DOGS KNOW IT, TREES WAIT FOR IT, AND THE WIND SNATCHES IT: VERBAL COLLEXEMES AND SEMANTIC DOMAINS OF NOUN + VERB PERSONIFICATIONS IN HUNGARIAN

GÁBOR SIMON

ELTE Eötvös Loránd University
simon.gabor@btk.elte.hu
https://orcid.org/ 0000-0001-5233-6313

Abstract

Verbal constructions of personification (i.e. nonhuman subject + predicate verb primarily used for human beings only) can be considered a key linguistic realization of personification. Although as grammatical metaphors they are rather invisible, the verb's selection restrictions have a crucial role in expressing conceptual personification. In other words, the nominal form of personification as cross-domain mapping (e.g. WIND IS A THIEF) is often realized via verb + argument structures (e.g. *the wind grabbed the papers out of my hand*). In a previous study, the domains of MOVEMENT, CONTROL and MENTAL ACT proved to be the most frequent conceptual categories of verbal personifications in a corpus of 20th-century Hungarian poems. However, despite these initial findings, we have relatively little knowledge about what are the typical verbal components of personifications in Hungarian on a more general level.

The present paper aims to extend the scope of personification research, adopting corpus linguistic methodology to explore the semantic domains of verbal personification in Hungarian. In a collostructional analysis, I investigate the significant verbal collexemes of noun + verb personifications in the huTenTen12 corpus, using three categories of nominal keywords: animals (*dog, horse, fish*), plants (*tree, fruit, flower*) and inanimate natural phenomena (*water, air, fire*). The study hypothesises that verbs of MENTAL ACT will dominate the first category, MOVEMENT and CONTROL will be prominent among the verbal collexemes of nouns denoting personified natural phenomena, while the semantic domains of verbal personifications will be more heterogeneous in the realm of plants.

Keywords: personification, corpus, collostruction, collexeme, conceptual domain

1. Introduction

Despite recent developments in research on personification (see Dorst 2011, Dorst–Mulder–Steen 2011, Melion–Ramakers 2016), we have relatively little knowledge about the linguistic organization of personifying expressions. Clearly, one reason for this is that personification does not have a specific linguistic structure. Its variability extends from verbal expressions to adverbial, nominal and adjectival structures to mention only one linguistic factor, part of speech category (for an overview of the variability of personification in Hungarian, see Sájter 2008). Another likely reason is the preference for the exploration of conceptual modelling in cognitive linguistics, instead of focusing on the linguistic realization of personification. This approach assumes that one can distinguish the linguistic manifestation of personifications from their conceptual organization, and this is exactly what is emphasized by Aletta G. Dorst (2011: 122): a human-related but conventional meaning of a verb (e.g., *run, take, give* or *make*) does not count automatically as personification at the level of conceptualization, whereas at the level of lexical meaning it can be identified as personification, and in a specific context (e.g., in poetry) such personifying effects may become stronger. Thus, the literature emphasizes that the definition of personification may vary with the perspective and scope

DOI: 10.54888/slh.2022.34.5.21

of the analysis. However, we can also agree with Dorst (2011: 132) that "[t]he literature suggests a bias in favour of conceptual personification", which means that in the last decades researchers of the field were interested in shedding light on the conceptual organization of personification, without making observations about its linguistic realization. (An important exception is the work of Aletta G. Dorst and her colleagues, see Dorst–Mulder–Steen 2011.) The overall aim of the present study is to change this situation by initiating a thorough corpus linguistic examination of noun + verb personification in Hungarian.

Due to the fact that contemporary cognitive linguistics puts a premium on explaining the conceptual background of figurative language use, there are numerous alternative models for personifying meaning generation in the literature. Without going into the details, personification can be described as a kind of conceptual metaphor, either as an ontological metaphor (see Kövecses 2010) or as the specification of the EVENTS ARE ACTIONS generic metaphor (see Lakoff 2006); as motivated by conceptual metonymy (Low 1999, Dorst–Mulder–Steen 2011);[1] or as the functioning of more than one type of conceptual integration (Long 2018). Moreover, Bocharova (2016) suggests a structured connectionist approach to personifying meaning, in which embodied experience, basic-level categories, and primary and complex metaphors contribute to the process of meaning generation through projection and cultural associations. And though there is no consensus on whether one or more conceptual operations are essential for personification, the researcher can encounter an abundance of models and analyses in the field of cognitive linguistics.

The result of an overview is much more modest when we turn to the linguistic realization of personifying conceptualizations. There is a scarcity of studies focusing on the language of personification. In her aforementioned paper, Dorst (2011) highlights the difference between the conceptual organization of verbal personifications on the one hand and personifying meaning generation with nominal expressions on the other. The main difference lies in the process of elaborating the meaning: whereas in the case of verb + argument constructions the argument violates the selectional restrictions of the verb (or at least some of them, Dorst 2011: 119), with nominal expressions (e.g., body-part personifications) two conceptual structures are mapped onto each other, initiating the reconceptualization of an entity with the use of another (human) one. Thus, verbal, adjectival, and adverbial personifications are based on the elaboration of one piece of a conceptual representation (for instance the primary schematic figure) with a non-expected entity, while the conceptualization of nominal personifications extends to the whole concepts involved in the process of construal. According to this proposal, we can agree with Dorst (2011: 117) that "at the linguistic level, the role of word class cannot be ignored" in the examination of personifications.

Another striking aspect of linguistic personifications is that they are not confined to only one word in the discourse: as noted by Long (2018: 25), personifying expressions can be considered "extended units of meaning", including a node word and its collocation (frequently co-occurring context), colligation (a grammatical category associated with the node word), semantic preference (the meaning categories frequently occurring in the context of the node word) and semantic prosody (the positive or negative evaluation associated with the whole unit). In other words, linguistic personifications are multi-word expressions in general, and this statement is supported by the empirical result of the study reported by Dorst–Mulder–Steen (2011: 192): 62% of personifications recognized by non-expert readers were word-combinations.

As regards the make-up of these combinations, Long (2018: 23) observes that they can be varied but they usually consist of a "predicate verb (used for human beings only)", a "nonhuman subject" and "others" (i.e., other potential elements of the clause). Although this description seems to be an appropriate starting point for further analysis, it is clear that the expressions instantiating one variation of this pattern are verbal personifications, therefore this proposal cannot be extended to the description of other non-verbal expressions of personification. In addition, Long's observations

[1] Although Low (1999) regards metonymically motivated personification as "weak" personification, the results of the empirical study performed by Dorst and her colleagues contradict this wiew: according to their identification test metonymic personifications behave similar to novel (i.e., creative) personifications (see Dorst–Mulder–Steen 2011).

come from some illustrative examples of personification in English, lacking any solid and large-scale empirical basis.

There is a similar problem with Sájter's (2008) comprehensive description of Hungarian personifications. Although it offers a rich list of typical linguistic patterns of personification in Hungarian, including e.g. verb + argument structures, the possessive construction, nominal and adverbial cases, and the vocative, this yields a rather heterogeneous category (extending to morphological, syntactic, and even pragmatic markers). Moreover, the list is the result of individual observations, it is not based on systematic empirical investigation. These lists, thus, serve at most as a baseline for further research, and they cannot be evaluated as exhaustive descriptions of the linguistic manifestation of personifications in a language.

Finally, even the most specified grammatical pattern does not ensure that personification unfolds. As highlighted by Dorst (2011: 122–123), when the verbal element of a personifying expression is polysemous, it can have a conventionalized personifying meaning that is directly accessible to the language user. Thus, the basic human meaning remains inactive in the process of meaning generation, and though the verb + argument structure fails to satisfy the original selectional restrictions of the verb, the frequent personifying use of the verb makes it possible to override this process, and treat the actual occurrence of the verb as conventional. Dorst's example is the verb *run*, which is highly polysemous, and according to the Macmillan dictionary, 10 out of its 19 senses are non-human, hence inherently personifying. These "dead personifications" instantiate the grammatical schema without initiating personifying meaning in the discourse.

To sum up, we know that the participating elements' part of speech category is an important factor in generating personifying meaning, but we do not know what instances of a given grammatical category are involved in personifications.[2] We know that personification is generally a multi-word expression in discourse, but we do not know the exact patterns of such expressions. And we know that personification is an extended unit of meaning including semantic preference as well, but we do not know what these preferences are in the discourse. Put differently, previous research on linguistic personification has brought important aspects to light, without turning the qualitative analysis of personifying expressions into a full-fledged empirical description.

Corpus linguistics seems to be the next step in the research since it can provide a solid empirical foundation and new methods for analysing personifying structures. The present study demonstrates the benefits of using corpus linguistic methodology in the analysis of figurative language use. Not on the level of corpus building and annotation (for these issues see Simon 2022), but on pattern extraction from existing corpora and examination of the observed patterns. The investigation focuses on verb + argument structures. This kind of construction is known to be central in expressing personifying meaning, however, we have relatively little knowledge about what verbs and what arguments participate in personification frequently in Hungarian. Accordingly, the research questions of the study are the following: (i) How can we explore typical noun + verb personifications with the use of corpus linguistic methodology? (ii) What is the difference between typical verbal personifications related to different personified entities?

As a possible answer to the second question, I assume that the semantic preferences of various classes of nouns will differ, and this difference can be observed by examining the typical, recurring verbal components of these personifications. Specifically, my hypothesis is that while nouns referring to animals are personified dominantly in the domain of MENTAL ACT (i.e., mental capacity is frequently attributed to animals), in the case of plants the preference lies in the MOVEMENT domain (i.e., the typical personification of plants describes them as moving agents or agents being able to move their body), and nouns referring to natural phenomena prefer the domain of CONTROL (i.e., intentional and self-performed manipulation of objects; about the domains see Dorst–Mulder–

[2] The dictionary-based method for personification identification elaborated by Dorst and her colleagues (Dorst–Mulder–Steen 2011, for Hungarian see Simon 2022) can prove a solution to this problem. However, it requires manual analysis of each and every component of a potentially personifying expression, therefore it is really time-consuming and usable only in small-scale research.

Steen 2011, for their application to Hungarian data see Simon 2021). In the following, I discuss the material and methodology of my analysis (2). Then the results are presented in the form of a thorough examination of verbal collexemes of one keyword from each category (3.1) as well as a broader overview of the extracted patterns (3.2). The paper ends with some concluding remarks (4).

2. Material and method

The first, merely technical assumption of the present study is that corpus linguistic methods can shed new light on linguistic personification. Particularly, I refer here to the ways of extracting and analysing a large sample of data, within the broader framework of a corpus-driven approach to language (Tognini-Bonelli 2001, Biber 2010). For such an investigation one needs an existing corpus of a language composed of oral, written and/or online texts. I have chosen the Hungarian Web Corpus (huTenTen12), which is a corpus of online Hungarian texts with a size of almost 2.5 billion words. The corpus is available in the Sketch Engine web-based corpus analysing platform,[3] and its linguistic material is tokenized, lemmatized, PoS-tagged, and morphologically preprocessed. Although the Hungarian Web Corpus is not the most modern linguistic database of Hungarian, it is relatively new, and it is a member of the so-called TenTen corpus family (see Jakubícek et al. 2013), which is a collection of corpora designed with the same theoretical and methodological principles for different languages. (Image 1 illustrates the dashboard surface of the enTenTen15 corpus in Sketch Engine: it is the English equivalent of the Hungarian Web Corpus, compiled in 2015.)

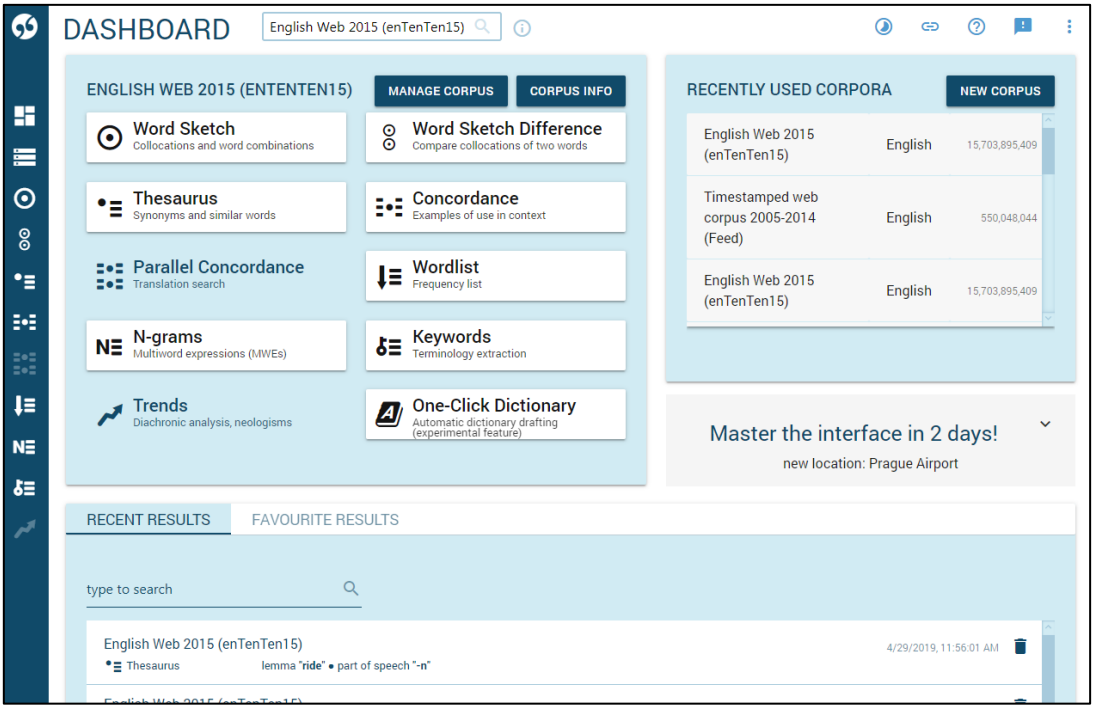

Image 1. The online surface of the enTenTen15 corpus in Sketch Engine

One of the advantages of the huTenTen12 corpus is that it was annotated both with the tagset of the Hungarian National Corpus (MSD code tagset)[4] and with an independent morphological tagset

[3] https://www.sketchengine.eu/hutenten-hungarian-corpus/?gclid=EAIaIQobChMI37Hrqe65-gIV0_V3Ch3SJgh3EAAYASAAEgL5VvD_BwE (last access: 29/09/2022)

[4] The tagset is available here: https://www.sketchengine.eu/hungarian-msd-code/ (last access: 22/11/2022).

(emorph tagset),[5] thus, its annotational precision outperforms other available corpora in Hungarian.[6] In addition, the huTenTen12 was carefully composed of crawled online texts, avoiding duplication in the source material, therefore the analysis of co-occurrence data in this corpus results in more precise patterns than for instance in the HNC corpus. And finally, the complete Sketch Engine toolkit is available for those working with this corpus, which means that the researcher can generate not only concordances, word lists and frequency lists but keyword analysis, n-gram extraction, and collocation measurement are also performable in huTenTen12.

The large size of the corpus makes it possible to extract a vast amount of data with noun + verb patterns. To limit the scope of the study, I concentrated on three categories of nominal referents, namely animals, plants and natural phenomena. The main reason for this is that the personification of these kinds of entities is presumably frequent in everyday online discourse represented in the corpus. Additionally, these categories are distinct enough from each other to test the hypothesis that the semantic preferences of their personification are different. As a preliminary step of sampling, a frequency list of nominal lemmas was generated, from which the following keywords of the analysis were chosen:

- animals: *kutya* ('dog'), *ló* ('horse'), *hal* ('fish')
- plants: *fa* ('tree'), *gyümölcs* ('fruit'), *virág* ('flower')
- natural phenomena: *víz* ('water'), *levegő* ('air'), *tűz* ('fire').

It is worth noting that the keywords are not always the most frequent members of the given group, as other nouns referring to plants or natural phenomena have a higher token number in the corpus. However, these additional candidates would not have been ideal keywords, either because of their nominal homonymy (e.g., *szél*$_1$ refers to the wind and *szél*$_2$ to the edge of something),[7] or due to their more complex or general meaning (e.g., the Hungarian noun *erdő* 'forest' refers to plants, but not individual specimens; the Hungarian noun *növény* 'plant' has a more general meaning, therefore its personified use is presumably not so frequent).

After selecting frequent nominal components of a construction, all the verbal lemmas occurring immediately after the particular nouns (in the R1 position) were needed. This step of sampling was performed by obtaining those occurrences of the nominal keywords in which they are followed by a verb form in the past or present tense, in 3Sg or 3Pl. I used a CQL form to perform this filtering. The specification of the contextual position of the verbs made it possible to omit data in which the verb occurs in the closer context of the keyword, but the latter is not the verb's subject argument. However, this means also that due to the word order limitation of the query, alternative realizations of the noun + verb structure are omitted from the sample. The last phase was to generate a list of verbal lemmas in the R1 position. For the sake of rendering the amount of data perceptible, in table 1 I specify the number of total occurrences of the keywords and the number of verbal lemmas following them.

Keyword	**Frequency**	**Number of verbal lemmas following it**
kutya ('dog')	452146	20915
ló ('horse')	224227	8061
hal ('fish')	137746	1330

[5] The tagset is available here: https://www.sketchengine.eu/hungarian-emmorph-based-part-of-speech-tagset/ (last access: 22/11/2022).

[6] Since the former tagset is more limited, it covers only one part of PoS categories and morphological phenomena. The latter tagset is more extended, thus, using both of them results in a large range of identified and labelled grammatical categories. Moreover, because of the adoptation of the tagset used in the Hungarian National Corpus, the two corpora can be compared with each other regarding PoS categories and morphological phenomena.

[7] Although the Hungarian word *tűz* is also homonymous, it has a nominal and a verbal meaning ('fire' and 'pin', respectively), thus the in-built PoS-analysis can sort the data according to word classes.

fa ('tree')	428044	9661
gyümölcs ('fruit')	178338	2475
virág ('flower')	176991	4846
víz ('water')	888590	31950
levegő ('air')	330502	9617
tűz ('fire')	187858	9473

Table 1. The total frequency of the nominal keywords and the number of the verbal lemmas following them in the corpus

As a result of careful sampling and querying, all the verbs adjacent to the nouns became available for analysis. But this means a vast amount of information that cannot be examined manually. In addition, for a semantic analysis of typical personifications, one needs not only the verbs occurring in the context of the keywords but those verbs that occur significantly frequently with the keyword. To extract the most relevant lexemes occurring in a given construction, corpus linguistics offers the method of collostructional analysis (Stefanowitsch 2013, Stefanowitsch 2020). To put it simply, collostructional analysis explores which words are associated with a particular construction. These words are called the collexemes of the construction. Collostructional analysis has three different but theoretically and methodologically related ways of collexeme analysis: simple collexeme analysis (measuring the association of words with a construction), distinctive collexeme analysis (measuring and comparing the association of words with two or more constructions), and covarying collexeme analysis (identifying words that co-occur frequently in two slots of a given construction). In the present study, I adopted only the first method, but of course, the other two are also good candidates for a sophisticated corpus-driven analysis of figurative language.

At the heart of collostructional analysis is statistical testing of the contingency of a word in a construction. In other words, the analyst examines not only the frequency of a word in a construction: this data is compared to the overall frequency of the word in the corpus (i.e., its frequency in any other constructions) and to the total number of constructions in the corpus.[8] Therefore, a collostructional analysis is based on the data of contingency tables, and the most widespread function for contingency testing is the Fisher-Yates exact test (see Levshina 2015 and Stefanowitsch 2013 for other optional tests).

To explore the verbal collexemes of personifying noun + verb patterns, I performed a simple collexeme analysis with the RStudio statistical testing software (version 4.1.0, R Core Team 2021). For implementing the contingency test, the collostructions package developed by Suzanne Flach (2021) was used. This package offers multiple ways of statistical testing, including the canonical Fisher-Yates exact test, with the logarithmic transformation of the results recommended by the literature (Levshina 2015: 232). Image 2 illustrates the platform used for the analysis.

[8] Fort the sake of simplicity, the last number is calculated by counting the occurrences of verbs in the corpus. Since constructions are mainly centred around verbs, the absolute frequency of verb forms in a corpus can be considered to be roughly equal to the number of constructions (see Levshina 2015: 227).

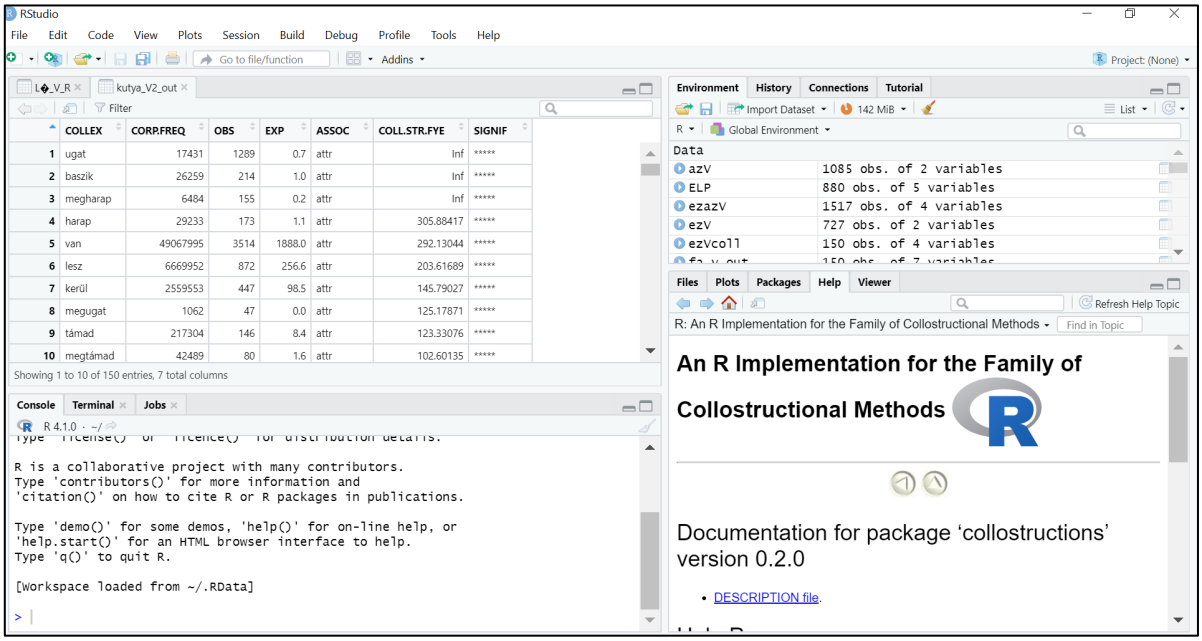

Image 2. The platform of RStudio

With the use of RStudio, simple collexeme analysis can be implemented easily. Even so, it requires considerable effort to test all the verbal lemmas occurring after the keywords in the corpus, since for example in the case of 20150 verbs, we need exactly the same number of contingency tables. Therefore, I limited the scope of the empirical measurement again, taking only the first 150 members of the verb lists into consideration. This limitation, however, does not distort the overall picture strongly, because after the first 150 the individual occurrences of verbal collexemes in the construction are so infrequent that these verbs cannot be considered typical collexemes, but rather hapaxes.[9]

One advantage of this analysis is that it provides information not only about the associated collexemes of a construction but also about the repelled words, i.e., those lexemes that significantly avoid the given grammatical pattern. Since personification is not an infrequent phenomenon in everyday language use either, we can have the expectation that there will be personifying (i.e., basically human-oriented) verbal collexemes among the associated words.

3. Results and discussion

I will discuss the results of the analysis in two steps. First, I demonstrate the data type of collexemes extractable from a simple collexeme analysis. As we will see, the aforementioned expectation about the correlation between personification and associated collexemes was not borne out. Secondly, by zooming out from the individual keywords toward the categories, a broader picture can be provided about personifying meanings that unfold with the use of a noun + verb pattern.

[9] The limitation of the analysed material may, of course, hide personifying instantiations of the construction, since among the verbal collexemes with low frequency there can be interesting data. However, the study is interested in the typical, frequently recurring instances of personification in the construction. Therefore, the potential strength of the analysis is to shed light on this pattern, whereas its weakness resides in excluding non-typical but theoretically relevant expressions. In future research, it will be worth comparing personifying occurrences among strongly and weakly associated verbs of the construction.

3.1. Verbal collexemes of nominal keywords

The aim of this section is to illustrate what kinds of data can be obtained as a result of collostructional analysis. This method is used to assess the lexical and/or grammatical variability of constructions, thus it is not self-evident how it can be adapted in the cognitive linguistic investigation of figurative language use. The main idea in the background of this endeavour is that figurative expressions show construction-like behaviour, or at least they can be described with recurring patterns. Consequently, even though the noun + verb pattern cannot be considered a specific marker of personification, moreover, its description as a construction requires more in-depth examination, using the collected dataset of nominal keywords and their neighbouring verbal lemmas as the input of collexeme analysis will shed light on some typical semantic patterns of personification in Hungarian.

COLLEX	CORP.FREQ	OBS	EXP	ASSOC	COLL.STR.FYE	SIGNIF
ugat 'bark'	17431	1289	0,7	attr	Inf	*****
baszik 'fuck'	26259	214	1	attr	Inf	*****
megharap 'bite'	6484	155	0,2	attr	Inf	*****
harap 'snap'	29233	173	1,1	attr	305,88417	*****
van 'be, exist'	49067995	3514	1888	attr	292,13044	*****
lesz 'will be'	6669952	872	256,6	attr	203,61689	*****
kerül 'get (somewhere)'	2559553	447	98,5	attr	145,79027	*****
megugat 'bark at'	1062	47	0	attr	125,17871	*****
támad 'attack'	217304	146	8,4	attr	123,33076	*****
megtámad 'assault'	42489	80	1,6	attr	102,60135	*****
fekszik 'lie'	204888	122	7,9	attr	97,20096	*****
csahol 'bay'	1133	38	0	attr	96,73201	*****
kóborol 'roam'	4496	46	0,2	attr	92,99589	*****
őriz 'watch over'	154078	100	5,9	attr	83,37047	*****
megtanul 'learn'	189522	107	7,3	attr	83,06096	*****
él 'live'	1640815	268	63,1	attr	81,44237	*****
vonít 'howl'	689	30	0	attr	80,02327	*****
szalad 'run'	57135	71	2,2	attr	78,67071	*****
marcangol 'lacerate'	2926	36	0,1	attr	75,87754	*****
megeszik 'eat'	68454	72	2,6	attr	74,71934	*****
vezet 'lead'	918577	30	35,3	rep	0,67798	ns
dolgozik 'work'	1185360	39	45,6	rep	0,73636	ns
kezd 'start'	1713455	57	65,9	rep	0,82821	ns
jár 'move, go'	1597749	51	61,5	rep	1,00658	ns
hagy 'leave'	980044	29	37,7	rep	1,06386	ns
megy 'go'	2723229	89	104,8	rep	1,19231	ns
ért 'understand'	1089696	28	41,9	rep	1,83262	*
néz 'look'	1837588	50	70,7	rep	2,22949	**
ismer 'know, be familiar with'	1014583	22	39	rep	2,66273	**
fog 'take'	3350463	92	128,9	rep	3,43736	***
csinál 'do'	1191689	21	45,9	rep	4,4897	****

hoz 'bring'	1553475	29	59,8	rep	5,12516	*****
tesz 'put'	4065887	95	156,4	rep	7,15571	*****
akar 'want'	3024986	57	116,4	rep	9,16751	*****
lát 'see'	3679625	58	141,6	rep	14,98278	*****
tud 'know, can'	9142019	202	351,8	rep	17,91354	*****
ad 'give'	2697447	25	103,8	rep	19,83909	*****
tart 'hold'	2525031	20	97,2	rep	20,82327	*****
kell 'need, must'	9961946	159	383,3	rep	38,70377	*****
mond 'say, tell'	5293753	25	203,7	rep	56,36016	*****

Table 2. Associated and repelled verbal collexemes of the noun + verb pattern with the *kutya* ('dog') keyword

Table 2 represents the first 20 and the last 20 collexemes of the investigated pattern, from the total amount of 150 verbs.[10] The first column contains the collexemes and their English meanings. The second tells us the number of occurrences of the individual verbs in the entire corpus, whereas the third and fourth columns specify the observed and the expected frequency of the verbs in the pattern under investigation. The most important information can be found in the last three columns. In the fifth, the status of the collexeme is given (i.e. whether it is associated or repelled), in the sixth there is the statistical result of the Fisher-Yates exact test, and finally, the last column informs the reader about the significance of association/repellence. (The number of asterisks indicates the degree of significance: the higher their number is, the more significant is the relationship.)

Focusing on the keywords, an unexpected result can be observed in the table. Among the associated collexemes, we can hardly find any good examples of personification. The verbs *őriz* ('watch over') and *megtanul* ('learn') personifies the animal subject, attributing human (or prototypically human) mental capacities and intentions to it.[11] However, the rest of the verbs refer to the typical behaviour of dogs (e.g. barking, howling, biting, attacking) or to the general processes of animal life (e.g. living, eating, running, lying). And even the potentially personifying verbs denote activities that are frequently associated with dogs. Thus, on the one hand, simple collexeme analysis works in general: it demonstrates what are the most salient processes performed by dogs. On the other hand, however, it fails to extract personifications with dogs, since the verbal collexemes associated with the keyword the most are used literally, or their figurative meanings can be considered rather typical (or even "dead") personification.

But if we turn to the list of repelled verbs, a surprising fact comes to the fore: 6 out of 20 are personifications (*vezet* 'lead', *ért* 'understand', *ismer* 'know, be familiar with', *akar* 'want', *tud* 'know, can', *mond* 'say, tell'), and if we count the more conventional (and thus less human-centred or even "dead" personifying) meanings as well (e.g. *dolgozik* 'work', *ad* 'give'), roughly 40% of the repelled words count as potential personification. By way of explanation, verbs that are not associated with the investigated pattern (because they occur rarely in it, and they are more infrequent in the pattern than in the whole corpus) instantiate relatively strong personifications.

[10] Although it is typical in collostructional analyses to focus only on the strongest collexemes of a construction, in the case of personification the analysis may benefit from observing the repelled words as well, on the basis of the assumption that personification itself is not necessary typical instantiation of the construction. However, there is also a limitation even in this decision: there can be data deserving more attention among the collexemes between the first 20 and the last 20 verbs, but they remain invisible from the perspective of the analysis. Adapting collostructional analysis to the investigation of figurative meaning needs further methodological considerations, thus in future research, it will be worth observing systematically the distribution of personification in the whole sample.

[11] The basic meaning of *megtanul* (according to the Concise Dictionary of Hungarian (Pusztai ed. in chief 2003)) is ' aquire knowledge or competence through studying', thus it refers to an intentional act of studying (accomplished by a human being). Similarly, the basic meaning of *őriz* ('keep an eye on somebody or something, in order to prevent them/it from an undesirable event') presupposes complex intentional and causal thinking characteristic of human minds.

It is worth checking whether the situation is similar with other keywords as well. Table 3 presents the first 20 associated and the last 20 repelled verbal collexemes of the noun *fa* ('tree') from the analysed 150 verbs.

COLLEX	CORP.FREQ	OBS	EXP	ASSOC	COLL.STR.FYE	SIGNIF
áll 'stand'	2234011	520	42,5	attr	Inf	*****
nő 'grow'	806740	360	15,3	attr	Inf	*****
dől 'fall/timber'	60634	198	1,2	attr	Inf	*****
szegélyez 'border'	5102	109	0,1	attr	287,4986	*****
kidől 'fall'	3880	52	0,1	attr	127,0341	*****
van 'be, exist'	49067995	1661	933,1	attr	121,0339	*****
kiszárad 'die/wither'	21390	60	0,4	attr	105,6764	*****
kizöldül 'come into leaf'	985	33	0	attr	94,21944	*****
övez 'surround'	24069	52	0,5	attr	85,85474	*****
eltakar 'hide'	20312	50	0,4	attr	85,40339	*****
elpusztul 'perish'	44884	57	0,9	attr	81,00264	*****
elszárad 'wither'	4064	32	0,1	attr	71,11725	*****
korhad 'moulder'	3017	29	0,1	attr	67,05141	*****
alkot 'create/make'	266622	81	5,1	attr	66,01048	*****
magasodik 'rise'	7403	33	0,1	attr	65,1592	*****
less 'will be'	6669952	355	126,8	attr	62,71369	*****
borít 'cover'	67291	51	1,3	attr	61,3565	*****
árnyékol 'shade'	5161	29	0,1	attr	60,28246	*****
pusztul 'die'	31141	41	0,6	attr	59,16456	*****
takar 'cover/hide'	64208	47	1,2	attr	55,92597	*****
enged 'let/allow'	518890	10	9,9	attr	0,27946	ns
fordul 'turn'	560231	9	10,7	rep	0,42124	ns
tetszik 'be liked'	760793	10	14,5	rep	0,8343	ns
választ 'choose'	1021860	14	19,4	rep	0,89126	ns
segít 'help'	1214134	17	23,1	rep	0,92515	ns
mutat 'show/point'	1045263	14	19,9	rep	0,96109	ns
számít 'count (as')'	822433	10	15,6	rep	1,04436	ns
válik 'become'	1306079	16	24,8	rep	1,39701	*
kezd 'begin'	1713455	22	32,6	rep	1,48776	*
fog 'take'	3350463	47	63,7	rep	1,76439	*
hoz 'bring'	1553475	16	29,5	rep	2,31979	**
kell 'need/must'	9961946	148	189,4	rep	3,04317	***
tart 'hold'	2525031	25	48	rep	3,7244	***
jön 'come'	2355900	20	44,8	rep	4,59225	****
jár 'go/move'	1597749	10	30,4	rep	4,77769	****
tesz 'put'	4065887	38	77,3	rep	6,30785	*****
megy 'go'	2723229	18	51,8	rep	7,29759	*****
kap 'get/receive'	2125871	9	40,4	rep	8,58189	*****

lát 'see'	3679625	14	70	rep	15,50282	*****
tud 'know/can'	9142019	37	173,8	rep	36,23789	*****

Table 3. Associated and repelled verbal collexemes of the noun + verb pattern with the *fa* ('tree') keyword

Again, the list of associated verbs does not abound with clear examples for personification: it is dominated by the biological processes of plants (e.g. coming into leaf, withering, falling, rising, growing) and the description of a landscape with trees (they border, surround, cover, shade or hide something). One weak candidate for personifying meaning generation is *alkot* ('create/make'), but it can refer again to the formation of trees (e.g., *Fák – rozsdabarna lombsátort alkotnak* 'Trees – creating a rusty brown leaf canopy'),[12] which can be considered neither an intentional act of creation nor an example of strong personification.

In contrast, 7 out of 20 repelled collexemes (35%) clearly exemplify personification, attributing the ability to move (coming, going, turning, moving) or other mental capacities (seeing, knowing/being able to, choosing) to trees. And, again, we can observe the fuzzy zone of conventional verbal meanings (e.g. beginning something, showing/demonstrating/pointing to something, putting something) with either conventional non-human primary figures (as the lexicalised extension of the basic human meaning) or a schematic meaning that can be elaborated as personification or not, depending on the context.

From the third category of keywords consider the 20 most associated and the 20 most repelled verbal collexemes of *levegő* ('air'), presented in Table 4.

COLLEX	CORP.FREQ	OBS	EXP	ASSOC	COLL.STR.FYE	SIGNIF
áramlik 'stream'	26866	868	0,5	attr	Inf	*****
érkezik 'arrive'	672725	465	13,4	attr	Inf	*****
lehűl 'cool down'	6561	125	0,1	attr	320,9718	*****
megtelik 'fill up'	32562	132	0,6	attr	249,8656	*****
csap 'whip'	113077	158	2,2	attr	226,3986	*****
árad 'flow'	29365	106	0,6	attr	195,484	*****
jut 'get to'	1355538	295	27	attr	193,891	*****
felmelegszik 'get warm'	7112	75	0,1	attr	173,4994	*****
beáramlik 'stream into'	1957	56	0	attr	154,2576	*****
átjár 'permeate'	14429	71	0,3	attr	140,7716	*****
okoz 'cause'	615419	173	12,2	attr	131,4505	*****
vibrál 'vibrate'	5114	51	0,1	attr	117,0458	*****
áraszt 'give out/radiate/exude'	25824	67	0,5	attr	114,3404	*****
kerül 'get somewhere'	2559533	270	50,9	attr	103,4662	*****
megcsap 'whip'	6219	47	0,1	attr	102,2687	*****
megfagy 'freeze'	13256	53	0,3	attr	100,5591	*****
távozik 'leave'	143681	88	2,9	attr	95,59648	*****
izzik 'glow'	8787	39	0,2	attr	76,0112	*****
felszáll 'fly up'	23606	47	0,5	attr	75,1318	*****
felemelkedik 'rise'	15191	39	0,3	attr	66,77728	*****

[12] http://drsoregistvan.lapunk.hu/galeria/novenyvilag-fak-cserjek-viragok/fak-rozsdabarna-lombsatort-alkotnak-783119 (Last access: 10/05/2022)

lép 'step'	701194	14	13,9	attr	0,27606	ns
ér 'get to/reach (a place)'	556035	11	11,1	rep	0,24221	ns
tesz 'put'	4065887	80	80,8	rep	0,30826	ns
nyújt 'stretch'	624953	11	12,4	rep	0,38341	ns
alakul 'take shape'	540150	9	10,7	rep	0,43278	ns
végez 'finish'	666380	10	13,3	rep	0,63725	ns
kell 'must/need'	9961946	187	198,1	rep	0,64926	ns
vár 'wait'	1521333	24	30,3	rep	0,83477	ns
ad 'give'	2697447	44	53,6	rep	0,98895	ns
hagy 'let/leave/allow'	980044	12	19,5	rep	1,31154	*
hoz 'bring'	1553475	21	30,9	rep	1,40652	*
számít 'count as'	822433	9	16,4	rep	1,44165	*
lesz 'will be'	6669952	109	132,6	rep	1,72092	*
tart 'hold'	2525031	32	50,2	rep	2,40321	**
vesz 'take'	2821985	36	56,1	rep	2,57166	**
nincs 'not exist'	3054168	31	60,7	rep	4,73453	****
jelent 'mean'	1879193	11	37,4	rep	6,40767	*****
fog 'take'	3350463	30	66,6	rep	6,43828	*****
van 'be/exist'	49067995	788	975,7	rep	10,74173	*****
tud 'know/can'	9142019	38	181,8	rep	38,32197	*****

Table 4. Associated and repelled verbal collexemes of the noun + verb pattern with the *levegő* ('air') keyword

As it can be observed, the associated verbs do not initiate personification in the third case either. Mainly physical (but not intentional) motion and the change of physical state are described by them. In other words, the construction's collexemes refer to the mechanics and thermodynamics of air as material. Only the verbs *csap/megcsap* ('whip') and maybe *távozik* ('leave') have also human-related meanings, but the first pair of words denotes the experience of perceiving an intense stimulus through the air (for example an olfactory perception of an unpleasant smell), and the second one seems to be a more formal alternative of describing the situation when something lets the air go. Thus, even if they may have some kind of figurative meaning, it is conventional and not primarily human-related.

Turning to the repelled lexemes a similar pattern can be recognized: there are potential personifications in the list with relatively clear human-oriented basic meanings (e.g. *lép* 'step', *nyújt* 'stretch', *végez* 'finish', tud 'know/can' or *vár* 'wait'), and some of the verbal collexemes can be considered as transitions between clear and conventional personifications (e.g. *ad* 'give', *hagy* 'let/leave/allow', hoz 'bring' or *vesz* 'take'). The proportion of potential personifying verbs is a little higher (10 out of 20, 50%) in this case than in the previous samples; however, there are more examples of conventional (and thus less recognizable) expressions among the collexemes. Put differently, the proportion of strong personifications is roughly equal or even lower with this keyword when compared to the other two.

Two conclusions can be drawn from the results of collostructional analysis. On the one hand, personifying verbs are not strongly associated with the investigated construction, since they can be found at the end of the collexeme list, among the repelled lexemes. This does not mean that the investigated grammatical pattern itself would not be used for expressing personifying meaning. But its main semantic profile (explored with statistical analysis) involves species-specific processes and behaviour or processes which are peculiar to the given entity. On the other hand, the

method of collostructional analysis can be adapted to the exploration of figurative language use, since good candidates of verbal personifications can be extracted with it. However, the method needs to be used in a reverse way, not focusing on the associated collexemes but rather examining the verbs that are not preferred by the pattern, i.e. which are very infrequent among its instantiations in the corpus.

3.2. The analysis of verbal collexemes of the keywords – conceptual domains of personification

Based on the previous findings, we can make an attempt at scrutinizing whether there are typical conceptual domains in the background of personifying animals, plants and natural phenomena. With more technical terms, the following semantic analysis answers the question whether the conceptual type of the target domain of personification has some correlation with typical source domains or not. As a set for optional source domains, the list proposed by Dorst and her colleagues (2011: 180) has been used in the analysis, namely HUMAN BODY, CHARACTER TRAITS, MOVEMENT and CONTROL) completed with the domains of MENTAL ACT and COMMUNICATION in accordance with the basic meaning of the verbal collexemes. (For an extended list of conceptual domains of personification see Simon 2021.) Since simple collexeme analysis produces verbal lemmata associated with the construction, their categorization into conceptual domains is based on their basic (human-related) meaning. In the following, I focus only on potential personifications, and the non-personifying collexemes are disregarded here.

3.2.1. Verbal personifications of animals

The keywords in this category occur in the context of the following personifying verbs. (After the English equivalent of collexemes, the abbreviation of the nature of the verb and the grammatical pattern can be found: "a" means associated and "r" means repelled.)

- *kutya* ('dog'): *megtanul* ('learn', a), *őriz* ('watch over', a), *ért* ('understand', r), *ismer* ('know/be familiar with', r), *akar* ('want', r), *tud* ('know/can', r), *néz* ('look', r), *vezet* ('lead', r), *mond* ('say/tell', r)
- *ló* ('horse'): *mutat* ('show/demonstrate/point', r), *segít* ('help', r), *néz* ('look', r), *szeret* ('love/like', r), *akar* ('want', r), *tud* ('know/can', r), *kezd* ('begin', r), *mond* ('say/tell', r)
- *hal* ('fish'): *beszél* ('speak/talk', r), *érez* ('feel', r), *vár* ('wait', r), *szeret* ('love/like', r), *akar* ('want', r), *kezd* ('begin', r)

The number of potential personifications is not the same for the three keywords, and the results clearly demonstrate that personification is a scalar phenomenon regarding both its strength (ranging from clear human-oriented meaning to less personifying, more conventional cases) and its lexical semantic richness. Although animals are on the same level on the scale of empathetic recognisability, which means that in general, they attract more attention from the perceiver than inanimate objects (Gibbons–Whiteley 2018: 153, based on Stockwell's cognitive poetics), therefore, their personification does not count as a marked departure from ordinary attentional arrangement, the collexemes nevertheless do tell us something about how different animal species tend to be personified. The more behavioural interaction with a species we have, the more diverse its profile is in personification.

Considering the conceptual domains, a relatively homogeneous pattern can be recognized in the sample. The majority of the verbs belong to the domain of MENTAL ACT (thinking, emotion and awareness), which means that it seems to be typical to attribute (intentional) mental capacity to animals in Hungarian. The other recurring domain is CONTROL (leading, showing, beginning something), which can be characterized by some kind of intentionality, too. Finally, the domain of COMMUNICATION also has verbs in the sample, referring to the human capacity of saying something or

speaking in general. These domains make it possible to conceptualize animals as higher-order entities since the activities belonging to them are not only volitional or involve some level of agency but they presuppose an intentional mind and conscious planning of actions.

3.2.2. Verbal personifications of plants

The pattern of verbal personifications of plants does not differ considerably from the previous one of animals, but there is a certain divergence between the samples. The keywords in this group are personified with the following verbs in Hungarian.

- *fa* ('tree'): *fordul* ('turn', r), *választ* ('choose', r), *segít* ('help', r), *mutat* ('show/demonstrate/point', r), *kezd* ('begin', r), *hoz* ('bring', r), *jön* ('come', r), *jár* ('move/go', r), *tesz* ('put/do', r), *megy* ('go', r), *lát* ('see', r), *tud* ('know/can', r)
- *gyümölcs* ('fruit'): *mutat* ('show/demonstrate/point', r), *játszik* ('play', r), *tesz* ('put/do', r), *kezd* ('begin', r), *talál* ('find', r), *hoz* ('bring', r), *jár* ('move/go', r), *tud* ('know/can', r)
- *virág* ('flower'): *szokik* ('do something regularly', r), *hoz* ('bring', r), *vesz* ('take', r), *jön* ('come', r), *ad* ('give', r), *segít* ('help', r), *jár* ('move/go', r), *szeret* ('love/like', r), *megy* ('go', r), *akar* ('want', r), *tud* ('know/can', r)

The keywords in this group do not refer to different species of plants: the noun *fa* ('tree') denotes a general type of plants and the other two keywords describe the parts of vegetal organisms. Therefore, there are no observable differences in the lexical diversity of personification between the nominal expressions. Both the number of potential personifications and their strength seem to be similar for the keywords. It is worth noting, however, that the verbs which have been considered conventional personifications in the case of animals (e.g., *ad* 'give', *tesz* 'put', *hoz* 'bring') previously, are evaluated as more clear personifications due to the fact that plants have different bodily organization and they cannot engage in motion or manipulation of objects, whereas some animal species are able to perform these activities (even if not with arms, for instance).

The frequent domains of personification are MOVEMENT, CONTROL and MENTAL ACT. Since plants are not able to move volitionally, the verbs from the first domain can constitute prominent personifications, despite the fact that motion itself is not a very complex and human-specific ability. It is even more interesting that the other two domains occur in the sample as well, rendering it possible to attribute higher-order mental capacity to plants, too. Consequently, the boundary between animals and plants is not strict from the perspective of personifications: according to the results, plants can also be construed as having minds and performing intentional and planned actions in Hungarian.

3.2.3. Verbal personifications of natural phenomena

Moving on along the great chain of beings (Kövecses 2010: 151), natural phenomena count as the least human-like entities: without bodies and intellectual capabilities, all the processes they participate in are conceptualized essentially figuratively. The question is whether there are some tendencies of personification peculiar to this group of keywords compared to animals and plants. The potential verbal personifications of natural phenomena are listed in the following.

- *levegő* ('air'): *lép* ('step', r), *tesz* ('put/do', r), *nyújt* ('stretch/provide', r), *végez* ('finish', r), *vár* ('wait', r), *ad* ('give', r), *hoz* ('bring', r), *vesz* ('take', r), *tud* ('know/can', r)
- *tűz* ('fire'): *vezet* ('lead', r), *segít* ('help', r), *játszik* ('play', r), *ad* ('give', r), *hoz* ('bring', r), *vár* ('wait', r), *megy* ('go', r), *tesz* ('put/do', r), *tud* ('know/can', r)

- *víz* ('water'): *halad* ('go', r), *visz* ('take', r), *nyújt* ('stretch/provide', r), *vár* ('wait', r), *szokik* ('do something regularly', r), *fogad* ('receive', r), *hoz* ('bring', r), *ad* ('give', r), *vesz* ('take', r), *választ* ('choose', r), *jár* ('move/go', r), *játszik* ('play', r), *tesz* ('put/do', r), *tud* ('know/can', r)

What may catch our attention first is that there are more potential personifications in this sample than in the other two. There are two possible explanations for this. On the one hand, natural phenomena do not have any agency and/or volitional capacities, thus, even the conventionally human-oriented meanings (which count as weak or "dead" personifications in other samples) can be considered relevant data. On the other hand, some of the prominent verbs have more than one function in discourse: they can behave as light verbs for example, like in the examples *a* […] *levegő nyújt biztonság-ot* (the […] air provide.PRS.3SG safety-ACC 'the […] air provides safety')[13] or *a víz játszik fontos szerep-et* (the water play.PRS.3SG important role-ACC 'the water plays an important role').[14] The latter observation demonstrates also the limitations of the study and the application of collexeme analysis to research on figurativity: as non-literal meaning generation always depends on the context, and collostructional analysis results in lemmata (without any specification), polysemy cannot be disregarded in the interpretation of the data, and a further check is needed for identifying genuine instances of figurative usage.[15]

Taking this limitation into consideration, and handling carefully the potentially personifying verbal collexemes, the conceptual domains of MOVEMENT, CONTROL and MENTAL ACT can be identified in the category. Whereas rather conventional and schematic verbs belong to the first two domains (e.g., *hoz* 'bring', *vesz* 'take', *megy* 'go', *ad* 'give'), the third domain is represented with only one or two verbs (e.g., *választ* 'choose' or *tud* 'know/can'), and they are not necessarily prototypical activities of the human mind.[16] Moreover, an overlap can be assumed between the domains of MOVEMENT and MENTAL ACT, or CONTROL and MENTAL ACT: for illustration, the verb *vár* ('wait') also denotes the meaning of 'having expectations', and *vezet* ('lead') describes not only a physical act of 'moving ahead' but also the intentional act of 'taking the role of a leader'.

In sum, there is only a superficial similarity between the conceptual backgrounds of personifying plants and natural phenomena: the potential verbal personifications of air, water and fire are more conventional (regarding their lexical meaning) and less clear instances of personification in Hungarian. Nevertheless, they are good candidates for further analysis, especially in the realm of literature.

4. Conclusions and future prospectives

After the overview of the broader pattern of collexemes and focusing on the conceptual domains of personification in the categories of the keywords, what lessons can be drawn from the collostructional analysis of personifying noun + verb patterns? The first research question (How can we explore typical noun + verb personifications using corpus linguistic methodology?) can be answered by saying that simple collexeme analysis can be adapted to the exploration of figurative language use only with restrictions. First of all, the chosen method is used for examining typical words in a given constructional pattern; however, the potential personifications occurred among the repelled verbs almost without exceptions.

[13] https://www.zeptermagazin.hu/tisztabb-levego-hosszabb-elet/ (last access: 10/07/2022)
[14] https://ma7.sk/tajaink/mesefilmet-forgatnak-a-tallosi-vizimalomnal (last access: 10/07/2022)
[15] This additional step can prove the personifying usage of the collexeme, too: in the example *a nap és a víz játszik egymással* (the sun and the water play.PRS.3SG together-INS'THE sun and the water plays with each other' https://hu-hu.facebook.com/akvariumklub/videos/im%C3%A1djuk-ahogy-a-nap-%C3%A9s-a-v%C3%ADz-j%C3%A1tszik-egym%C3%A1ssal-huppanj-le-%C3%A9s-%C3%A9lvezd-a-naps%C3%BCt%C3%A9st-/1727498907270894/, last access: 10/07/2022) the verb játszik ('play') functions as personification.
[16] The verb *választ* ('choose') can be integrated with verbal prefixes as *szét-* ('apart') or *el-* ('away') resulting the meaning of 'separate', which not necessarily an intentional mental act. And if the verb *tud* ('know/can') refers to a kind of (physical) ability in a context, it does not instantiate the schematic meaning of having knowledge about something.

Thus, a new question emerges from the analysis: how typical are these verbal collexemes as personifications? Obviously, they appear in the construction with a certain frequency in the corpus, but the absolute numbers of their occurrences are low, in addition to the fact that they are not associated with the pattern. A potential cause of this might be that the pattern itself is too restricted: remember that only one variation of word order has been queried, namely when the verb directly follows the noun. Consequently, with a more flexible pattern (for instance, in a three-word-sized window before and after the nominal keyword) one can expect both more personifying verbs and perhaps more typical ones. To put it simply, the typicality of personifications found with collostructional analysis depends partly on the way the researcher defines the pattern. Another explanation of the results is that personification itself is not so typical in online discourse represented in the chosen corpus, therefore personifying verbs are also underrepresented in the sample. This assumption can be tested easily, by changing the corpus and hence the scope of the analysis from general language use and its texts to literary discourse, i.e., by performing a simple collexeme analysis on a corpus of literary works of art.

Nevertheless, the analysis rendered it possible to extract potential personifying verbs in the context of the nominal keywords. This means that we have candidate expressions for studying how language users personify animals, plants and natural phenomena in Hungarian. In other words, we do have some clues about what the salient verbs are for generating personifying meaning. Collostructional analysis has a filtering role in research on figurative language: the emerging collexemes can serve as input for a more precise process of extracting genuine personifications from the corpus.

The answer to the second question of the study (Is there any difference in typical verbal personifications related to different personified entities?) is not so simple again. There is a clear dissimilarity between the investigated categories of keywords: while animals are personified mainly in the domain of MENTAL ACT and COMMUNICATION (overlapping with the domain of CONTROL), in the case of plants and natural phenomena MOVEMENT and CONTROL dominate the data. However, the representative verbs of the domain of MENTAL ACT do not disappear in the latter two categories, even if their proportion decreases and their strength as personifications weakens. Furthermore, the number of conventional personifications becomes higher as we move from animals to natural phenomena, which means that the distinction between linguistic and conceptual personifications becomes sharper at the non-animate end of the scale. As Dorst (2011: 122) claims, "[w]hen linguistic, conceptual and behavioural analyses of personification are kept separate, this creates the possibility of analysing examples as personifications at the linguistic level but not necessarily the conceptual or cognitive level." In the present study, we witnessed the reverse of this: a verbal collexeme does not count as good or strong personification at the linguistic level (based on its basic meaning and its polysemy or polyfunctionality), but at the conceptual level it has the potential of initiating personifying meaning in the discourse (which depends on the context as well).

Therefore, the main conclusion drawn from collexeme analysis is that there is a remarkable difference between verbal personifications regarding the personified target, but neither the conceptual nor the linguistic level of analysis is sufficient in itself to describe this difference. Collostructional analysis can shed light on recurring personifying instantiations of a pattern, and so it may open up new ways of mapping the diversity of personifications in a language. A distinctive collexeme analysis or another simple collexeme analysis using verbs as keywords seem to be promising next steps in this endeavour.

Acknowledgements

NEMZETI KUTATÁSI, FEJLESZTÉSI ÉS INNOVÁCIÓS HIVATAL

The research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and the ÚNKP-22-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

References

Biber, Douglas 2010. Corpus-based and corpus-driven analyses of language variation and use. In: Heine, Bernd – Narrog, Heiko (eds.): *The Oxford handbook of linguistic analysis*. New York: Oxford University Press. 159–191. https://doi.org/10.1093/oxfordhb/9780199544004.013.0008

Bocharova, Jean 2016. Personification allegory and embodied cognition. In: Melion, Walter S. – Ramakers, Bart (eds.): *Personification. Embodying meaning and emotion.* Leiden–Boston: Brill. 43–69. https://doi.org/10.1163/9789004310438_003

Dorst, Aletta G. 2011. Personification in discourse: Linguistic forms, conceptual structures and communicative functions. *Language and Literature* 20(2): 113–135. https://doi.org/10.1177/0963947010395522

Dorst, Aletta G. – Mulder, Gerben – Steen, Gerard J. 2011. Recognition of personification in fiction by non-expert readers. *Metaphor and the Social World* 1(2): 174–201. https://doi.org/10.1075/msw.1.2.04dor

Flach, Suzanne 2021. *Collostructions: An R implementation for the family of collostructional methods*. Package version v.0.2.0. https://sfla.ch/collostructions

Jakubícek, Miloš – Kilgarriff, Adam – Kovár, Vojtech – Rychlý, Pavel – Suchomel, Vit 2013. The TenTen Corpus Family. In: Hardie, Andrew – Love, Robbie (eds.): *Corpus Linguistics 2013*. Abstract book. Lancaster: Ucrel. 125–127.

Kövecses, Zoltán 2010. *Metaphor. A practical introduction*. Second Edition. New York: Oxford University Press.

Lakoff, George 2006 [1993]. The contemporary theory of metaphor. In: Geeraerts, Dirk (ed.): *Cognitive Linguistics: Basic readings*. Berlin – New York: Mouton de Gruyter. 185–238. https://doi.org/10.1515/9783110199901.185

Levshina, Natalia 2015. *How to do Linguistics with R. Data exploration and statistical analysis.* Amsterdam–Philadelphia: John Benjamins. https://doi.org/10.1075/z.195

Long, Deyin 2018. Meaning Construction of personification in discourse based on conceptual integration theory. *Studies in Literature and Language* 17(1): 21–28.

Low, Graham 1999. "This paper thinks…" Investigating the acceptability of the metaphor an essay is a person. In: Cameron, Lynne – Low, Graham (eds.): *Researching and applying metaphor.* Cambridge: Cambridge University Press. 211–248. https://doi.org/10.1017/CBO9781139524704.014

Melion, Walter S. – Ramakers, Bart 2016. Personification: An introduction. In: Melion, Walter S. – Ramakers, Bart (eds.): *Personification. Embodying meaning and emotion.* Leiden–Boston: Brill. 1–41. https://doi.org/10.1163/9789004310438_002

R Core Team 2021. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Sájter, Laura 2008. Megszemélyesítés [Personification]. In: Szathmári, István (ed.): *Alakzatlexikon. A retorikai és stilisztikai alakzatok kézikönyve* [Lexicon of figures of speech. A handbook of rhetorical and stylistic figures of speech]. Budapest: Tinta Könyvkiadó. 383–388.

Simon, Gábor 2021. The event structure of personification in the poetry of Attila József. In: Tóth, József – V. Szabó, László (eds.): *Event in language, literature and culture.* Berlin: Peter Lang. 67–79.

Simon, Gábor 2022. Identification and analysis of personification in Hungarian: the PerSECorp project. In: Calzolari, Nicoletta – Béchet, Frédéric – Blache, Philippe – Choukri, Khalid – Cieri, Chritopher – Declerk, Thierry – Goggi, Sara – Isahara, Hitoshi – Maegaard, Bente – Mariani, Joseph – Mazo, Hélene – Odijk, Jan – Piperidis, Stelios (eds.): *Proceedings of the 13th Language Resources and Evaluation Conference.* Paris: European Language Resources Association (ELRA), 2730–2738.

Stefanowitsch, Anatol 2020. *Corpus linguistics. A guide to methodology.* Berlin: Language Science Press.

Tognini-Bonelli, Elena 2001. *Corpus linguistics at work*. Amsterdam–Philadelphia: John Benjamins. https://doi.org/10.1075/scl.6