

*Tamás Farkas*

***A Surname Typology Project:  
The Lessons Learnt from the Distribution of the Most  
Frequent Hungarian Surnames***

**1. The program of the typological-statistical processing  
of the Hungarian surname stock**

There are several methods for the comprehensive and representative analysis of surname systems. These include examining surname frequency, surname typologies, while utilising certain geolinguistic approaches, and also contrastive analysis.

The European Surname Typology Project (ESTP) combines all these approaches, the details of which were described comprehensively at the *International onomastic projects and cooperation* symposium of the 26<sup>th</sup> International Congress of Onomastic Sciences (CHAREILLE–DARLU et al. 2017). The major characteristics of this initiative are the following: 1. It sets out to study the surname stock of several languages and countries. 2. It intends to process not complete surname stocks but is based on the analysis of the 100 most frequent surnames representing those. 3. It studies the internal structure of the surname corpuses by name type distribution. 4. It carries out the analyses in territorial units and thus connects it with the geographic dimension. 5. It also aims to analyse its findings in a contrastive way.

Following this Western European initiative, I set myself the task of processing a Central European surname system, namely the Hungarian. Preliminary attention was dedicated to relevant methodological questions: overviewing arising problems, testing and evaluating these, because clarifying and following them is especially important in the case of research conducted within an international cooperation. This was followed by a comprehensive analysis of the complete picture of the surname stock of Hungary, based on complete (or at least representative) name corpuses. Several subsets of the complete surname stock were examined: alongside the contemporary subset, the historical, namely the early 18<sup>th</sup> century subset, and the subset of artificially created Hungarian surnames (results of surname changes) of the 19<sup>th</sup> and 20<sup>th</sup> centuries. Findings were interpreted in comparison with each other as well as with those in other countries or languages. I gave a detailed account of this research at the ICONN3 international onomastic conference in Baia Mare in 2015 (published as FARKAS 2015).

Later phases of this research were expanded to include the geolinguistic aspects of the matter, i.e. the regional distribution of the contemporary Hungarian surname stock (in detail, in Hungarian see FARKAS 2016, 2017). This paper provides a summary of the results of this research.

## 2. The sources and methods of data processing

Regarding the contemporary Hungarian surname stock, work was carried out with several different name corpuses (see also in FARKAS 2015: 121–125).

Data on the surname stock of the current population of Hungary was collected based on official registries. Data for the entire country was processed based on the full dataset of the 2007 national registry. Regional distribution was examined on a dataset from 2009, containing only the 100 most frequent surnames of each region (19 counties and the capital city, Budapest). (HAJDÚ 2010, DHS. 2007 and 2009; I am indebted for this data to the late MIHÁLY HAJDÚ and FERENC VÖRÖS.) In the case of the complete surname stock sources allowed for the merging of name variants in spelling and pronunciation (that is, a higher degree of lemmatisation) as well. While in the case of regional distribution, orthographic variants had to be examined separately.

The study was extended to the surname stock of ethnic Hungarians native to Romania. An appropriate source was available for such an overview: a representative (if not absolute) onomastic survey based on the surname stock of the students of the schools in Romania where Hungarian is the language of tuition (1994–2002, 45 thousand students; MURÁDIN 2005). The area analysed by the study falls between that of the county level and the national level surname stock in Hungary, but the source does not provide appropriate information on the territorial distribution of its data. The database of Hungarian names in Romania merges orthographic variants (thus represents a certain degree of lemmatisation).

Thus, slightly different types of sources were available for the geolinguistic study. The methodology of dealing with surname variants separately versus taken together makes little or no difference, as earlier analyses (FARKAS 2015: 125–126) have demonstrated – at least in the case of the Hungarian surname stock and in its typological analysis. (In fact, typically there are relatively few variants of names in the Hungarian surname stock, and these regularly tend to be borne by significantly fewer individuals than base versions.) As a result, relevant conclusions can be reached based on these datasets of somewhat different backgrounds – and, especially for lack of better ones.



The analysis used the following methods (in detail see FARKAS 2015: 125–128). Surname variants were treated according to the possibilities offered by sources.

The analysis considered not individual names (lemmas), but their frequency – in other words, not the number of lemmas, but the number of name-bearers. The typological composition of lists of the top 100 surnames was also examined. A list like this includes approximately one third of the name stock of the given population in Hungary. Surnames of Hungarian origin were categorised in the following four main types (allowing for multiple categorisations): a) patronymics, b) occupational names (and titles, dignities), c) nicknames (personal characteristics), and d) names referring to origin (toponymic or ethnic).

However, before moving on to the findings of this typologisation, a closer look should be taken at surnames of foreign origin among the most common surnames in Hungary.

### 3. Surnames of foreign origin in the Hungarian surname stock

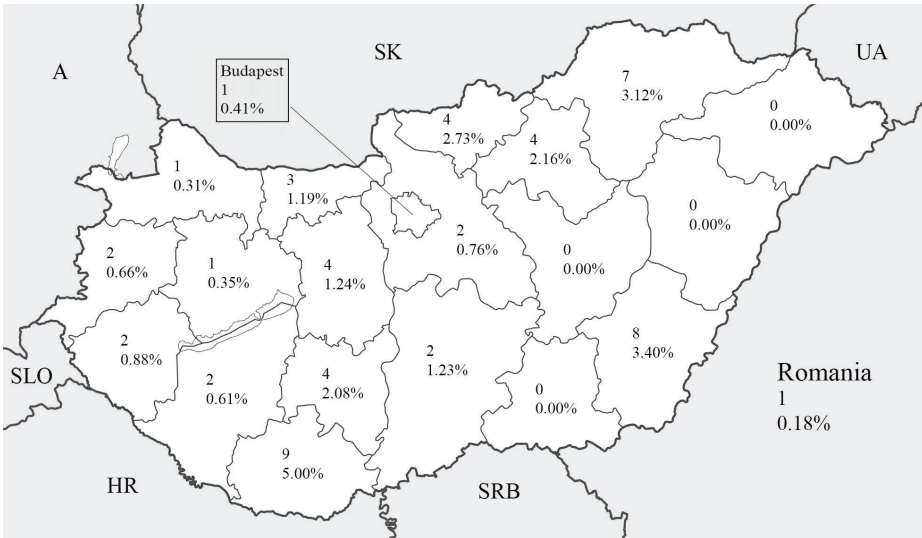
For practical reasons, and because this could be considered the most exact method, only names which originated without a doubt from a foreign language were considered surnames of foreign origin here. Thus, the surnames *Kovács* 'smith', *Kocsis* 'carter' and *Polyák* 'Pole' (which names, apart from their plausible Hungarian origin can also come, with more or less certainty, from Slavic languages) were considered Hungarian surnames (along with certain surnames that can be of Hungarian or Romani origin in the lists). It was also necessary to separate the surnames of foreign origin because they could not be categorised along with the surnames of Hungarian linguistic origin. The proportion of surnames of non-Hungarian origin is relatively low among the most frequent surnames, and their frequency is also typically a lot lower than that of surnames of Hungarian origin (cf. also HAJDÚ 2010: 528).

In Hungary, one single surname of non-Hungarian origin is found on the collated list of surname frequency: *Novák*, of Slavic origin (with 0.11% of the entire population bearing this name, which thus comes in at 83<sup>rd</sup> place if variants are considered separately and 84<sup>th</sup> if lemmatised).

A look at the different counties in Hungary shows that the number of surnames of non-Hungarian origin ranges from 0 to 9 within the top 100. The majority of these is Slavic, the minority is German, with one Romanian surname found in one single county. Names of other origins do not feature among the most frequent surnames (apart from a few that might or might not also be of Romani origin). The surnames of non-Hungarian origin featuring multiple times on the regional top lists are: *Novák* (8 regions); *Kollár* (6); *Radics* (5); *Schmidt* (4); *Lipták*, *Mayer*, *Müller*, *Sztojka* (2-2 regions). There are 25 surnames that occur in one region each. (In detail see FARKAS 2016: 50–56.)

*Novák* is the only surname that found its way onto the top 100 list of surnames in the capital city, Budapest. This is hardly surprising, as the Budapest surname stock tends to be a good representation of the national average. The geographic distribution of surnames of non-Hungarian origin is, obviously, not independent of the traditional regions of residence for respective ethnic minorities. Thus, for example, German names tend to feature predominantly in the Transdanubian region, while Slovak surnames in Békés county in the South East, which is also home to the only Romanian surname (*Árgyelán*) on the list.

As for the corpus of Hungarian surnames in Romania, there is a single name of non-Hungarian origin in it: *Ráduly* (0.18% of all name-bearers). This data, however, is not to be seen as absolute, due to the unique principles guiding the compilation of the corpus used (i.e. excluding names with foreign spellings). Yet, in comparison with the data from Hungary, it still shows typical regional characteristics, as *Ráduly*, a typical name in the Romanian surname stock (cf. BENŐ 2012: 123–126), is a lot less frequent in Hungary (where, even if all its variants are collapsed, it only reaches the 1159<sup>th</sup> position on the frequency list).



**Map 1:** Surnames of non-Hungarian origin among the regional top 100 surnames (number of names and proportion of name bearers)

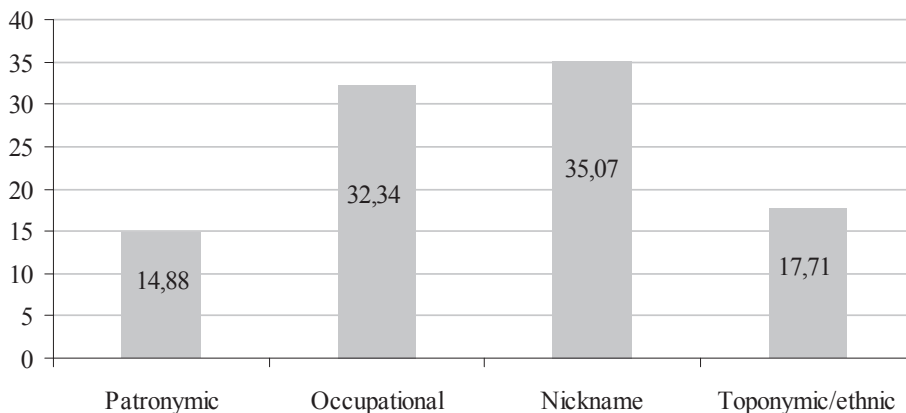
#### 4. The spatial distribution of the Hungarian surname types

The analysis of the typological distribution of surnames of (plausible) Hungarian origin was conducted following the methods mentioned above. Lists that provided different name variants separately were prioritised in my analysis here, apart from the list for Romania, where orthographic variants



were merged in the original source. Surnames of clearly non-Hungarian origin were excluded.

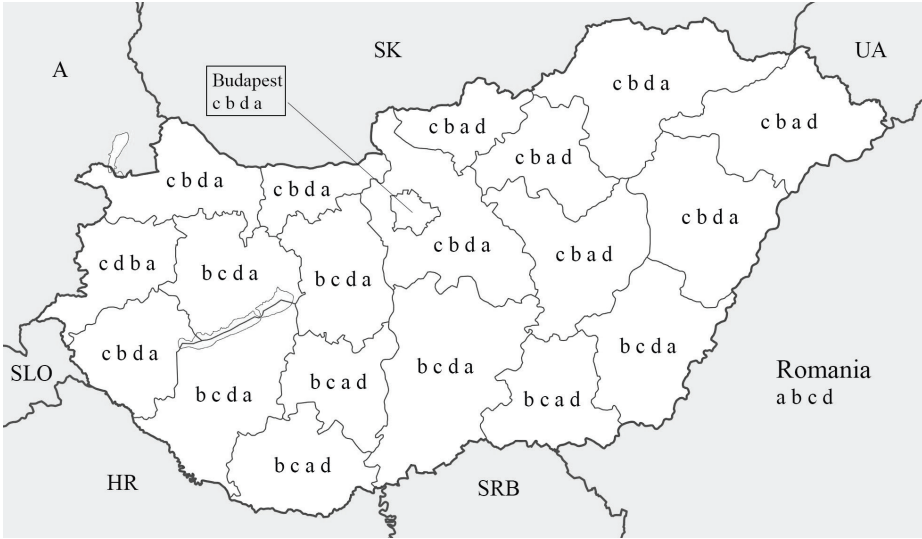
The following will provide a picture of the complete Hungarian surname stock for the top 100 list in Hungary (based on non-lemmatised names; for a slightly different picture, based on lemmatised names, see FARKAS 2015: 126, 129, 130–131). It can be seen that the patronymic and toponymic/ethnic categories on the one hand, and the occupational and nickname categories on the other hand show similar proportions.



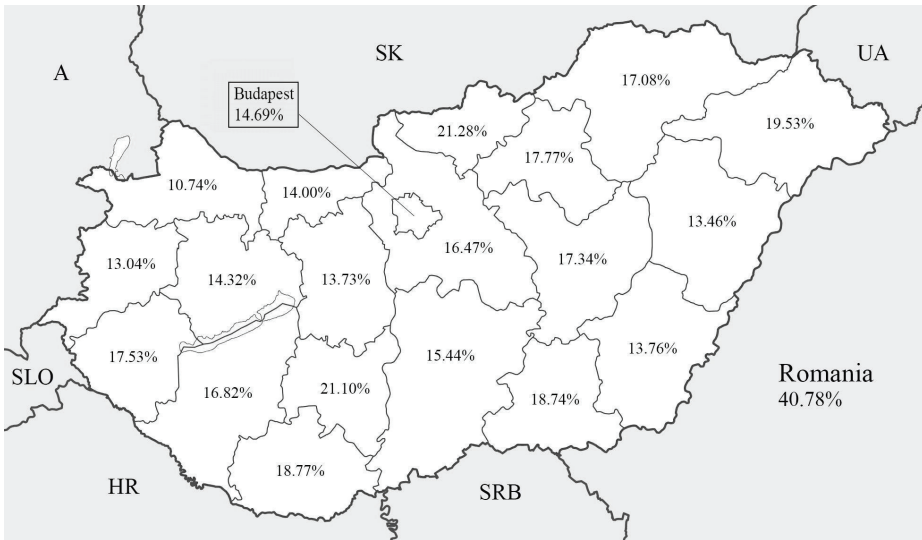
**Graph 1:** The frequency of different surname types within the top 100 surnames in Hungary (proportion of name bearers, non-lemmatised names)

Knowing the national level, the regional distribution of surname types should also be studied. The maps showing the geographic distribution of surname types also include data for Hungarian surnames in Romania.

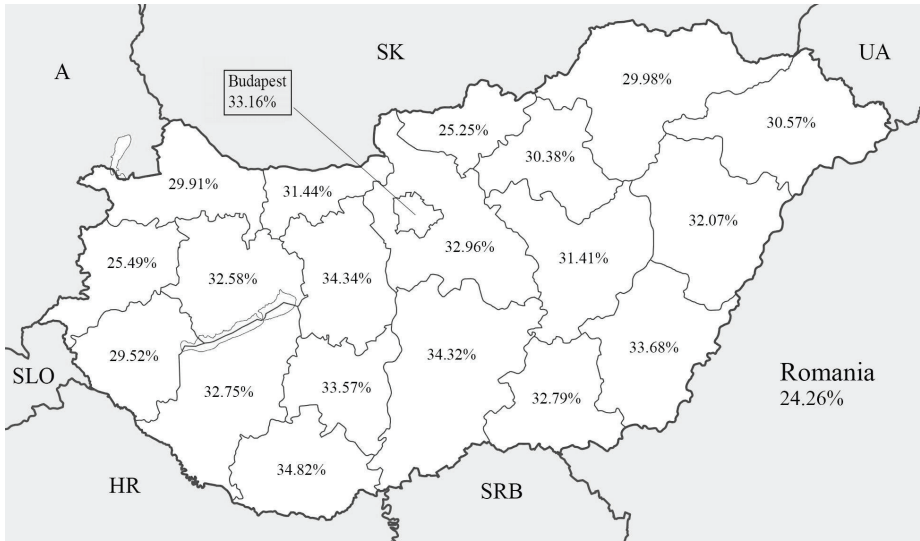
The most characteristic surname type in each region is always one of the two which are dominant on a national level: occupational or nicknames.



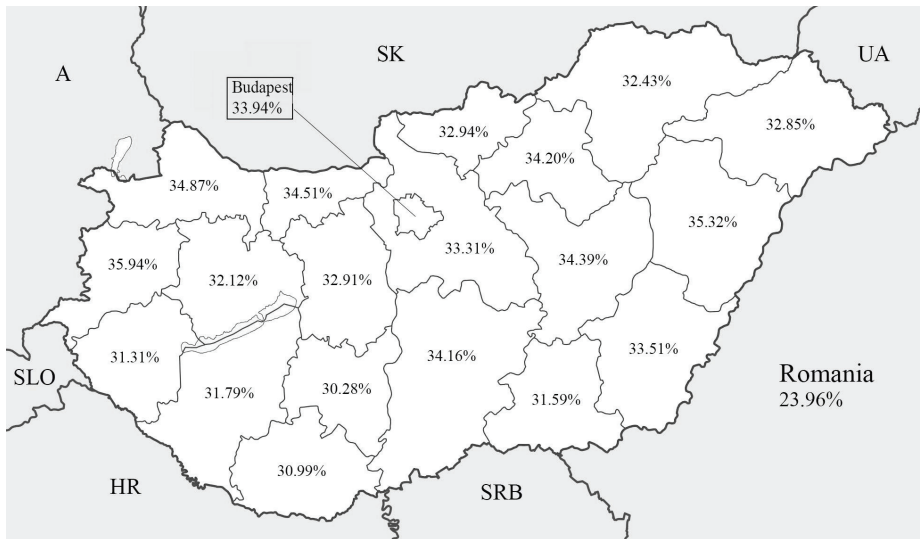
**Map 2:** The types of surnames by region and in order of frequency (proportion of name bearers): a) patronymic, b) occupational, c) nicknames, d) referring to origin (toponymic/ethnic)



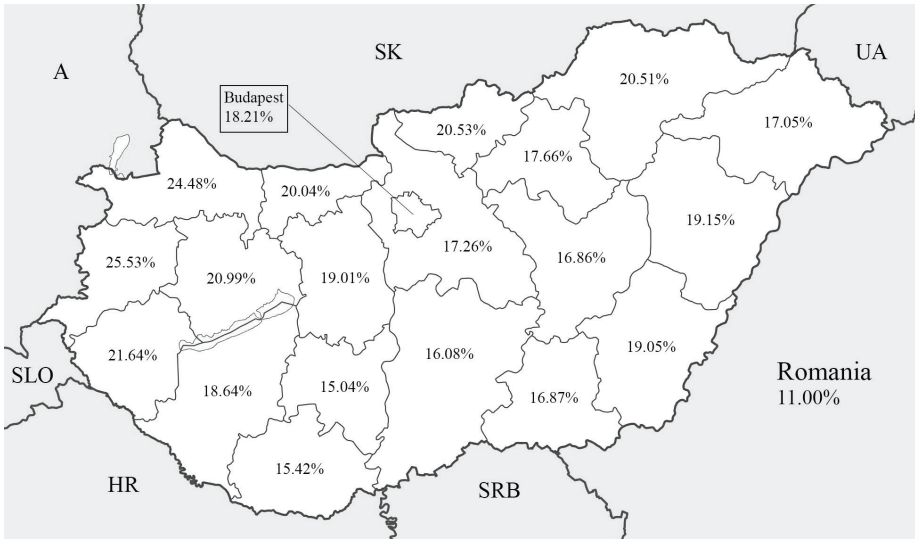
**Map 3:** The regional frequency of surname type a) patronymic (proportion of name bearers)



**Map 4:** The regional frequency of surname type b) occupational (proportion of name bearers)



**Map 5:** The regional frequency of surname type c) nicknames (proportion of name bearers)



**Map 6:** The regional frequency of surname type d) referring to origin (toponymic/ethnic) (proportion of name bearers)

The maximum difference between any two regions is only about 10% (and only half of that for the nickname category). However, it is again noteworthy that the proportions of Budapest are almost identical to the national average, which can be explained by the population history of the capital city of the country.

Surname type	Hungary				Romania
	Budapest	National	County min.	County max.	
a) patronymic	14.69	14.88	10.74	21.28	40.78
b) occupation	33.16	32.34	25.25	34.82	24.26
c) nickname	33.94	35.07	30.28	35.94	23.96
d) origin	18.21	17.71	15.04	25.53	11.00

**Table 1:** The frequency of surname types on different top 100 lists (proportion of name bearers)

The typological composition of the Hungarian surname stock in Romania shows a picture very different from that of Hungary. Here the distribution numbers fall outside the range of distribution in Hungary. The most obvious difference is how much higher the proportion of the patronymic category in the Hungarian surname stock of Romania is. The explanation for this is the higher proportion of such names, and their altogether higher positions on the respective surname



lists (i.e. their higher relative frequency). The frequency of the patronymic category is almost double that of the most frequent surname category in Hungary. Furthermore, partly because of this, all other Hungarian surname types in Romania are below their respective minimum regional frequencies in Hungary. The strong presence of patronymic surnames is characteristic in certain regions of Transylvania, especially in the so-called Seklerland, where there is an especially large population of ethnic Hungarians within Romania. This is known from both synchronic and diachronic studies (cf. FARKAS 2017: 114–115), however, the source used for the current study does not allow the creation of a more detailed geolinguistic picture of this phenomenon.

These findings also highlight the fact that when looking at the surname stock of a given language, research should not stop at the given country's borders. Any analysis should be extended to the whole linguistic area, especially native ethnic minorities. Geolinguistic differences can be significant even concerning basic surname types.

## 5. The three most frequent surnames

The three most frequent surnames in the surname stock of Hungary are *Nagy* 'big, large', *Kovács* 'smith' and *Tóth* 'Slav, Slovak' (both separating and merging the name variants of each), and there is little deviation from this pattern on the regional podiums. The *Tóth* 'Slav, Slovak', *Horváth* 'Croat', *Németh* 'German' (ethnonymic surnames); *Kovács* 'smith' and *Szabó* 'tailor' (occupational surnames); and *Nagy* 'big, large' (a nickname) are the surnames that make it into the top 3 multiple times in regional lists, alongside additional surnames that occur among the top three in one region each. The three most frequent surnames of the Hungarian surname stock in Romania (*Szabó*, *Nagy*, *Kovács*) correlate with the picture of Hungary. In certain regions of Romania these surnames are the absolute most frequent ones, that is, if the surnames of the whole Romanian population are included they remain the most frequent (cf. LIPAN 2012).

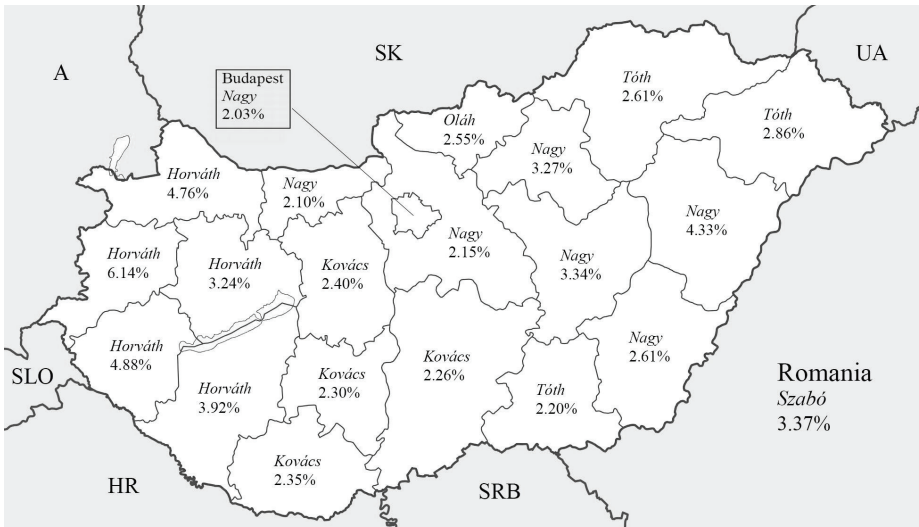
Thus, the regional top 3 surname lists mostly feature surnames that refer to ethnicity, occupation or personal characteristics. The surprising frequency of ethnonymic surnames, a conspicuous feature of the Hungarian surname system (cf. FARKAS 2013), can be observed in the current findings as well, while their specific regional distribution reflects specific instances of population history. These surnames were considered both in the categories of those referring to a person's origin and personal characteristics. This is because the categorisation is based on the supposed motivation behind name giving (rather than the primary meaning of the linguistic unit forming the base of the surname). In the case of

surnames based on ethnonyms motivational backgrounds can point in various directions (descriptive, metaphorical, metonymic).



**Map 7:** The 3 most common surnames in each region

This overview can be detailed by examining the numerical frequency of each surname, next to their order on the name lists. There can be significant (as large as threefold) differences among the frequency of the most frequent surnames in different regions. Thus, it is possible for example, that a surname is only the third most frequent in a given county, but its actual frequency surpasses that of the most frequent names of other counties.



**Map 8:** The most common surnames in each region (proportion of name bearers)



## 6. The regional characteristics of the occurrence of specific surnames

Almost half of the surnames (190 of 403) on the frequency lists for the 19 counties and Budapest only appear amongst the most frequent names in a single region. However, there are 24 surnames on these lists, which feature among the top 100 in all 20 regions. This in itself demonstrates the great regional differences in surname use among the different regions. If Hungarian surnames in Romania are also considered, this list shrinks to 21 surnames, which are also usually featured in the most frequent third of the national top 100 list. Thus, all in all, these surnames can be considered now the most typical of the entire Hungarian surname stock.

The differences between specific regional top 100 lists can highlight regional differences especially well. It is best to start this analysis with a comparison of the Hungarian top 100 list in Romania with the same list for Hungary, as differences from a typological-statistical aspect are most conspicuous here. 17 surnames can be found here that are absent from the list for Hungary, and thus are characteristic of the Hungarian surname stock in Romania. Most of these occur in the middle or towards the end of the top 100 list for Romania. The vast majority of these names – in accordance with the general characteristics of the Hungarian surname stock in Romania – are patronymic; in the basic form of the name (e.g. *Benedek* 'Benedict'), with a diminutive (e.g. *Tankó*: *Tam-* + *-kó* < *Tamás* 'Thomas') or a patronymic suffix (e.g. *Ferenci*: *Ferenc* 'Francis' + *-i*). There are also pronunciation (dialectal) variants (both belong to occupational surnames) that are, in this form, characteristic of Transylvanian Hungarian, which are only found in their base forms on the top 100 list in Hungary: *Fazakas* (~ *Fazekas* < *fazekas* 'potter') and *Szócs* (~ *Szücs* < *szücs* 'furrier'). However, an inverse study is also possible, in the form of the list of surnames absent from the Transylvanian 100 list, yet ubiquitous on the lists in Hungary. This list is a lot shorter, with only *Juhász* ('shepherd'), *Mészáros* ('butcher'), and the standard form of the above mentioned dialectal occupational names (*Fazekas*, *Szücs*), and a nickname (*Vörös* 'red').

The surnames found on just one specific region's top list supply interesting examples for various onomastic phenomena, and for the regional characteristics of the Hungarian surname stock, too.

The above mentioned example of the only surname of Romanian origin featuring on the top lists for Hungary, *Árgyelán* 'Transylvanian' is among the most frequent names in the South Eastern county, Békés, which is the epicentre of the ethnic Romanian community in Hungary (cf. FARKAS–N. FODOR 2016).

Another, lexical example is a surname of Hungarian origin with strong regional ties: *Koplányi*. This surname only features on the top 100 list in Nógrád county,



in the North of Hungary (with 300 name bearers, in the 89th position, while there were only 433 people with the surname in the whole country, and the name is not listed in the surname dictionaries for Hungary at all). The scarcity and strong regionality can both be explained easily, as the surname is based on a toponym: *Koplány* (+ *-i* ‘from, of’ suffix), the name of a village in the region that was demolished during the time of the Ottoman conquests.

An example of morphological phenomena can be found in patronymic surnames with the *-fi* suffix (‘-son’). Surnames with this suffix are characteristic of the Hungarian surname stock, albeit not very frequent, thus only two examples can be found on the regional top 100 lists in two different counties. The first is *Pálfi* (‘Paul’ + *-fi*) in Zala county in the South West, where this morphological surname type is the most widespread. The second is *Bánfi* (*Bán* ‘an obsolete title’ + *-fi*) in the South Eastern county Csongrád, and especially common in the surname stock of one specific town (Hódmezővásárhely), for currently unknown reasons of population history.

Finally, an example of a phonetic phenomenon, the surname *Kámán*, which has proven remarkably frequent in the South Western county, Zala (31st position, 746 name bearers). The surname *Kálmán*, ‘[given name] Kálmán, Coloman’, a variant of which *Kámán* is, can be found on the top 100 list in seven further counties, mainly in the western part of Hungary. Surnames originating from given names are relatively frequent in Zala county, with the mentioned lexical type (*Kálmán* ~ *Kámán*) being especially so. However, it is also this unique, local dialectal variant which has come to dominate here.

The regional frequency of individual surnames (here: how many county’s top 100 lists feature them) usually roughly correlates with their national frequency. Despite this, the surname *Orsós* ‘reelmaker or -seller’ shows an interesting deviation from this general observation. It is in the 31st position of the national top 100 list, while it is only found on 6 regional lists, but is exceptionally frequent in these (other names of this frequency are found on 17–19 regional lists, cf. VÖRÖS 2011: 195–201, FARKAS 2016: 47–48). As it is a surname characteristically borne by Roma people its unusual spatial distribution patterns reflect the special demographic features and surname characteristics of the Roma population in Hungary. The majority of Roma in Hungary have surnames of Hungarian linguistic origin, with a handful of surnames being outstandingly frequent among them.

## 7. Summary

This study has aimed to look at the questions of the etymological, typological and regional distribution of the Hungarian surname stock, in a comprehensive way, using large and representative name corpuses. To connect to the European

Surname Typology Project, and based on its methodology, the analysis was conducted using the stock of the one hundred most frequent surnames, which, within its own limits, provides a picture that can be considered representative. However, this overview should be studied in a wider, European context in the future.

### References

- BENŐ, ATTILA 2012. Gyakoribb román eredetű családneveink földrajzi elterjedtségének tanulságai. [Some Hungarian surnames of Romanian origin. Their geographical distribution and possible diachronic trends.] In: VÖRÖS, FERENC ed. *A nyelvföldrajztól a névföldrajzig III. Nyelvjárás – néprajz – művelődéstörténet*. Szombathely, Savaria University Press. 123–131.
- CHAREILLE, PASCAL–DARLU, PIERRE et al. 2017. *The European Surname Typology Project: A progress report*. Paper presented at the 26th International Congress of Onomastic Sciences, Debrecen.
- DHS. 2007 = *Database of Hungarian Surnames (2007)*. Project leader: HAJDÚ, MIHÁLY.
- DHS. 2009 = *Database of Hungarian Surnames (2009)*. Project leader: VÖRÖS, FERENC.
- FARKAS, TAMÁS 2013. Surnames of Ethnonymic Origin in the Hungarian Language. In: FELECAN, OLIVIU ed. *Name and Naming. Proceedings of the Second International Conference on Onomastics. Onomastics in Contemporary Public Space. Baia Mare, May 9–11, 2013*. Cluj-Napoca, Editura Mega–Editura Argonaut. 504–517.  
URL: [http://onomasticafelecan.ro/iconn2/proceedings/5\\_04\\_Farkas\\_Tamas\\_ICONN\\_2.pdf](http://onomasticafelecan.ro/iconn2/proceedings/5_04_Farkas_Tamas_ICONN_2.pdf)
- FARKAS, TAMÁS 2015. The most frequent Hungarian surnames. A study of some aspects of contrastive surname typology. In: FELECAN, OLIVIU ed. *Name and Naming. Proceedings of the Third International Conference on Onomastics “Name and Naming”. Conventional / Unconventional in Onomastics. Baia Mare, September 1–3, 2015*. Cluj-Napoca, Editura Mega–Editura Argonaut. 122–136.  
URL: [http://onomasticafelecan.ro/iconn3/proceedings/2\\_6\\_Farkas\\_Tamas\\_ICONN\\_3.pdf](http://onomasticafelecan.ro/iconn3/proceedings/2_6_Farkas_Tamas_ICONN_3.pdf)
- FARKAS, TAMÁS 2016. A magyar családnévanyag területi és tipológiai tagolódása. Vizsgálatok a leggyakoribb magyarországi családnevek alapján. [The geographical and typological distribution of the surname stock in Hungary today. Research on the basis of the most frequent surnames.] In: VÖRÖS, FERENC ed. *A nyelvföldrajztól a névföldrajzig VII. Névregeionalizmusok*. Szombathely, Savaria University Press. 39–66.



- FARKAS, TAMÁS 2017. Az erdélyi magyar családnévanyag tipológiai-statisztikai szempontú vizsgálata. [The typological-statistical analysis of the Hungarian surname stock in Romania.] In: VÖRÖS, FERENC ed. *A nyelvöldrajztól a névföldrajzig VIII. A nyelvterület keleti peremén*. Kolozsvár–Szombathely, Erdélyi Múzeum-Egyesület–Savaria University Press. 107–120.
- FARKAS, TAMÁS–N. FODOR, JÁNOS 2016. Surnames of Romanian origin in the Hungarian surname stock: *Árgyelán, Marosán, Moldován*. In: FELECAN, OLIVIU–FELECAN, DAIANA eds. *Magistri et alumni, amore scribendi. Studia. In honorem Professoris Nicolae Felecan*. Cluj-Napoca, Editura Mega–Editura Argonaut. 424–438.
- HAJDÚ, MIHÁLY 2010. *Családnévek enciklopédiája. Leggyakoribb mai családnéveink*. [Encyclopedia of surnames. Our most frequent surnames.] Tinta Kiadó, Budapest.
- LIPAN, FLORIN 2012. *Romanian names*. URL: <http://nume.ottomotor.ro>
- MURÁDIN, LÁSZLÓ 2005. *Erdélyi magyar családnévek*. [Hungarian surnames in Transylvania.] Nagyvárad, Europrint Könyvkiadó.
- VÖRÖS, FERENC 2011. Hol tartanak a mai magyar családnévátlasz munkálatai? [How far have we got in the work of compiling 'The Atlas of Contemporary Hungarian Surnames?'] *Névtani Értesítő* 33: 191–202.

## Abstract

The goal of this paper is, based on its specific methodology, to continue the research on the structure of the Hungarian surname stock today, focusing on its etymological, typological and geographical distribution. The analysis is based on the one hundred most frequent surnames regionally in Hungary (in the 19 counties and in the capital, Budapest), and, additionally, in the Hungarian surname stock in Romania. The paper initially presents the background for this study: (1) the framework of the European Surname Typology Project, as well as (2) the sources and the methodology of the typological-statistical processing. It analyses the geolinguistic characteristics of the regionally most frequent surnames: the distribution of the surnames of (3) non-Hungarian and (4) Hungarian linguistic origin, as well as (5) the pool of the most frequent (three) surnames in the various regions. Finally, the paper demonstrates (6) the possible different types of regional characteristics of the stock of the most frequent surnames. Within its own limits, the analysis provides an appropriate onomastic methodology, and a representative picture for the given surname stock, to be interpreted also in the broader context of European surname studies today.



**Keywords:** geonomastics, surname geography, surname typology, Hungarian names, language contacts