



## Увид у израду вишејезичног репозиторијума фраземских конструкција

Кристина Илић<sup>1</sup> — Катрин Шлунд<sup>2</sup>

<sup>1</sup>  <https://orcid.org/0009-0002-7426-5530>, <sup>2</sup>  <https://orcid.org/0000-0001-9130-0233>

<sup>1</sup> Универзитет у Београду – Филолошки факултет (Србија)

<sup>2</sup> Универзитет Мартина-Лутера Хале-Витенберг (Немачка)

<sup>1</sup> kristina.ilic997@gmail.com (✉), <sup>2</sup> katrin.schlund@slavistik.uni-halle.de

### Сажетак

У раду се пружа увид у основе израде вишејезичног репозиторијума такозваних фраземских конструкција (ФК) у оквиру COST акције PhraConRep *Вишејезични репозиторијум фраземских конструкција на језицима централне и источне Европе* („A Multilingual Repository of Phraseme Constructions in Central and Eastern European Languages”). Фраземске конструкције (ФК) су шаблони који се састоје од фиксних лексичких елемената, тзв. сидра, и одређеног броја празних места која се попуњавају тзв. филерима. Један од главних циљева COST акције је стварање конструктивно ограниченог опсега, односно дигиталног, вишејезичног речника у којем ће се инвентарисати ФК немачког језика и њихови еквиваленти у четрнаест европских језика, укључујући српски. ФК представљају изазов за лингвисте, професионалне преводиоце и ученике страних језика, а чак и преводачки програми базирани на вештачкој интелигенцији не нуде увек адекватне преводе. На основу аутентичних примера из конструктографске праксе, у раду се објашњава који су граматички, семантички и прагматички елементи описа ФК и како се одређују њихови еквиваленти.

**Кључне речи:** фраземске конструкције, идиоматске конструкције, конструкциона граматика, конструкцион, лексикографија, немачки језик, српски језик, превођење.

### Abstract

This paper presents the basics for the creating of a multilingual repository of so-called Phraseme Constructions (PhraCons) in the framework of the COST Action PhraConRep „A Multilingual Repository of Phraseme Constructions in Central and Eastern European Languages”. PhraCons are multiword units consisting of fixed elements (the anchor), and open slots for fillers. The COST Action PhraConRep aims to create a construction of limited scope, i.e., a digital, multilingual repository in which PhraCons will be stored and described. Based on German as the source language of the repository, PhraCons are translated into fourteen languages of Central, Eastern and South-Eastern Europe, including Serbian. PhraCons represent a challenge for linguists, professional translators and students of foreign languages, and even translation programs based on artificial intelligence do not always offer adequate translations. Based on authentic examples from constructographic practice, the paper explains the grammatical, semantic and pragmatic elements of description of PhraCons in the repository and how their equivalents are determined.

**Keywords:** phraseme constructions, constructional idioms, construction grammar, constructicon, lexicography, German language, Serbian language, translation.

### 1. Увод

Међу најзначајније разлике између матерњих и нематерњих говорника спадају учесталост и спретност употребе формулаичног језика (TOMASELLO 2005: 101–102; MILLER 2018: 23). Зато је овладавање формулаичним језиком један од најважнијих циљева у настави страног језика. Вишечлане јединице (енгл. *multiword units*, нем. *Mehrworteinheiten*) с лексички неодређеним компонентама представљају посебан

тип формулаичног језика, чије је истраживање у прошлој деценији довело до приближавања фразеологије и конструкционе граматике (уп. ЗЕМ 2018).

Овде описани лексикографски рад спроводи се у оквиру COST акције 22115 PhraConRep „Вишејезични репозиторијум фраземских конструкција на језицима централне и источне Европе”. Доброволски (DOBROVOL'SKIЈ 2011: 114) дефинише фраземске конструкције (ФК) као лексички делимично специфич-

не парове облика и значења (тј. конструкције), чија семантика произлази не само из лексичког значења њихових компоненти и начина њихове синтактичке повезаности, већ и из саме конструкције као целине. Другим речима, значење ФК је некомпозиционо. Ради се, дакле, о конструкцијама у смислу конструкционе граматике (DOBROVOL'SKIЈ 2011: 114; MELLADO BLANCO et al. 2022: 4). ФК се састоје из лексички фиксних елемената (тзв. *речи-сидра* које заједно чине тзв. *сидро*) и из празних места која се тек при коришћењу шаблона попуњавају конкретним лексичким материјалом, тзв. *филерима*. ФК и сличне појаве такође су познате под другим називима, међу којима су најважнији следећи: рус. *фразеосхеме* (уп. ШМЕЛЕВ 1977), нем. *Phraseoschablonen* (уп. FLEISCHER 1982), енгл. *constructional idioms* (уп. BOOIJ 2022; MELLADO BLANCO, у штампи), *syntactic idioms* (уп. MEL'ČUK 2022), *semi-preconstructed phrases* (WRAY 2002: 7). Филери морају задовољити специфичне критеријуме (семантичке, граматичке, прозодијске), који се морају одредити за сваку појединачну ФК<sup>1</sup>. ФК се дакле разликују и од варијанти класичних фразема, јер могући филери нису ограничени на одређени скуп могућности<sup>2</sup>. Између ФК и фразеологизама постоји динамичан однос јер најчесталије реализације ФК могу постати идиоми, а многе ФК потичу из варијанти идиома (MELLADO BLANCO 2023).

COST акција 22115 PhraConRep посвећена је контрастивном истраживању ФК у немачком језику као изворном и у једанаест словенских језика (босански, бугарски, хрватски, чешки, македонски, пољски, руски, словачки, словеначки, српски и украјински), те у мађарском, румунском и албанском као циљним језицима. Други циљеви COST акције укључују допринос теоријском описивању појма ФК, типолошке анализе ФК у сродним и несродним европским језицима, развијање метода дигиталне лексикографије, те развијање дидактике проучавања ФК у страним језицима, пре свега у немачком и другим пројекатским језицима.

Бележење јединица као што су ФК представља посебан изазов за лексикографски рад (FILIPOVIĆ PETROVIĆ 2018: 83). Следећи примери<sup>3</sup> илуструју реализације ФК у немачком и њихове преводе на српски језик помоћу српских ФК. У угластим заградама су дате

<sup>1</sup> Детаљан опис карактеристика ФК може се наћи у Шлунд – Павлова (SCHLUND – PAVLOVA, у штампи).

<sup>2</sup> За дискусију о опису варијанти фразема у речницима, в. Filipović Petrović (2018: 88–90).

<sup>3</sup> Примери и наводи који се односе на ФК у овом подељку преузети су из самог репозиторијума, који се налази на платформи *Github*, али још увек није јавно доступан. У случају примера из белетристике дају се и аутор и назив дела. Извори (нпр. корпуси, линкови) дати су само за примере који не потичу из репозиторијума.

апстрактне схеме (енгл. *patterns*).

- (1) а. *Du bist mir ja ein schöner Gentleman!* [ФК: PERS.PRON ist mir ja ein schön- N<sub>NOM</sub>]<sup>4</sup>  
б. *И ти си ми неки џентлмен!* [ФК: и PERS.PRON ми је нек- N<sub>NOM</sub>]
- (2) а. *Der Mann, der auf dem Richtplatz stand, war die Unschuld in Person.* (SÜSKIND 1985: 285) [ФК: DET N<sub>NOM</sub> in Person]  
б. Човек који је стајао на губилишту био је *оличење невиности*. [ФК: оличење N<sub>GEN</sub>] (ZISKIND 2001)
- (3) а. *Netflix kann einpacken:* Hier gibt es das beste Streaming-Angebot des Jahres. [ФК: N<sub>NOM</sub>/PERS.PRON kann einpacken]  
б. *Нејџфликс може да се пакује:* Овде се налази програмска понуда године. [ФК: N<sub>NOM</sub>/PERS.PRON може да се пакује];
- (4) а. *Das ist so was von bescheuert!* [ФК: so was von ADV/ADV]  
б. Ово је *лудости неопевана*. [ФК: N<sub>NOM</sub> неопеван-]

С појавом алата за превођење заснованих на вештачкој интелигенцији, наде су биле велике да ће превођење формулаичних израза такође бити олакшано. Међутим, у области ФК преводилачки софтвери често наилазе на ограничења. Ово посебно важи за мање језике, попут оних који су у фокусу пројекта PhraConRep, а од којих неки нису ни представљени на платформама као што су *DeepL*. Међутим, чак и када је реч о преводима на светске језике попут енглеског и руског, програми подржани вештачком интелигенцијом као што су *DeepL* или *ChatGPT* и даље могу давати искривљене предлоге превода за текстове у којима се појављују ФК. На пример, веома учестала руска ФК *хоть (не) им* погрешно је преведена софтвером *DeepL* у следећим преводима на енглески и немачки:

- (5) Давно желал с тобом встречу иметь. А дальше — хоть голова с плеч! (Руски национални корпус: Борис Евсеев, *Евстиней*, 2010)<sup>5</sup>

‘Давно сам хтео да се састанем с тобом. А даље – могу да ми скину и главу са рамена.’

а. Енгл.: I've long wished to meet you. *And then I'll be head over heels!* (*DeepL*, 8. 8.2024)

б. Нем.: Ich habe mir schon lange gewünscht, dich kennenzulernen. *Und dann werde ich mich Hals über Kopf verlieben!* (*DeepL*, 8. 8. 2024)

Утисак је да софтвер једноставно измишља значење оригинала. Пошто се речи *лава* и *раме* појављују у руском тексту, у енглеском и немачком преводу користе се фразеологизми са сличним лексичким компонентама, при чему оба значе „лудо се заљубити”, а не „брига ме шта ће бити даље, па макар ми и главу скинули са рамена”, што би био тачан превод руске ФК.

<sup>4</sup> ФК често имају неколико варијанти због опционих филера и ситних варијација речи сидра. Овде смо се одлучили за једноставнију, репрезентативну варијанту, како би се олакшала читљивост.

<sup>5</sup> В. Савчук и др. (2024).

Циљ овог чланка је да пружи опис тренутног стања израде вишејезичног репозиторијума ФК. Након кратке презентације двају теоријских модела на којима је пројект заснован (одељак 2), у трећем одељку преставља се макроструктура репозиторијума и могућности оријентисања корисника у речнику. У четвртном одељку описује се микроструктура лематских описа. У закључку се осврћемо на нека још нерешена питања и перспективе за даљи рад.

## 2. Теоријска оријентација

Аналогно појму „лексикографија”, „конструктографија” означава праксу развоја (пре свега, дигиталних) ресурса у којима се не бележе лексеме или фразеологизми, већ се инвентаришу и описују целокупне конструкције према начелима конструкционе граматике (LYNGFELT 2018). У зависности од циља, конструктографски речници описују одређене врсте конструкција, док други обухватају целокупни фонд конструкција појединог језика. Као резултат настаје тзв. „конструктикон” као структурисани опис (одређених) конструкција датог језика. Иако се конструктикони могу конципирати тако да обухвате све врсте конструкција датог језика, од најапстрактнијих и схематских, до лексички потпуно специфичних, у пракси се већина конструктикона посвећује конструкцијама на средњем нивоу схематичности. Један од централних елемената конструктографске праксе представља емпиријско утемељење описа конструкција на основу лингвистичких корпуса (VOAS 2019). Инспириран овим концептом, циљ нашег пројекта је израда вишејезичног конструктикона таквих конструкција, то јест ФК. Конкретан и драгоцен узор за наш рад чини конструктикон руског језика (*русский конструктикон*, BAST et al. 2021), у којему су укључене и ФК и друге конструкције.

Други лексикографски модел који је инспирисао чланове пројекта представља концепт „активног речника” Московске семантичке школе (уп. APRESJAN 2024). Активни речник омогућава корисницима не само да стекну пасивно знање, већ и да активно користе леме описане у речнику. Зато лексикографска обрада ФК у репозиторијуму иде даље од обичног семантичког објашњења, ослањајући се на искуства Московске семантичке школе у њиховом раду на *Активном речнику руског језика* (Апресјан 2014). Речнички чланци таквог речника не укључују само детаљне семантичке и граматичке, већ и прагматичке податке, уз аутентичне примере употребе. Док је *Активни речник* Московске семантичке школе једнојезичан и није посебно посвећен ФК, наш репозиторијум располагаће детаљним описима семантике, прагматике, морфологије, синтаксе и стила ФК

на немачком језику и пружати аотиране преводе на друге, горенаведене језике.

## 3. Општа својства и функције репозиторијума

Репозиторијум је намењен лингвистички обученој и заинтересованој публици, као што су студенти лингвистике, лингвисти, школски и универзитетски наставници језика или професионални преводиоци (PAVLOVA 2022: 596). Објашњења су дата на лингвистичко-стручном језику, који додуше не захтева специјална знања одређене синтаксичке школе. Метајезик на којем се описују леме јесте немачки.

Инвентар немачких ФК већ више од пет година утврђује се бележењем немачких ФК из различитих врста писаних текстова, али и из усмених извора (нпр. филмови, лична комуникација) (PAVLOVA 2022: 601). Тренутни циљ је забележити и превести укупно 300 немачких ФК. У овом подухвату помаже и веома напредан, али још увек јавно недоступан двојезични репозиторијум са руским као изворним и немачким као циљним језиком.

Корисници могу у нашем репозиторијуму пронаћи ФК на разне начине. Прва могућност је да се ФК нађу по апстрактној схеми (тј. конструкцији), која укључује само речи на месту сидра и морфолошки опис филера. Такав уопштени запис лако може да постане доста комплексан, зато што елементи који се налазе на месту сидра често нису ограничени на одређен облик конкретне речи, него се могу мењати, на пример, по лицу и броју када је у питању лична заменица. Са друге стране, понекад је класа речи које могу попуњавати слот у ФК тако ограничена да би се филери могли схватити и као варијанте сидра. Граница између елемената који стоје на месту сидра и филера који попуњавају слотове у таквим случајевима постаје замагљена, а елементи многих ФК налазе се заправо у прелазној зони између речи које представљају сидро и филера (SCHLUND – PAVLOVA, у штампи). Пример ФК са речима разних нивоа флексибилности које престављају сидро ФК је нем. *Du bist mir ein schön- N<sub>NOM</sub>* ‘ти си ми нек- *N<sub>NOM</sub>*’ (в. пример (1)), у којој ниједан елемент сидра није везан за само један облик реализације. Заменица у номинативу не може се заменити само заменицама других лица у једнини и множини, него и именичним фразама, а нарочито фразама са властитим именима. Копула се, разуме се, исто мења у складу са субјектом. Заменица у етичком дативу, иако углавном ограничена на прво лице једнине (*mir*), може да се изостави, а уместо ње (или заједно с њом) типично се појављује једна или чак и више партикула (нпр. *ja*, *vielleicht*, *mal*). Наравно, и придеви атрибуту испред именичког филера варирају, а *schön-* ‘леп’ само је један



од најчешће употребљених, поред других придева који имају значење позитивне оцене. Такође, може се десити да се атрибут сасвим изостави. У таквим случајевима, мора се пронаћи компромис између апстракције и читљивости. Запис дате ФК могао би да изгледа овако:

das/PERS.PRON/ $N_{\text{NOM}}$  / KOP.PRES. [mir] [ja/vielleicht] ein(e) [schön-/nett-/toll-/fein-]  $N_{\text{NOM}}$ !

Подаци о лексичким ограничењима филера могу се додати у витичастим заградама. У датом примеру, могло би се додати да придев на месту атрибута мора да има позитивно-евалутивну семантику: {+evaluation}. Каталог скраћеница за таква ограничења још је у изради и биће укључен у метаинформације о речнику. Као што горњи пример показује, питање нивоа апстракције од кључног је значаја за бележење лема. Тежи се средњем нивоу апстракције, који јасно разликује речи што представљају сидра и филере, а који је је истовремено прилагођен корисницима, те стога не садржи превише апстрактне нотације (MELLADO BLANCO – STEYER 2018: 266).

Друга је могућност да се ФК нађе по конкретној реализацији (тј. по једном одређеном конструкту). Ово ће функционисати пре свега код ФК у којима су неки филери нарочито чести<sup>6</sup>. Чест конструкт ФК дат је испод уопштеног апстрактног записа (в. пододељак 4.1).

Трећа могућност за проналажење ФК је слободно претраживање конкретних облика речи и фраза у целом репозиторијуму. На овај начин може се претраживати и на једном од других језика представљених у речнику, док су прве две могућности доступне само за претрагу на немачком језику.

Поред тога, могу се вршити и ономасиолошка претраживања, према семантичким особинама, комуникативној сврси и илокуцијама које ФК изражава. Стога је такође могуће идентификовати групе функционално сличних ФК.

#### 4. Микроструктура речничких чланака

Речнички чланак сваке ФК у репозиторијуму састоји се од одређеног броја елемената с посебним ознакама (енгл. *tags*), који описују различите аспекте конструкције – семантичке, синтаксичке, прозодијске и друге. Најважније ознаке за опис фраземске конструкције чине: (1) функција <function><sup>7</sup>, (2) лема <lemma>, (3) значење <meaning>, (4) примери за лему <examples for lemma> (5) семантичка

ограничења <semantic restrictions>, (6) морфологија <morphology>, (7) синтакса <syntax>, (8) употреба <usage>, (9) стил <style>, (10) прозодија <prosody>, (11) преводи примера на појединачне језике <example translation>, (12) коментари уз превод на појединачне језике <general translation comment>, (13) синоними <synonyms> и <full synonyms>, (14) хомоними <homonyms> и (15) пароними <paronyms>.

У оквиру пројекта настаје велики број појединачних студија о ФК у немачком језику, врло често у поређењу с најмање једним циљним језиком. Уколико такве студије постоје, биће наведене на крају описа леме и, ако је могуће, директно повезане.

#### 4.1. Функција, лема, значење, семантичка ограничења и примери

Под функцијом ФК подразумевају се семантичка поља која карактеришу значење ФК. Појам семантичког поља користи се у широком смислу, обухватајући семантичке, прагматичке и дискурзивне димензије. Према томе, под ознаком *функција* бележе се илокуције (нпр. упозорење, утеха), дискурзивне функције (нпр. топикализација, коментар), модална значења (нпр. (не-)вероватност, (не-)жељеност), логична значења (нпр. разлог и последица), те и лексичка значења (нпр. изобиље, дуго трајање) (PAVLOVA 2024: 160–161). Код ФК *Du bist mir ein schön-  $N_{\text{NOM}}$*  ('Ти си ми нек-  $N_{\text{NOM}}$ '), ради се о говорном чину процене.

Лема представља уопштени запис ФК, при чему се користе симболи као што су (1) велико слово *икс* ( $X$ ), за филер чија се врста речи не може унапред дефинисати; (2) апостроф (') код првог и другог филера, уколико је у питању иста реч у оба случаја или (3) затворени знаци навода (") код другог филера, уколико се он разликује од првог; (4) симболи за различите врсте речи, као нпр.  $N$  за именицу,  $V$  за глагол, *Pron* (нем. *Pronomen*) за заменицу итд.; (5) угласте заграде ( [ ] ) за факултативне елементе ФК; као и (6) коса црта (/) за могуће варијанте у оквиру једне ФК. На пример, у ФК  $X'_{\text{NOM}}$  *bleib(s)t/bleibe/bleiben*  $X'_{\text{NOM}}$  (' $X'_{\text{NOM}}$  остаје/остају  $X'_{\text{NOM}}$ '), чија је једна од могућих реализација, односно конструкт, *Räuber bleibt Räuber* ('лопов остаје лопов'), реч *Räuber* понавља се као филер и представљена је симболима  $X'$  на оба места зато што врста речи филера није ограничена на именице. Опис леме укључује и једну типичну реализацију дате ФК, тј. типичан конструкт, одмах испод апстрактног описа конструкције. У идеалном случају, типична реализација дата у репозиторијуму подударара се с најфреквентнијом реализацијом (тј. најчешћим конструктом). Најчешћи конструкт утврђује се помоћу претрага у корпусима. Осим немачких корпуса на платформи *SketchEngine*

<sup>6</sup> Често коришћени типови једне ФК одражававају високу учесталост типова (енгл. *type frequency*) и могу постати идиоми (MELLADO BLANCO, у штампи).

<sup>7</sup> У оштрим заградама су наведени називи елемената на енглеском језику.

(KILGARRIFF et al. 2004; KILGARRIFF et al. 2014), за анализе фреквенције највише се користи корпус *DeReKo* (Deutsches Referenzkorpus – Немачки референтни корпус) (KUPIETZ et al. 2018). На пример, за ФК  $N'_{pl}$  *über N'\_{pl}* испоставило се да је најчешћа реализација *Fragen über Fragen* ‘веома много питања’ у немачким корпусима на платформи *Sketch Engine*. Међутим, за неке ФК није могуће поуздано утврдити најчешће филере, на пример, ако је ФК врло ретка или ако постоји много конкурената са честим филерима који су по броју међусобно веома слични. У таквим случајевима бира се типичан пример који што краће и разумљивије илуструје одговарајућу конструкцију. У случају ФК  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben X'\_{NOM}* ‘ $X'_{NOM}$  остаје/остају  $X'_{NOM}$ ’, један такав илустративни пример може бити *Kind bleibt Kind* ‘дете остаје дете’ као кратка реализација са лако разумљивом лексиком и за чије разумевање није потребан контекст.

Трећа ознака, ознака <meaning> уводи опис општег значења ФК. Уколико је то неопходно за разумевање семантике, овде се такође разматра и дискурзивна ситуација у којој се дата ФК може користити. Неке ФК, на пример, користе се као реплика претходној изјави саговорника. ФК *selber X* ‘сам-си/сте  $X$ ’, односно њена реализација *selber doof* ‘сам си глуп’) функционише као усклик након критике или увреде од стране саговорника у дијалозима. Ако се разликује више значења за неку ФК, прво следи општи опис укупног значења, а затим различита подзначења, заједно с примерима из аутентичног материјала, без превода на друге језике. Примери за лему потичу из различитих извора – књижевних дела писаца с немачког говорног подручја, песама, дигиталних издања дневне и недељне штампе, интернет-страница радија, телевизије, блогова, као и других извора. Под ознаком *семантичка ојраничења* наводе се, на пример, лексичке класе речи које се не могу јавити на месту филера дате ФК.

#### 4.2. Морфологија и синтакса

Код морфолошких одлика ФК, описане су морфолошке карактеристике филера, а наводе се и примери употребе одређеног или нултог члана, уколико је то случај код те ФК. Код филера је одређена врста речи филера, падеж и врста именице, као на пример *власниша*, *заједничка*, уколико је у питању именица, као и време или начин код глагола. У ФК  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben X'\_{NOM}* ‘ $X'_{NOM}$  остаје/остају  $X'_{NOM}$ ’, као оба филера превасходно се јавља *именица у номинативу која се йонавља*, а у појединачним случајевима може се јавити и *йридегв*, *инфинитив* или *лично име*. Уколико се ради о *ајсйрактиним именицама* или приликом употребе која није рефе-

ренцијална, члан се елиминише, као код реализација *Kind bleibt Kind* ‘Дете остаје дете’), примера за други случај. Некада члан остаје само код прве именице, као код реализације *Ebbe und Flut wechseln, aber das Meer bleibt Meer* ‘Плима и осека се мењају, али море остаје море’), а неодређени члан може се јавити код обе именице, као, нпр., у конструкту *Ein Esel bleibt ein Esel* ‘Магарац остаје магарац’).

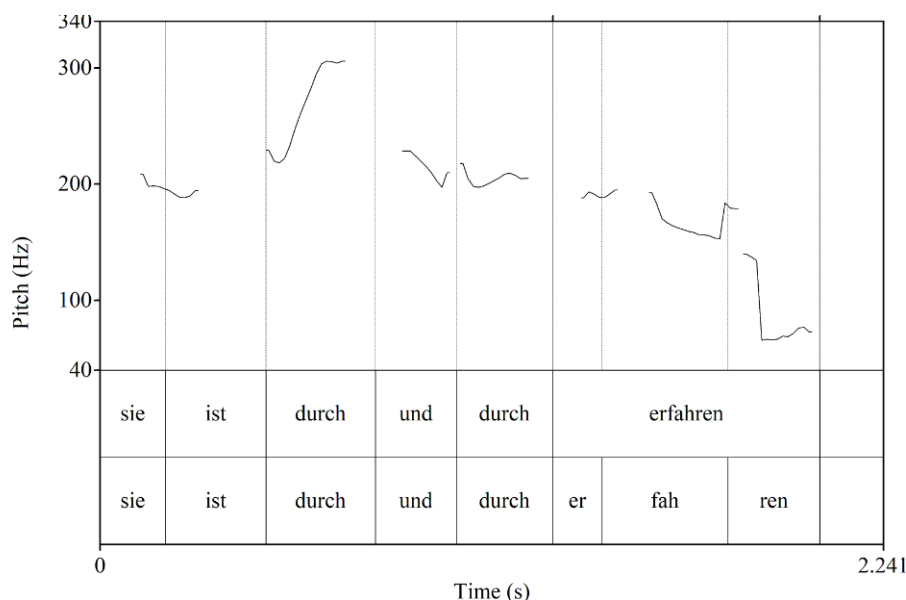
Синтаксичке одлике ФК наводе се на нивоу целе конструкције, али и на нивоу њених саставних филера. ФК може, примера ради, бити самостална реченица или као целина вршити одређену функцију у реченици. ФК  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben X'\_{NOM}* ‘ $X'_{NOM}$  остаје/остају  $X'_{NOM}$ ’, на пример, може бити самостална реченица или независна реченица унутар низа реченица или реченичне структуре. Што се тиче њених филера, први филер је субјекат реченице, а други заједно с глаголом чини њен предикат.

#### 4.3. Употреба, стил и прозодија

Под ознаком <usage> подразумевају се комуникативне ситуације, врста дискурса и/или врсте текста у којима се дата ФК најчешће користи, као на пример у дијалозима, у наративу, у усменом и/или писаном дискурсу, у публицистичким текстовима и другим. Наглашава се и где се одређена ФК не може наћи, као на пример у научним, правним и другим стручним текстовима, у случају ФК  $N'_{GEN PL}$  *der N'\_{GEN PL}* ‘ $N'_{GEN PL}$  над  $N'_{GEN PL}$ ’, чија је једна од реализација *Ring der Ringe* ‘прстен над прстеновима’.

Када је у питању стил неке ФК, он представља њену припадност одређеном регистру, на пример (1) неутралном, (2) неформалном или (3) узвишеном. Уз то се дају напомене и о томе да ли је конструкција шалвива, иронична, реторичка, филозофска, књижевна и слично. Код примера  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben X'\_{NOM}* ‘ $X'_{NOM}$  остаје/остају  $X'_{NOM}$ ’ ради се о неутралном до узвишеном стилу зато што је ФК реторички и филозофски обојена.

Код неких ФК, акцентовање конструкције, односно наглашавање одређене речи или одређених речи игра веома важну улогу у преношењу поруке. У примеру *Ein Esel bleibt ein Esel* ‘Магарац остаје магарац’), ради се о *консйрукцији с двосйруким акцентом*, што значи да су оба филера *најлашена*, и то *изразишо*, како би се истакла реч *Esel* ‘магарац’) на оба места, а самим тим и особине које та животиња има, односно порука која треба да се пренесе употребом ове конструкције. Из тог разлога, како би што боље и пријемчивије био приказан акценат одређене конструкције, у репозиторијуму ће бити дат и звучни запис и визуализација звучног записа сваке конструкције.



Приказ 1. Визуализација звучног записа немачке ФК *durch und durch* 'сасвим/скроз/изузетно X' помоћу конструкта *sie ist durch und durch erfahren* 'она је изузетно искусна'.

#### 4.4. Синоними и преводи

ФК могу бити и синоним једна другој, те се стога као такве наводе у речничком чланку у репозиторијуму. Притом се обично не ради о апсолутној еквиваленцији, него о делимичној еквиваленцији различитог степена. У случају ФК  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben*  $X'_{NOM}$  (' $X'_{NOM}$  остаје/остају  $X'_{NOM}$ ', конструкције које долазе у обзир за синониме су (1) ФК  $X' ist X'$  ('X је X'): *Kind ist Kind* ('дете је дете'), (2) ФК *Einmal X', immer X'* ('једном X', увек X'): *Einmal Richter, immer Richter* ('Једном судија, увек судија') и (3) ФК  $X' ist und bleibt X'$  ('X' јесте и остаје X'): *ein Dieb ist und bleibt ein Dieb* ('лопов јесте и остаје лопов'). У овом једнојезичном делу репозиторијума, дати синоними искључиво су ФК, а не лексеме или фразеологизми. Поред ознаке <synonym> постоји и ознака <full synonym> 'потпуни синоним', који се, додуше, само ретко користи. У случају ФК  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben*  $X'_{NOM}$  једино ФК  $X' ist X'$  ('X' је X') може се сматрати потпуним синонимом. Обе ФК структурно су идентичне и чини се да су и заменљиве у свим контекстима. Чак и овај редак случај потпуне синонимије може се, међутим, разрешити ако се узме у обзир да се заправо ради о једној ФК са две различите речи на месту сидра (*bleiben* 'остати' и *sein* 'бити'). Друге синонимне ФК могу заменити лему ФК тек у неким случајевима. Тиме долазимо и до проблематике превођења немачких ФК на циљне језике.

Приликом утврђивања преводилачких еквивалената, пројект је такође посвећен емпиријском приступу, јер се на тај начин могу циљано утврдити функционални еквиваленти (MELLADO BLANCO – STEYER 2018). Појам функционалног еквивалента значи да је као централни критеријум за утврђивање међујезичких

еквивалената важна њихова функционална и комуникативна подударност, утврђена на основу аутентичних језичких података (ĐURČO 2018; MELLADO BLANCO – STEYER 2018). При томе, идеално би било када би се добио прототипичан еквивалент у циљном језику – еквивалент који се може користити у што више контекста у оба језика (ĐURČO 2018: 286–288). Прототипичан еквивалент може бити структурно и лексички сличан изворном језику, али и не мора да буде.

Ако се у циљном језику нађе еквивалентна ФК чија структура, семантика и функција у највећој мери подсећају на ФК из изворног језика, онда је реч о ретком случају потпуне еквиваленције (ČAGALJ – SVÍTKOVÁ 2014: 34). За ФК то значи да се о потпуној еквиваленцији може говорити само ако у оба језика постоји ФК. Поред захтева за семантичком и структурном сличношћу, морају бити упоредиви и учесталост употребе, стилски ниво и могућности варијације код филера како би се две ФК могле сматрати потпуним еквивалентима (PAVLOVA 2024: 167–169). У случају одступања у погледу структуре, функције или семантике, реч је само о делимичној еквиваленцији (ČAGALJ – SVÍTKOVÁ 2014: 34–36; PAVLOVA 2024: 170–171). Кандидати за могућу потпуну еквиваленцију су ФК  $X'_{NOM}$  *bleib(s)t/bleibe/bleiben*  $X'_{NOM}$  (' $X'_{NOM}$  остаје/остају  $X'_{NOM}$ ') и ФК  $X' ist X'$  ('X је X'), јер су структурно идентични и подударују се у семантичком и прагматичком погледу. ФК *Einmal X', immer X'* ('једном X', увек X'), међутим, није им потпуни еквивалент, зато што она укључује процену (често, али не увек, негативну), која није својствена другим ФК.

До еквивалената ФК може се доћи на неколико различитих начина, у зависности од доступности ресурса



са: (1) претрагом у паралелним корпусима немачког и циљног језика; (2) у преводима књижевних дела с немачког говорног подручја на циљни језик, ручно или уз помоћ паралелних корпуса књижевних дела, или (3) самосталним преводом од стране учесника пројекта. Све ове стратегије имају своје предности и мане.

Приликом потраге за еквивалентима у паралелним корпусима изворног и циљног језика, углавном се бира онај еквивалент у циљном језику који се најчешће користи у паралелном корпусу и који је најпрототипичнији (MELLADO BLANCO – STEYER 2018: 266f). При изради репозиторијума, чланови пројекта користе различите паралелне корпусе, као што су корпуси на платформи *Sketch Engine* (KILGARRIFF et al. 2004; KILGARRIFF et al. 2014), као и паралелни корпуси на платформи *InterCorp* (ČERMÁK – ROSEN 2012; ROSEN 2016; ČERMÁK 2019; ROSEN et al. 2022), која је настала као део пројекта *Чешкој националној корпусу*. За велики број циљних језика пројекта, међу којима је и српски, на платформи *Sketch Engine* може се користити *OpenSubtitles* (LISON – TIEDEMANN 2016), збирка паралелних корпуса филмских титлова, која се састоји од 60 корпуса на 58 језика<sup>8</sup>. *InterCorp* садржи паралелне текстове на чешком језику и 60 циљних језика, међу којима су и српски и други циљни језици пројекта. Текстови потичу из различитих извора, као што су књижевна дела, медији или титлови за филм. За разлику од већих циљних језика пројекта, српски је међу оним језицима за које још увек не постоји велики број паралелних корпуса, те се најчешће користе ова два корпуса. Претрага у овим корпусима врши се помоћу опције *Parallel Concordance*. Добијени резултати веома су разноврсни у погледу тачности, с обзиром на то да долази до грешака приликом паралелизације корпуса, као и начина на који је ФК из изворног језика преведена на циљни језик и морају се даље ручно анализирати, како би се дошло до еквивалената. У репозиторијум се првенствено уносе они преводи који представљају ФК у циљном језику која је потпуни еквивалент ФК у изворном језику. Тако је еквивалент немачке ФК *X der X*, односно њене реализације *König der Könige*, у српском језику такође ФК, и то *краљ краљева*, што се закључује на основу претраге у корпусу *Open Parallel Corpus (Opus)* на платформи *Sketch Engine*. С обзиром на то да се и у изворном и у циљном језику ради о ФК која се састоји од филера у номинативу и филера у генитиву, са значењем „најбољи/најбоља/најбоље или најгори/најгора/најгоре у групи”, која се користи у сврху појачавања исказа, може се рећи да се ради о потпуној еквиваленцији.

Пошто је репозиторијум замишљен као ресурс за професионалне преводиоце и за научнике у области превођења, поред прототипичних еквивалената, инвентаришу се и посебни случајеви, па понекад чак и погрешни преводи који служе за илустративне сврхе. Такође су понекад забележени случајеви у којима превод одговара конкретној употреби неког језичког конструкта, иако се не може лако пренети на друге контексте. Овакви случајеви не би били узети у обзир само на основу прототипског приступа, али пружају релевантне информације о типу еквиваленције одређених ФК.

Ручно претраживање превода састоји се од проналажења конструкције у оригиналном делу, а затим и њеног еквивалента у преводу. То је процес који, услед недостатка претраживих дигиталних издања, може захтевати већи труд и уложено време од других начина потраге за еквивалентима. Претрага у књижевним преводима може се вршити и у паралелним корпусима. Преводи књижевних дела са немачког говорног подручја на циљне језике могу бити део општих паралелних корпуса који се састоје од различитих извора, као што је то случај са корпусом *InterCorp*, или могу сачињавати посебан корпус, који обухвата паралелне текстове књижевних дела и њихових превода на циљни језик, као што је то случај са српско-немачким литерарним корпусом *SrpNemKor* (ANDONOVSKI 2019) у оквиру алата *Bibliša* (STANKOVIĆ et al. 2015). Проблем који се јавља приликом употребе корпуса за потрагу за преводима ФК у књижевним делима јесте недовољна заступљеност паралелних корпуса српског језика који садрже књижевна дела, као и мали број књижевних дела који је њима обухваћен. Предност коришћења превода из корпуса који садрже различите изворе, у односу на преводе из књижевних дела, јесте савременост језика, с обзиром на то да се употреба неких ФК у међувремену променила, уколико се ради о старијим књижевним делима. Са друге стране, превод књижевних дела често је далеко квалитетнији од оног у корпусу. Међутим, многа дела с немачког говорног подручја нису преведена на све циљне језике пројекта.

Добар пример делимичног еквивалента немачке ФК *X der X*, односно њене реализације *Ring der Ringe*, може се пронаћи у делу *Тако је говорио Заратустра* Фридриха Ничеа, у преводу Бранимира Живојиновића из 2007. године, у издању издавачке куће *Feniks Libris*, као што се види у примеру (6), који представља оригинал на немачком језику, и примеру (6.а), преводу на српски језик.

- (6) Oh wie sollte ich nicht nach der Ewigkeit brünstig sein und nach dem hochzeitlichen *Ring der Ringe*, — dem Ring der Wiederkunft! (NIETZSCHE 1884: 112)

<sup>8</sup> Доступно на:

<<https://www.sketchengine.eu/opensubtitles-parallel-corpora/>>.

а. Ох, како да страсно не жудим за вечношћу и за свадбеним *ірсиѐном над ирсиѐновима* – прстеном враћања! (Ниче 2007: 247)

С обзиром на то да у циљне кориснике репозиторијума спадају и преводиоци и лингвисти који се баве теоријом превођења, у репозиторијум могу бити уврштени и различити преводи истог књижевног дела, како би се и они могли међусобно упоређивати, јер превод заправо представља једно од могућих решења за које се тај преводилац одлучио. Међутим, у неким случајевима, нарочито код класика, неретко долази до преузимања истог превода у више издања и од стране различитих издавача. На пример, то је случај с романом *Jagi mlađoia Veriјера* Ј. Волфганга фон Гетеа, који је са немачког на српски језик превео С. Винавер.

Превод ФК оригинала датог у примеру (7), који је унет у репозиторијум, превод је С. Винавера, који можемо видети у примеру (7а).

- (7) “Mein Freund”, rief ich aus, “*der Mensch ist Mensch*, und das bißchen Verstand, das einer haben mag, kommt wenig oder nicht in Anschlag, wenn Leidenschaft wütet und die Grenzen der Menschheit einen drängen.” (ГОЕТНЕ 2016: 36) [наглашавање наше – К. И. и К. Ш.]

а. Мој пријатељу! – узвикнем ја на то – *човек је човек*, и оно мало разума што коме допадне у део, слабо долази у обзир или баш и ни мало, кад се разбесни страст а притесне нас међе и границе човечанске! (ГЕТЕ 1949: 62)

Уместо употребом паралелних корпуса и књижевних дела, чланови пројекта до превода могу доћи и самостално, или уз употребу вештачке интелигенције. На пример, у случају да се ФК јавља у извору који нема директан превод, било да се ради о непреведеном књижевном делу или извору којег нема у паралелном корпусу, чланови пројекта могу самостално преводити примере с немачког језика. Приликом самосталног превођења ФК, на основу бројних примера у којима се он јавља у немачком језику, сарадник на пројекту ослања се на своју билингвалну компетенцију, с обзиром на то да се по правилу преводи на матерњи језик. Како би се проверило да ли је превод ФК који члан пројекта предлаже заиста еквивалент, врши се претрага у корпусу циљног језика или интернет-претраживачу као што је *Гуил* и посматрају се контексти у којима се он користи. Еквиваленција у циљном језику мора се, дакле, проверити уз помоћ претраге у корпусу циљног језика (MELLADO BLANCO – STEYER 2018: 266–267). Вештачка интелигенција може послужити као извор идеја на основу којих се члан пројекта одлучује за адекватан превод. У случају самосталног превођења уз употребу вештачке интелигенције или без ње, с преводом се морају сложити сви чланови пројекта који превод на дати циљни језик, како би се избегла

субјективност приликом доношења одлуке који преводи су адекватни и улазе у репозиторијум. Пример за самостални превод може бити превод ФК *X ist das neue Y*, која се јавља у наслову књиге Лоте Кин *Sechzig ist das neue Vierzig: Selbst-Entfaltung für die Frau in den besten Jahren*. Иако ова књига није преведена на српски језик, ФК која се налази у наслову може се самостално превести уз помоћ ФК *szeszeseite су нове чеијгесезеи*. Једноставном претрагом помоћу интернет-претраживача *Гуил* установљено је да ова ФК структурно, семантички и функционално одговара немачкој ФК, те да се може сматрати њеним потпуним еквивалентом.

Осим разлике између потпуне и делимичне еквиваленције, при изради репозиторијума важно је узети у обзир то да је еквиваленција којој теже лексикографи негде између еквиваленције на нивоу система и на нивоу текста. Долазак до еквивалената путем претраге у корпусу базира се на еквиваленцији на нивоу текста, на нивоу *живе речи*<sup>9</sup>, односно *parole* (MELLADO BLANCO – STEYER 2018: 266). Приликом израде репозиторијума, до тог нивоа између системског нивоа и нивоа текста долазимо управо тако што анализирамо квалитет превода из корпуса и дајемо своје предлоге адекватних превода.

У репозиторијуму, уз сам пример и превод, наводи се и извор одакле потиче, уз библиографске податке, односно стратегија доласка до тог превода, као на пример *самостални превод, превод уз помоћ вештачке интелигенције* или *Sketch Engine*. С обзиром на то да је један од главних циљева пројекта опис ФК у циљним језицима, у репозиторијум се уноси и коментар уз превод, у коме се наводи да ли је могуће превести ФК уз помоћ ФК у циљном језику, да ли је у питању потпуна или делимична еквиваленција, као и структурне, семантичке или функционалне појединости о еквиваленту или еквивалентима.

#### 4.5. Хомоними и пароними

Хомоними фраземских конструкција, као и хомонимне речи, имају исти гласовни склоп, али различита значења. За разлику од наведених синонима, који су и сами ФК, хомоними обично нису идиоматске, него слободне комбинације речи. Често је у питању валентност глагола или именице, што значи да хомоними уз себе захтевају одређен морфолошки облик, који се пак јавља и у хомонимној ФК. Добрих примера за хомонимију има код ФК *N' über N' PL* и њене реализације *Regeln über Regeln*, чије је фразеолошко значење ‘веома много правила’. Нефразеолошко је значење истог облика буквално, просторно

<sup>9</sup> Овај превод за термин *parole* употребила је Р. Драгићевић (2018).



значење, тј. ‘правила изнад правила’. У следећем случају, хомоним ове ФК заправо је настао услед валентности глагола *herrschen* ‘владати’, који такође захтева предлог *über* ‘над’ и именицу у дативу: *Hier herrschen Menschen über Menschen* (‘Овде владају људи над људима’).

За исту ФК може се навести и пример за пароним. За разлику од хомонима, пароними се разликују од дате ФК не само по значењу и, понекад, по прозодији, него и по структури. Пример попут *Gesetze über Gesetzen* је један од ретких случајева паронима. Разлог за то лежи у чињеници што *Gesetze über Gesetzen* није реализација ФК, него има буквално, просторно значење ‘закони над законима’. Форма ФК била би *Gesetze über Gesetze*, с фразеолошким значењем ‘веома много закона’. Разлог за формалну разлику је у томе што се датив именице *Gesetz* ‘закон’, који мора следити након предлога *über* ‘изнад’ у тој буквалној употреби, разликује од других падежа (*Gesetze*<sub>NOM/ACC/GEN.PL.</sub> vs. *Gesetzen*<sub>DAT.PL.</sub>).

Често се хомоними разликују од ФК и по прозодији. ФК *so was von* *ADJ./ADV.* [+*PAST.PART*]/*NP* означава висок степен неке особине (нпр. *so was von bescheuert* ‘изузетно глуп’ или *so was von Mainstream* ‘апсолутно мејнстрим’). У фразеолошком значењу, реч-сидро *so* носи главни акценат, што у нефразеолошким употребама низа речи *so was von* + *NP* није случај:

- (8) Und wer redet bei *so was von Mord*? Es war ein Unfall. (REMARQUE 1989: 190)

‘А ко би у таквом случају говорио о убиству? То је била несрећа.’ (самостални превод)

Морфосинтакса примера није условљена фразеолошки, већ произлази из приложне ознаке *bei so was* и рекције глагола *reden über* + датив ‘говорити о’. За разлику од ФК, главни акценат у неутралном контексту лежи на именици *Mord* ‘убиство’, а партикула *so* може носити други акценат, али и не мора.

## 5. Закључна разматрања

Рад на вишејезичном репозиторијуму лексички само делимично специфицираних јединица, као што су ФК, представља лексикографски изазов у многим аспектима. То се посебно односи на разликовање сидра и филера при нотацији лема. До сада су аутори репозиторијума углавном поступали интуитивно у овом погледу. У будућности ће бити потребно одредити максималан број варијанти које могу да имају речи на месту сидра. Уколико број варијанти коју има нека реч која стоји на месту сидра прелази тај праг, сматраће се филером. Код утврђивања типичних реализација чини се да концепт прототипа достиже своје границе (ĐURČO 2018: 286–288). Ако се под

прототипом подразумева најчешћа реализација неке ФК, онда се прототип односи на реализацију типа са највећом фреквенцијом токена. Као што је поменуто, најчешћи филери често се не могу емпиријски утврдити, јер велики број варијација могућих филера представља једну од конститутивних особина ФК (SCHLUND – PAVLOVA, у штампи; MELLADO BLANCO, у штампи). Другим речима, типичне ФК управо су окарактерисане тиме да се њихова продуктивност не огледа у врло високој фреквенцији појединачних токена, већ у продуктивности шаблона (енгл. *pattern*) као целине<sup>10</sup>.

Још једно питање, које досада није решено ни у конструкционој граматички, ни у традиционалној фразеологији, односи се на разликовање између варијанти једне ФК и сродних, а различитих ФК. Ако се ради о различитим ФК, поставља се питање у којем су односу сродне ФК. У лексикографској пракси ради се о кључном питању да ли је потребно увести засебну лему или варијанту укључити у постојећу лему. И овде ће пројекат, у суштини, морати да иде својим путем, који пре свега мора бити практичан и прилагођен корисницима. У случају разликовања између варијанти и одвојених ФК, морају се исто развити граничне вредности. На пример, могло би се размотрити да се одвојене ФК уводе тек када се њихова сидра разликују у више од два елемента и/или када постоји одређени минимални број структурних разлика. Друга димензија сличности односи се на семантичку и прагматичку сличност и преплиће се са подручјем синонима.

У оквиру пројекта, да би се подржао тим који ради на креирању речничких чланака, постоје и тимови који се баве теоријским питањима и одговорни су за техничку подршку речнику. Од њих такође у наредним годинама очекујемо важне подстицаје који могу унапредити и развој дигиталне лексикографије и теорију конструкционе граматике.

У будућности ће структура речника такође бити прилагођена потребама различитих корисника, тако да неће бити могуће само прецизно претраживати еквиваленте на одређеном језику, већ, на пример, и одређене типове текста. У следећој фази пројекта биће испробане методе које се могу користити за усвајање ФК у школској и у универзитетској настави страних језика. По узору на руски конструктикон (BAST et al. 2021), могу бити укључене и интерактивне вежбе различитих нивоа тежине у репозиторијум (планиране су само вежбе за немачки као страни језик). Како се рад на репозиторијуму може наставити

<sup>10</sup> Продуктивност се у овом смислу може описати као „продуктивност као опшност” („productivity as generativity”, IVORRA ORDINES, у штампи).

и након његове објаве, у будућности ће такође бити могуће укључити материјале за вежбање ФК на другим језицима репозиторијума.

На крају, још један фокус нашег рада јесу типолошке сличности и разлике у области ФК, које се могу истраживати на основу репозиторијума. До сада има врло мало знања о овом питању. Из перспективе аутора овог рада, који је ограничен на немачки, руски, српски и хрватски језик, чини се да су разлике између немачког и руског на подручју ФК значајно веће него између немачког и српског или немачког и хрватског. То се такође види у примерима у овом раду, где често постоји велика структурна и функционална сличност између немачких ФК и њихових српских еквивалената, што код руских ФК често није случај. И типолошка питања ће бити истраживана у наредним годинама пројекта.

### Напомена

This article is based upon work from the COST Action „PhraConRep – A Multilingual Repository of Central and Eastern European Languages” CA22115, supported by COST (European Cooperation in Science and Technology). Овај чланак заснован је на раду у оквиру COST акције „PhraConRep – A Multilingual Repository of Central and Eastern European Languages” CA22115, коју подржава COST (Европска сарадња у науци и технологији).

### Литература

- АПРЕСЈАН, Юрий Д. (2014). *Активный словарь русского языка. Том 1 А-Б*. Москва: Языки славянской культуры.
- ДРАГИЋЕВИЋ, Рајна (2018). *Српска лексика у прошлости и данас*. Нови Сад: Магица српска.
- ШМЕЛЕВ, Дмитрий Н. (1977). *Современный русский язык. Лексика*. Москва: Просвещение.
- \* ANDONOVSKI, Jelena (2019). Mreža otvorenih podataka i jezički resursi u procesu izgradnje srpsko-nemačkog literarnog korpusa [doktorska disertacija]. Beograd: Univerzitet u Beogradu, Filološki fakultet.
- АПРЕСЈАН, Јуриј (2024). Active Dictionary. In: Marc L. Greenberg (ed.), *Encyclopedia of Slavic Languages and Linguistics Online*. Leiden: Brill. Doi: <[https://doi.org/10.1163/2589-6229\\_ESLO\\_COM\\_032490](https://doi.org/10.1163/2589-6229_ESLO_COM_032490)>.
- BOAS, Hans C. (2019). Zur empirischen Grundlage der Konstruktikografie. In: Dániel Czicza, Volodymyr Dekalo, Gabriele Diewald (ed.), *Konstruktionsgrammatik VI. Varianz in der konstruktionalen Schematizität* (p. 237–263). Tübingen: Stauffenburg.
- BOON, Geert (2002). Constructional Idioms, morphology, and the Dutch lexicon. *Journal of Germanic Linguistics*, 14(4), 301–329.
- ČAGALI, Ivana, SVÍTKOVÁ, Milina (2014). Tipologija frazeološke ekvivalencije na primjeru hrvatskih i slovačkih frazema s ihtionimskom sastavnicom. U: Ivana Vidović Bolt (ur.), *Životinje u frazeološkom ruhu* (str. 1–15). Zagreb: FF-press, Filozofski fakultet.
- ČERMÁK, František, ROSEN, Alexandr (2012). The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics*, 17(3), 411–427.
- ČERMÁK, Petr (2019). InterCorp: A parallel corpus of 40 languages. In: Irene Doval, M. Teresa Sánchez Nieto (ed.), *Parallel Corpora for Contrastive and Translation Studies: New resources and applications* (p. 93–101). Amsterdam: John Benjamins Publishing Company.
- DOBROVOL'SKII, Dmitrij (2011). Phraseologie und Konstruktionsgrammatik. In: Alexander Lasch, Alexander Ziem (ed.), *Konstruktionsgrammatik III. Aktuelle Fragen und Lösungsansätze* (p. 110–130). Tübingen: Stauffenburg.
- ĐURČO, Peter (2018). Faktoren der konvergenten und divergenten Äquivalenz von präpositionalen Wortverbindungen: Deutsch – Slowakisch. In: Kathrin Steyer (ed.), *Sprachliche Verfestigung. Wortverbindungen, Muster, Phrasem-Konstruktionen* (p. 285–306). Tübingen: Narr Francke Attempto.
- FILIPOVIĆ PETROVIĆ, Ivana (2018). *Kada se sretnu leksikografija i frazeologija*. Zagreb: Srednja Europa.
- FLEISCHER, Wolfgang (1982). *Phraseologie der deutschen Gegenwartssprache*. Leipzig: Bibliographisches Institut.
- IVORRA ORDINES, Pedro (у штампии). Productivity and creativity of constructions. In: Xu Wen, Chris Sinha (ed.), *The Cambridge Encyclopedia of Cognitive Linguistics*. Cambridge: Cambridge University Press.
- KILGARRIFF, Adam, BAISA, Vít, BUŠTA, Jan, JAKUBÍČEK, Miloš, KOVÁŘ, Vojtěch, MICHELFEIT, Jan, RYCHLÝ, Pavel, SUCHOMEL, Vít (2014). The Sketch Engine: ten years on. *Lexicography*, 1, 7–36.
- KILGARRIFF, Adam, RYCHLÝ, Pavel, SMRŽ, Pavel, TUGWELL, David (2004). The Sketch Engine. In: Geoffrey Williams, Sandra Vessier (ed.), *Proceedings of the 11th EURALEX International Congress* (p. 105–116). Lorient: Université de Bretagne-Sud, Faculté des lettres et des sciences humaines.
- KUPIETZ, Marc, LÜNGEN, Harald, KAMOČKI, Paweł, WITT, Andreas (2018). The German Reference Corpus DeReKo: New Developments – New Opportunities. In: Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, Takenobu Tokunaga (ed.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (p. 4353–4360). Miyazaki: European Language Resources Association (ELRA).
- LISON, Pierre, TIEDEMANN, Jörg (2016). OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles. In: Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis (ed.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC-2016)* (p. 923–929), Portorož, Slovenia: European Language Resources Association (ELRA).
- LYNGFELT, Benjamin (2018). Introduction: Constructicons and constructicography. In: Benjamin Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent (ed.), *Constructicography: Constructicon Development Across Languages* (p. 1–18). Amsterdam – Philadelphia.
- MEL'ČUK, Igor (2022). Russian reduplicative surface-syntactic relations in the perspective of general syntax. *Russian Journal of Linguistics* 26(4), 881–907.
- MELLADO BLANCO, Carmen (2023). From idioms to semi-schematic constructions and vice versa. The case of [a un paso de X]. In: Evelyn Wiesinger, Inga Hennecke (ed.), *Constructions in Spanish* (p. 103–128). Amsterdam – Philadelphia: John Benjamins.
- MELLADO BLANCO, Carmen (у штампии). Constructional Idioms. In: Xu Wen, Chris Sinha (ed.), *The Cambridge Encyclopedia of Cognitive Linguistics*. Cambridge: Cambridge University Press.
- MELLADO BLANCO, Carmen, MOLLIĆA, Fabio, SCHAFFROTH, Elmar (2022). Phrasem-Konstruktionen in der heutigen Forschung: ein Überblick. In: Carmen Mellado Blanco, Fabio Mollica, Elmar Schafroth (ed.), *Konstruktionen zwischen Lexikon und*

- Grammatik: Phrasem-Konstruktionen monolingual, bilingual und multilingual* (p. 1–18). Berlin – Boston: De Gruyter. <<https://doi.org/10.1515/9783110770209-001>>.
- MELLADO BLANCO, Carmen, STEYER, Kathrin (2018). Auf der Suche nach Äquivalenz. Lexikalisch geprägte Muster kontrastiv: Deutsch – Spanisch. In: Kathrin Steyer (ed.), *Sprachliche Verfestigung. Wortverbindungen, Muster, Phrasem-Konstruktionen* (p. 265–284). Tübingen: Narr Francke Attempo.
- MILLER, Julia (2018). Research in the pipeline: Where lexicography and phraseology meet. *Lexicography*, 5, 23–33. <<https://doi.org/10.1007/s40607-018-0044-z>>.
- PAVLOVA, Anna (2024). Äquivalenz bei Übersetzung von Phrasem-Konstruktionen. In: Anna Gondek, Alina Jurasz, Alexandra Kamińska, Przemysław Staniewski, Joanna Szczęk (ed.), *Interkulturelles und Interdisziplinäres in der Phraseologie und Parömiologie II* (p. 159–178). Hamburg: Verlag Dr. Kovač.
- PAVLOVA, Anna (2022). Mehrsprachige Datenbank der Phrasem-Konstruktionen. In: Annette Klosa-Kückelhaus, Stefan Engelberg, Christine Möhrs, Petra Storjohann (ed.), *Dictionaries and Society. Proceedings of the XX EURALEX International Congress* (p. 594–604). Mannheim: IDS.
- ROSEN, Alexandr (2016). InterCorp – a look behind the façade of a parallel corpus. In: Agnieszka Leńko-Szymańska, Ewa Gruszczyńska (ed.), *Polskojęzyczne korpusy równoległe. Polish-language Parallel Corpora* (p. 21–40). Warszawa: Instytut Lingwistyki Stosowanej.
- SCHLUND, Katrin – PAVLOVA, Anna (у штампни). Core and boundaries of the notion of Phraseme Construction. *Quaderns de Filologia - Estudis lingüistics. Special Issue: New approaches to Construction Grammar(s)*, 29.
- STANKOVIĆ, Ranka, KRSTEV, Cvetana, LAZIĆ, Biljana, VORKAPIĆ, Dalibor (2015). A bilingual digital library for academic and entrepreneurial knowledge management. In: John Christopher Spender, Giovanni Schiuma, Vito Albino (ed.), *10th International Forum on Knowledge Asset Dynamics – IFKAD* (p. 1764–1777). Bari: Institute of Knowledge Asset Management.
- TOMASELLO, Michael (2005). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, Mass.: Harvard University Press.
- WRAY, Alison (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- ZIEM, Alexander (2018). Construction Grammar meets phraseology. *Linguistik Online*, 90(3), 1–18. doi: 10.13092/lo.90.4316
- NIETZSCHE, Friedrich (1884). *Also sprach Zarathustra*. Chemnitz: Verlag von Ernst Schmeitzner.
- REMARQUE, Erich Maria (1989). *Der schwarze Obelisk*. Köln: Kiepenheuer & Witsch.
- ROSEN, Alexandr, VAVŘIN, Martin, ZASINA, Adrian Jan (2022). InterCorp, Release 15 of 11 November 2022. Institute of the Czech National Corpus, Charles University. <<http://www.korpus.cz>>. [22.11.2024]
- SÜSKIND, Patrick (1985). *Das Parfum: Die Geschichte eines Mörders*. Zürich: Diogenes Verlag.
- ZISKIND, Patrik (2001). *Parfem. Hronologija jednog zločina*. Novi Sad: Solaris.

## Корпуси и књижевни извори

- ГЕТЕ, Јохан Волфганг (1949). *Јаги млагоја Верџера*. Нови Сад: Матица српска.
- Национални корпус руског језика. [www.ruscorpora.ru](http://www.ruscorpora.ru) [22.11.2024]
- НИЧЕ, Фридрих (2007). *Тако је говорио Заратустра*. Београд – Нови Сад: Feniks Libris – Будућност.
- \*
- BAST, Radovan, ENDRESEN, Anna, JANDA, Laura A., LUND, Marianne, LYASHEVSKAYA, Olga, McDONALD, James, MORDASHOVA, Daria, NESSET, Tore, RAKHILINA, Ekaterina, TYERS, Francis M., ZHUKOVA, Valentina (2021). The Russian Constructicon. An electronic database of the Russian grammatical constructions. Available at: <<https://constructicon.github.io/russian/>>. [24. 11. 2024]
- COSMAS I/II (Corpus Search, Management and Analysis System), <http://www.ids-mannheim.de/cosmas2/>, © 1991–2022 Leibniz-Institut für Deutsche Sprache, Mannheim. [22. 11. 2024]
- ГОЕТНЕ, Johann Wolfgang (2016). *Die Leiden des jungen Werther*. [s. l.]: Berliner Ausgabe.



**Kristina Ilić — Katrin Schlund**

**Insights into the Development of  
a Multilingual Repository of Phraseme Constructions**

---

(Summary)

The creation of a multilingual online repository of German Phraseme Constructions (PhraCons) represents a novel development in lexicography. Since PhraCons are lexically only partially specified units, their underlying abstract schemas must be determined based on concrete realizations. In the next step, equivalents in the target languages must be identified based on these realizations. The result will be a construction of German PhraCons in the sense of Construction Grammar, which also provides translations into the target languages of the project. In line with the concept of the “Active Dictionary” of the Moscow Semantic School, functional explanations of the use of German PhraCons will be provided, enabling users of the repository to employ PhraCons in their own text production. The article describes the macro- and microstructure of the repository, with special attention given to how equivalents in the target languages are identified. The project’s collaborators deploy various strategies to find the most suitable equivalent of each PhraCon from the repository. The most frequently used translation strategies by the project members are parallel corpora of German and at least one of the project languages as the target language, the literary translations, either independently or as a part of a parallel corpus of literary works and own translations, with or without AI assistance. The best possible outcome is to detect a full equivalent of a PhraCon in the target language and then describe it in the repository. In some cases, full equivalents do not exist, and then partial equivalents are taken for the repository. Numerous practical issues must be addressed in the work that have not yet been satisfactorily treated on a theoretical level—neither in Construction Grammar nor in the field of traditional phraseology. This includes, for example, the question of which quantitative and, possibly, also qualitative, criteria can be used to distinguish (typically, minimal) variation in the anchors of PhraCons from (typically, maximal) variation in fillers. This problem also involves the question of the elusive concept of productivity of Constructions: A large number of different fillers, as it is typical for PhraCons, represents the general productivity of a PhraCon. High token frequency, on the other hand, allows for the identification of the most frequent fillers for a PhraCon, that is, of a prototypical realization. However, such a prototype often does not exist for PhraCons with great general productivity. Also, variants of constructions must be distinguished from separate constructions and integrated into networks of constructions based on structural and/or semantic similarity—a notorious discussion within Construction Grammar.