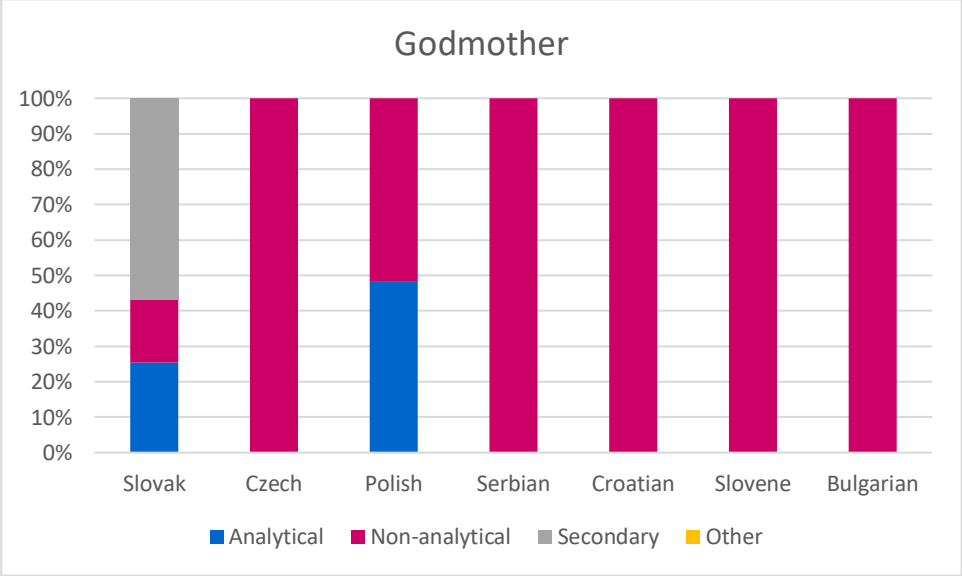


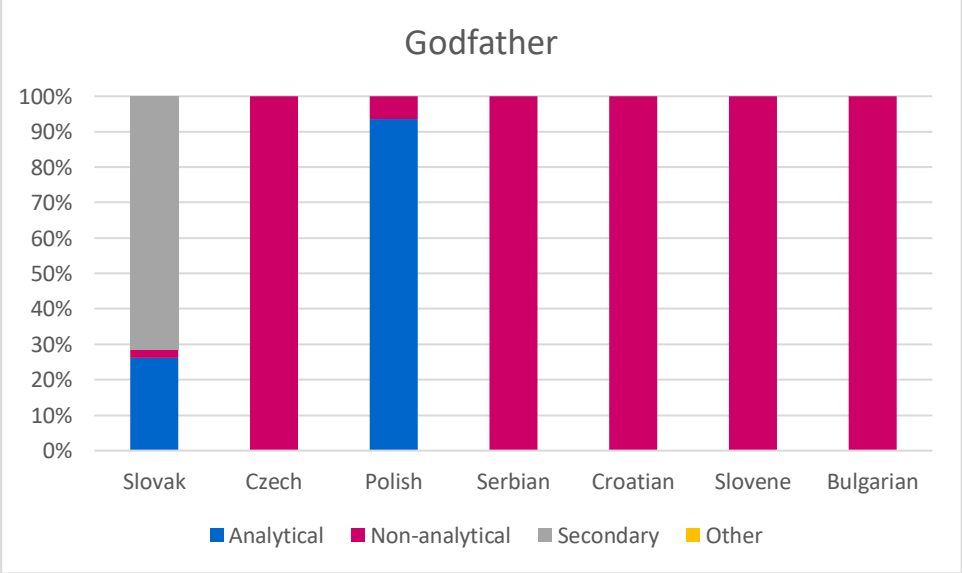
Godparent terminology

In godparent terminology analytic terms can be found exclusively in West Slavic languages, more precisely in Slovak and Polish. Terms for Polish reach an incidence of up to 50% for godmother and up to 100% for godfather (where numbers might be influenced by the famous

movie). For Slovak, the situation is even more explicit, however, here we need to count secondary terms among the analytic ones. These include *krstný* (*krsný*) and *krstná* (*krsná*), which are surely just shortened versions of these analytic terms. If we do so, Slovak shows an incidence of over 80% of analytic terms for godmother and almost 100% for godfather.

Reasons for the Slovak and Polish deviation above need to be detected, however, assumptions about the influence of Hungarian and German can be proposed, as both Hungarian and German use compounds for these kinship terms (the same situation as for grandparent terms).



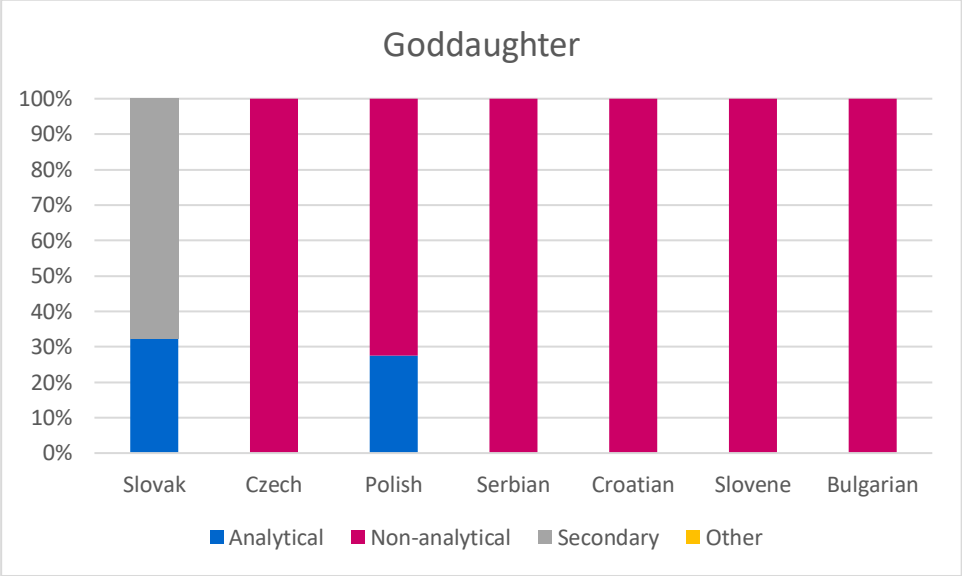


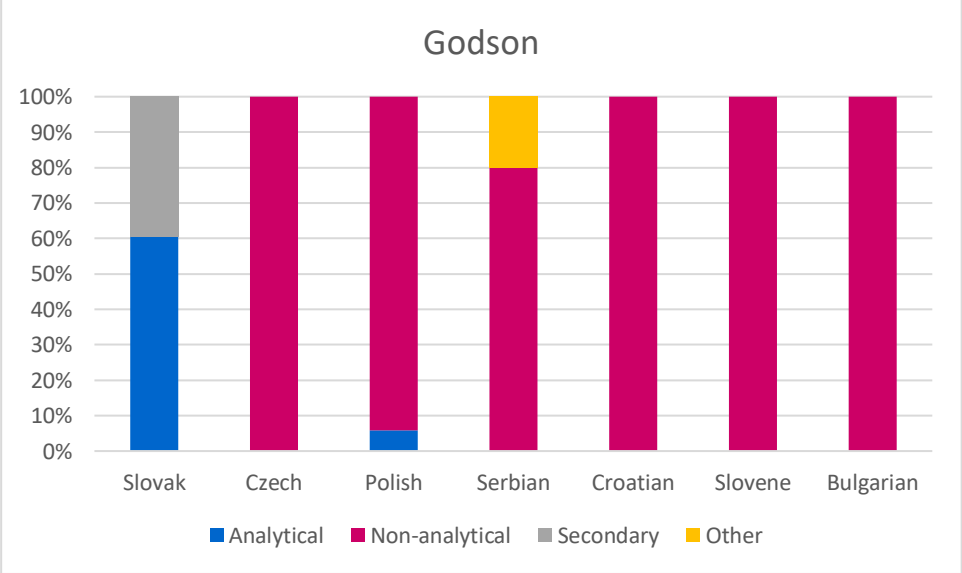
Godchild terminology

Godchild terminology mostly copies that for godparents, with several exceptions. Non-analytic terms are even more rare in Slovak, but widespread in Polish, where analytic terms reach an incidence of up to 30% only for goddaughter and less than 10% for godson. We cannot find any

analytic terms in Czech or any of the South Slavic languages again, but in the case of Serbian we need to point out a special term derived from the original term *kum*, *kumašin*, naming godson in a relation created by wedding instead of christening. Even though special terms in Serbian exist for goddaughter in this relation, its existence in modern language was not confirmed in corpora.

Here again, we can only assume that the reasons for differences in attitude among Slovak and Polish speakers towards this part of kinship terminology are caused by influence from different languages. However, further examination of the reasons needs to be conducted in all cases discussed in this paper.





Conclusions

As we confirmed above, Slovak is the most prone to analytical terms in kinship terminology of the West Slavic languages (10 out of 10 occurrences). Even though Polish is viewed as the most diverting from the fusional type, our research confirms that (at least in this area) it is the most fusional of the West Slavic languages, having significant evidence of analytical terms in naming stepchildren and godparent, partly in stepparents and godchildren, too (6 significant occurrences). Czech stands somewhere in the middle, using analytical terms to name stepparents and stepchildren (4 occurrences) in great extent. On the other hand, there are several other examples of Polish analyticism, e.g. in terms for cousin (*kuzyn, kuzynka*), which cannot be found in any other Slavic language and therefore were not included in this study.

South Slavic languages differ a lot, where Slovenian (2 analytical occurrences) and Serbian use in small extent two-world (analytical) terms, while using some one-word loans from German for naming grandmother at the same time, again of analytical nature. Bulgarian adopts different attitudes , either rejecting analytical terms of accepting them fully (in case of stepfather and stepchildren; 4 significant occurrences in total). Serbian and Croatian also show innovations in differing between adopted child and child of a partner (and vice versa), as well as in different, however derived, term for godchild gained via marriage.

The question of the reasons behind the above innovations remains, as innovations that might be influenced by surrounding languages took place only in some of the studied languages. Due to geographical differences, researching East Slavic languages might bring new insight.

Bibliography

Beekes 1995: Beekes, R. S. P. Comparative Indo-European linguistics: an introduction. Amsterdam: John Benjamins Publishing Company, 1995.

Damborský 1977: Damborský, J. Studia porównawcze nad słownictwem i frazeologią polską i czeską. Warszawa: Państwowe Wydawnictwo Naukowe, 1977.

Dolník 2009: Dolník, J. Všeobecná jazykoveda: opis a vysvetľovanie jazyka. Bratislava: VEDA, 2009.

Itkonen 1979: Itkonen, E. Qualitative vs. Quantitative Analysis in Linguistics. – In: Perry (ed.). Evidence and Argumentation in Linguistics. Berlin/Boston: De Gruyter, 1979, 339 – 366.

Lotko 1981: Lotko, E. Polština a čeština z hlediska typologického. Olomouc: Univerzita Palackého, 1981.

Lotko 1986: Lotko, E. Čeština a polština v překladatelské a tlumočnické praxi. Ostrava: Profil, 1986.

Skalička and Čermák 2004: Skalička, V. a F. Čermák. Vladimír Skalička: souborné dílo. II. díl, (1951-1963). Praha: Univerzita Karlova v Praze, nakladatelství Karolinum, 2004.

Skalička 1972: Skalička, V. Űber die typologische Eingliederung der Balkansprachen. – In: Jiřina Smrčková (eds.). Les Études balkaniques tchécoslovaques IV. Praha: Univerzita Karlova v Praze, 1972, 27 – 33.

Skalička 1974: Skalička, V. Typologischer Vergleich der Balkansprachen. – In: Bohuslav Havránek (eds.). Les Études balkaniques tchécoslovaques V. Praha: Univerzita Karlova v Praze, 1974, 5 – 14.

Corpora

(accessed via Sketch Engine application https://www.sketchengine.eu/, 27. 9. 2018)

- Slovak language: Slovak Web 2011 (skTenTen11),
- Czech language: Czech Web 2012 (csTenTen12 v9),
- Polish language: Polish Web 2012 (plTenTen12, RFTagger),
- Bulgarian language: Bulgarian Web 2012 (bgTenTen12, TreeTagger v2),
- Serbian language: Serbian Web (srWaC 1.2),
- Croatian language: Croatian Web (hrWaC 2.2, RFTagger),
- Slovenian language: Slovenian Web (slWaC 2.1).