sciendo
DOI 10.2478/jazcas-2023-0013

GENDER ASYMMETRY OF VISEGRÁD GROUP LANGUAGES AS REFLECTED BY WORD EMBEDDINGS[1]

RADOVAN GARABÍK – JANA WACHTARCZYKOVÁ
Ľudovít Štúr Institute of Linguistics, Slovak Academy of Sciences, Bratislava, Slovakia

GARABÍK, Radovan – WACHTARCZYKOVÁ, Jana: Gender asymmetry of Visegrád group languages as reflected by word embeddings. Jazykovedný časopis (Journal of Linguistics), 2022, Vol. 73, No. 3, pp. 354–379.

Abstract: Today, word embeddings have become a standard method in natural language processing, largely due to the availability of large language corpora. The models effectively reflect the semantic relationships between words without any additional linguistic input. Recently, more emphasis has been placed on interpreting the seemingly discriminatory results of some queries, with the goal of de-biasing language models.

However, if we consider the vector space to be a reasonably valid model of a linguistic semantic space, does not the asymmetry and subsequent discrimination in word embeddings reflect the (average) discriminatory tendencies inherent in the language? This article explores word embedding models for the Visegrád group languages and we apply basic vector arithmetic to demonstrate the basic language asymmetry present in the models.

It is well known that in English models, vector transfers result in eerily accurate predictions when swapping genders (the famous *king* – *man* + *woman* = *queen*), but these transfers also result in rather uncomplimentary roles for certain occupations (*doctor* – *man* + *woman* = *nurse*, or *computer programmer* – *man* + *woman* = *homemaker*). The article explores similar transfers in models of V4 languages – Slovak, Czech, Polish, and Hungarian. With Hungarian gender neutrality, Polish strong generic masculine, and close parallels between Slovak and Czech, we hope to uncover interesting similarities and differences in gender asymmetry in these languages, based on real language data.

Key words: word embeddings, discrimination, NLP, grammatical gender, gender stereotypes, generic masculine, gender symmetry, gender asymmetry

1. INTRODUCTION

Word embeddings (as introduced by Mikolov et al. 2013) have become a crucial and indispensable component of advanced natural language processing (NLP) research and language analysis. It is widely recognized that the relationships between vectors capture the semantic values of language, although the exact reason for this is not fully understood (Şenel et al. 2018). This semantic relationship also extends to

[1] The paper was supported by the project VEGA 2/0014/19 *Discriminatory instrumentalization of language*.

proper names. Informally, we will refer to the "semantic closeness" and "synonyms" (also regarding proper names) as the closeness of their vectors in our models. Similarly, we will use the term "closeness in semantic space" to mean small angle between vectors in our word embedding models and the "semantic surroundings of a word" to mean the n-dimensional hypersphere (of a certain radius) surrounding the vector in our model. Another improvement on the basic word embedding method is the addition of substring vectors to the vector of the word, in the form of the fastText algorithm (Bojanowski et al. 2017). This improves the analysis of inflected languages without the need for "traditional" lemmatization and related NLP processes.

Our article aims to explore the possibilities of vector models in examining deep semantic relationships when expressing gender equivalence in the V4 languages. This is motivated by the pan-European trend towards bias-free language, gender-sensitive expressions and gender-balanced language. This requirement is mandatory for the creation of official EU documents and is part of the *Gender Equality Strategy 2020–2025*[1] and most recently also part of the agenda of the Slovak Academy of Sciences in the form of the *Plan of Gender Equality in the Slovak Academy of Sciences*.[2] The specific guidelines for gender-neutral expression in the Slovak language are contained in the handbook *Inclusive Communication in the General Secretariat of the Council of the European Union*.[3] It follows from the handbook that two parallel and legitimate paths lead to inclusive language in Slovak.[2] Namely gender neutralisation (the use of gender-neutral terms to refer to persons and groups, e.g. *osoba* [person], *skupina, kolektív* [group, team] and feminisation (the use of feminatives to express the feminine gender in the names of professions, social roles, functional, etc.) e.g. *hosteska* [hostess], *športovkyňa* [sportswoman]. Gender-balanced language in Slovak is based on the explicit parallel naming of male and female objects, while gender pairs are formed, e.g. *herec* ↔ *herečka* [actor ↔ actress]. The goal of this explicit binarization is to induce gender symmetry. However, the question arises whether the existence of a female name in a gender pair is a sufficient expression of symmetry? Is formal symmetry a reflection of the full semantic equivalence of the members of the gender pair, or do the language users somehow differentiate the members of the pair and use them a little differently in practical communication? Furthermore, if differences and asymmetries arise here, what are they, and what do they tell?

[2] Handbook *Gender neutral language in European Parliament* (2018, p. 6)[8] states: "Genderless languages (such as Estonian, Finnish and Hungarian) do not generally need a particular strategy to be gender-inclusive, save for the very specific cases that are discussed in the particular guidelines for those languages". In Hungarian, gender-sensitive language is partly reflected in the formation of gender pairs in the names of professions (e.g. *tanár* [male teacher] – *tanárnő* [female teacher]), but the dominant tendency is to use gender neutral terms. The aim of the tendencies towards gender sensitivity (***away*** [emphasis ours] from gender neutrality) in Hungarian is "that the language will not be sexist and at the same time will use its grammatical genderlessness" (Satinská 2018, p. 102).

Linguistic research based on word embeddings can perhaps be the first step towards the answers. Word embeddings are not the first, neither the only method of representing words by vectors in a multidimensional vector space, but they are computationally tractable with (by contemporary standards) just modest computing equipment and they conveniently reflect several important linguistic features (such as a combination of semantic and grammatical relations). We try to obtain an insight into how the users of a given language express their semantic space regarding gender pairs. We emphasize that this is just a probe into corpus data and the results should be understood as such, reflecting the corpora[3] and not necessarily the society nor the language. Equivalence of relationships between gender pairs can be observed at the level of one language and interlingual comparisons as well. Thanks to the relative objectivity of word embeddings, we can observe more or less symmetric phenomena in the verbal manifestation of the genus of people and animals in individual languages. We aimed to show what type of results the models bring in monitoring the degree of gender equivalency of words and represented by related vectors.

Although well established, word embedding models still require clarification of their functionality and material base for the linguistic audience. Therefore, in the structure of our paper, we will first very briefly describe the basics of word embeddings and the foundations of our web interface to the word embeddings used, as well as the corpora used to train the models. Later we will present the language material and a selection of examples, on which we follow the way of manifestation of gender equivalence. We define gender equivalence and its types. We also deal with interpreting and comparing individual examples between V4 languages. At the end of the article, we present a summary of our results.

1.1 ARANEA corpora family

ARANEA family of web corpora is a group of web corpora collected, annotated and compiled using the same methodology and procedures, thus creating a set of comparable corpora of similar composition. Currently, there are 24 languages available, of various sizes and quality. Most of the corpora are lemmatized and POS tagged, and several moderately inflected languages represented therein have full morphological annotation. However, there are corpora that are not lemmatized, even for languages with rich morphology, due to technical or other reasons (e.g. lack of sufficiently advanced NLP tools for the language). In general, corpora are available in three different sizes: *Maximum* (all the crawled text, after deduplication), *Maius* (1.2 billion[4] token subset of the *Maximum*, created only if *Maximum* is bigger than

[3] We should note that the corpora had been collected before the widespread public availability and usage of Large Language Models; we expect the internet space to become soon "infested" by automatically generated texts virtually indistinguishable from human output and the reliability of web corpora to decrease, perhaps even significantly.

[4] We are using 1 billion = 10^9.

1.2 billion tokens) and *Minus* (120 million tokens, accessible without a registration, created only if *Maximum* is bigger than 120 million tokens). *Maius* and *Minus* are therefore comparable across the languages.

The sizes of Slovak, Czech, Polish and Hungarian (i.e. the *Visegrád group*, later in the article *V4*) corpora, as well as the English corpus used in this article are:

Table 1. Overview of corpora (and word embeddings trained thereon) used in this article

Language	*Corpus*	*Size [Gtokens]*	*Crawl time*
Czech	A. Bohemicum IV Maximum	7.1	2013–2018
English	A. Anglicum II Maximum	11.4	2013–2017
Hungarian	A. Hungaricum Maius	1.2	2014
Polish	A. Polonicum Maius	1.2	2013
Slovak	A. Slovacum V Maximum + prim-8.0-juls-all	5.5	2013–2019

Slovak word embeddings are different from the other languages, though – they are based on a bigger corpus, a union of the *Araneum Slovacum V Maximum* (Slovak web corpus of the Aranea family) and *prim-8.0-juls-all*, a representative corpus of contemporary Slovak, part of the Slovak National Corpus project. There is also a different Slovak word embedding model (called *sk-ll*) that differs in a significantly lower threshold for token frequency (10 occurrences), aimed for specific lexicography use, but we will not discuss this model here.

The corpora (and vector models) are therefore not really comparable, since we aim for the best (i.e. biggest) possible corpus for the language in question, but we can still get valuable information and make comparisons from the models. There are three models for most of the languages; the first model is trained on lemmas, the second one on the original raw word forms (case-normalized, to correctly capture e.g. sentence-initial words), the third one uses fastText algorithm. Unless stated otherwise, in this article we will use the model trained on lemmas. The composition of the corpora is as expected from web-crawled texts – generic webpages, business webpages, online discussion, news articles etc. The languages in question share many similar characteristics on their web usage, we feel the differences are not crucial for our purposes. We will highlight several features of the V4 corpora:

- Most of the text come from the respective top level domain (.cz, .hu, .pl, .sk);
- There is a non-negligible amount of Slovak texts in the .cz domain and vice versa; these texts are removed by language filtering;
- Compared to other languages, the Hungarian corpus contains bigger proportion of texts from minorities in neighbouring countries;

- The languages are mostly monocentric; written colloquial Czech is somewhat popular on the internet, but dwarfed by the literary standard; Cyrillic-script Eastern Slovak as used in Vojvodina does not come into the mainstream consideration at all;
- Czech, Polish and Hungarian possess a lot of classic literature present on the internet (regarding the expiration/exemption from copyright protection) while Slovak literature appears in noticeable (for corpora and language models) quantity only in the 20th century (but the inclusion of the *prim-8.0* corpus could somewhat compensate for the presumed lack of Slovak classical fiction on the internet).

The English corpus, on the other hand, is collected without any territorial limitations. Its composition is expected to be different, given the worldwide span of English, but in any case, English serves as a comparison and is not the focus of this article (and the results we obtained conform to those reported in the literature).

2. WORD EMBEDDINGS

2.1 Web Interface

Word embeddings are quite easy to use, with several mature OpenSource software frameworks, libraries and packages in major programming languages. However, this approach can still be cumbersome for casual users, such as language teachers or learners, or in linguistic research. To address this, we have developed a web interface to query the models, with the intention to be make word embeddings accessible to both experienced linguists (or lexicographers) and language enthusiasts. The interface is described in (Garabík 2020), and we just summarize the main points used in this article:

- A query will display a table of nearest words from the embedding model and a visualisation graph, displaying the surroundings of the result, in either 2D, 3D or 4D projection, using ISOMAP dimensionality reduction.
- We define the "closeness" of words as $\sqrt{(1-\cos^2\varphi)}$, where φ is the angle between the vectors corresponding to the two words). In the article, we will define the function $|word1, word2| = \sqrt{(1-\cos^2\varphi(word1,word2))}$.
- We support simple vector arithmetic, consisting of addition and subtraction. The result of the expression is used as a vector around which we look for semantically close words and display the table of them in a similar manner to the previous usage cases.
- In this article, we adopt the following convention: in an arithmetic expression composed of words, we understand the words to stand for their respective vectors[5]; the equal sign = means the resulting vector is closest to

[5] By writing a *word* directly in these expressions we of course mean the corresponding vector as a function of the *word*.

the word after the equal sign; the ≈ sign means the resulting vector is not the closest one, but is near enough to be relevant in the expression (e.g. the closest one is a typo or a mis-lemmatized word; or a different part of speech, etc.). We might include several words separated by commas (and spaces): this means that they are in the vicinity of the result, in the order of increasing semantic distance.

2.2 Vector transfer

The use of vector arithmetic over word embedding models is well known – subtracting "masculinity" from and adding "femininity" to a word vector ideally gives us a feminine variant of the word; the iconic example of *king – man + woman = queen* is already mentioned in (Mikolov et al. 2013). Close (in the sense of their angle being small) vectors signify there is some kind of semantic or syntactic similarity between the words, which even allows us to quantify the level of synonymy between the words – this is however not comparable across different vector models.

Of the languages we focus on, Hungarian completely lacks grammatical gender; Polish, Slovak and Czech follow a typical pattern of Slavic languages with three main genders (masculine, feminine, neutrum), with the masculine further divided into animate and inanimate. These three languages also make a heavy use of generic masculine for gender-neutral or gender-agnostic roles, a feature that is increasingly being targeted by various gender-equality activists (however, examples of other "non-sexist" languages, such as Hungarian, are discussed very seldom, if at all). Among these languages, Polish exhibits a strong generic masculine on the lexical level, especially in names of professions, where feminine alternatives often do not even exist, or they are uncommon and perceived as marked. In contrast, in both Czech and Slovak, feminine names of professions are productively derived and commonly used, while generic masculines are also common.

3. LANGUAGE EXAMPLES AND CATEGORIAL OPERATORS

3.1 Gender pairs

The subjects of our interest are gender pairs, that is, gender equivalents in Slovak, Czech, Polish and Hungarian. In common sense, we consider a gender pair to be a pair of words, one of which names a male representative and the other a female representative.[6] A gender pair can be singular if it expresses gender

[6] Thanks to their nature, word embedding models conflate biological, semantic and grammatical genders (reflecting the use of genders in V4 languages). The results from gender transfer within the vector space(s) should be interpreted with this in mind.

opposition between individuals or animals e.g. *žiak – žiačka* [schoolboy – schoolgirl], *kohút – sliepka* [rooster – hen] or collective if it expresses gender opposition between groups of people or animals e.g. *žiaci – žiačky* [schoolboys – schoolgirls], *kohúty – sliepky* [roosters – hens]. Gender pairs in terms of expressive complexity in Slovak can be full-expressions if both members of the gender pair are expressed as heteronymous lexemes, e. g. *chlapec ↔ dievča*[7] [boy ↔ girl] or symbolic if the gender of the other member of the gender pair is signalled only by the gender suffix, e.g. *boh ↔ bohyňa* [god ↔ goddess]. Defining gender pair is primarily meaningful when referring to animate objects (living beings or beings considered living metaphorically, e.g. mythical beings, AIs, etc.) that have a biological genus, but of course we can consider masculine or feminine components of vector representations for any word.

The tendency to use gender pairs in the interest of inclusive language is promoted, for example, in Slavic languages, which use male → female transition[8] to express their gender affiliation or in languages using generic masculinity. Both ways – male→female transition and generic masculine are considered discriminatory by some groups of activists. The dominant tendency in Slovak and Czech is feminization – derivation of female wordforms (feminatives) from the names of persons in the masculine gender e.g. *riaditeľ → riaditeľka* [director → directress], *novinár → novinárka* [newsman – newswoman]. However, the phenomenon of gender equivalence also includes the formation of male counterparts to the names of persons in the female gender e.g. *letuška ↔ steward* [stewardess ↔ steward]. On the other hand, in the name of gender sensitivity, there are also some efforts in Slovakia and Czech Republic to abolish or loosen the strict use of female counterparts to male surnames (Komanická 2016; Valdrová 2001).

The situation in Polish when naming persons of the feminine gender is more complicated, because two ways of expressing gender are used in parallel in the names of professions, titles and social functions: lexemes with the grammatical (word-forming) exponent of the gender (suffixes) e.g. *-ka: dziennikarka* [newswoman] *-owa*: *szefowa* [female boss], *-yni/-ini*: *twórczyni* [female creator], *mistrzyni* [female master], but also through the generic masculine: (*pani*) *doktor* [doctor], (*kobieta*) *szef* [boss], (*kobieta*) *twórca* [creator], (*kobieta*) *mistrz* [master] (Nowosad-Bakalarczyk 2020, p. 215). The dominant tendency of decades past – to use the gender-neutral generic masculine to refer to persons of the feminine gender – is expressed in the *Nowy słownik poprawnej polszczyzny* (1999, cited by Nowosad-

[7] This is an interesting example of a disparity between grammatical gender and (biological or social) sex. *Dievča* [girl] is grammatically neutrum but biologically (or socially) feminine. Yet we perceive it as a feminine equivalent of the masculine *chlapec* [boy].

[8] The Slovak term *prechyľovanie* (and Czech *přechylování*) does not have a well-established English translation. In Lipková 2008 the term *transition* is introduced; Kolek and Valdrová 2020 use the term *feminization*. Other authors use other various terms or circumlocutions.

Bakalarczyk 2020, p. 215): "In the Polish language of the last decades, female names with the exponent *-ka* have been considered not very official, disrespectful, and inconsistent with the seriousness, rank and social position of the indicated people. There was a massive retreat even from the already adopted terms, such as *dyrektorka* [female director], *kierowniczka* [female manager], *profesorka* [female professor], in favour of the expressions of the (*pani*) *dyrektor* [director], (*pani*) *kierownik* [manager], (*pani*) *profesor* [professor]. In the female form, only the names of trades traditionally performed by women remain, e.g. *aktorka* [actress], *malarka* [female painter], *nauczycielka* [female teacher], *pisarka* [female writer], or regarded as unattractive, of low social rank, e.g. *ekspedientka* [saleswoman], *fryzjerka* [hairdresser], *sprzątaczka* [cleaning lady]. In today's Polish language there is no name of a prestigious position, degree or academic title that would have a female word-formation form".

3.2 Categorial operators

Relations between vectors in word embedding models reflect semantic or syntactic relations between words in the language. Unfortunately, it is difficult to arrive at a linguistically sound interpretation of the axes (or coordinates or the basis of the vector space). Therefore, there is not a straightforward operation that replaces one semantic category with a different one (e.g. swapping masculinity for femininity), and we have to resort to a transfer achieved by a difference of another two vectors with well-known semantic roles, preferably perfectly symmetrical and balanced. Usually, the equivalents of *man* and *woman* or personal pronouns *he* and *she* are used, implicitly assuming that other semantic categories of the words (apart from masculinity or femininity) are equivalent ("of the same strength"). We will consider these equivalents as categorial operators.

In our case, these operators express a gender category, namely the category of masculinity: Slovak and Czech *muž*, *on*, Polish *mężczyzna*, Hungarian *férfi* [man], the category of femininity: Slovak and Czech *žena,* Polish *kobieta,* Hungarian *nő* [woman], the superior category of humanity: Slovak *človek,* Czech *člověk,* Polish *człowiek,* Hungarian *férfi* [man] or the category of animality: Slovak *zviera,* Czech *zvíře,* Polish *zwierzę*, Hungarian *állat* [animal]. In arithmetic operations with categories of words, we define categorial inhibitors by "subtracting" a certain category from the examined expression. For example, if we "subtract" the masculinity category called *man* from the term *king*, the term *man* will be a categorial inhibitor of the operation. At the same time, if we add the category of femininity represented by the term *woman* in the same equation (*king – man + woman*), the term *woman* will be categorial activator of the operation (see section 4.1.1). We performed these operations to determine what are the equivalents of the word in question, if by the arithmetic operation we put one category into the background (e.g. masculinity) and another category to the forefront (e.g. femininity).

3.3 Trans-categorial result types

3.3.1 Partial symmetry

The result of the arithmetic operation in our word embedding model represents a trans-categorial result and creates the equivalent of the word, e.g. the well-known English example: *doctor – man + woman = midwife, nurse*. It is evident that the trans-categorial equivalents *midwife* and *nurse* to *doctor* express a typical professional configuration (*doctor ↔ nurse*), but they are not functional equivalents regarding gender equality. We consider this situation to be a case of a gender asymmetry, i.e. the results *midwife* or *nurse* express partial asymmetry to the term *doctor*. However, it is an asymmetry from a comparative point of view, applicable to the case when we consider the correspondence of form and content in configurations such as *king ↔ queen, doctor ↔ doctor* as a standard of gender symmetry, while we compare languages with a different gender system, a language with a grammatical gender and a language with a natural gender (Czech, Slovak, Polish versus English). The topics of our article are the languages of the V4 countries, i.e. three Slavic languages, mostly inflected and strongly gendered (Slovak, Czech, Polish) and one primarily agglutinative and grammatically genderless language (Hungarian).

3.3.2 Full symmetry

The same example yields a different result when we look into an equivalent in the Slovak language word embeddings model. In the expression *doctor – man + woman* (see section 4.3.), *muž* [man] is a categorial inhibitor and *žena* [woman] is a categorial activator. The result is the Slovak feminative *doktorka* [female doctor] as the full equivalent, and it represents the situation of full gender symmetry in the gender pair in Slovak *doktor ↔ doktorka* [male doctor ↔ female doctor]. Male ↔ female word transition is the language mechanism by which a complete symmetry is achieved in inflected languages. This transition – derivation via gender suffixes enables a consistent creation of gender-equivalent forms (feminatives) to appellate male forms e.g. *kuchár ↔ kuchár-ka* [cook – female cook], or to proprial male words e.g. in the surnames *Balog ↔ Balog-ová*.

3.3.3 Asymmetry – nonequivalency

An example of asymmetry is when the investigated expression in the vector model has no opposite-gender equivalent. For example, the term *doctor* in the English language does not have a feminative. Of course, word embeddings show partial context equivalents, already mentioned *nurse* and *midwife*. In the English language environment, the term *doctor* is gender non-equivalent because it represents a syncretic form; *doctor* is a syncretic invariant. Such an "invariant" is a linguistic sign that combines or merges several linguistic meanings. It is one of the most prominent manifestations of asymmetric dualism (Vobořil 2017, p. 429). In the case of the *doctor*, it is not the absence of a feminine equivalent but its inherent presence

in the syncretic invariant (i.e. the generic masculine). Another type of non-equivalence is a situation where the observed expression has no equivalent because there is no equivalent object to it in a real (non-linguistic) world. This happens e.g. in the case of *priest*, when there are no direct gender equivalents to the Polish term *proboszcz* or to the Hungarian term *plebános* (cf. 4.2.2.).

3.3.4 Equivalence continuum

We use word embeddings to show semantic distance between words; the semantic equivalents are displayed in a table, arranged according to their semantic proximity to the result (Table 2). Thus, if there is a semantic equivalent, it will be at the beginning of the table, with a small semantic distance (first column). For example, if we look for the gender equivalent to the word *king*, we see that after performing a vector transfer from the male to the female, the closest expression is the word *queen*, which forms a symmetric gender pair with it. In the result table, words that are semantically close to the resulting word/vector may constitute partial equivalents, e.g. in our example case these are *prince, princess, monarch*. The highest degree of equivalence in terms of gender transfer of the word *king* is represented by the word *queen*, and the interface displays related words according to their semantic distance where, in addition to the most typical gender equivalents, less typical equivalents may occur.

Vector models show the semantics distance between the words table, while the equivalent of the search term is located in the second row of the result table (Table 2). We look for the gender equivalent of the word *king*, and find it in the term *queen*, which is most closely associated with the term *king*, and forms the symmetric gender pair with it in English. Looking at the hierarchy of items in the result table, we see that in lower positions are placed other terms related to the searched word; these may in specific contexts appear as partial equivalents, e.g. *prince, princess, monarch*. Thus, the vector model statistically evaluates the contextual distribution of related expressions to the searched word and creates their hierarchy regarding the degree of equivalence. The term *queen* represents the highest degree of equivalence to the word *king*, but the contextual, gender-differentiated equivalents in certain situations may also be the terms *prince* and *princess*. Vector models display a continuum of related expressions where, in addition to the most typical gender equivalents, less typical equivalents can and do occur. The degree of their proximity to the key expression is quantified by the semantic distance in the vector models for individual languages.

4. TRANS-GENDER EQUIVALENTS

In following examples, we show the table of nearest (i.e. having minimal semantic distance to the result of the arithmetic expression) words (i.e. vectors) and

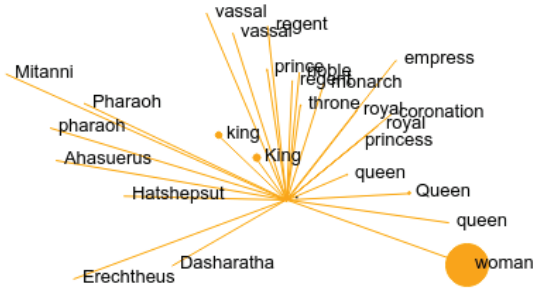
a dimensionality-reduced picture showing spatial arrangement of the nearest vectors. Since the original 200-dimensional space is reduced to two or three dimensions, a certain (and huge) distortion is inevitable and so the rotation and viewing angle were chosen to demonstrate key points of the resulting vector relations, if possible. We also keep the dimensions of the graphs down to three, since because of certain technical difficulties, displaying four-dimensional graphs in printed media is difficult.

4.1 Personal transgender equivalents, case "king"

4.1.1 English vector model: *king* ↔ *queen* (gender symmetry)

In Table + Figure 2, we display the result of the expression *king* – *man* + *woman* = *queen* in the English language model. Note the inconsistent capitalization, common for high-ranking titles in English, and the semantic closeness of the capitalized and uncapitalized variants. Semantic space surrounding the result visibly contains a delineated region of "feminine rulers", with the *queen* being closest to the feminine equivalent of *king*. The left part of the vector model in Fig. 2 represents a proprietary scene in which appear: specific personifications of various ancient rulers,[9] the name of the ancient country in Mesopotamia (*Mittani*), the common name *pharaoh* and the female pharaoh *Hatshepsut*.

Table + Figure 2. Result of the arithmetic expression *king* − *man* + *woman* = *queen* in the English language model.



☒	word	count	
0.326	king	706726	G a A W
0.547	queen	125493	G a A W
0.606	King	830838	G a A W
0.666	monarch	47566	G a A W
0.666	princess	61717	G a A W
0.696	Queen	300230	G a A W
0.713	prince	136930	G a A W
0.717	queen	2017	G a A W

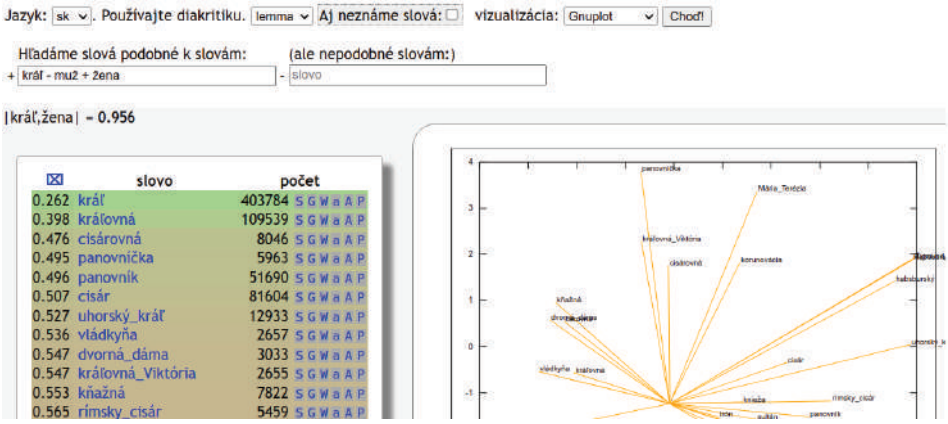
4.1.2 Slovak vector model: *kráľ* ↔ *kráľovná* (gender symmetry)

In Table + Fig. 3, we see the equivalent expression in the Slovak language *kráľ* – *muž* + *žena* = *kráľovná*. Compared to English, the semantic region of "feminine

[9] *Ahasuerus* refers to the fictional ruler of Persia, King Xerxes; *Dasharatha* denotes the ruler of an empire in ancient northern India; *Erechtheus* refers to the King of Athens in Greek mythology; *Hatshepsut* refers to the female pharaoh.

rulers" is more clearly separated, but otherwise the results are expected and unsurprising. In addition to the *kráľovná* [queen] equivalent, the terms *cisárovná, panovníčka, vládkyňa* [empress, female monarch, female ruler] also appear in Slovak as contextual equivalents. The proprietary scene is represented by the names of historical figures such as *Mária Terézia* [Maria Theresa] and *kráľovná Viktória* [Queen Victoria]. In Slovak, the titles of female power holders are consistently feminized.

Table + Figure 3. Result of the arithmetic expression *kráľ – muž + žena = kráľovná* (Slovak for *king – man + woman = queen*).



Compared to English, the semantic region of "feminine rulers" is more clearly separated. Unlike other examples, we bring a screenshot of the (almost) complete search interface here, to demonstrate the features present in the interface.

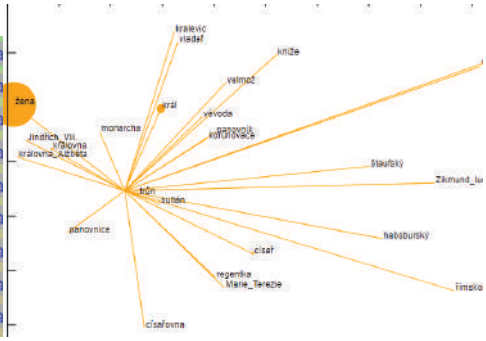
4.1.3 Czech vector model: *král* ↔ *královna* (gender symmetry)

Similarly for Czech *král – muž + žena = královna*. Contextual equivalents are also the terms *panovnice* [female monarch], princezna [princess], následnice [female successor]. The proprietary sphere consists of the names of female historical (real and fictional) personalities like *královna Alžběta Sophia, Brunhilda, Diana, Sibyla.*[10]

[10] The name *Sibyla* can refer to the mythical soothsayer *Sibyl* or to the real historical figure *Sibyla Libštejnská from the House of Kolowrat.*

Table + Figure 4. Result of the arithmetic expression *král – muž + žena = kráľovná* (Czech for *king – man + woman = queen*)

☒	slovo	počet	
0.143	král	533546	s G a
0.380	královna	137948	s G a
0.406	panovník	60418	s G a
0.469	panovnice	3401	s G a
0.472	císař	176151	s G a
0.517	korunovace	11616	s G a
0.519	kníže	99737	s G a
0.529	vévoda	28574	s G a
0.533	trůn	82467	s G a



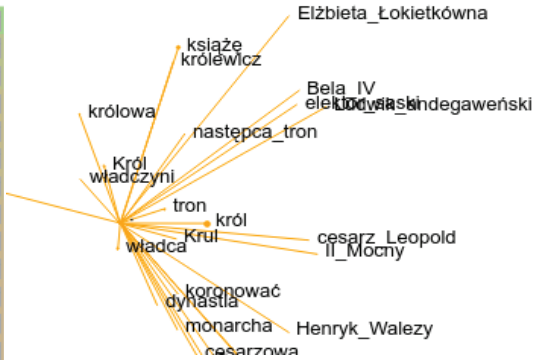
4.1.4 Polish vector model: *król* ↔ *królowa* (partial gender asymmetry)

In Polish (Table + Fig. 5), the result is much less clear: *król – mężczyzna + kobieta* ≈ *władca, monarcha* (i.e. [ruler, monarch], both grammatically masculine), with much less amount of feminine ruler titles. Although one of the equivalents in the vector model is *królowa* [queen consort], it is relatively far from the result of the arithmetic expression (semantic distance of 0.521, cf. Table 5). In Polish, there is a frequent and neutral referencing of female persons through male forms. It often happens in the designation of prestigious professions, social roles and functions, such as: "95-letnia królowa Elżbieta, najstarszy i najdłużej panujący monarcha na świecie, w przyszłym miesiącu będzie obchodzić swoje siódme dziesięciolecie jako władca". [The next month, the 95 years old Queen Elisabeth, the oldest and longest reigning monarch in the world, will celebrate her 75th anniversary as a ruler].[5] Or: "Historia Polski zna dwa przypadki, w których o kobiecie możemy powiedzieć "król": Jadwiga Andegaweńska i Anna Jagiellonka" [The history of Poland knows two cases where we can call woman a "king": the two being Jadwiga Andegaweńska and Anna Jagiellonka].[11] The lower representation of feminatives and the higher competitiveness of male forms in referring to feminine persons in Polish is a feature of this language that is noticeably different from Slovak and Czech. One or the other reference method is used depending on the context or attitude. In Polish, there is a correlation between the expression of social prestige, high importance or historical factuality and the generic masculine (i.e. non-feminatives word form) referring to female persons (*król, władca, monarcha* [king, ruler, monarch], but also *doktor, prezydent, prezes* [doctor, president, chairman] etc.). In addition to this method, feminatives such as *władczyni, cesarzowa, królowa, królewna, caryca* [female ruler, emperor, queen consort, queen, tsarina] are also applied contextually. The proprietary sphere in the Polish vector model (Fig. 5) is expressed by the (rare) name *Elżbieta Łokietkówna* [Queen of Poland and Hungary].

[11] The history of Poland knows two cases in which we can give the title "king" to a woman: Jadwiga Andegaweńska and Anna Jagiellonka.[6]

Table + Figure 5. Results of the arithmetic expression *król – mężczyzna + kobieta* (Polish for *king – man + woman*)

	word	count
0.222	król	88183
0.473	władca	31513
0.478	monarcha	4780
0.521	królowa	15932
0.529	cesarzowa	1235
0.534	cesarz	17168
0.537	cesarz_rzymski	846
0.560	Henryk_Walezy	221
0.571	II_Mocny	301
0.573	książę	52563
0.575	następca_tron	1534
0.576	Cesarz	3752
0.577	Elżbieta_Łokietkówna	54

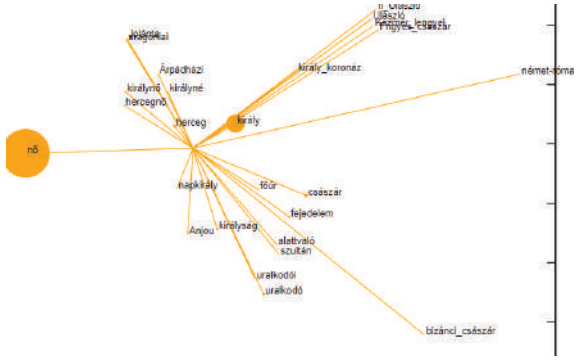


4.1.5 Hungarian vector model: *király* ↔ *királynő, királyné* (gender asymmetry)

Hungarian makes lexical difference between *királyné* [queen consort] and *királynő* [queen monarch], the result of an equivalent expression is *király – férfi + nő = királyné*, i.e. *queen consort*. Let us note that the difference in semantic distance from the word *király* is quite comparable and the preference of *királyné* can be just a result of random corpus composition and model learning process: |*király*, *királyné*| = 0.426 (Table 6); |*király*, *királynő*| = 0.521, while |*királyné*, *királynő*| = 0.356 (i.e. understandably quite close).

Table + Figure 6. Result of the arithmetic expression *király – férfi + nő* (Hungarian for *king – man + woman*)

	word	count
0.212	király	166733
0.426	királyné	7971
0.461	császár	31296
0.483	uralkodó	23201
0.516	herceg	26923
0.521	királynő	18288
0.544	fejedelem	16917
0.545	László	926
0.546	Kázmér_lengyel	137
0.553	királyság	23931
0.571	német-római_császár	879
0.586	Jolánta	102
0.595	uralkodói	3225
0.597	napkirály	589



4.2 Personal transgender equivalents: case "priest"

We can look into some words that are clearly gender-asymmetric (with no change in sight, despite all the prominent public talking about gender neutrality). Let us take the equivalent of two arithmetic expressions in our languages (for brevity, we will omit the graphs and tables of semantically close words), the first one being *priest – man + woman* and the second being *parson – man + woman*.[12] We distinguish the following results according to the degree of equivalence to the phenomena of gender symmetry and asymmetry.

4.2.1 Gender symmetry

In Slovak, *kňaz – muž + žena = rehoľníčka, rehoľná sestra* [priest – man + woman = religious sister]; *farár – muž + žena = farárka* [parson – man + woman = female parson].

In Czech, the results are similar: *kněz – muž + žena ≈ řeholnice* [priest – man + woman ≈ religious sister]; *farář – muž + žena = farářka* [parson – man + woman = female parson].

4.2.2 Gender asymmetry

In Polish, the search for a female equivalent of the terms *ksiądz* [priest] and *proboszcz* [parson] results in gender asymmetry. In the case of the expression *ksiądz – mężczyzna + kobieta*, the semantically closest term is the phrase *błogosławionej pamięci* (or the acronym *bp, BP*) [of Blessed Memory], which used to occur in relation to the deceased of the Jewish faith, but nowadays is extended towards other religions, as the corpus occurrences show). Our interpretation of the result is that there is no female equivalent of the word *ksiądz* (Tab. 7a).

In the case of *proboszcz – mężczyzna + kobieta*, the resulting vector remained close to the word *proboszcz* (*proboszcz, ksiądz proboszcz, proboszcz, parafia*, Tab. 7b). *Parafia* [parish] just happened to be a feminine gender noun closest to the result of the expression, similar to the phrase *błogosławionej pamięci* (combination of two feminines in Polish) being close to the term *ksiądz.*

In both cases, the asymmetry is also well visible on the visualization (Fig. 7a, 7b). There are no results even remotely relevant present, meaning that the vector transfer did not move us to anything resembling the "feminine region". This implies that the given expressions do not have a female counterpart in the Polish language. Here, language transgender non-equivalency corresponds to non-language asymmetry. In the absence of denotation in the subject reality, the most frequent word or phrase contextually linked to the search term is often nonsense or irrelevant (*błogosławionej pamięci, parafia*).

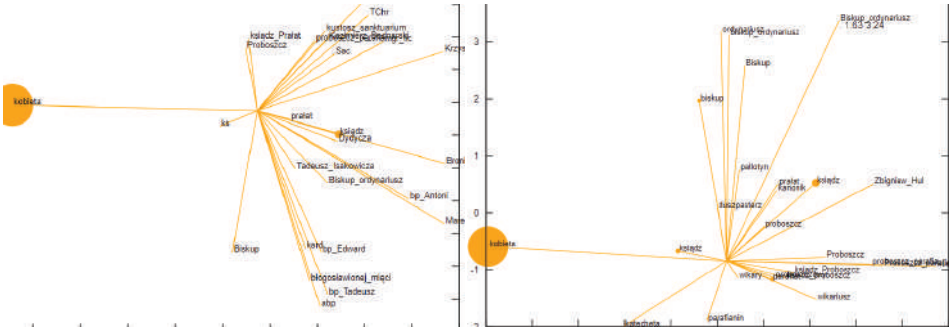
[12] There are no exact semantic equivalents for these religious terms between these languages and English; the glosses we give in the text are approximate.

Table 7a. Result of the search scheme *ksiądz – mężczyzna + kobieta*
Table 7b. Result of the search scheme *proboszcz – mężczyzna + kobieta*

0.189	ksiądz	129825	G a A
0.474	błogosławionej_mięci	3450	G a A
0.480	Proboszcz	4293	G a A
0.503	Biskup	12002	G a A
0.503	Zbigniew_Hul	75	G a A
0.514	bp_Antoni	481	G a A
0.516	Sac	2879	G a A
0.521	bp_Edward	375	G a A
0.522	ks	3407	G a A

0.414	Proboszcz	4293	G a A W
0.471	proboszcz_parafia	5796	G a A W
0.490	ksiądz_proboszcz	1973	G a A W
0.510	parafia	73621	G a A W
0.519	wikary	1912	G a A W
0.527	ksiądz	129825	G a A W
0.532	Biskup	12002	G a A W
0.540	parafianin	9428	G a A W
0.558	biskup	64293	G a A W
0.563	wikariusz	3778	G a A W

Figure 7a. Result of the arithmetic expression *ksiądz – mężczyzna + kobieta* (Polish for *priest – man + woman*)
Figure 7b. Result of the arithmetic expression *proboszcz – mężczyzna + kobieta* (Polish for *priest – man + woman*)



In Hungarian, *pap – férfi + nő = szerzetesnő* ([religious sister] feminine); *plébános – férfi + nő = ferences_atya* ([Franciscan father] shows no relevant results, there is no feminine specific region in the semantic space.

In the result table (Tab. 8) for the word *pap* [priest], the word *szerzetesnő* [nun] appears as an equivalent and refers to a conventionalized pair of *priest* ↔ *nun, religious sister*. But the nearest semantic equivalent *szerzetesnő* [nun] cannot be considered a gender equivalent of the term *pap*. However, symmetric family pairs denoting *monks* and *nuns* have the same denotation value, e.g. *szerzetes* ↔ *szerzetesnő*, [religious brother ↔ religious sister]. Thus, neither the term *pap* (nor *plebános*) itself has a gender equivalent in Hungarian and as such, it is an example of gender asymmetry.

Table + Figure 8. Result of the arithmetic expression *pap – férfi + nő* (Hungarian for *priest – man + woman*)

0.220	pap	66114	G a A W
0.487	szerzetesnő	657	G a A W
0.509	papság	6750	G a A W
0.516	szerzetes	20894	G a A W
0.532	hívő	20635	G a A W
0.536	világi_hívő	384	G a A W
0.543	katolikus	35977	G a A W
0.552	klérus	1179	G a A W
0.552	keresztény	30652	G a A W
0.553	püspök	42411	G a A W

We hypothesize that there is a (historical) religious difference among the countries (Catholics vs. Protestant) – there is no shift to female semantic region in countries without protestant tradition for the word *parson*, while the *priest* has a *religious sister* as an equivalent (with the exception of Polish). This can be considered a case of clear gender related asymmetry.

4.3 Personal transgender equivalents: the well-known "doctor" case

4.3.1 Gender asymmetry

It is a common knowledge that in English (Bolukbasi et al. 2016), vector transfer from masculine to feminine while starting from the word *doctor* gives less prestigious jobs, such as *nurse* or *midwife* as the feminine equivalents. This is also confirmed in our English language model: *doctor – man + woman = midwife, nurse, physician, pharmacist, lactation consultant, pediatrician, gynecologist.* Apart from *gynecologist* (and perhaps *physician*), all these words are indeed less prestigious – the transfer took us to the semantic region of less prestigious job titles. And the presence of *gynecologist* is easily explained by the gender transfer being relevant not to the gender of the doctor, but to the gender of the patients. Since the word *doctor* is theoretically gender neutral, we can reverse the transfer and arrive at the result of *doctor – woman + man = physician, surgeon, dentist, neurosurgeon, cardiologist, urologist*. We moved to a semantic region of more prestigious job titles, with the *urologist* perhaps filing the same role as *gynecologist* in the previous expression, i.e. catering for exclusively male patients.

In the V4 languages there are two different words for this job title – both used in the general sense of *practitioner of the medical profession*, but the second one is in addition used also as an academic title (situation quite similar to the use of *doctor* in English).

As for the Polish vector model, we have *doktor* [doctor], *lekarz* [physician].

In Polish, *doktor – mężczyzna + kobieta = profesor, prof*; and *lekarz – mężczyzna + kobieta = onkolog, pediatra, ginekolog, neonatology* etc. There is no transfer to a feminine region in any of the cases, while *ginekolog* represents the transfer to the gender of the patient.

In the Hungarian vector model, we can look at the words *doktor* [doctor], *orvos* [physician]: *doktor – férfi + nő = orvos, professzor, sebész…* [physician, professor, surgeon…]; *orvos – férfi + nő = nőgyógyász, szakorvos, családorvos, onkológus…* [gynecologist, specialist doctor, family doctor, oncologist…]. We see no transfer to a specifically female semantic region present in the latter two languages; neither is there any transfer to a region of less prestigious job titles in any of the V4 languages.

4.3.2 Gender symmetry

For the Slovak and Czech vector model, we have *doktor* [doctor], *lekár/lékař* [physician].

In both Slovak and Czech, it reads *doktor – muž + žena = doktorka* [female doctor]; in Slovak, *lekár – muž + žena = gynekológ*; in Czech, *lékař – muž + žena = pediatr, lékařka, gynekolog, diabetolog*. If fact, the result of the gender vector transfer moved us somewhere between *lekárka, gynekológ* and a region of female job titles (productively derived from the male ones); equivalently in Czech. Accidentally, *gynekológ* in Slovak or *pediatr* [paediatrician] in Czech were slightly closer than other words.

Getting back to the English model, we have to realize that a typical web content in English is quite different as compared to V4, not being limited to a group of geographically and culturally close and connected countries. In addition, there is a lot of diachronical text (e.g. fiction not covered by copyright protection, old digitized newspaper articles etc.), where the medical profession was male dominated. This is not the case of V4 languages, where old newspaper archives are rarely publicly available (and usually only in the form of scanned pictures, that did not make their way to the web corpus), with older fiction being rather rare.

4.4 Personal transgender equivalents: the case of "steward"

We can look at some examples where we expect some lexical gender asymmetry – in both Slovak and Czech, *letuška* is the term for *flight attendant*, but the word (of a feminine gender) refers only to female members of that profession, with the male version being a rather notable lexical lacuna, especially considering that modern trends in gender equality reached this profession already many years ago. Although still not quite equal in the numbers, male employees are quite common and not a remarkable occurrence any more.

In Slovak, we get *letuška – žena + muž = steward, stevard*; in Czech, *letuška – žena + muž = stevard* (and *steward* still present among the closest words, just not at the first nor second place).

Both Polish and Hungarian models are not equivalent at all – in Polish, the feminine *stewardesa* is regularly derived from the masculine *steward*, so there is no lexical lacuna to be probed; nevertheless, we see that *steward – mężczyzna + kobieta = assistant, stewardesa, kelner, stewardessa.* The result of the transfer is however quite far away from any of those words.

4.5 Impersonal transgender equivalents (animal ancestral pairs): the case of "dog"

In Slovak, it is notoriously difficult to arrive at the feminine form of the word *pes* [dog], which is a masculine gender, used either specifically for males or as a generic term applied indiscriminately for all the members of the *Canis familiaris* species. The widely used *fena* or *fenka* is fought against by prescriptivists (as a loanword from Czech); the prescribed "correct" *suka* is perceived as very offensive and vulgar. Since prescriptivism used to dominate the Slovak lexicographic scene for decades, the term *suka* is continued to be widely used by linguistically self-conscious authors anyway.

Investigating transgender equivalents in word embedding models, we can observe how language users perceive gender equivalence in a group of non-personal[13] life objects, e.g. animals. The most common species of animals have different names in Slovak for males and females of the same species, e.g. *kocúr* ↔ *mačka* [tomcat ↔ female cat, *kohút* ↔ *sliepka* [rooster ↔ hen]. However, the results of vector arithmetic operations show that gender equivalence is not so straightforward and that people's perception of the equivalence of expressions is determined not only by the factual relationship between denotations (biological gender) but also by formal relationship between expressions (grammatical gender) and other circumstances (cf. Wachtarczyková – Garabík 2022, p. 85). An example of this is the word *dog*. The equivalence in our Slovak word embedding model expressed by *pes – on + ona = mačka* ([dog] grammatical masculine) *– he + she* = ([cat] grammatical feminine) (Tab. 9a) represents a case of a generic feminine. We used the fact that 3rd person pronouns are lemmatized as their base form (nominative singular) but without the change of the gender.

Thus, in our word embedding model, the gender equivalent of a *dog* is a *cat*. The perception of the gender equivalent of a certain masculine expression in Slovak seems to be also influenced by the grammatical gender of the corresponding word (such an in the Slovak terms for the *ferret* or the *chihuahua*, which possess feminine grammatical gender). Other symptoms of femininity such as smallness, animal resilience can be derived from some recurring items in the result table of Slovak and Hungarian equivalents of *ferret* (Slovak *fretka*, Hungarian *görény*), *guinea pig*

[13] We do not insinuate animals are not or should not be treated as persons; we are using "person" as a linguistic term here.

(Slovak *morča,* Hungarian *tengerimalac), small dog* (Slovak *psík*, *šteňa,* Hungarian *kutyus*, *kölyökkutya*).

Table 9a. Result of the arithmetic expression *pes – on + ona* (Slovak for *dog – he + she*)
Table 9b. Result of the arithmetic expression *kutya – férfi + nő* (Hungarian for *dog – man + woman*)

0.110	pes	740820	S	G	W	a	A	P	0.207	kutya	149265	G	a	A	W
0.369	mačka	211899	S	G	W	a	A	P	0.359	macska	43713	G	a	A	W
0.468	psík	131885	S	G	W	a	A	P	0.410	cica	16203	G	a	A	W
0.471	fretka	7596	S	G	W	a	A	P	0.459	kutyus	12412	G	a	A	W
0.484	čivava	4497	S	G	W	a	A	P	0.475	gazdi	12266	G	a	A	W
0.499	šteňa	33832	S	G	W	a	A	P	0.508	tengerimalac	2073	G	a	A	W
0.514	zviera	584333	S	G	W	a	A	P	0.510	kölyökkutya	1451	G	a	A	W
0.530	králik	39779	S	G	W	a	A	P	0.511	kiskutya	6816	G	a	A	W
0.534	morča	8152	S	G	W	a	A	P	0.537	görény	3040	G	a	A	W

Table 9c. Result of the arithmetic expression *pes – on + ona* (Czech for *dog – he + she*)
Table 9d. Result of the arithmetic expression *pies – on + ona* (Polish for *dog – he + she)*

0.356	pes	333197	S	G	a	A	W	0.327	pies	47808	G	a	A	W
0.448	kočka	92082	S	G	a	A	W	0.483	sunia	2147	G	a	A	W
0.485	fretka	4092	S	G	a	A	W	0.491	suczka	2693	G	a	A	W
0.495	čubina	490	S	G	a	A	W	0.543	suka	2339	G	a	A	W
0.513	psík	6815	S	G	a	A	W	0.565	kotka	4992	G	a	A	W
0.521	jezevčík	5645	S	G	a	A	W	0.585	jamniczka	125	G	a	A	W
0.523	krysařík	653	S	G	a	A	W	0.589	piesek	5049	G	a	A	W
0.524	fenka	28132	S	G	a	A	W	0.598	szczeniak	1815	G	a	A	W
0.525	borderka	772	S	G	a	A	W	0.609	psinka	130	G	a	A	W
0.534	jezevčice	663	S	G	a	A	W							

The same approach would work neither in Czech nor Polish models since their lemmatization is different – personal pronouns are lemmatized as the nominative singular masculine: *ona* is lemmatized as *on*. However, we can use the "word" model (not the "lemma" one) and get the results for Czech *pes – on + ona = kočka* ([cat] feminine) and for Polish *pies – on + ona = sunia* ([bitch] diminutive feminine). Slovak and Czech models share the cat as the feminine variant of dog. In the Czech model (Table 9c) the lower positions also include terms denoting generally *psík* ([dog] deminutive), specific representatives of small breeds e.g. *jezevčík, krysařík*, a slang term *čubina* [bitch], and way down in the list, we find the neutral expression *fena*.

An equivalent approach is not possible in Hungarian because the language lacks gendered pronouns.

Nevertheless, we might try a different arithmetic expression, *kutya – férfi + nő = macska, cica* [dog – man + woman = cat, kitty]. We are subtracting (human) male and adding (human) female, which is not the best approach, but we hope the other semantic categories (i.e. vector components) carried by the words *férfi* and *nő* cancel themselves out. And indeed, the (rather unexpected) result (Tab. 9b) strongly suggests that there are other factors at play, not just the grammatical gender (which the Hungarian lacks); or the biological one. We hypothesize that the result reflects a typical gender of the *owner* of the pet animal – men are usually more likely perceived as dog owners and women as cat owners (at least based on the texts in the corpus). Or, equally likely, dogs are often described as more masculine, strong, alpha-males; cats are described as effeminate, refined, emotional.

In the Polish vector model, gender (and species) symmetry manifests itself in the word *dog*. The first three equivalents in the Polish model (Tab. 9d) denote the *female dog* (*sunia, suczka* [bitch]). They only differ in the emotional value, as *sunia* and *suczka* (also the term *psinka*) are diminutives of the neutral expression *suka* [bitch].

4.6 An example with humanity inhibitor and animality activator

We can move beyond simple gender related lexical asymmetry and explore other types, e.g. geographical (see Garabík 2020). We will not discuss these asymmetries here, just give an interesting gender related example – the equivalents of *man (male human) – human + animal* and *woman – human + animal*, i.e. the transfer from humans to animals, while keeping the gender category intact.

In Hungarian, *férfi – ember + állat = hím* [animal male]; *nő – ember + állat = nőstény* [animal female]; this is what we expect from simple ontological arguments.

In Polish, *mężczyzna – człowiek + zwierzę = ciężarny, czworonóg* [pregnant, quadruped]; *kobieta – człowiek + zwierzę = ciężarny* [pregnant]; i.e. the transfer took us somewhere within the same semantic region, but there is no noticeable "gendered animal region" around, and the nearest words are there just by chance.

In Slovak, *muž – človek + zviera ≈ králik*, *šteňa* [rabbit, puppy]; *žena – človek + zviera = mačka* ([cat] female or generic feminine). There is an equivalent animal region in the semantic space, but the perception of an animal equivalent is somewhat unfavourable for both of the genders.

In Czech, *muž – člověk + zvíře ≈ kanec* [boar]; *žena – člověk + zvíře ≈ lvice* [lioness]. Given the use of *kanec* and *lvice* in Czech to denote (human) sexual prowess, the authors notice certain unpleasant cross-language discrimination greatly favouring the Czech language.

Table 10a. Result of the arithmetic expression man – human + animal (Czech *muž – člověk + zvíře*)

Table 10b. Result of the arithmetic expression woman – human + animal (Czech *žena – člověk + zvíře*)

0.453	zvíře	851859	S G a A W
0.654	muž	2162680	S G a A W
0.695	tygřice	2420	S G a A W
0.708	lvice	8933	S G a A W
0.709	kanec	11853	S G a A W
0.726	býk	62104	S G a A W
0.739	skot	38683	S G a A W

0.459	zvíře	851859	S G a A W
0.621	žena	2746127	S G a A W
0.689	lvice	8933	S G a A W
0.702	gravidní	1839	S G a A W
0.708	tygřice	2420	S G a A W
0.713	kojící	26492	S G a A W
0.735	těhotný	123744	S G a A W

These examples just show how the categories of masculinity, femininity, humanity and animality are projected into individual expressions in different languages.

5. CONCLUSION

We examined gender vector transfer for selected words in word embedding models for the Visegrád group languages: Slovak, Czech, Polish and Hungarian. Unlike the well-known transfer to less prestigious professions present in English models (when transferring from masculine to feminine, e.g. in the medical field), these languages do not exhibit such a phenomenon. We attribute this absence to the composition of the English corpora which the word embedding models are typically trained on (often web corpora) and which do contain a significant amount of diachronic or otherwise biased text.

Slovak and Czech show results unsurprisingly similar to each other (productive use of feminine equivalents, without corresponding decrease in prestige); Polish exhibits functioning generic masculine and Hungarian is gender-neutral, though specific female professions are lexically distinguished. In the Slovak and Czech vector models, gender equivalence and feminization were common in most of the examples (*king*, *doctor*, *priest*), while the perception of animal gender equivalents is similar (*dog*).

In terms of gender symmetry, Polish is different from Slovak and Czech. In the category of nomina agentis (names of professions), gender asymmetry is apparent by the widespread use of generic masculine (e.g. *doktor, filolog* [doctor, philologist]) as a gender neutral term for both males and females. The gender is indicated analytically, by prepending the terms with the words (honorifics) *Pan/Pani*, which – unlike the generic masculine – are inflected. Feminatives are only the names of traditionally "female" professions (*aktorka, nauczycielka* [actress, female teacher]) or not very attractive and socially non-prestigious professions (Lipková 2008, p. 45).

A correlation between generic masculinity and prestige that seems to appear in the Polish language was noted in (Novosad-Bakalarczyk 2006, p. 136; Lipková 2008, p. 42). Gender asymmetry in Polish, manifested by the use of the generic masculine to denote women's professions, is dominant in formal and official communication, where it indicates social status. In informal communication and referring to women from a personal perspective, gender pairs are regularly created and gender symmetry is applied.

The dominance of gender asymmetry in texts from the public sphere is reflected in word embeddings, as demonstrated by the words *doctor, lekarz* [doctor, physician]. Another manifestation is the absence of denotation in reality (and thus also the absence of a female equivalent) for the term *ksiądz* [priest] and *proboszcz* [parson]. Nevertheless, transfer to a less prestigious "semantic region" by transferring the gender from masculine to feminine was not exposed by our word embeddings model.

Slovak, Czech and Polish are distinguished from Hungarian by the presence of a grammatical category of gender attributed to each noun. In contrast, Hungarian distinguishes by gender only those nouns that refer to objects with a notable biological genus, i.e. persons and some animals. Hungarian derivation of feminine variants of masculine words bears only a superficial similarity to the Slavic languages and is more a reflection of the agglutinative characteristics of the language. The second component of these compound words is often the noun *nő* ([woman] e.g. *király* ↔ *királynő*), sometimes *asszony* [wife] (e.g. *szomszédasszony*) or *lány* [girl] (e.g. *diáklány*) (Misadová 2011, p. 68). There are also gender pairs expressed heteronymously (*bácsi* ↔ *néni*) especially in family relations and animals. An example of a gender symmetry in Hungarian is manifested by word *doctor*, with the gender pair *orvos* ↔ *orvosnő*. The partial symmetry *király* ↔ *királyné/királynő* appears for the pair *king* ↔ *queen*. On the contrary, gender asymmetry was demonstrated by the absence of denotation for the Hungarian equivalent of the word *priest*. The word embeddings model does not have an equivalent (only surrogate) gender equivalents for the terms *pap* and *plebános*. There were also similarities between Slovak/Czech and Hungarian in the perception of family pairs of animals for the equivalent of the word *dog* (*pes, kutya*), where the female equivalent of *dog* (*kutya*) is also a *cat* (*macska, cica*).

Finally, when comparing typologically different languages, the question arises whether the idea of an inclusive and gender-balanced language is universal or a challenge only for certain types of languages? The question is also whether feminization is the most important indicator of a gender-balanced language, if the generic masculine has an irreplaceable place in referring to holders of certain professions in some languages (Polish). In the Polish language, the names of professions with higher social prestige which refer to persons of the female gender require the use of the generic masculine with a feminine honorific (*pani profesor, pani architekt*).

On the other hand, in most cases in Slovak and Czech (*doctor, priest, queen, steward*) vector models show consistent feminization and no drop in prestige for feminatives (at least compared to widely reported drop in prestige in English language models). In these languages of the V4 group, there is the highest degree of gender symmetry in the nomina agentis category.

References

BOJANOWSKI, Piotr – GRAVE, Edouard – JOULIN, Arnaud – MIKOLOV, Tomas (2017): Enriching word vectors with subword information. In: *Transactions of the Association for Computational Linguistics*, Vol. 5, pp. 135–146.

BOLUKBASI, Tolga – CHANG, Kai-Wei – ZOU, James Y. – SALIGRAMA, Venkatesh – KALAI, Adam T. (2016): Man is to computer programmer as woman is to homemaker? debiasing word embeddings. In: *Advances in Neural Information Processing Systems*, pp. 4349–4357.

GARABÍK, Radovan (2020): Word Embedding Based on Large-Scale Web Corpora as a Powerful Lexicographic Tool. In: *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, Vol. 46, No. 2, pp. 603–618.

KOLEK, Vít – VALDROVÁ, Jana (2020): Czech gender linguistics: topics, attitudes, perspectives. In: *Slovenščina 2.0*, Vol. 8, No. 1, pp. 35–65.

LIPKOVÁ, Zuzana (2008): *Přechylování v českém a polském jazyce*. Bakalářská diplomová práce. Olomouc: Univerzita Palackého v Olomouci.

MIKOLOV, Tomas – YIH, Wen-tau – ZWEIG, Geoffrey (2013): Linguistic Regularities in Continuous Space Word Representations. In: *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics*, pp. 746–751.

MISADOVÁ, Katarína (2011): *Kapitoly z morfológie maďarského jazyka. Kontrastívny opis niektorých morfologických javov maďarského jazyka*. Bratislava: Univerzita Komenského v Bratislave. Avalaible at: https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/Misad_Kapitoly_z_morfologie_madarskeho_jazyka.pdf [cit. 09-02-2023].

KOMANICKÁ, Daniela (2016): Patronymia a patrilinearita vo vzťahu k nerovnému postaveniu žien v spoločnosti. In: L. Molnár Satinská – I. Valentová (eds.): *Prechyľovanie: áno – nie?* Bratislava: Veda, vydavateľstvo SAV, pp. 55–67.

NOVOSAD-BAKALARCZYK, Marta (2006): Tendencje w sposobie wyrażania żeńskości we współczesnej polszczyźnie. In: *Język Polski*, Vol. 86, No. 2, pp. 126–136.

NOVOSAD-BAKALARCZYK, Marta (2020): O kategorii płci i rodzaju w polszczyźnie. In: *Open series Vilnius University 2020*, Vol. 1, pp. 209–228. DOI: https://doi.org/10.15388/VLLP.2020.10 [cit. 09-02-2023].

ŘEHŮŘEK, Radim – SOJKA, Petr (2010): Software framework for topic modelling with large corpora. In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 46–50.

SATINSKÁ, Lucia (2014): Problematika rodovo citlivého vyjadrovania v maďarčine (v Maďarsku a na Slovensku). In: J. Cviková (ed.): *Analýza významu a možností používania*

rodovo vyváženého jazyka. Bratislava: Centrum vzdelávania MPSVaR, pp.101–109. Available at: http://www.ruzovyamodrysvet.sk/chillout5_items/1/6/0/1/1601_5755f0.pdf cit. [03-02-2023].

ŞENEL, Lüfti Kerem – UTLU, Ihsan – YÜCESOY, Veysel – KOÇ, Aykut – ÇUKUR, Tolga (2018): Semantic structure and interpretability of word embeddings. *IEEE/ACM Trans.* In: *Audio, Speech and Lang*. Vol. 26, No. 10, pp. 1769–1779.

ŠTEFAŇÁKOVÁ, Jana (2021): Rodová symetria v slovenských publicistických textoch na pozadí medzinárodnej politiky zameranej na inkluzívny/rodovo vyvážený jazyk. In: *Slovenská reč*, Vol. 89, No. 3, pp. 292–327.

URBANCOVÁ, Lujza (2018): Slovenčina a diskurz o rodovej rovnosti. In: J. Pekarovičová – M. Vojtech (eds.): *Studia Academica Slovaca 47*. Bratislava: Univerzita Komenského v Bratislave, pp. 91–306.

VALDROVÁ, Jana (2001): Stereotypy a klišé v mediální projekci genderu/Gender Stereotypes and Clichés as Presented in the Media. In: *Sociologický Časopis/Czech Sociological Review*, Vol. 37, No. 2, pp.183–205. Available at: http://www.jstor.org/stable/41131617 [cit. 13-01-2023].

VOBOŘIL, Ladislav (2017): Synkretismus jako jazykový jev, lingvistická kategorie a lingvistický termín. In: *Slavica Vratislaviensia,* Vol. 165, pp. 429–440. DOI: 10.19195/0137-1150.165.39. [cit. 09-02-2023].

WACHTARCZYKOVÁ, Jana – GARABÍK, Radovan (2022): Vlastné mená a vektorové priestory ako výsledky korpusovej inštrumentalizácie. In: T. Bánik – J. Bauko (eds.): *Vlastné mená v interdisciplinárnom kontexte*. Nitra: Univerzita Konštantína Filozofa 2022, pp. 78–102.

Internet sources

[1] Gender Equality Strategy 2020 – 2025 (Európska stratégia rodovej rovnosti 2020 – 2025). Available at: https://eur-lex.europa.eu/legal-content/SK/TXT/PDF/?uri=CELEX:52020DC0152 [cit. 09-02-2023].

[2] Plan of Gender Equality SAS 2021 – 2024 (Plán rodovej rovnosti SAV 2021 – 2024. Available at: https://www.sav.sk/?lang=sk&doc=sas-gender [cit. 09-02-2023].

[3] Inclusive Communication in the General Secretariat of the Council of the European Union (Inkluzívna komunikácia na Generálnom sekretariáte Rady Európskej únie). https://www.consilium.europa.eu/media/35439/sk_brochure-inclusive-communication-in-the-gsc.pdf [cit. 09-02-2023].

[4] Glosár rodovej terminológie. Available at: http://glosar.aspekt.sk/default.aspx?smi=1&ami=1&vid=175 [cit. 09-02-2023]

[5] https://www.rp.pl/polityka/art19270361-70-lat-krolowej-elzbiety-na-tronie-republikanie-chca-zniesienia-monarchii [cit. 09-02-2022].

[6] https://pch24.pl/kobieta-krol-dwa-przypadki-w-polskiej-historii [cit. 09-02-2023].

Resumé

RODOVÁ ASYMETRIA JAZYKOV VYŠEHRADSKEJ SKUPINY A JEJ ODRAZ VO VEKTOROVOM PRIESTORE

Do problematiky diskriminačnej inštrumentalizácie jazyka patria aj stratégie a spôsoby vyjadrovania rodu v jazyku a sledovanie prejavov rodovej ekvivalencie. V článku využívame metódu vektorovej reprezentácie slov, ktorá ako štandardná metóda spracovania prirodzeného jazyka poskytuje pohľad na jazykovú realitu prostredníctvom modelov odrážajúcich sémantické vzťahy medzi slovami. Vektorový priestor, ktorý tieto modely vytvárajú, odráža mieru a podobu rodovej symetrie, resp. ekvivalencie, pri skúmaní rodových párov v kategórii pomenovaní osôb alebo zvierat. Skúmame vektorové reprezentácie vybraných slov v jazykoch Vyšehradskej skupiny a aplikujeme základnú vektorovú aritmetiku na demonštráciu základnej jazykovej asymetrie prítomnej v modeloch. Nadväzujeme na známy príklad z anglických vektorových modelov, v ktorých vektorový transfer pri hľadaní rodového ekvivalentu smeruje k úplnej a symetrickej ekvivalencia (*kráľ* ↔ *kráľovná*: king – man + woman = queen), alebo k neúplnej, resp. nesymetrickej ekvivalencii s diskriminujúcim vyznením ako napr. v prípade *programátor*, ktorého náprotivkom je *žena v domácnosti* (programmer – man + women = homemaker). V článku skúmame podobné transfery v modeloch jazykov V4 – slovenčiny, češtiny, poľštiny a maďarčiny. Rozlišujeme typy ekvivalencie, pričom z hľadiska miery vyjadrenia feminity a zároveň adekvátnosti referencie uvažujeme o úplnej symetrii, čiastočnej asymetrii a úplnej asymetrii v rámci sledovaných rodových párov. Výsledky analýzy korešpondujú so známymi jazykovými faktami: feminizácia (rodová symetria) sa dominantne prejavuje vo všetkých flektívnych jazykoch, v poľštine je však evidentná zóna rodovej asymetrie v názvoch profesií, daná historicky a sociokultúrne. Čiastočná rodová asymetria je aj dôsledkom rodovej neutrality maďarčiny. Vektorové reprezentácie slov však odrážajú aj menej zjavné pôsobenie faktorov, ktoré sa podieľajú na spôsobe manifestácie rodu (feminity) v jazykoch. Ide napr. o rôzne medziparadigmatické analógie a korelácie dané vecnou blízkosťou denotátov alebo pôsobením gramatického rodu súvzťažných slov (v prípade absencie rodového ekvivalentu k výrazu *kňaz* v poľštine). Cieľom článku bolo načrtnúť možnosti ďalšej analýzy a interpretácie jazykového materiálu prostredníctvom vektorových modelov, ktoré práve pri typologicky odlišných jazykoch relevantne zachytávajú kontextuálnu sémantiku výrazov a uľahčujú orientáciu vo vzťahoch ich vzájomnej podobnosti, a tým aj vo vzťahoch rodovej ekvivalencie.