# Cross-linguistic research into affixation processes in European languages[1]

Lívia Körtvélyessy, P.J. Šafárik University, Košice, Slovakia
Ján Genči, Technical University, Košice, Slovakia

*The paper summarizes the fundamental principles of research into 73 European languages examined and evaluated on the basis of 100 word-formation characteristics. The focus of this paper is on affixation, in particular, the role of suffixation, prefixation, suffixal-prefixal derivation, circumfixation and infixation in forming new complex words. This objective is achieved by means of two main parameters, the parameter of structural complexity whose quantitative representation is saturation value (both absolute and relative) and the parameter of the measure of occurrence. Based on the evaluated data the authors identify the SAE core and periphery for suffixation, prefixation and prefixal-suffixal derivation.*

***Keywords:*** *word-formation, typology, affixation, structural complexity, measure of occurrence, Standard Average European*

## 1. Introduction

Körtvélyessy, Štekauer & Genči. (2017) present the basic theoretical and methodological principles as well as some the of results of a large-scale project aimed at typological research into word-formation systems of European languages. In this paper, we briefly summarize the method of research (Section 2). This will establish the scene for a detailed presentation of the affixation processes employed in the formation of new complex words in European languages (Section 3). The findings are summarized in Section 4.

## 2. Method of research

In this section, we briefly summarize the fundamental principles underlying the research project implementation.

### 2.1 *Word-formation comparables*

Körtvélyessy, Štekauer & Genči (2017) is the first typological analysis of word-formation *systems*. It is based on an analysis of 100 word-formation characteristics that serve a comprehensive comparison of word-formation systems in various languages. They are labeled as 'comparables', i.e., *linguistic features that are used to compare prototypical (theory-independent) manifestations of word-formation systems in sample languages.* We distinguish two types of comparables, *basic comparables* that include basic word-formation techniques of coining new words, and *complex comparables* that include word-formation processes. Our 100 basic comparables represent 12 complex comparables, including prefixation, suffixation,

[1] This research has been implemented with the financial support from the APVV-16-0035 and VEGA 1/0002/17 grants.

prefixal-suffixal derivation, circumfixation, infixation, postfixation, compounding, conversion, reduplication, blending and internal modification. The basic comparables of affixation are listed in 3.1.

2.2 *Sample languages*

Our sample covers 73 European languages that represent 22 genera and 7 language families (Appendix 1). Any selection of languages faces the problem of unequal classifications in various sources (e.g., compare the classifications in WALS vs. Ethnologue). Therefore, our language sample relies on the most recent and representative source of data on word-formation in European languages edited by Müller, Ohnheiser, Olsen & Rainer (2016).

2.3 *Data evaluation*

For the sake of data evaluation we follow the principles introduced in Körtvélyessy (2015; 2016), in particular, we analyze word-formation systems of sample languages in terms of their *structural richness* the quantitative expression of which is the *saturation value*. The parameter of structural richness indicates the degree to which a particular word-formation system makes use of all the available word-formation options, i.e., of the basic word-formation comparables.

The *absolute saturation value* indicates the absolute number of word-formation comparables used in a particular language, language genus and language family. Since, however, the individual word-formation processes are characterized by unequal numbers of basic comparables (27 in compounding, 17 in both suffixation and prefixation, 12 in conversion, 6 in reduplication, etc.), the absolute saturation value does not enable us to compare the relative role of various word-formation processes. This is achieved by the parameter of *relative saturation value* calculated for a particular word-formation process according to the following formula:

$$(1) \qquad \frac{AC}{\sum comparables} \times 100 = RSV\ (\%)$$

where *AC* stands for the number of actually realized comparables of a given WF process, ∑ *comparables* stands for the sum total of comparables available for a given word-formation process. *RSV* refers to relative saturation value expressed as a percentage.

Another parameter used for the evaluation of word-formation systems is the *measure of occurrence* of a comparable in an examined sample of languages. The *Maximum Comparable Occurrence* (MCO) identifies those basic word-formation comparables that occur in all languages of a sample, of a language genus or a family.

## 3. Affixation in European languages

The research covers the following affixation processes: suffixation (Section 3.1), prefixation (3.2), prefixal-suffixal derivation (3.3), circumfixation (3.4) and infixation (3.5). Section 3.6 evaluates the data from the point of view of relative saturation. Section 3.7 looks at the data from the perspective of Standard Average European.

## 3.1 *Suffixation*

### 3.1.1 *Suffixation comparables*

Both suffixation and prefixation are evaluated by means of 16 basic comparables defined by the word-class of the input and the output, and one comparable examining the possibility to form a new complex word by suffixation/prefixation of an already suffixed/prefixed word:

Table 1 *List of suffixation and prefixation comparables*

| | |
|---|---|
| Class-maintaining | N>N |
| | V>V |
| | A>A |
| | Adv > Adv |
| Class-changing | A>N |
| | V>N |
| | Adv>N |
| | A>V |
| | N>V |
| | Adv>V |
| | N>A |
| | V>A |
| | Adv>A |
| | N>Adv |
| | V>Adv |
| | A>Adv |
| multiple suffixation/prefixation | |

### 3.1.2 *Structural richness*

Suffixation is used in all 73 languages of our sample. The use of a word-formation process is acknowledged by the occurrence of at least one comparable of that WF process in a given language.

The sample includes 22 genera. They are obviously represented by different numbers of languages. This has two reasons: (i) genetic, and (ii) areal (restriction of our research to European languages). While, for example, the sample includes 13 Slavic, 9 Germanic, 8 Romance, 8 Turkic languages, 6 Lezgic languages, 5 Avar-Andic-Tsezic languges, the Baltic, Lak-Dargwa, Finnic, Permic and Iranian genera are represented by two languages each, and Basque, Albanian, Greek, Ugric, Mordvin, Mari and Semitic are represented by only one language each. For this reason, we will evaluate only the genera with 5 and more languages.

The data suggests that in the Slavic genus the Slovene word-formation system has the highest saturation value in terms of suffixation processes (16), followed by Bulgarian, Croatian, Slovak and Czech with 14 comparables. The lowest saturation value of suffixation is found in Upper Serbian (10).

In the Germanic genus, the structurally richest languages are German, Dutch, Swedish and Faroese (12 comparables); Frisian makes use of only 9 suffixation comparables.

In the Romance genus, the top position is assumed by Catalan (13), followed by Portuguese, Spanish, French and Romanian (11). On the other hand, Sardinian and Ladin use only 7 suffixation options.

The highest saturation value in the Turkic genus has been identified for Turkish (12), Crimean Tatar and Gagauz (11). The lowest saturation value characterizes Bashkir (7).

Budugh clearly dominates among the Lezgic languages (13); none of the other languages of this genus exceed the saturation value of 7.

Finally, the Avar-Andic-Tsezic languges are dominated by Avar and Bezhta (14). Akwakh is at the bottom rank with as many as 11 suffixation comparables, which in itself witnesses to a crucial role of suffixation in this genus. This situation is similar to that in Slavic languages.

3.1.3 *Measure of occurrence*

The most common suffixation comparable is the class-maintaining N>N suffixation that is used for the formation of new complex words in each of the 73 sample languages. It can therefore be considered a Euroversal, i.e., a universal word-formation feature of European languages. This is the only suffixation comparable with the maximum occurrence. The most common suffixation comparables are listed in Table 2:

Table 2 *Ten most common suffixation comparables in European languages*[2]

| **Suffixation comparable** | **No. of languages** |
|---|---|
| Class-maintaining N>N suffixation | 73 |
| Class-changing V>N suffixation | 71 |
| Class-changing A>N suffixation | 71 |
| Class-changing N>A suffixation | 70 |
| Class-changing V>A suffixation | 63 |
| Class-maintaining A>A suffixation | 59 |
| Class-maintaining V>V suffixation | 55 |
| Class-changing N>V suffixation | 54 |
| Class-changing A>Adv suffixation | 53 |
| Class-changing A>V suffixation | 52 |

At the opposite end of the scale we find the the Adv>V and Adv>N comparables that have the lowest occurrence in our European sample. They are used in 11 and 12 languages, respectively, primarily in the Avar-Andic-Tsezic genus (4 and 3 languages, respectively, out of 5 languages). The V>Adv type occurs in 15 languages, mainly in the Slavic (6) and North-Caucasian ones (3). The only class-maintaining type of suffixation at this end of the scale is Adv>Adv with 14 occurrences, mainly in Slavic languages (6). It is remarkable that there are no zero-occurrence WF comparables within suffixation.

3.2 *Prefixation*

3.2.1 *General*

The prefixation comparables are listed in Table 1.

Prefixation is not so widespread in the world's languages as suffixation. According to

[2] For a more detailed overview of and comments on the *Maximum Comparable Occurrences* by language genera and language families see Körtvélyessy et al. (2017).

the data in Štekauer, Valera & Körtvélyessy (2012), there are 53 out of 55 sample languages (96.37%) employing suffixation for word-formation in a productive way. Contrary to this, there are 'only' 39 languages (70.9%) employing derivational prefixes. An analogical picture is found in our sample of European languages. The use of suffixation by 73 languages (100%) is opposed to 52 languages (71.23%) employing prefixation.

The situation, however, differs from language family/genus to language family/genus. There are, for example, languages with an extremely large capacity of prefixes to express various semantic subtleties through the system of Aktionsart (Slavic languages, German, Latin, etc.). On the other hand, there are languages without derivational prefixes or with their minimum number, for example, languages of the Uralic family. For illustration, as pointed out by Wagner-Nagy (2016: 3206), in Nenets, there are two minor exceptions from the total absence of prefixation: the emphasizer prefix *wu-* and the interrogative prefix *xa-* attach to the negative auxiliary *ńiisʲ* to yield Tundra Nenets *wuńiisʲ* and Forest Nenets *wińiš.* These prefixes never attach to any other verb.

In Estonian, another Uralic language, there are only two genuine derivational prefixes, *eba-* and *mitte-*. Apart from them, there are semiprefixes combined with verbs, for example, *taas-* 're-; lit. again', *eel-* 'pre-', *üle-* 'over-', *ala-* 'under-', *järel-* 'after-' (*eel+hinda-ma* 'to pre-evaluate', *järel+küpse-ma* 'to ripen after picking; lit. after-mellow-INF', *taas+esita-ma* 'to represent; lit. again+present-INF', *üle+rahvasta-ma* 'to overpopulate) (Kerge 2016: 3240). However, views of derivational prefixes in Estonian differ. Thus, Kilgi (p.c.) maintains that there is only one prefix in Estonian, *eba-* 'false-, pseudo-, quasi-'. *eba* can also function as a stem: there are words made by adding suffixes, such as *eba-rd* 'monster, freak', *eba-le-ma* 'to hesitate'. These facts do not, however, affect the conclusion that Estonian has the capacity to form new complex words by prefixation.

Only two native prefixes can also be found in Finnish. They are *epä-* and *ei-*; both of them express negation. Moreover, according to Laakso (pers. comm.), the former of them is not very productive.

In Mordvinic there is only one (negative) prefix *a-*: *lamo* 'a lot' vs. *alamo* 'a few', *sati* 'enough' vs. *asati* 'not enough' (Maticsák 2016: 3289).

The Uralic family is a good example of the advantage of the method of evaluation based on the principle of structural complexity that does not depend on uncertainties concerning the number of prefixes and the degree of their productivity. What matters is the ability of the word-formation system to produce new complex words by prefixation.

### 3.2.2 *Structural richness*

The structural richness of prefixation is lower than that of suffixation in the majority of languages, language genera and language families. The structurally richest prefixation in the Slavic genus is found in Croatian (8 comparables), followed by Slovak and Czech with 6 comparables. In the Germanic genus, it is German with 9 and Dutch with 7 prefixation comparables. The Romance genus is dominated by Romanian (8) and Catalan (7). On the other hand, the only Semitic language in our sample, Maltese, employs 7 prefixation comparables. The Iranian language Ossetic has also a relatively high structural richness of prefixation (8 comparables). Aghul with 2 comparables is a sort of exception to an almost prefixation-less Lezgic languages. Almost all languages of the Uralic, Nakh-Daghestanian and Altaic language families are totally or almost totally devoid of the capacity to form new complex words by prefixation. It is especially true of the Altaic family featuring a total absence of prefixation.

From the point of view of genera, it is absent in Ugric, Mordvinic, Mari, Permic, Mongolian and Turkic.

It follows from this overview that the highest structural richness of prefixation in European languages is found in German. 9 comparables means about 53% saturation value, which is much lower compared to the maximum saturation value of suffixation, this being 94 % for Slovene. In general, the relative saturation value of suffixation in individual languages, genera and families is much higher than that of prefixation.

3.2.3 *Measure of occurrence*

The situation in terms of the measure of occurrence is parallel to the situation in structural complexity of prefixation. Unlike suffixation, there is no maximum comparable occurrence. The most common prefixation comparables are listed in Table 3. The field is dominated by the class-maintaining types of prefixation:

Table 3 *Most common prefixation comparables*

| **Prefixation comparable** | **No. of languages** |
|---|---|
| Class-maintaining V>V prefixation | 48 |
| Class-maintaining N>N prefixation | 45 |
| Class-maintaining A>A prefixation | 41 |
| Multiple prefixation | 39 |

The other prefixation comparables are much less common. Interestingly, there are four prefixation comparables that do not occur in any of European languages. They include the following comparables:

Table 4 *Zero occurrence prefixation*

| **Prefixation comparable** |
|---|
| Class-changing Adv>N prefixation |
| Class-changing Adv>A prefixation |
| Class-changing V>Adv prefixation |
| Class-changing A>Adv prefixation |

If the situation is analyzed by genera the results are analogical. In the Indo-European genera (Slavic, Germanic, Romance, Baltic and Celtic) the most common are three class-maintaining prefixation comparables. The other families are short of prefixation. Therefore, this aspect cannot be further evaluated.

3.3 *Prefixal-suffixal derivation*

3.3.1 *General*

It is necessary to distinguish between prefixal-suffixal derivation and circumfixation. The former is the process of double affixation, i.e., the attachment of two affixes simultaneously, the latter means the attachment of one single affix split in two parts, one of them placed to the beginning, the other to the end of the word-formation base.

This word-formation process, also labeled as parasynthesis, occurs in 28 languages (38.36%) of our sample. It is totally absent in the Uralic and Altaic families. On the other hand, it is fairly common in Slavic and Romance languages in which it is characterized by a

multiplicity of its formal and semantic manifestations. It is, for example, productive in the formation of Portuguese verbs from nouns and adjectives through a number of patterns (Villalva 2003: 955):

(2) Denominal verbs
*a*-N-*e*-: *assenhorear-se* 'to take possession of' < *senhor* 'master'
*a*-N-*ec*-: *anoitecer* 'to dawn' < *noite* 'night'
*en*-N-*iz*-: *encolerizar* 'to exasperate' < *cólera* 'fury'
*des*-N-*e*-: *descantear* 'to smoothen the edges' <*canto* 'edge'

(3) Deadjectival verbs
*a*-A-*e*-: *aformosear* 'to embellish' < *formoso* 'beautiful'
*a*-A-*ec*-: *amadurecer* 'to ripen' < *maduro* 'mature'
*a*-A-*ej*: *anegrejar* 'to turn black' < *negro* 'black'
*en*-A-*e*-: *engalhardear* 'to make elegant, graceful' < *galhardo* 'spruce, elegant'

In Sardinian, the productive patterns include *a(d)-...-are*/*ai*, *in-...-are*/*-ai*, and *(i)s-…-are*/*-ai*. Semantically, the resulting complex words can be classified into two macro-categories, ingressive (*a(d)-...-are*/*ai* and *in-...-are*/*-ai*) and egressive (*(i)s-…-are*/*-ai*):

(4) Ingressive: *abbrayare*$_{V}$ 'to become embers' < *bráya*$_{N}$ 'embers'
Egressive: *illanare*$_{V}$ 'to shear the flock' < *lána*$_{N}$ 'wool' (Pinto 2016: 2702)

Apart from these two semantic categories, prefixal-suffixal derivation covers a number of other semantic categories, including abstract concepts, human beings, animals, things, place names, similarity, quality, result of action, etc. The enormous capacity of this word-formation process can be illustrated by a range of meanings, mostly, but not exclusively, within the semantic category of Cause or Result of Action (especially when the base is an adjective) that can be expressed by verbal prefixal-suffixal derivation.

(5)

| | | |
|---|---|---|
| a. to remove N | Slovak | *od-vod-niť* 'to drain'< *voda* 'water' |
| b. to make N: | Russian | *s-grud-i-t'* 'to make a heap' < *gruda* 'heap' |
| c. to act with N: | Russian | *s-bolt-it'* 'to bolt' < *bolt* bolt' |
| d. to approach N: | Russian | *pri-gub-it*' 'to take a little sip' < *guba* 'lip' |
| e. to get covered with N: | Russian | *za-mš-et'* 'to get covered with moss' < *moh* 'moss' |
| f. to be deprived of N: | Russian | *obes-pamât-et'* 'to lose memory' < *pamât'* 'memory' (Uluhanov 2016: 2965) |
| g. not without N | Kabardian | *mə-aqəλ-ə-nša* 'not without intelligence' (Colarusso 1992: 150) |
| h. to cause to be A | Belarussian | *z-bâdn-i-c´* 'to make sb. poor' < *bednych* 'poor' |
| i. to cause V | Belarussian | *us-kryk-vac´* 'to give a cry' < *krychac'* 'to cry' |
| j. duration of V | Belarussian | *pry-gavor-vac´* 'to keep talking' < gavaryc' 'to talk' (Lukašanec 2016: 2947) |

Adverbs, too, are within the realm of prefixal-suffixal derivation:

(6) a. Location Croatian *na-leđ-ice* ‘on one’s back’ < *leđa* ‘back’
Croatian *po-sam-ce* ‘individually’ < *sam* ‘alone’
(Grčevič 2016: 3012)
b. Manner Russian *v-pravd-u* ‘indeed; lit. in truth’; *po-nov-omu* ‘anew’
c. Time Russian *v-načal-e* ‘in the beginning’ < *nachalo* ‘beginning’
*iz-davn-a* ‘from time immemorial’ < *davno* ‘long ago’
(Uluhanov 2016: 2970)

Close to this derivational process is productive combination of prefixation and conversion (zero derivation) in some languages. The following examples come from Portuguese:

(7) *a*-N-0 *ajoelhar* ‘to kneel down’ < *joelho* ‚knee‘
*des*-N-0 *descafeinar* ‘to decaffeinate’ < *cafeína*
*en*-N-0 *enterrar* ‘to bury’ < *terra* ‘earth’ (Pöll 2016: 2612)

and Croatian

(8) bez-N-0 *bez-imen* ‘nameless’ < *ime* ‘name’ (Grčevič 2016: 3009)

3.3.2 *Structural richness*
Prefixal-suffixal derivation is evaluated according to the output word-class. By implication, we distinguish four basic types, nominal, verbal, adjectival and adverbial. A high structural richness in Slavic languages is evidenced by the fact that 9 out of 13 Slavic languages make use of all four types. Exceptions include Bulgarian, Macedonian and Upper Sorbian (3 types – without adverb-formation) and Kashubian (1 type – only verb-formation).

The formation of new words by prefixal-suffixal derivation in Romance languages is limited to two word-classes (verbs and adjectives). Other languages of the sample do not employ this word-formation processes in a significant way, with the exception of the North-Caucasian language Kabardian that makes use of it for the formation of nouns and adjectives.

3.3.3 *Measure of occurrence*
The highest occurrence of prefixal-suffixal derivation in European languages has been observed for verb-formation (26 languages) followed by adjective-formation (22 languages). Nouns are formed in 15 languages and adverbs in 10 languages. If we go by genera, we identify the maximum comparable occurrence for Slavic verbs that are formed in all 13 languages, adjectives and nouns are formed in 12 languages (absent in Kashubian). Adverbs have the occurrence of 9. In Romance languages both verbs and adjectives have the occurrence of 8. The values for all the other languages are negligible.

3.4 *Circumfixation*

This is not a common WF process in European languages. It is productively used only in two Germanic languages, German and English. Examples are given in (9):

(9) German *Ge-red-e* ‘gossip’ < *reden* ‘speak’ (Barz 2016: 2395)
Dutch *ge-berg-te* ‘mountain chain’ < *berg* ‘mountain’ (Booij 2016:2439)

### 3.5 *Infixation*

There is no productive infixation in word-formation of European languages.

### 3.6 *Relative saturation*

The relative saturation value (RSV) enables us to compare word-formation processes as a whole. It tells us which of the word-formation processes is structurally richest or, in other words, which of them most significantly contributes to the structural richness of the word-formation system of a given language. In the following, we provide the relevant data on suffixation, prefixation and suffixal-prefixal derivation in selected language genera.

#### 3.6.1 *Slavic genus*

In the Slavic genus, the highest RSV is identified for Slovene (94.12%) suggesting that this language makes of 94.12% of the available suffixation possibilities. Its RSV for prefixal-suffixal derivation reaches 100% while the role of simple prefixation in the Slovene system is rather low with mere 23.53%. This is the lowest prefixation RSV in the Slavic genus (identical to Bulgarian, Polish and Upper Sorbian).

The gap between the RSV of suffixation and prefixation is smaller in Czech and Slovak. Both of these languages feature the proportion of suffixation to prefixation of 82.35% to 35.29%. The smallest proportion of this sort is found in Croatian which has the same RSV of suffixation as Czech and Slovak but the highest RSV from among all Slavic languages (47.06%).

In general, the RSV of prefixation in Slavic languages might seem surprisingly low, but it should be noted that this parameter does not reflect the already mentioned enormous capacity of expressing various semantic subtleties. The RSV reflects the degree of utilization of formal possibilities of word-formation.

#### 3.6.2 *Germanic genus*

None of the Germanic languages reaches the RSV of the top Slavic languages in suffixation. The highest RSV is identified for German, Dutch, Swedish, and Faroese (70.59%). In contrast to Slavic languages, however, the RSV of prefixation is fairly high in German (52.94%) and Dutch (41.18%). It is especially in German that the relative contribution of suffixation and prefixation to the overall structural richness is close. This is different from Swedish where this gap is considerably high (70.59% vs. 29.41%). A similar proportion between suffixation and prefixation, at a lower level, though, can be observed in Icelandic (58.82% vs. 41.18%). The minimum role of prefixal-suffixal derivation in these languages is also reflected in their RSVs. The RSV of German and English is 25%, the other Germanic languages do not employ this word-formation process.

#### 3.6.3 *Romance genus*

With the exception of Catalan (76.47%), the suffixation RSV does not exceed 65%. The prefixation values range from fairly high (47.06% for Romanian and 41.18% for Catalan) to extremely low (5.88% for Sardinian, 11.76% for Ladin, and 17.65% for Italian. In addition, three other languages have the prefixation RSV under 30% (French, Spanish and Potuguese). All languages of this genus have a 50% RSV of prefixal-suffixal derivation.

3.6.4 *Lezgic languages*

Budugh assumes a special position among these languages because its suffixation RSV of 76.47% significantly exceeds the RSVs of the other languages – all of them are below 50%. The RSVs for prefixation are extremely low: 11.76% for Rutul and Udi, 5.88 for Budugh, and 0% for the other three languages. The big gap between suffixation and prefixation RSVs for Budugh clearly show the great difference between their respective contribution to the overall affixation RSV. None of the Lezgic languages employs prefixal-suffixal derivation.

3.6.5 *Turkic languages*

The suffixation RSV in Turkic languages is fairly high because only one of these languages (Bashkir) falls below 50%. Turkish has the highest suffixation RSV (76.47%) and two other languages have their RSVs above 70% (Gagauz and Crimean Tatar). Neither prefixation nor prefixal-suffixal derivation are used in Turkic languages.

3.7 *SAE perspective*

We have pointed out on several occasions (e.g., Körtvélyessy 2015, Körtvélyessy, Štekauer & Genči 2017) that the identification of SAE core depends on the specific parameters selected for the evaluation. Consequently, the conclusions of, for instance, Kortmann (1998), Haspelmath (1998), de Auwera (1998), Körtvélyessy (2015), and Körtvélyessy, Štekauer and Genči (2017) differ.

In the following, we evaluate the core and the peripheral languages for suffixation, prefixation and prefixal-suffixal derivation on the basis of relative saturation value.

3.7.1 *Suffixation*

Table 5 gives languages with the highest RSV in suffixation:

Table 5 *European languages with the highest RSV for suffixation*

| **Language** | **Genus** | **Family** | **SV** | **RSV (in %)** |
|---|---|---|---|---|
| Slovene | Slavic | IE | 16 | 94.12 |
| Basque | Basque | Basque | 15 | 88.24 |
| Kabardian | Northwest Caucasian | Northwest Caucasian | 15 | 88.24 |
| Kalmyk | Mongolian | Altaic | 14 | 88.24 |
| Croatian | Slavic | IE | 14 | 82.35 |
| Bulgarian | Slavic | IE | 14 | 82.35 |
| Czech | Slavic | IE | 14 | 82.35 |
| Slovak | Slavic | IE | 14 | 82.35 |
| Estonian | Finnic | Uralic | 14 | 82.35 |
| Abkhaz | Northwest Caucasian | Northwest Caucasian | 14 | 82.35 |
| Bezhta | Avar-Andic-Tsezic | Nakh-Daghestanian | 14 | 82.35 |
| Avar | Avar-Andic-Tsezic | Nakh-Daghestanian | 14 | 82.35 |

Table 5 suggests that the highest relative contribution of suffixation to word-formation in European languages should be sought primarily among Slavic languages, partly among

Northwest Caucasian and Avar-Andic-Tsezic languages. This finding enables us to identify the SAE core. The periphery is situated on the territories spoken by the Iranian language Ossetic and the Lezgic languages Udi, Archi and Aghul whose RSVs lie under 40%.

As a next step, we can compare the RSV in language genera as a whole. Only those genera are taken into account for this purpose that are represented by at least 5 languages in our sample. They include the Slavic, Germanic, Romance, Lezgian, Avar-Andix-Tsezic and Turkic genera. The results are specified in Table 6:

Table 6 *Comparison of suffixation RSV in selected language genera*

| **Language genus** | **RSV (%)** |
|---|---|
| Avar-Andic-Tsezic | 75.29 |
| Slavic | 72.40 |
| Germanic | 64.71 |
| Turkic | 63.87 |
| Romance | 58.09 |
| Lezgic | 43.14 |

Genus-wise, the 'centre of European suffixation lies in the non-IE Avar-Andic-Tsezic and the IE Slavic genus. What makes these two genera substantially different from each other is the almost total reliance of the former genus on suffixation, with the RSV of prefixation approaching zero and the RSV of prefixal-suffixal derivation being zero, on one hand, and a much more balanced contribution of these three affixation processes in the Slavic genus, on the other hand.

3.7.2 *Prefixation*

The topmost prefixation RSVs are given in Table 7:

Table 7 *European languages with the highest RSV for prefixation*

| **Language** | **Genus** | **Family** | **SV** | **RSV%** |
|---|---|---|---|---|
| German | Germanic | IE | 9 | 52.94 |
| Croatian | Slavic | IE | 8 | 47.06 |
| Romanian | Romance | IE | 8 | 47.06 |
| Ossetic | Iranian | IE | 8 | 47.06 |
| Dutch | Germanic | IE | 7 | 41.18 |
| Icelandic | Germanic | IE | 7 | 41.18 |
| Catalan | Romance | IE | 7 | 41.18 |
| Breton | Celtic | IE | 7 | 41.18 |
| Maltese | Semitic | Afro-Asiatic | 7 | 41.18 |

Maltese is the only non-IE language among the top prefixation languages. This word-formation process is evidently concentrated on the territory spoken by IE languages. On the other hand, the SAE periphery in terms of prefixation is formed by languages in which prefixation is not used at all (or used by a very restricted number of prefixation options) for word-formation. They include the majority of Uralic and Nakh-Daghestanian languages.

What is striking at first sight is the fact that – when we compare the RSVs for the topmost suffixation and prefixation languages, the relative contribution of prefixation to the overall structural complexity of European word-formation systems is about half the RSV of suffixation. This indicates a *strong suffixation preference* of European languages in the field

of derivation. This is supported by the fact that has already been mentioned above: in contrast to suffixation there are a number of European languages without derivational prefixes.

These observations are also supported by a comparison of the genera identical to those listed in Table 6. First, the RSVs of three topmost genera are less than a half of their RSVs for suffixation. Second, the prefixation RSV of three language genera approaches or is 0%.

Table 8 *Comparison of prefixation RSV in selected language genera*

| **Language genus** | **RSV (%)** |
|---|---|
| Germanic | 33.99 |
| Slavic | 29.86 |
| Romance | 25.00 |
| Lezgic | 4.9 |
| Avar-Andic-Tsezic | 1.18 |
| Turkic | 0 |

3.7.3 *Prefixal-suffixal derivation*

Prefixal-suffixal derivation reaches very high values in the Slavic genus where 9 out of 13 languages feature a 100% RSV, and only Kashubian has its RSV under 50%. All Romance languages as well as Kabardian have a 50% RSV. None of the other European languages exceed 25%. The majority of them do not employ this kind of derivation. The SAE centre for prefixal-suffixal derivation thus unambiguously lies on the territory spoken by Slavic languages.

## 4. Conclusions

The results of our research data analysis enable us to draw the following conclusions:

(i) Suffixation is the dominant affixation process used for the formation of new complex words in European languages. In some of its options it is used by all sample languages.

(ii) The data on the use of suffixation and prefixation in European languages map the data of the world's sample.

(iii) The use of prefixation across European languages is not distributed evenly. Prefixation is primarily used by Indo-European languages.

(iv) The structural complexity of prefixation is in general about half the structural complexity of suffixation.

(v) The parameter of comparable occurrence provides additional evidence of the dominant position of suffixation. The top ten suffixation comparables feature a higher occurrence in European languages than the most widespread prefixation comparable.

(vi) There is one suffixation comparable with maximum occurrence, i.e., occurrence in all sample languages. No such prefixation comparable has been identified. On the other hand, while there are four zero-occurrence comparables in prefixation no such phenomenon has been identified for suffixation.

(vii) Circumfixation and infixation do not contribute to structural complexity of word-formation systems in Europe.

(viii) The centre of SAE depends on the parameter chosen. From the point of view of suffixation, the core of SAE is represented by Slavic languages, partly by Northwest Caucasian and Avar-Andic-Tsezic languages. From the point of view of prefixation, the SAE core is constituted by German, Croatian, Romanian and Ossetic. The SAE core for prefixal-suffixal derivation lies on the territory spoken by Slavic languages.

**References**


Barz, Irmhild. 2016. German. In P.O. Müller et al., 2387-2410.

Booij, Geert. 2016. Dutch. In P.O. Müller et al., 2427-2451.

Colarusso, John. 1992. *A Grammar of the Kabardian language*. Univgersity of Calgary Press.

Grčevič, Mario. 2017. Croatian. In P.O. Müller et al., 2998-3016.

Haspelmath, Martin. 1998. How young is Standard Average European? *Language Sciences* 20(3). 271-287.

Kerge, Krista. 2016. Estonian. In P.O. Müller et al., 3228-3259.

Kortmann, Bernd. 1998. Adverbial subordinators in the languages of Europe. In: J. van der Auwera & D. P. Baoill (eds.), *Adverbial Constructions in the languages of Europe. Empirical Approaches to Language Typology - EUROTYP 20-3,* 457-561. Berlin/New York: Mouton de Gruyter.

Körtvélyessy Lívia. 2015. *Evaluative Morphology from a Cross-Linguistic Perspective.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Körtvélyessy Lívia & Štekauer, Pavol & Genči, Ján & Zimmermann, Július. 2017.Word-formation in SAE languages. Manuscript.

Lukašanec, Alaksandr. 2016. Belarussian. In P.O. Müller et al., 2932-2953.

Maticsák, Sándor. 2016. Mordvinic. In P.O. Müller et al., 3288-3307.

Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.). 2016.*Word-Formation. An International Handbook of the Languages of Europe.* Berlin: Mouton.

Pinto, Immacolata. 2016. Sardinian. In P.O. Müller et al., 2693-2712.

Pöll, Bernhard. 2016. Portuguese. In P.O. Müller et al., 2600-2619.

Štekauer, Pavol & Valera, Salvador & Körtvélyessy, Lívia. 2012. *Word-Formation in the World's Languages. A Typological Survey*. Cambridge: Cambridge University Press.

Uluhanov, Igor. S. 2016. In P.O. Müller et al., 2953-2978.

Van de Auwera, Johan. 1998. Conclusion. In J. van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*, 813-833. Berlin/New York: Mouton de Gruyter.

Villalva, Alina 2003 Formação de palavras: afixação. In: M. H. Mira Mateus & A. M. Brito & I. Duarte & I. Hub Faria (eds.), *Gramática da Língua Portuguesa*, 939–967. Lisboa: Caminho.

Wagner-Nagy, Beáta. 2016. Nenets. In P.O. Müller et al., 3197-3209.

## Appendix 1 List of sample languages

| Family | Genus | Language |
|---|---|---|
| INDO-EUROPEAN | SLAVIC | Macedonian |
| | | Slovene |
| | | Croatian |
| | | Serbian |
| | | Bulgarian |
| | | Sorbian, Upper |
| | | Polish |
| | | Kashubian |
| | | Czech |
| | | Slovak |
| | | Ukrainian |
| | | Belarusian |
| | | Russian |
| | GERMANIC | English |
| | | German |
| | | Dutch |
| | | Frisian |
| | | Faroese |
| | | Danish |
| | | Swedish |
| | | Icelandic |
| | | Norwegian |
| | ROMANCE | Portuguese |
| | | Spanish |
| | | French |
| | | Sardinian |
| | | Italian |
| | | Romanian |
| | | Ladin |
| | | Catalan |
| | BALTIC | Latvian |
| | | Lithuanian |
| | CELTIC | Breton |
| | | Irish |
| | | Welsh |
| | IRANIAN | Ossetic |
| | | Tat |
| | ALBANIAN | Albanian |
| | GREEK | Greek |
| BASQUE | BASQUE | Basque |
| URALIC | SAMOYEDIC | Nenets |

| | | |
|---|---|---|
| | FINNIC | Estonian |
| | | Finnish |
| | UGRIC | Hungarian |
| | MORDVIN | Mordvin |
| | MARI | Mari |
| | PERMIC | Udmurt |
| | | Komi |
| NORTHWEST CAUCASIAN | NORTHWEST CAUCASIAN | Abkhaz |
| | | Adyghe |
| | | Kabardian |
| NAKH-DAGHESTANIAN | LEZGIC | Rutul |
| | | Budugh |
| | | Udi |
| | | Aghul |
| | | Archi |
| | | Khinalug |
| | LAK-DARGWA | Lak |
| | | Dargwa |
| | AVAR-ANDIC-TSEZIC | Bezhta |
| | | Botlikh |
| | | Akhwakh |
| | | Avar |
| | | Khwarshi |
| ALTAIC | MONGOLIAN | Kalmyk |
| | TURKIC | Turkish |
| | | Bashkir |
| | | Tatar |
| | | Crimean Tatar |
| | | Gagauz |
| | | Karaim |
| | | Chuvash |
| AFRO-ASIATIC | SEMITIC | Maltese |

*Lívia Körtvélyessy*
*Department of British and American Studies*
*P.J.Safarik University*
*Kosice*
*Slovakia*
*livia.kortvelyessy@upjs.sk*

*Ján Genči*
*Technical University*
*Kosice*
*Slovakia*
*jan.genci@tuke.sk*

In SKASE Journal of Theoretical Linguistics [online]. 2017, vol. 14, no.2 [cit. 2017-11-02]. Available on web page http://www.skase.sk/Volumes/JTL35/pdf_doc/02.pdf. ISSN 1336-782X.