

JAZYKOVEDNÝ ČASOPIS

2

ROČNÍK 76, 2025

JAZYKOVEDNÝ ÚSTAV EUDOVÍTA ŠTÚRA
SLOVENSKÁ AKADEMIA VIED

JAZYKOVEDNÝ ČASOPIS
VEDECKÝ ČASOPIS PRE OTÁZKY TEÓRIE JAZYKA

JOURNAL OF LINGUISTICS
SCIENTIFIC JOURNAL FOR THE THEORY OF LANGUAGE

Hlavná redaktorka/Editor-in-Chief: doc. Mgr. Gabriela Múcsková, PhD.

Výkonní redaktori/Managing Editors: Mgr. Jaroslava Rusinková, PhD., Mgr. Roman Sooky, PhD.,
Mgr. Miroslav Zumrík, PhD.

Redakčná rada/Editorial Board: PhDr. Klára Buzássyová, CSc. (Bratislava), prof. PhDr. Juraj Dolník, DrSc. (Bratislava), PhDr. Ingrid Hrubaničová, PhD. (Bratislava), prof. Mgr. Martina Ivanová, PhD. (Prešov), Mgr. Nicol Janočková, PhD. (Bratislava), PhDr. Alexandra Jarošová, CSc. (Bratislava), prof. PaedDr. Jana Kesselová, CSc. (Prešov), PhDr. Ľubor Králik, DSc. (Bratislava), Univ. Prof. Mag. Dr. Stefan Michael Newerkla (Viedeň – Rakúsko), Prof. Mark Richard Lauersdorf, Ph.D. (Kentucky – USA), prof. Mgr. Martin Ološtiak, PhD. (Prešov), prof. PhDr. Slavomír Ondrejovič, DrSc. (Bratislava), prof. PaedDr. Vladimír Patráš, CSc. (Banská Bystrica), prof. PhDr. Ján Sabol, DrSc. (Košice), prof. PhDr. Juraj Vaňko, CSc. (Nitra), prof. PhDr. Pavol Žigo, CSc. (Bratislava).

Technický redaktor/Technical editor: Mgr. Vladimír Radik

Vydáva/Published by: Jazykovedný ústav Ľudovíta Štúra Slovenskej akadémie vied, v. v. i.

- v tlačenej podobe vo vydavateľstve SAP – Slovak Academic Press, s. r. o., Bratislava
- elektronicky v Paradigm Publishing Services, De Gruyter Brill Sp. z o. o., Warsaw, Poland
<https://reference-global.com/journal/JAZCAS>

Adresa redakcie/Editorial address: Jazykovedný ústav Ľ. Štúra SAV, Panská 26, 811 01 Bratislava
e-mail: jazykovedny.casopis@juls.savba.sk

web: <https://journals.savba.sk/index.php/jc/>, https://www.savba.sk/?lang=sk&doc=journal-list&journal_no=26

Odvodzďavanie príspevkov/Article submission: <https://journals.savba.sk/index.php/jc/about/submissions>

Vychádza trikrát ročne/Published triannually

Dátum vydania aktuálneho čísla (2025/76/2) – december 2025

Quartile ranking 2024: Q3

CiteScore 2024: 0,6

SCImago Journal Rank (SJR) 2024: 0,150

Source Normalized Impact per Paper (SNIP) 2024: 0,508

JAZYKOVEDNÝ ČASOPIS je evidovaný v databázach/JOURNAL OF LINGUISTICS is covered by the following services: Baidu Scholar; Cabell's Journalytics; CEEAS (Central & Eastern European Academic Source); CEJSH (The Central European Journal of Social Sciences and Humanities); CEEOL (Central and Eastern European Online Library); CNKI Scholar (China National Knowledge Infrastructure); Crossref database; Dimensions; EBSCO; ERIH PLUS (European Reference Index for the Humanities and Social Sciences); ExLibris; Google Scholar; IBR (International Bibliography of Reviews of Scholarly Literature in the Humanities and Social Sciences); IBZ (International Bibliography of Periodical Literature in the Humanities and Social Sciences); International Medieval Bibliography; J-Gate; Jisc - Open Policy Finder; JournalGuide; JournalTOCs; KESLI-NDSL (Korean National Discovery for Science Leaders); Linguistic Bibliography; Linguistics Abstracts Online; MLA International Bibliography; Naver Academic; Naviga (Softweco); OpenAlex; ProQuest; ReadCube; ScienceON/AccessON; SCILIT; SCImago (SJR); Scite_; SCOPUS; Semantic Scholar; TDNet; Ulrich's Periodicals Directory/ulrichsweb; WorldCat (OCLC); X-MOL.

ISSN 0021-5597 (tlačená verzia/print)

ISSN 1338-4287 (verzia online)

MIČ 49263

JAZYKOVEDNÝ ČASOPIS

2

ROČNÍK 76, 2025

JAZYKOVEDNÝ ÚSTAV EUDOVÍTA ŠTÚRA
SLOVENSKÁ AKADEMIA VIED

OBSAH CONTENTS

Štúdie Studies

- 393 **Jana SOKOLOVÁ: Koncept ‘*oboch*’ v zmysle jednoty súhlasnosti a rozdielnosti *dvoch*. Sonda do jazykového a diskurzívneho priestoru súčasnej slovenčiny**
Jana SOKOLOVÁ: The concept of ‘*both*’ in the sense of the unity of agreement and difference between the *two*. A probe into the linguistic and discursive space of contemporary Slovak
- 412 **Václav BLAŽEK – Martina ŠMEJKALOVÁ: Z historie evropské medicínské terminologie: řecké kólov, jeho etymologie a deriváty**
Václav BLAŽEK – Martina ŠMEJKALOVÁ: From the history of European medical terminology: the Greek word kólov, its etymology and derivatives
- 433 **Peter BEDNÁR – Marek DOBEŠ – Radovan GARABÍK: Training of large language model Mistral on Slovak language data**
Peter BEDNÁR – Marek DOBEŠ – Radovan GARABÍK: Training of large language model Mistral on Slovak language data
- 452 **Zuzana GAŽÁKOVÁ: Slovenská právna dokumentácia pre utečencov a žiadateľov o azyl a otázka jej zrozumiteľnosti v prekladoch do arabčiny**
Zuzana GAŽÁKOVÁ: Slovak legal documentation for refugees and asylum seekers and the question of its comprehensibility in Arabic translations
- 468 **Katarína WELNITZOVÁ – Daša MUNKOVÁ: Persisting grammatical errors of machine translation**
Katarína WELNITZOVÁ – Daša MUNKOVÁ: Persisting grammatical errors of machine translation

Rozhľady Horizons

- 493 **Tomáš BRATINA: Purism and nationalism in language norms: A comparative analysis of the Norwegian and Slovak spelling reform adopted during World War II**
TOMÁŠ BRATINA: Purism and nationalism in language norms: A comparative analysis of the Norwegian and Slovak spelling reform adopted during World War II
- 509 **Roman VOJTECHOVSKÝ: Fonologické procesy v prúde posunkovania**
Roman VOJTECHOVSKÝ: Phonological processes in the stream of signing

Recenzie

Book Reviews

- 534 **PUCHOVSKÁ, Zuzana: Slovesný prítomný a minulý čas z francúzsko-slovenskej kontrastívnej perspektívy. (Martina VLADOVIČOVÁ)**
- 537 **LIPTÁKOVÁ, Ľudmila (ed.) – DZIAK, Dávid – GOGOVÁ, Eva – BRESTOVIČOVÁ, Alexandra – VUŽŇÁKOVÁ, Katarína – LUKÁČ, Juraj – KOPČÍKOVÁ, Marta: Výskum porozumenia textu v mladšom školskom veku. (Roman SOÓKY)**

KONCEPT ‘OBOCH’ V ZMYSLE JEDNOTY SÚHLASNOSTI A ROZDIELNOSTI *DVOCH*. SONDA DO JAZYKOVÉHO A DISKURZÍVNEHO PRIESTORU SÚČASNEJ SLOVENČINY

JANA SOKOLOVÁ

Filozofická fakulta Univerzity Konštantína Filozofa v Nitre, Nitra, Slovensko

SOKOLOVÁ, Jana: The concept of ‘both’ in the sense of the unity of agreement and difference between the *two*. A probe into the linguistic and discursive space of contemporary Slovak. Jazykovedný časopis (Journal of Linguistics), 2025, Vol. 76, No. 2, pp. 393–411.

Abstract: The study focuses on the semantics and linguistic pragmatics of the expression *both* (*obidvaja*, *obidva*, *obaja*, *oba*, *obe*, *obidve*, *obidvoje*, *oboje*), which in Slovak can be classified as numerals, and marginally even as deictic parts of speech. The interpretation focuses on the aspects of following and varying the principles of maximality, homogeneity and salience. We point out that the cognitive definition of paired, generic and quantitatively bound multiplicity is strongly anchored in Slovak. In addition to the synthetic method, the singularity – multiplicity opposition is realized through lexical intervention with special lexical means. According to the nature of the composition of entities and lexical-semantic types of nouns, plurality is formed by word combinations of numerical, paired, set-based, generic and quantitatively defined multiplicity. From the point of view of logical semantics, the expressions behave not only as determiners of uncertainty, but also as generalized, general, existential and floating quantifiers. From the point of view of linguistic pragmatics, the analyzed expressions are egocentric. Due attention is therefore paid to the concept of ‘both’ in the context of the speaker’s verbal actions, relational projections and plurality of word combinations.

Keywords: concept, set, quantifier, projection, homogeneity, maximality, phrase

1. ÚVOD A ZÁKLADNÉ VÝCHODISKÁ

Koncept ‘*oboch*’ má v slovenčine výrazné lexikálne zastúpenie, porov. *obidvaja*, *obidva*, *obaja*, *oba*, *obe*, *obidve*, *obidvoje*, *oboje*, *obojaký*, *obojako*. V porovnaní s inými slovanskými jazykmi, napr. ruštinou (*оба, обе*), ukrajinčinou (*обá, обудва, обудви*), bieloruštinou (*а́бодва, а́бедзве, а́боe*), češtinou (*oba, obě, oboje, obojí, obé, (po)obokrát*), poľštinou (*obaj, obydwa, obie, obydwie, oba, obydwa, oboje, obydwoje*), v slovenčine je evidentná vyššia početnosť, za ktorou stoja tieto skutočnosti: (a) synonymia výrazov *obidvaja* – *obaja*, *obidva* – *oba*, *obidve* – *obe*, *obidvoje* – *oboje*; (b) v singulári zámen opozícia v gramatickom rode: mužský rod oproti ženskému a strednému (*obidvaja*, *obaja*, *obidva*, *oba*: mužský rod; *obe*, *obidve*, *obidvoje*, *oboje*: ženský a stredný rod); (c) opozícia v mužskom rode: životné a neživotné

(*obidvaja, obaja*: životné; *obidva, oba*: neživotné); (d) konkurencia kvantitatívneho a deiktického parametra (*obidvaja, obidva, obaja, oba, obe, obidve, obidvoje, oboje*: číslovky; *obidvoje, oboje*: zámená); (e) konkurencia v kvantitatívnom parametri: numerický – súborový – druhový obraz entít (*obidvaja, obidva, obaja, oba, obe, obidve*: základné číslovky; *obidvoje, oboje*: skupinové číslovky; *obojaký, obojako*: druhové číslovky; (f) konkurencia kvantitatívneho a kvalitatívneho parametra (*obojaký* – druhová číslovka adjektívna; *obojaký* – blíži sa významu adjektíva); (g) konkurencia kvantitatívneho a okolnostného parametra (*obojako* – druhová číslovka príslovková; *obojako* – blíži sa významu príslovky); (h) homonymia súborového a súhrnového parametra¹ (*obidvoje, oboje*: skupinové číslovky; *obidvoje, oboje*: vymedzovacie zámená). Ako vidieť, interlexikálne sémantické vzťahy výrazov sú na paradigmatickej rovine založené jednak na identite významov a jednak na binárnom aj nebinárnom kontraste.

Základné gramatikalizované parametre a slovnodruhové vymedzenie výrazov zachytávajú lexikografické výklady. Uvádzajú tie vlastnosti, ktoré sú jazykovou prezentáciou kognitívneho spracovania skutočnosti a výsledkom jazykového videnia sveta slovenských užívateľov. Interpretujú jednak konceptuálny rámec výrazov a tým, že signalizujú ich kontextuálny charakter, zároveň naznačujú funkčné uplatnenie. Porovnajme: výraz *oba, obidva, obaja, obidvaja, obe, obidve* ako základná číslovka „vyjadruje počet dva vyčerpávajúci všetky dané veci, javy, osoby, ktoré sa chápu ako dvojica alebo sa vyskytujú obyčajne v páre, jeden aj druhý“ (Jarošová (ed.) 2021, s. 26), napr.² *Návšteva mala prehliť nadštandardné vzťahy oboch krajín po vstupe do Európskej únie*. Výraz *obojako, obidvoje* patrí medzi skupinové číslovky, keďže „vyjadruje počet dva ako vyčerpávajúci pri veciach, javoch, osobách, ktoré sa vyskytujú obyčajne v páre, pri pomnožných podstatných menách, jedny i druhé“ (tamže, s. 108), napr. *Od novembra pribudli pre chodcov tabule so zákazom vstupu na obidvoje schody vedúce na most*. Výraz *obojaký* ako druhová číslovka znamená „pozostávajúci z obidvoch, oboch druhov niečoho; jestvujúci v dvoch rozličných, často protichodných podobách, vyznačujúci sa vnútornou dvojakosťou alebo odlišnosťou, dvojaký“ (tamže, s. 108), napr. *U zvlášť nadaných básnikov významové a zvukové ozdoby niekedy splývajú, čím vznikajú tzv. obojaké ozdoby*. Výraz *obojako, obidvoje* sa zaraďuje aj medzi vymedzovacie substantívne zámená, lebo „vymedzuje súhrn dvoch odlišných vecí alebo javov, ktoré spoločne vytvárajú jednu situáciu, vyskytujú sa súčasne v jednom priestore a čase, jedno i druhé, to i ono“ (tamže,

¹ *Poznámka*: V záujme sprehľadnenia a zjednotenia výkladu budeme rozlišovať obsah pomenovaní *súbor* a *súhrn*. *Súbor* je ponímaný ako zoskupenie jednotlivín kreujúcich celok a majúci spoločné vlastnosti. Premenné súboru majú príznak homogenity. *Súhrn* je posudzovaný v zmysle početnosti jednotlivín tvoriacich celok. Tvoria ho premenné s príznakom nehomogenity.

² Ilustratívne príklady uvádzame podľa *Slovenského národného korpusu* (verzia: prim-9.0-public-sane). Doklady majú pôvodnú podobu aj v prípadoch, keď nezodpovedajú kodifikácii spisovnej slovenčiny.

s. 109), napr. *Máme nároky na život, i povinnosti k životu. K obojemu sa hlásime. Ako vidieť*, posudzované výrazy sú v súčasnej slovenčine slovnodruhovo zaradené ako číslovky (základné, skupinové a druhové) a výrazy *obidvoje, oboje* sa ponímajú aj ako vymedzovacie substantívne zámená. Lexikografické výklady približujú charakter výrazov, ktoré svoj sémantický obsah naplňajú v konkrétnom kontexte použitia (k rozlíšeniu medzi obsahom výrazu a jeho charakterom pozri Mráz 2008).

Vzniká otázka, či príslušnosť k dvom slovným druhom má poznávacie alebo viac diskurzívne pozadie. Nazdávame sa, že rozdielna slovnodruhová príslušnosť môže súvisieť s posudzovaním *číslo* už v zmysle dvoch epistemologických stratégií – atributívnej a substantívnej stratégie. Atributívna stratégia sa spája s menom J. S. Milla (1973), podľa ktorého *číslo* je vlastnosť vonkajších vecí a číselný údaj je získaný na základe pozorovania empirických entít. Substantívna stratégia vyplynula z koncepcie G. Fregeho (2011), ktorý *číslo* nepovažuje za vlastnosť vonkajších vecí, ale za samostatne množinovo existujúci teoretický predmet (citované podľa Sousedík – Svoboda 2014). Uvedené skutočnosti naznačujú, že číslovky môžu byť výrazom atributívnej stratégie a vymedzovacie substantívne zámená sú výrazom substantívnej stratégie posudzovania čísla. Obe stratégie vykazujú medzi výrazmi ako lexikálnymi jednotkami a gramatickou kategóriou čísla vzťah koligácie.³ Atributívna stratégia značí, že numeráliá sa spájajú s plurárovými tvarmi substantív a tvoria s nimi syntakticky jednotné celky signalizujúce ukotvenosť kvantifikátorového výrazu⁴ v syntaxi slovného spojenia. Referentmi slovného spojenia sú triedy predmetov tvoriacich množiny, ktorých kardinalita (t. j. počet prvkov množiny) je vždy limitovaná počtom dva. Pluralita je v nich kalkulovaná na sémantickej, pragmatickej aj syntaktickej úrovni. Sémantický obsah výrazov závisí od kontextu použitia a o referencii výrazov rozhodujú aj pragmatické faktory. Číslovky tvoria slovné spojenia: (i) s počítateľnými substantívami⁵ (životnými aj neživotnými): *Obaja muži sa vtom rozbehli smerom k výťahu. – Obe dohody sú výsledkom plánovanej kooperácie oboch krajín vo viacerých sférach.*; (ii) s párovými substantívami⁶ (životnými aj neživotnými): *Mláďatá krmia obaja rodičia rovnakým podielom. – Pohla sa, chcela odísť. Chytil ju za obe plecia.*; (iii) substantívami pluráliá tantum: *Potom sa rozhodneme, či*

³ Koligácia v zmysle princípu predurčeného výberu. Týka sa vzťahov medzi lexikálnou jednotkou a gramatickou kategóriou a/alebo textovou štruktúrou (porov. Cvrček 2017).

⁴ V texte štúdie budeme používať výraz *kvantifikátor* v zmysle *kvantifikátorový výraz*. Vo filozofickej praxi sa rozlišujú *kvantifikátory* a *kvantifikátorové výrazy*. Pod kvantifikátormi sa zvyčajne rozumejú určité sémantické *objekty*, nie výrazy. Kvantifikátory sú objekty, na ktoré sa vzťahujú kvantifikátorové výrazy, t. j. kvantifikátory budú sémantickým obsahom kvantifikátorových výrazov (porov. Zouhar 2006a).

⁵ Kvantita substantív disponuje aj aspektom počítateľnosti a nepočítateľnosti (porov. M. Sokolová 2007, s. 23).

⁶ Párové substantíva majú v tvaroch plurálu význam prirodzeného páru. Ako píše M. Majtánová (1976, s. 79), „vo význame párovosti sa silnejšie zdôrazňuje tesný vzťah obidvoch kvalitatívne rovnakých členov dvojice, t. j. ich vzájomné spoločenstvo.“

pôjdeme na Giro alebo na Okolo Kalifornie, oboje preteky sú v máji.; (iv) substantívami singuláriá tantum, ktoré význam plurálu vyjadrujú najmä pomocou kategoriálnych substantív *typ* a *druh*: *Súčasná medicína ponúka účinné očkovacie látky pre oba typy žltacky*. (Ku kolokačnému profilu čísloviek sa ešte vrátíme v 3. časti príspevku.) V atributívnej stratégii je tvar číslovky a substantíva ovplyvnený pravidlami kongruentnej syntagmy, porov. *obaja muži, obe dohody, obaja rodičia, oboje preteky* a pod. Substantívne pomenovania sú v nich vo funkcii dominujúceho člena určujúceho podobu dominovaného člena z hľadiska čísla, rodu a životnosti/neživotnosti.

V substantívnej stratégii sú deiktiká *obidvoje, oboje* vetnočlensky autonómne. Odkazujú na slovné spojenia prezentované výhradne singulárovými pádovými tvarmi substantív, ktoré sú vo vzťahu slabej logickej disjunkcie (porov. Daneš 2009). Vyjadrujú sumarizujúci vzťah medzi dvoma eventualitami. Hovoriaci potvrdzuje možnosť pôsobenia oboch faktorov súčasne v zmysle tretej eventuality. V poradí antecendentov sa uplatňuje faktor preferencie, ktorý umožňuje ich relativizovanie a subjektívne radenie: *Máš radšej leto alebo zimu? Oboje. = Máš radšej zimu alebo leto? Oboje*. Na rozdiel od základných čísloviek nekvantifikujú predmety mimojazykovej reality, ale textové pozície (Veselý 2011, s. 192).

Epistemický rozmer posudzovaných výrazov sa v minulosti fixoval gramatickými tvarmi dvojného čísla (duálu). V súčasnom slovenskom jazyku je ponímaný v rámci kategórie množného čísla. Vychádza z konceptu *mnohosti*, resp. kvantovosti, kvantitatívnosti, ktorý V. Filkorn definoval ako „mnohosť rovnako rozličných vecí“ (Filkorn 1998, s. 239). Parafrázujúc túto definíciu, koncept ‘*obidvoch*’ možno ponímať ako „úplnú mnohosť dvoch rovnako rozličných vecí“. Pre porovnanie, v *Mluvnici češtiny 2* sa výrazy *oba, obě, oboje, obojí* charakterizujú ako „vyčerpání množství prvků množiny o dvou prvcích“ (Mluvnice češtiny 2 1986, s. 111). Z prirodzenej podstaty veci vyplynula úzka prepojenosť duálu a párových substantív, rezíduá ktorej zostávajú konceptuálne ukotvené a v súvislosti s párovými substantívami vysvetľujú preferenciu číslovky *oba/obe* pred číslovkou *dva/dve* a ohraničenosť používania totalizátorov.⁷

Je teda nesporné, že vo výrazoch *obidvaja, obidva, obaja, oba, obe, obidve, obidvoje, oboje* sa spája predovšetkým faktor úplnosti, ktorý súvisí s uplatňovaním princípu maximality, a faktor rovnakosti⁸, ktorý predpokladá princíp homogenity. V ďalšom výklade sa preto zameriame na stručné priblíženie rešpektovania a variovania oboch princíпов v priesečníku hľadísk logickej sémantiky a lingvopragmatiky.

⁷ Totalizátory sa vymedzujú v rámci viacerých slovných druhov. V *Novom encyklopedickom slovníku češtiny* (Osolobě 2017) s odvolaním sa na *Mluvnici češtiny 2* (1986) sa do triedy totalizátorov zaraďujú zámená (*všetok, všetko, každý, sám*), adverbá (*vždy, všade, zakaždým*) a číslovka *oba*.

⁸ Poznámka: výrazy *totožnosť, rovnakosť, identickosť* používame v ich relatívnom zmysle.

2. KONCEPT ‘OBOCH’ Z HEADISKA LOGICKEJ SÉMANTIKY

Numerické výrazy *obidvaja*, *obidva*, *obaja*, *oba*, *obe*, *obidve*, *obojaký*, *obojako* sa správajú ako jazykové kvantifikátory (kvantory),⁹ ktoré sú, na rozdiel od predikátovologických, neurčité a neostre (porov. Oliva 2017). Referentom oznamovacích viet, resp. propozícií s uvedenými výrazmi je stav vecí, resp. ich pravdivostná hodnota. Riešenie otázky, za akých podmienok sú vety pravdivé alebo nepravdivé, sa spravidla realizuje negáciou implikujúcou odlišnú pravdivosť viet. Porovnajme vety: *Kamil zjedol oba chleby*. – *Kamil nezjedol oba chleby*. Prvá veta implikuje skutočnosť, že *Kamil zjedol dva chleby*. Druhá veta implikuje skutočnosť, že *Kamil zjedol jeden chlieb*. Logická vlastnosť definovaná pomocou pravdivostných hodnôt a vyplývania je v logickej sémantike ponímaná ako *veridikalita* a jej náprotivky *neveridikalita*, resp. *antiveridikalita*. Veridikalita je citlivosť lexém na určitú vlastnosť kontextu, v ktorom sa vyskytujú. Citlivosť sa prejavuje buď ako pozitívna kontextová závislosť, alebo ako negatívna závislosť (výraz je neústrojný v kontexte s negatívne polarizovaným výrazom, t. j. nesmie byť v rozporujúcom kontexte). Z pohľadu veridikalit je výraz *oba* **neveridický determinátor** (porov. Dočekal 2017b). Porovnajme: **Obaja muži prehovorili ani slovo*. $\approx$ výraz *obaja* sa nachádza v rozporujúcom kontexte, keďže častica *ani* je negatívne polarizovaná. Nerozporujúci kontext si v tomto prípade vyžaduje neveridickú (nerozporujúcu) slovesnú negáciu: *Obaja muži neprehovorili ani slovo*. Výraz *obaja* umožňuje ďalšie implikácie len prostredníctvom veridických determinátorov – kvantifikátorov *žaden* alebo *nikto*: *Žaden z mužov neprehovoril ani slovo*. – *Nikto z mužov neprehovoril ani slovo*. Výpovede ukazujú, že výraz *obaja* má vlastnosti vyplývania dole (porov. Dočekal 2017a), a potvrdzujú, že kvantifikátory *žaden* a *nikto* patria do množiny klesajúcich kvantifikátorov (sprava aj zľava).¹⁰

Z hľadiska logickej sémantiky sa výrazy javia aj ako zvláštny druh **všeobecného (univerzálneho) kvantifikátora**, t. j. výrazu $(\forall x)(Px)$, ktorý značí *pre každé x platí P(x)*: *Obaja rodičia prispievajú na výživu svojich detí podľa svojich schopností, možností a majetkových pomerov*. Determinátor *obaja* sa považuje za univerzálny, lebo ide o variant *všetci/každý* v prípadoch, keď reštriktor obsahuje práve dva elementy (porov. Šimík – Dotlačil 2017): [QP *obaja* [NP *rodičia*]] [VP *prispievajú na výživu*] = je pravda, pokiaľ $\forall x$ [*rodič(x) & |rodič| = 2*] $\rightarrow$ [*prispievajú na výživu(x)*]. Hovoriaci predostiera individualizované čítanie kvantifikátora *obaja* – stotožňuje sa so všeobecne prijatou mienkou, že *každý* z rodičov (aj matka aj otec) má prispievať

⁹ Ku kvantifikátorom sa zaraďujú základné číslovky, zámená (*všetci*, *všetky*, *všetko*, *každý*, *žiadny*, *nikto*, *niekto*, *niektorý* a pod.) a časti nominálnych skupín (napr. *všetci ľudia*).

¹⁰ Ďalšie kvantifikátory logického štvorca sú prvkami týchto množín: *všetci*, *každý* – rastúcich kvantifikátorov sprava a klesajúcich zľava; *niekto*, *niektorý* – rastúcich kvantifikátorov sprava aj zľava; *nie všetci*, *nie každý* – rastúcich kvantifikátorov zľava a klesajúcich sprava (porov. Zouhar 2006b, s. 552).

na výživu svojich detí. Ide o to, že pri použití výrazu „všetky x “ máme na mysli všetky predmety s danou vlastnosťou *spoločne*, kým pri výraze „každé x “ ide o všetky takéto predmety *jednotlivo* (porov. Zouhar 2006a, s. 245). Negácia vety so všeobecným kvantifikátorom je pravdivá, keď je nepravdivý jej negovaný náprotivok: *Žiaden/ani jeden z rodičov neprispieva na výživu svojich detí podľa svojich schopností, možností a majetkových pomerov*. Rovnako nepravdivý výrok je aj v prípade použitia totalizátora *všetci* namiesto *obaja*: *Všetci rodičia prispievajú na výživu svojich detí podľa svojich schopností, možností a majetkových pomerov*. Univerzálne kvantifikátory sú modálne výrazy, ktoré podliehajú overovaniu vo vzťahu k vonkajšiemu svetu.

V prípade, že výrazy sú sémanticky aplikovateľné na každý element reštriktora jednotlivo, majú povahu **existenčného kvantifikátora**, t. j. výrazu, ktorý znamená $\exists xP(x)$: *existuje x , pre ktoré platí $P(x)$* . Existenčný kvantifikátor vyjadruje relatívne presný počet z určitého počtu. Znamená „existuje presne n objektov“: $\exists_{=n}$. Výraz *obaja* ako existenčný kvantifikátor značí „existujú presne dva objekty“. Sémanticky je teda podmienený prítomnosťou dvoch subjektov a slovesom prevažne s významom rezultatívnosti, jednorazovosti a uzavretosti deja. Napríklad veta *Madonna a Michael Jackson sa obaja narodili v roku 1958* znamená, že každý z oboch spomínaných spevákov sa narodil v roku 1958: aj *Madonna sa narodila v roku 1958* aj *Michael Jackson sa narodil v roku 1958*. Existenčné kvantifikátory sú možnostné modálne výrazy.

Výrazy nadobúdajú povahu **plávajúceho kvantifikátora** vtedy, keď signalizujú neukotvenosť kvantifikátora v syntaxi vety, napr. *Obaja budúci manželia Rogonoví sú z Ardanoviec*. = *Budúci manželia Rogonoví sú obaja z Ardanoviec*. = *Budúci manželia Rogonoví sú z Ardanoviec*. *Obaja*. Plávajúci kvantifikátor umožňuje signalizovať subjektívny postoj hovoriaceho, vyplývajúci z možnosti vyjadrenia stupňa aktivovanosti daného výrazu v aktuálnom členení výpovede. Výraz *obaja* sa dokonca poníma ako prototypický plávajúci kvantifikátor (porov. Šimík 2017), ktorý sa môže používať v iniciálovej, stredovej, finálnej pozícii a aj ako osamostatnený vetný člen.

3. KONCEPT ‘OBOCH’ Z HĽADISKA LINGVOPRAGMATIKY

Pri posudzovaní obsahových a funkčných stránok výrazov *obidvaja*, *obidva*, *obaja*, *oba*, *obe*, *obidve*, *obidvoje*, *oboje*, *obojaký*, *obojako* a pri analýze ich gramatikalizovaných parametrov vychádzame zo skutočnosti, že z lingvopragmatického hľadiska uvedené lexémy patria medzi **egocentriká**, t. j. jazykové výrazy, ktoré signalizujú implicitnú prítomnosť hovoriaceho, porov. *obidvaja* = dvaja + ‘niečo navyše’ (hovoriaci zdieľa obsah ‘*obidvoch*’ v zmysle ‘už spomínaní dvaja’; ‘tí dvaja, o ktorých bola reč’; ‘jeden aj druhý’; ‘ten i onen’; ‘úplný súbor tvoria dvaja’ a pod.). Lexémy teda obsahujú prvky, ktoré možno špecifikovať len pragmaticky. Sú výrazom abstrakčnej flexibili-

ty hovoriaceho (k termínu pozri Dolník 1999) a s číslovkou *dvaja* sú v asymetrickom vzťahu: *obidvaja* sú vždy *dvaja*, ale *dvaja* nie sú vždy *obidvaja*.

V prípade výrazov *obidvaja*, *obidva*, *obidve*, *obidvoje* poukázanie na počet *dvoch* je v slovenčine priamo súčasťou slovného tvaru (lexémy sa vyznačujú morfo-sémantickou transparentnosťou svojho tvaru), v prípade výrazov *obaja*, *oba*, *obe*, *oboje* počet dvoch entít ako konečného počtu sa implicitne predpokladá. Dôležitá je ale skutočnosť, že všetky výrazy spája ich deiktická podstata a dôraz na úplnosť. Deixa tu znamená späťnosť s predchádzajúcim kontextom a úplnosť je ponímaná v zmysle kompletnosti danej dvojčlennej množiny entít. Inými slovami, výrazy synkreticky evokujú okrem kvantovosti (kvantitatívnosti) a pronominálnosti¹¹ aj úplnosť istých faktov. Konzekvenciou je, že príznak kvantovosti zaraďuje výrazy do triedy čísloviek, príznak odkazovania spája výrazy so zámenami, príznak úplnosti zaraďuje výrazy medzi tzv. delimitátory (totalizátory) a príznak anaforickosti naznačuje efekt kontextovej implikácie, t. j. záveru vyvoditeľného z informačného vstupu a z kontextu. Sémantický obsah výrazov závisí od kontextu použitia. Výrazy sú štandardnými jazykovými prostriedkami endoforického odkazovania. Ako deiktické identifikátory nemajú pojmovo vymedzenú referenčnú oblasť (o ich referencii rozhodujú pragmatické faktory), a tak sa môžu zakaždým vzťahovať na aktualizované entity danej rečovej situácie, daného kontextu. Posudzované výrazy súvisia so situácnou sémantikou, základnou myšlienkou ktorej je skutočnosť, že výpovede sa týkajú určitých situácií, nie sveta ako celku (Zouhar 2007, s. 534). Sú založené na (a) relatívnej identite pomenovaní¹² odkazujúcich k dvom referentom alebo na (b) alterácii (vzťahu sémantickej obdoby, porov. *Mluvnice češtiny 3* 1987)¹³ realizovanej prostredníctvom pomenovaní dvoch odlišných entít z rovnakej sémantickej triedy a tiež vzťahujúcim sa k rôznym referentom. Numerické výrazy *obidvaja*, *obidva*, *obaja*, *oba*, *obe*, *obidve* sú vo funkcii totalizátora (lexikálnej jednotky označujúcej úplnosť celku majúceho dva rovnaké členy) a spájajú príznak homogenity a maximality, porov. *Obidve výstavy potrvajú do 30. marca*. Deiktické výrazy *oboje*, *obidvoje* sú vo funkcii alterátora a spájajú príznak nehomogenity a maximality, porov. *Dôchodok alebo nemocenská dávka. Ale nie obidvoje*.

Implicitná prítomnosť hovoriaceho navodzuje aspekt projekcie, inými slovami postoja, videnia, posudzovania a hodnotenia dvoch entít v danej rečovej situácii. Pri tom naznačuje, že počet entít je limitovaný počtom *dva* (projekcia konečného počtu dvoch entít majúcich niečo spoločné). V slovenčine sa projekcia konečného počtu

¹¹ Na skutočnosť, že samotný deiktický systém je egocentrický, upozorňoval už J. Lyons (1969, s. 275).

¹² Spojenie *obe ruky* referuje na *pravú* a *ľavú ruku*, ktoré nie sú síce absolútne identické, ale vyhovujú kritériu homogenity a keďže tvoria jeden pár, vyhovujú aj kritériu maximality.

¹³ *Alterácia* je vzťah medzi pomenovaniami predmetu reči, ktoré majú nejakú spoločnú vlastnosť (napr. príslušnosť k skupine, triede, druhu) a nikdy nie sú referenčne totožné (porov. *Mluvnice češtiny 3* 1987, s. 701 – 702).

prejavuje v koncepte ‘*oboch*’, ktorý odráža: (a) numericko-deiktické videnie sveta (v odkazovaní na počet dvoch relatívne identických entít jednej sémantickej triedy ako konečného a úplného počtu; úplnosť celku zahrňujúceho počet dva): *obidvaja, obidva, obaja, oba, obe, obidve*: *Obaja súperi bojujú o tretiu pozíciu v tabuľke.*; (b) súborovo-deiktické videnie sveta (v odkazovaní na súbor dvoch relatívne identických entít z jednej sémantickej triedy ako kompletnej skupiny): *obidvoje, oboje*: *Vstal a zamkol oboje dvere.*; (c) druhovo-deiktické videnie sveta (v odkazovaní na počet dvoch často protikladných entít jednej skupiny ako úplného počtu; druhovú úplnosť celku zahrňujúceho počet dva): *obojaký*: *Ťažko určiť, čo je sladké, čo slané a čo obojaké.*; (d) súhrnovo-deiktické videnie sveta (v deixe súhrnu dvoch rôznych entít z dvoch odlišných sémantických tried): *obidvoje, oboje*: *Ak sa spýtate obyvateľa Fínska, čo mu chýba najviac z jeho domova v zahraničí, odpoveď najčastejšie znie: ražný chlieb a sauna. Prípadne oboje.*

S konceptom ‘*oboch*’ sa teda v slovenčine spájajú dve entity, ktoré vyhovujú kritériu $\pm$ homogenity a maximality.¹⁴

Slovné spojenia s číslovkou *dva* sa prirodzene spájajú s totalizátormi *všetok, všetka, všetko, všetci, všetky*: *Všetky dve kiná pracovali nonstop*. Reflektujú videnie sveta v zmysle úplného súboru entít. V tomto prípade je spojenie *všetky dve* plne nahraditeľné výrazom *obe*: *Obe/obidve kiná pracovali nonstop*. Spojenie dvoch totalizátorov **všetky obe* sa nepripúšťa: **Všetky obe kiná pracovali nonstop*. Podobne sa spája číslovka *dva* s totalizátormi *každý, každá, každé*: *Hasiči sa stretávajú každé dva roky, vždy v inej obci*. Prezentuje videnie sveta v zmysle diskretnosti v súbore entít a tiež vylučuje spojenie dvoch totalizátorov **každé oba*: **Hasiči sa stretávajú každé oba roky, vždy v inej obci*.

3.1 Koncept ‘*oboch*’ v kontexte verbálneho konania hovoriaceho

Jadrom verbálneho konania hovoriaceho je egovzťahná interpretácia situácie, udalosti, faktu (k problematike pozri Sokolová 2023). Ide o obraz situácie vo vzťahnom rámci ega. Vzťahný systém, ktorý sa vyprofiloval ako skúsenostný komplex, znamená sformovaný a sprostredkovaný vzťah hovoriaceho k vonkajšiemu svetu. Skúsenostný komplex nasmerúva hovoriaceho uviesť výraz v adekvátnom kontexte, ktorý je jazykovo-znalostnou projekciou udalosti. Verbálne konanie hovoriaceho je založené na znalostiach tohto lexikálneho súboru a možnostiach a pravidlách používania jeho jednotiek.

¹⁴ Na porovnanie uvedme projekciu individuálnych predmetov daných počtom *dva*. Ukazuje sa, že v slovenčine sa koncept ‘*dvoch*’ prezentuje: (a) v numerickom obraze sveta (v počte dvoch entít): *dva, dvaja, dve*: *Uvedomujeme si, že na nás čakajú dvaja silní súperi.*; (b) v numericko-súborovom videní sveta (v súbore dvoch entít ako kompletnej skupiny): *dvoje*: *Z veľkej haly viedli chodby rozličnými smermi a dvoje schodov do vyšších poschodí.*; (c) v numericko-druhovom videní sveta (v súhrne dvoch druhov ako kompletnej skupiny): *dvojaký*: *Podľa vinárov hrozno zraje dvojakým spôsobom – nielen na slniečku, ale aj tým, že v chlade stráca vodu*. S číslovkou *dva* sa tradične dostávajú do súvislosti dve entity, ktoré vyhovujú kritériu $\pm$ homogenity a nemaximality.

Vo verbálnom konaní hovoriaceho sa egovzťažná interpretácia udalosti prejavuje už vo výbere prezentácie účasti participantov komunikácie, ktorý predostiera tieto možnosti: (a) hovoriaci udalosť vzťahuje aj na seba: *My obe vieme, že jedinou cestou môže, ba musí byť návrat k starým hodnotám. – Toto je však šport a obe chceme ísť na olympiádu.*; (b) hovoriaci udalosť vzťahuje na adresátov: *Vy obaja hráte v mojom živote dôležitú úlohu. – Obe ste začali svoju kariéru v mladom veku.*; (c) hovoriaci udalosť vzťahuje na osoby neprítomné v bezprostrednej komunikačnej situácii: *Oni obaja celý život chodili do práce a fyzicky drelí. – Obaja si doliečovali ťažké zranenia z predchádzajúcej sezóny.* Participantská účasť sa vždy vyjadruje osobným tvarom slovesa v pluráli. Výrazy *oba/obe* sa spájajú s príslušnými osobnými zámenami prvej, druhej alebo tretej osoby buď explicitne alebo implicitne. Vo verbálnom konaní hovoriaceho môžeme teda identifikovať jednak zdieľanie faktov na základe jeho internej alebo externej participantskej spoluúčasti.

V prípade zdieľania faktov na základe internej participantskej spoluúčasti hovoriaci zdieľa subjektívny názor na fakty v introspektívnej reflexii: *Nebol to dnes žiadny veľký zápas. Obidve sme hrali nervózne, pričom ja som bola o trochu lepšia.* Koncept internej participantskej spoluúčasti hovoriaceho explicitne vyjadruje komitatívna konštrukcia: *Myslím, že obidvaja s Michalom sme cítili, že naše cesty sa rozchádzajú.* ≈ ja a Michal; – *Toto bola pre mňa – pre nás obe – skúsenosť, ktorá nám zmenila život.* ≈ pre mňa a pre teba, resp. pre mňa a pre ňu. Konštrukcie *obidvaja/obidve s Y* a *my obe/obaja* patria do sféry egoprezentácie hovoriaceho, ktorý v nej vystupuje ako inkorporovaný participant₁. Konštrukcia *obidvaja/obidve s Y* fakticky označuje *ja* a *Y*, čím sa zaraďuje do typu koordinatívneho komitatívu. Formálne konštrukcia najprv uvádza, že počet účastníkov je viac ako jeden, t. j. hovoriaci a participant₂, a potom lexikálne konkretizuje len participanta₂, ktorý je započítaný akoby dvakrát – raz vo výraze *obidvaja/obidve* a druhý raz v konštrukcii *s Y*. Interná participantská účasť sa viaže na použitie 1. slovesnej osoby vo výpovedi.

V prípade zdieľania faktov na základe externej participantskej spoluúčasti, hovoriaci spravidla naplňa vlastnú potrebu vyjadriť sa k faktom na vecnej rovine, a to buď ako: (i) všeobecný, generalizovaný názor na fakty, ktoré označuje ako reálne: *Žena a muž sú obaja ľudia, ale funkcie, ktoré plnia v živote, nie sú totožné.*; (ii) vlastný, subjektívny názor na fakty (hodnotí, analyzuje, vysvetľuje, zdôvodňuje, vyvodzuje záver) z pohľadu pozorovateľa: *V príjemnom jesennom počasí obidve mužstvá príliš hernou úroveňou prítomných divákov nenadchli.* Externá participantská účasť sa viaže na použitie 3. slovesnej osoby vo výpovedi.

V súvislosti s verbálnym konaním vyvstáva otázka zdroja informácií, ktorým hovoriaci odôvodňuje svoje tvrdenie a na základe ktorého realizuje výpoveď. Ukazuje sa, že výrazy *obidvaja, obidva, obaja, oba, obe, obidve, obidvoje, oboje* sú prostriedkami funkčno-sémantickej kategórie evidenciality potvrdzujúcej, že hovoriaci interpretuje udalosť egovzťažne podľa primárnych alebo sekundárnych zdrojov. Primárnym zdrojom je pre hovoriaceho jeho vlastná percepčná činnosť a jeho znalosť

sveta, ktorá predstavuje egovzťazne generalizovaný skúsenostný komplex, ego-vzťazne organizovaný obraz sveta, individuálne skúsenosti hovoriaceho. Priame zdroje zahŕňajú: (i) informácie o udalosti, ktorej bol účastníkom: *So sestrou máme spoločné to, že obe sme učiteľky angličtiny.*; (ii) znalosti, skúsenostný komplex hovoriaceho, na základe ktorého vyjadruje vlastné stanovisko, vlastné presvedčenie o platnosti faktov: *O týždeň sa celkom určite obaja súperi predstavia znovu v Prešove.* Sekundárne zdroje znamenajú informácie o faktoch, ktoré sú v egovzťahnom rámci inej osoby. Keďže hovoriaci sa odvoláva na cudziu mienku, tvrdenie tretej osoby, reportívna projekčnosť je založená na deduktívnom usudzovaní, ktoré je argumentačne založené na autorite, resp. svedectve. Spravidla je signalizované pomocou autorizačného adverbialu zreteľa: *Podľa vyjadrenia pyrotechnika boli obidve delostrelecké strely schopné výbuchu. – V stanovisku mesta sú obaja kandidáti označení ako rovnocenní.,* resp. prostredníctvom kognitívnych slovies: *Myslí si, že obe krajiny môžu mať problémy. – Predpokladá sa, že obe jaskyne sú len časťou veľkého dosiaľ neprebádaného jaskynného systému.*

3.2 Koncept ‘oboch’ vo vzťahovej projekcii

Z egocentrickej podstaty výrazov vyplýva, že vyhovujú podmienkam presupozícií a majú vzťahovú podstatu. Z pohľadu extenzionálnej sémantiky totiž vyjadrujú binárne relácie medzi dvoma množinami – množinou dvoch konkrétnych entít a množinou všetkých entít danej triedy.

Pri posudzovaní vzťahovej projekcie ‘*oboch*’ vychádzame najprv z predpokladu, že východiskovým aspektom je perspektivizácia hovoriaceho ohľadne usporiadania entít, ktoré má charakter statickej alebo procesuálnej konceptuálnej schémy.¹⁵ Hovoriaci prezentuje udalosť so zameraním buď na fakty (ako fotografiu), alebo na procesuálnu stránku (ako film). Výpoveď *Spadla a zlomila si obe ruky v zápätí.* naznačuje dve perspektívy. V statickej perspektivizácii sa realizuje vzťahová projekcia faktov: $\approx$ zlomená je pravá a zlomená je aj ľavá ruka, v dynamickej perspektivizácii sa uplatňuje vzťahová projekcia dejov: $\approx$ zlomila si pravú a zlomila si aj ľavú ruku. V súvislosti s polyinterpretačným charakterom výpovedí s analyzovanými výrazmi budeme v ďalšom výklade rozlišovať dve vzťahové hľadiská – hľadisko vzťahu rovnako rozličných vecí a hľadisko vzťahu rovnako rozličných dejov. Korešponduje to s faktom, že kvantifikátory determinujú statické alebo dynamické príznaky.

3.2.1 Vzťahová projekcia rovnako rozličných vecí

Vzťahová projekcia dvoch premenných anticipuje mentálnu možnosť rozloženia (parcelácie) na komponenty alebo možnosť zloženia (kompletácie) komponentov v rámci daného súboru. Hovoríme o vzájomnej podmienenosti a súvislosti sub-

¹⁵ Statické schémy sa vzťahujú na objekty a štruktúry objektov, procesuálne schémy súvisia s priebehom deja, fázami deja a pod.

stancií a vzťahov v súbore. Ak hovoriaci vychádza zo vzťahu, zúži sa súbor možných substancií, keď sú východiskom substancie, zúži sa nevyhnutne súbor vzťahov (porov. Krupa 1980, s. 68).

Vzťahovosť dvoch entít z pohľadu posudzovaných výrazov je pragmaticky orientovaná a hovoriaci ju prezentuje z troch hľadísk: ako komplementárnosť, totálnosť a súhrnnosť premenných.

Komplementárnosť premenných je možnosť individuálneho posudzovania dvoch premenných súboru v zmysle 'jedna aj druhá jednotlivina navzájom'. Vzťahuje sa na numericko-deiktickú projekciu párových a počítateľných substantív, ktoré v rámci kvantitatívnych opozícií vyjadrujú buď proporcionálny alebo izolovaný protiklad (porov. Šefčík 2017).

(a) Proporcionálny protiklad značí analogický pomer medzi premennými jednej triedy. Vychádza z princípu komplementarity, čo znamená, že dve entity sa navzájom podmieňujú a dopĺňajú. Podstatou komplementarity dvoch prvkov v určitom systéme je teda to, že navzájom podmieňujú možnosť svojej pochopiteľnosti a existencie. Zároveň platí, že ak sú prvky určitého systému komplementárne, nevyhnutne to znamená, že nie sú absolútne totožné (porov. poznámku 7 o relatívnej podstate *totožnosti, rovnakosti, identickosti*). Proporcionálny protiklad sa týka párovej komplementarity a asociovej komplementarity.

Párová komplementarita si pri premenných označujúcich ľudí vyžaduje znalosti hovoriaceho o partnerských, príbuzenských, medziľudských vzťahoch: *Obe jej babičky totiž pochádzajú zo Slovenska.* ≈ teda nie akékoľvek babičky, ale babička z matkinej strany a babička z otcovej strany. Párové premenné, ktoré neoznačujú ľudí predpokladajú znalosti prototypového charakteru: *Pri páde si zlomil obe nohy a ruku.* ≈ pravú aj ľavú nohu; – *Starý most spájajúci oba brehy Dunaja nahradí nový.* ≈ pravý aj ľavý breh. Párová projekcia aktivuje aspekt prototypu. Nezameriava sa na počet *dva/dve*, ale na inherentný protiklad vychádzajúci zo sociálneho alebo priestorového usporiadania.

Asociovaná komplementarita sa vzťahuje na také počítateľné premenné, ktoré si vyžadujú konštantné znalosti hovoriaceho týkajúce sa diferencovania dvoch stránok jedného pomenovania: *Súboje oboch tímov sú väčšinou veľmi vyrovnané a končia sa najtesnejšími výsledkami.* ≈ tím domácich a tím hostí; – *Voličský preukaz sa dá vybaviť pre obe kolá volieb naraz.* ≈ prvé kolo a druhé kolo.

(b) Izolovaný protiklad znamená nepomer počítateľných premenných v zmysle dve konkrétne entity vs. ostatné entity jednej triedy: *Očakávané príjmy sa však nezaplnili, obe médiá vykazujú straty.* ≈ zo všetkých médií vykazujú straty (tieto) dve (konkrétne) médiá; – *Mne sa páčia obe mená.* ≈ zo všetkých mien sa mi páčia (tieto) dve (konkrétne) mená. To znamená, že nie všetky v zmysle nerozčleneného súboru, ale dve ako konkrétne jednotliviny. Číslovka *obe* sa zúčastňuje na vyčlenení časti z celého súboru. Podmnožiny *obe médiá*, *obe mená* tvoria binárne relácie s množinami *všetkých médií* a *všetkých mien*, pričom podmnožiny *obe médiá*, *obe mená* sú

spoločnými prvkami množín *všetkých médií* a *všetkých mien*. Zároveň reflektujú implicitnú opozíciu demonstratív (*tí, títo, tie, tieto*) a determinatív (*všetci, všetky*).

Totálnosť (úplnosť) premenných implicitne naznačuje význam 'úplný počet v súbore'. Vzťahuje sa na súborovo-deiktickú projekciu substantív pluráliá tantum a počítateľných substantív označujúcich dve entity z identickej sémantickej triedy: *Obidvoje dvere boli zamknuté*. ≈ všetky dvojce dvere (v izbe) boli zamknuté; – *Po druhej svetovej vojne som pravidelne navštevoval obidve kiná v Banskej Bystrici*. ≈ všetky dve kiná v Banskej Bystrici; – *Oba vchody magistrátu sa na noc zamykajú, rovnako aj všetky kancelárie*. ≈ všetky dva vchody magistrátu. To znamená, že všetky v zmysle rozčleneného súboru (jeden aj druhý) a priestorového ukotvenia, teda jednoznačnej lokalizácie. V prípade výrazov *oba, obidva, obe, oboje* význam 'všetkého' je inherentnou súčasťou lexikálneho významu.¹⁶ Spojenia *obidvoje dvere, obidve kiná, oba vchody* sú kontextovo senzitivne v zmysle nutnosti už spomínanej priestorovej lokalizovateľnosti.

Súhrnnosť premenných implicitne naznačuje význam 'súhrnu dvoch'. Vzťahuje sa na deikticko-súhrnovú projekciu a predpokladá integráciu dvoch odlišných premenných v ad hoc vzťahu: *Niektorí si zabudli zobrať peniaze, iní zas kartu a nájdu sa aj takí, ktorí v bankomate nechajú oboje*. ≈ niektorí v bankomate zabudnú aj peniaze aj kartu. – *Máme nároky na život, i povinnosti k životu. K obojemu sa hlásime*. ≈ hlásime sa aj k povinnosti aj k životu. Vo výrazoch *obojce, obidvoje* sa spájajú sémanticky rôzne pomenovania odkazujúce na súhrn rozčlenených entít. Súhrnnosť má kontingentnú povahu. Integrácia je komutatívna, lebo má nezávislé poradie premenných (operandov, entít), t. j. zmena poradia argumentov nemení pravdivostnú hodnotu výpovede: *Videl som, že sa chveje – hnevom, bolesťou či obidvojím*. = *Videl som, že sa chveje – bolesťou, hnevom či obidvojím*. Výraz *obojce* je tu anaforickou premennou k výrazom *hnev* a *bolesť*, ktoré sú, ako sme spomínali, vo vzťahu alterácie dvoch antecedentov.

Projekcia komplementárnosti a totálnosti premenných súvisí s epistemologickým myšlienkovým rámcom jednoty súhlasnosti (homogenity a maximality), súhrnnosť premenných je daná epistemologickým myšlienkovým rámcom jednoty rozdielnosti (t. j. nehomogenity a maximality).

3.2.2 Vzťahová projekcia rovnako rozličných dejov

Dynamická projekcia udalostí¹⁷ značí jej procesuálnu stránku a v súvislosti s výrazmi *oba/obe* predpokladá buď dvoch pôvodcov (tvoriacich definovanú dvojprvkovú množinu) a identitu dvoch rôznych dejov (a použitie plurálovej formy slo-

¹⁶ Na rozdiel od čísloviek *dva/dve* a *dvaja*, ktoré, ako sme už spomínali, význam 'všetkého' vyjadrujú adherentne, pomocou totalizátora *všetci, všetky*: *V Košiciach študovali všetci dvaja československí kozmonauti, ktorí boli vo vesmíre*.

¹⁷ *Udalosť* vymedzujeme ako základnú jednotku dejovosti. Okrem generických viet hovoriaci referuje o konkrétnej udalosti.

vesa), alebo jedného pôvodcu deja a dva objekty v projekcii dvoch časových úsekov jedného deja (nutnosť použitia singulárovej formy slovesa a možnosť použitia adverbii *najprv* – *potom*).

Tvar prísudku obsahuje ukazovatele postoja hovoriaceho k výpovedi a ukazovatele časovej a priestorovej zaradenosti. Sú to kategórie modálne (spôsob) a deiktické (časová a priestorová charakteristika deja). Udalosť hovoriaci prezentuje buď ako reálnu alebo ireálnu. Reálnosť značí posúdenie dejov ako skutočných, faktívnych.¹⁸ Realizuje sa prevažne v indikatíve prézenta a préterita. V indikatíve prézenta sa stretáva imperfektívna aspektová forma, indikatívna modusová forma a prézentná temporálna forma. Indikatív préterita spája perfektívnu/imperfektívnu aspektovú formu, indikatívnu modusovú formu a préteritnú temporálnu formu. Ireálnosť, t. j. nefaktívnosť znamená vyhodnotenie dejov buď ako hypotetických, ktoré prezentuje kondicionál, alebo nehypotetických, ktoré vyjadruje imperatív a futúrum. Ireálnosť realizácie deja často signalizuje prirovnávací konjunktív *akoby*: *Po zmene strán akoby obaja tréneri zakázali útočiť.* ≈ *akoby zakázal útočiť tréner₁* a *akoby zakázal útočiť tréner₂*.

Dvaja pôvodcovia a identita dvoch rôznych dejov umožňujú simultánnu, recipročnú a rezultatívnu projekciu. Sukcesívna projekcia naznačuje jedného pôvodcu deja a dva objekty v projekcii dvoch časových úsekov jedného deja (singulárová forma slovesa).

Simultánnosť dejov znamená súbežnosť, časovú a priestorovú zhodu priebehu dvoch dejov: *Obidvaja pacienti sa zotavujú doma.* ≈ *doma sa zotavuje pacient₁* a zároveň sa doma zotavuje pacient₂. Pri verifikácii vety sa berú do úvahy len pacienti, ktorí sa zotavujú doma. Negácia sa vzťahuje na predikát a vyžaduje si prototypicky distributívnu interpretáciu oboch premenných: *Ani jeden z (oboch) pacientov sa nezotavuje doma.* Simultánnosť sa môže vzťahovať na pôvodcu deja (*obidve*) aj na samotný dej (predikát v podobe lexikalizovaného spojenia *uprieť oči*): *Obidve sme upreli oči na dvere šéfovej kancelárie.*

Reciprocita dejov značí, že deje sú v projekcii obojstrannej vzájomnosti: *Napokon si obaja muži podali ruky.* ≈ muž₁ podal ruku mužovi₂ a (zároveň) muž₂ podal ruku mužovi₁; – *Obe ženy sa objali.* ≈ žena₁ objala ženu₂ a (zároveň) žena₂ objala ženu₁. Význam reciprocitu si vyžaduje recipročné reflexívne slovesá. Reciprocita sa spája so sukcesívnosťou alebo simultánnosťou.

Rezultatívnosť dejov anticipuje dovŕšenosť, ukončenosť realizácie dejov: *Obe lode boli vážne poškodené.* ≈ poškodená bola loď₁ a poškodená bola aj loď₂. – *Výšetrovatelia predpokladajú, že obaja muži zomreli na podchladenie.* ≈ na podchladenie zomrel muž₁ a na podchladenie zomrel aj muž₂.

Simultánna, recipročná aj rezultatívna projekcia dvoch identických dejov súvisí s epistemologickým myšlienkovým rámcom jednoty súhlasnosti. Slovné spojenia sú vo vetách v pozícii podmetu.

¹⁸ *Faktívnosť* je posúdenie realizácie deja na základe skúsenosti ako skutočného (Sokolová – Žigo 2014, s. 90).

Duratívna povaha deja umožňuje členiť líniu deja v zmysle sukcesívnosti ako-by dvoch identických dejov zacielených na dva objekty: *Pobozkal mamu na obe líca a rozlúčil sa.* ≈ najprv matku pobozkal na jedno líce₁, potom pobozkal na druhé líce₂. Sukcesívna projekcia nešpecifikovaných úsekov deja značí chronologickú a konvergentnú postupnosť s konečnou limitou *dvoch*. Tiež súvisí s epistemologickým myšlienkovým rámcom jednoty súhlasnosti. Slovné spojenia sú vo vetách spravidla v pozícii adverbálie.

4. KONCEPT ‘OBOCH’ V KONTEXTE PLURALITY SLOVNÝCH SPOJENÍ

V sémantickom protiklade jednot’ – mnohosť sa rozlišuje plurál limitovaného (malého) množstva a plurál nelimitovaného (veľkého) množstva.¹⁹ Výrazy *obidvaja, obidva, obaja, oba, obe, obidve, obidvoje, oboje* sa vyznačujú príznakmi homogenity a maximality napriek tomu, že sa asociujú s predstavou malého a limitovaného množstva entít. So substantívami tvoria empiricky zmysluplné slovné spojenia, ktoré sú spravidla zakotvené v paradigme štyroch členov: singulár mena a singulár slovného spojenia ↔ [plurál slovného spojenia] a plurál mena. Na škále medzi singulárom a plurálom mena sa uplatňuje interná aj externá diferenciácia slovných spojení, a to z hľadiska počtu, páru, súboru a druhu/týpu. Empirická zmysluplnosť vyplýva z charakteru zloženosti entít a lexikálno-sémantického typu substantív: (a) počítateľných substantív; (b) párových substantív; (c) substantív pluráliá tantum; (d) substantív singuláriá tantum.

(a) Počítateľné substantíva význam množstva realizujú gramaticky, alebo adherentne, a to buď pomocou lexikálneho kvantifikátora, alebo pomocou lexikálneho klasifikátora. Lexikálne kvantifikátory typu *skupina, pelotón, zväzok, kartón, krádeľ, črieda*²⁰ a pod. denotujú ohraničené počty objektov. Zónu jednotlivín tvorí singulár mena (*dieťa*) a singulár slovného spojenia (napr. *skupina detí* – kvantifikátor je v tvare singuláru, meno má tvar plurálu), zónu mnohosti vyjadruje plurál slovného spojenia (*obe deti* – *obe skupiny detí*: kvantifikátor aj meno sú v tvare plurálu)²¹ a plurál mena (*deti*). Ako vidieť, počítateľné substantíva tvoria päťčlenné paradigmy pozostávajúce zo singulárových tvarov mena (*dieťa*), singulárových tvarov slovného spojenia (*skupina detí*), dvoch plurálových tvarov slovného spojenia (*obe deti, obe*

¹⁹ Plurál malého množstva je gramaticky špecifikovaný a ohraničený na slovné spojenia s číslovkami *dva, oba, tri a štyri* a tvarmi substantív v nominatíve plurálu (na rozdiel od ruštiny, kde tvary substantív sú v genitíve singuláru: *два, оба, три, четыре мальчика*).

²⁰ *Skupinovosť* je vlastným lexikálnym významom daných substantív.

²¹ Keďže aplikovateľnosť výrazov *obidvaja, obidva, obaja, oba, obe, obidve, obidvoje, oboje* je viazaná na konkrétnu udalosť, ich prítomnosť v paradigmách ako systémovom jave budeme prezentovať čisto teoreticky, bez ohľadu na potvrdenie vo výpovediach. Spojenia s danými výrazmi uvádzame v hranatých zátvorkách.

skupiny detí) a plurálového tvaru mena (*deti*). Podobne: *pretekár* – *pelotón pretekárov* ↔ [*obaja pretekári* – *oba pelotóny pretekárov*] – *pretekári*; *kľúč* – *zväzok kľúčov* ↔ [*oba kľúče* – *oba zväzky kľúčov*] – *kľúče*; *cigareta* – *kartón cigariet* ↔ [*obe cigarety* – *oba kartóny cigariet*] – *cigarety*; *vrana* – *krdel' vrán* ↔ [*obe vrany* – *oba krdle vrán*] – *vrany*; *kôň* – *črieda koní* ↔ [*oba kone* – *obe čriedy koní*] – *kone*. Lexikálne klasifikátory *druh*, *typ* denotujú druhovo alebo typovo ohraničené počty objektov. Druhový singulár slovného spojenia počítateľných substantív umožňuje variantné tvary substantív, a to v genitíve singuláru alebo genitíve plurálu: *text* – *druh textu/ druh textov* ↔ [*oba texty* – *oba druhy textov*] – *texty*; *hrad* – *typ hradu/typ hradov* ↔ [*oba hrady* – *oba typy hradov*] – *hrady*; *zvier* – *druh zvieraťa/druh zvierať* ↔ [*obe zvieraťá* – *oba druhy zvierať*] – *zvieraťá*.

(b) Párové substantíva vzťah jednotlivina ↔ mnohosť realizujú gramaticky, čo znamená, že síce tvoria opozíciu tvarov v singulári oproti plurálu, ale jednotlivina je v párovej projekcii vždy ponímaná vo vylučovacom zmysle – buď jeden alebo druhý člen páru. Plurál v párovej projekcii značí jedny aj druhé členy párov. Plurál limitovaného množstva v párovej projekcii sa realizuje prostredníctvom lexikálneho kvantifikátora *pár*, *obaja*, *oba*, *obe*: *ruka* (= pravá alebo ľavá ruka) ↔ *obe ruky* (= pravá aj ľavá ruka) – *oba páry rúk* (dve pravé a dve ľavé ruky) – *ruky*. Podobne: *topánka* ↔ [*obe topánky* – *oba páry topánok*] – *topánky*; *lyža* ↔ [*obe lyže* – *oba páry lyží*] – *lyže*; *oko* ↔ [*obe oči* – *oba páry očí*] – *oči*; *krídlo* ↔ [*obe krídla* – *oba páry krídel*] – *krídla* a pod. Plurál párových substantív umožňuje tri projekcie: (i) párovú projekciu: *Položil som obe ruky na stôl*. ≈ pravú aj ľavú ruku;²² (ii) číselnú projekciu dvoch: *Zo stodvadsať ľudí sa dvíhajú dve ruky*. ≈ zdvihnutím ruky hlasujú dvaja ľudia; (iii) súborovú projekciu – s vymedzovacími zámenami *všetci*, *všetky*: *Na prekvapenie sa zdvihli takmer všetky ruky*. ≈ zdvihnutím ruky hlasovali takmer všetci prítomní. Plurálové tvary mena v číselnej a súborovej projekcii strácajú význam párovosti. Substantíva tvoria paradigmy singulárových tvarov mena a plurálových tvarov slovného spojenia a mena.

(c) Substantíva pluráliá tantum význam jednotliviny realizujú synteticky (*dvere*), resp. analyticky pomocou kvantifikátora *jedny* (*jedny dvere*) alebo druhového klasifikátora *typ/druh*: (*typ dverí/druh dverí*). Druhový plurál slovného spojenia s pomnožnými podstatnými menami umožňuje v nominatíve alebo genitíve plurálu variantné tvary substantív: *oboje dvere/dverí/dvier* – *oba typy dverí/dvier* – *oba druhy dverí/dvier*. Paradigmu završuje plurál mena *dvere*. Podobne: (*jedny*) *preteky* – *druh pretekov* ↔ [*oboje preteky/oboje pretekov* – *oba druhy pretekov*] – *preteky*; (*jedny*) *nožnice* – *druh nožníc* ↔ [*oboje nožnice/oboje nožníc* – *oba druhy nožníc*] –

²² Párová projekcia vylučuje spájateľnosť s číslovkami *dva/dve* (význam *dvoch* v zmysle konečného počtu je fixovaný a implicitne zahrnutý) a s vymedzovacími zámenami *všetci*, *všetky*, lebo počet *dvoch* sa neponíma v zmysle úplného súboru: **Položil som dve ruky na stôl*. – **Položil som všetky ruky na stôl*.

nožnice; (jedny) noviny – druh novín ↔ [oboje noviny/oboje novín – oba druhy novín] – noviny. Paradigmy pomnožných podstatných mien teda tvoria analytické tvary singuláru mena (jedny dvere), singuláru slovného spojenia (typ/druh dverí), plurálu slovného spojenia (oboje dvere/oboje dverí//dvier, oba typy/druhy dverí//dvier) a syntetický tvar plurálu mena (dvere).

(d) Substantíva singuláriá tantum prostredníctvom lexikálneho klasifikátora denotujú druhovú alebo typovú ohraničenosť objektov a pomocou lexikálneho kvantifikátora presne deklarujú ich množstvo.

(i) Abstraktá tvoria opozíciu: singulár mien – druhový singulár slovného spojenia – druhový plurál slovného spojenia: *poznanie* – druh *poznania* ↔ [oba druhy poznania] – druhy poznania; *múdrost'* – druh *múdrosti* ↔ [oba druhy múdrosti] – druhy múdrosti; *intimita* – druh/typ *intimity* ↔ [oba druhy/typy intimity] – druhy/typy intimity; *zločin* – druh *zločinu* ↔ [oba druhy zločinu] – druhy zločinu; *morálka* – druh *morálky* ↔ [oba druhy morálky]²³ – druhy morálky.

(ii) Látkové mená umožňujú druhovú aj kvantitatívnu špecifikáciu. Druhový plurál slovného spojenia u časti substantív je realizovaný spravidla analyticky: *múka* – druh *múky* ↔ [oba druhy múky] – druhy múky; *cukor* – druh *cukru* ↔ [oba druhy cukru] – druhy cukru; *fazuľa* – druh *fazole* ↔ [oba druhy fazule] – druhy fazule; u časti substantív umožňuje v genitíve singuláru alebo genitíve plurálu variantné tvary substantív a syntetický plurál mena: *syr* – druh *syra*/druh *syrov* ↔ [oba syry – oba druhy syra/oba druhy syrov] – syry; *víno* – druh *vína*/druh *vín* ↔ [obe vína – oba druhy vín] – vína. Látkové substantíva prostredníctvom lexikálneho kvantifikátora *šálka*, *pohár*, *kartón*, *plechovka*, *štipka*, *lyžička*, *sud*, *kilo*, *snop* a pod. denotujú ohraničený objem, váhu, množstvo objektov. Opozíciu tvorí singulár látkových mien – kvantifikovaný singulár slovného spojenia – kvantifikovaný plurál slovného spojenia: *káva* – šálka *kávy* ↔ [obe šálky kávy] – šálky kávy; *pivo* – sud *piva* ↔ [oba sudy piva] – sudy piva; *med* – lyžička *medu* ↔ [obe lyžičky medu] – lyžičky medu; *slama* – snop *slamy* ↔ [oba snopy slamy] – snopy slamy. Spojenia *dve kávy*, *dve piva* si zachovávajú význam viazaného množstva a implicitne predpokladajú príslušné kvantifikátory (*dve šálky kávy*, *dva poháre piva*).

(iii) Hromadné substantíva tvoria opozíciu singulár hromadných mien: *mužstvo* ↔ plurál slovného spojenia *obe mužstvá* – plurál hromadných mien: *mužstvá*. Podobne: *tím* ↔ [oba tímy] – tímy; *družstvo* ↔ [obe družstvá] – družstvá; *zväz* ↔ [oba zväzy] – zväzy; *súbor* ↔ [oba súbory] – súbory; *skupina* ↔ [obe skupiny] – skupiny; *trieda* ↔ [obe triedy] – triedy. Tieto podstatné mená majú aj tvary syntetického plurálu. Časť substantív s významom kolektívnosti umožňuje druhovú špecifikáciu: *hmyz* – druh *hmyzu* ↔ [oba druhy hmyzu] – druhy hmyzu; *obuv* – druh *obuvi* ↔ [oba

²³ Spojenie *oba druhy morálky* značí aj *obojakú morálku*, resp. *dvojakú morálku*. Druhovú číslovku *obojaký* je v slovenčine súčasťou ustálených spojení: *obojaké spoluhlásky*, *obojaké právo*, *obojaké pohlavie*, *obojaký spôsob prijímania*.

druhy obuvi] – *druhy obuvi*. Význam plurálu vyjadrujú dané podstatné mená len pomocou analytických tvarov.

Pluralitu slovných spojení podľa charakteru zloženosti entít a lexikálno-sémantických typov substantív tvoria: (a) slovné spojenia numerickej zloženosti: *oba programy*; (b) slovné spojenia párovej zloženosti: *obe oči*; (c) slovné spojenia súborovej zloženosti: *oboje noviny*; (d) slovné spojenia druhovo deklarovanej zloženosti: *oba druhy poistenia*; (e) slovné spojenia kvantitatívne deklarovanej zloženosti: *obe stáda oviec*.

5. ZÁVER

Výrazy *obidvaja*, *obidva*, *obaja*, *oba*, *obe*, *obidve*, *obidvoje*, *oboje*, *obojaký*, *obojako* tvoria v slovenčine koherentný a konzistentný inventár lexikálnych prostriedkov. V jeho priestore sa vyšpecifikovali najmä numerické a okrajovo deiktické slovnodruhovú určenia. Numeráliá sa v slovenčine profilujú v zmysle súboru, vymedzovacie zámená v zmysle súhrnu. Dominantné postavenie majú základné a skupinovú číslovky, ktoré vyjadrujú kontextovo viazanú, určitú a limitovanú početnosť ‘*oboch*’. Inými slovami, naplňajú princípy homogenity, maximality a salience (viazaného vzťahu medzi formou jazykových výrazov a ich referentmi s dôrazom na faktivnosť a konkrétnu skúsenosť mnohosti). Druhovú číslovky a vymedzovacie substantívne zámená sú maximatívne a akceptujú princíp nehomogenity.

Potvrdilo sa, že v rámci numerálií pôsobí súčinnosť gramatických, slovtvorných a lexikálnych spôsobov vyjadrenia mnohosti. Číslovky tvoria so substantívami kongruentné slovné spojenia. V ich kolokačnom profile sú tieto lexikálno-sémantické typy substantív: (a) počítateľné substantíva; (b) párové substantíva; (c) substantíva pluráliá tantum; (d) substantíva singuláriá tantum (abstraktá, látkové a hromadné substantíva).

Významy mnohosti okrem syntetického spôsobu vyjadrenia pomocou flexie umožňujú (a v prípade substantív singuláriá tantum si vyžadujú) lexikálnu intervenciu, inými slovami, použitie osobitných lexikálnych prostriedkov, ktoré majú schopnosť diferencovať opozíciu jednotlivosť – mnohosť. Ukázalo sa, že v slovenčine je silne ukotvené kognitívne vymedzenie druhovej a kvantitatívne viazanej mnohosti. Pluralitu podľa charakteru zloženosti entít a lexikálno-sémantických typov substantív tvoria slovné spojenia numerickej (*oba jazyky*), párovej (*obe topánky*), súborovej (*oboje preteky*), druhovo deklarovanej mnohosti (*oba druhy dôchodkov*) a kvantitatívne deklarovanej mnohosti (*oba poháre šampanského*).

V jazykovom a diskurzívnom priestore súčasnej slovenčiny koncept ‘*oboch*’ v číslovkovom vyjadrení súvisí s jednotou súhlasnosti *dvoch*, koncept ‘*oboch*’ v deikticko-substantívnom vyjadrení sa vzťahuje na integráciu rozdielnosti *dvoch*. To znamená, že kardinalita množiny či už v súborovom, súhrnovom alebo druhovom určení je vždy limitovaná počtom dvoch premenných, ktoré sú vo vzťahu relatívnej

identity, podobnosti (alterácie) a rozdielnosti. Spojitosť individuálnych entít v numerickom vyjadrení má viac-menej konštantnú povahu. Slovné spojenia s posudzovanými výrazmi sú divízne (členiteľnosť súboru *oboch kníh* navodzuje fakt aj jednej knihy), nekumulatívne (*obe knihy* + *obe knihy* ≠ *obe knihy*) a atomické (jazykovo sa dá vymedziť minimálna časť *oboch kníh* – kniha).²⁴ Spojitosť individuálnych entít v deiktickom vyjadrení má kontingentnú povahu. Je to dané tým, že závisia výhradne od komunikačno-pragmatických faktorov a obsah nadobúdajú iba v konkrétnom použití. *Oboje* vyjadruje sumarizujúci vzťah medzi dvoma explicitne vyjadrenými eventualitami. Substantíva sú vo vzťahu slabšej logickej disjunkcie. Obe výberové možnosti prichádzajú do úvahy, vzájomne sa nevylučujú a nie sú striktné inkompatibilné. Spojka *alebo* nestavia dve alternatívy proti sebe, ale kladie ich akoby vedľa seba. Hovoriaci potvrdzuje možnosť pôsobenia oboch faktorov súčasne v zmysle uplatnenia akoby tretej eventuality.

V súvislosti s možným polyinterpretačným charakterom výpovedí s analyzovanými výrazmi možno konštatovať, že pre vzťahové projekcie hovoriaceho je príznačná tolerancia nejednoznačnosti. To znamená, že hovoriaci na pozadí princípu kooperácie predpokladá kompetenciu adresáta rozpoznať nejednoznačnosť a tolerovať neistotu a rozpor, ktorý zo vzťahových projekcií môže vyplývať.

Bibliografia

CVRČEK, Václav (2017): Koligace. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: <https://www.czechency.org/slovník/KOLIGACE> [cit. 19. 06. 2024].

DANEŠ, František (2009): O spojce nebo. In: *Kultura a struktura českého jazyka*. Praha: Karolinum, s. 188 – 194.

DOČEKAL, Mojmir (2017a): Negativně polaritní výraz. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: [https://www.czechency.org/slovník/NEGATIVNĚ POLARITNÍ VÝRAZ](https://www.czechency.org/slovník/NEGATIVNĚ_POLARITNÍ_VÝRAZ) [cit. 31. 05. 2024].

DOČEKAL, Mojmir (2017b): Veridikalita. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: <https://www.czechency.org/slovník/VERIDIKALITA> [cit. 31. 05. 2024].

DOČEKAL, Mojmir (2017c): Počítatelné a nepočítatelné jméno. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: [https://www.czechency.org/slovník/POČÍTELNÉ A NEPOČÍTELNÉ JMÉNO](https://www.czechency.org/slovník/POČÍTELNÉ_A_NEPOČÍTELNÉ_JMÉNO) [cit. 15. 06. 2025].

DOLNÍK, Juraj (1999): Princíp ekvivalencie. In: *Princípy stavby, vývinu a fungovania slovenčiny*. Bratislava: Stimul – Centrum informatiky a vzdelávania FIF UK, s. 45 – 77.

FILKORN, Vojtech (1998): *Povaha súčasnej vedy a jej metódy*. Bratislava: Veda, vydavateľstvo SAV. 379 s.

FREGE, Gottlob (2011): Základy aritmetiky. In: Frege, G.: *Logická zkoumání. Základy aritmetiky*. Přeložil J. Fiala. Praha: Oikoymenh, s. 143 – 261.

²⁴ Divizitu, kumulativitu a atomicitu substantív bližšie ozrejmjuje M. Dočekal (2017c).

JAROŠOVÁ, Alexandra (ed.) (2021): *Slovník současného slovenského jazyka. O – Pn* [4. zv.]. Bratislava: Veda, vydavateľstvo SAV.

KRUPA, Viktor (1980): *Jednota a variabilita jazyka*. Bratislava: Veda, vydavateľstvo SAV. 166 s.

LYONS, John (1969): *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press. 519 p.

MAJTÁNOVÁ, Marie (1976): Postavenie duálu a používanie duálových tvarov v slovenčine predpisovného obdobia. In: *Slovenská reč*, roč. 41, č. 2, s. 77 – 84.

MILL, John Stuart (1963/1991): *Collected Works of J. S. Mill*. 33 Vols. J. M. Robson, et al. (eds.). Toronto: University of Toronto Press – London: Routledge and Kegan Paul.

Mluvnice češtiny 2 (1986). Tvarosloví. J. Petr (ved. red.). Praha: Academia. 536 s.

Mluvnice češtiny 3 (1987): Skladba. J. Petr (ved. red.). Praha: Academia. 748 s.

MRÁZ, Michal (2008): Kvantifikácia v prirodzenom jazyku (IX). In: *Organon F*, roč. 15, č. 1, s. 84 – 110.

OLIVA, Karel (2017): Kvantifikátor. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: <https://www.czechency.org/slovník/KVANTIFIKÁTOR> [cit. 22. 02. 2024].

OSOLSOBĚ, Klára (2017): Totalizátor. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: <https://www.czechency.org/slovník/TOTALIZÁTOR> [cit. 21. 04. 2024].

Slovenský národný korpus. Verzia prim-9.0-public-sane. Bratislava: Jazykovedný ústav Ľudovíta Štúra SAV 2018. Dostupný na: <http://korpus.juls.savba.sk>.

SOKOLOVÁ, Jana (2023): Pragmatické určenie *akoby* v egovzťažných súvislostiach. In: *Jazykovedný časopis*, roč. 74, č. 2, s. 459 – 476. DOI 10.2478/jazcas-2024-0004. [cit. 22. 02. 2024].

SOKOLOVÁ, Miloslava (2007): *Nový deklinačný systém slovenských substantív*. Prešov: Filozofická fakulta Prešovskej univerzity v Prešove. 338 s.

SOKOLOVÁ, Miloslava – ŽIGO, Pavol (2014): *Verbálne kategórie aspekt a tempus v slovenčine*. Bratislava: Veda, vydavateľstvo SAV. 320 s.

SOUSEDÍK, Prokop – SVOBODA, David (2014): Je číslo vlastnosť vnějších věcí? Mill – Frege – Kessler. In: *Organon F*, roč. 21, č. 1, s. 45 – 62.

ŠEFCÍK, Ondřej (2017): Opozice. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: <https://www.czechency.org/slovník/OPOZICE> [cit. 22. 04. 2024].

ŠIMÍK, Radek (2017): Plovoucí kvantifikátor. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: https://www.czechency.org/slovník/PLOVOUCÍ_KVANTIFIKÁTOR [cit. 21. 04. 2024].

ŠIMÍK, Radek – DOTLAČIL, Jakub (2017): Univerzální kvantifikátor. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupný na: https://www.czechency.org/slovník/UNIVERZÁLNÍ_KVANTIFIKÁTOR [cit. 21. 04. 2024].

VESELÝ, Vojtěch (2011): Obojí, oboje, obé, všechno. In: *Naše řeč*, roč. 94, č. 4, s. 181 – 193.

ZOUHAR, Marián (2006a): Kvantifikácia v prirodzenom jazyku (II). In: *Organon F*, roč. 13, č. 2, s. 232 – 251.

ZOUHAR, Marián (2006b): Kvantifikácia v prirodzenom jazyku (IV). In: *Organon F*, roč. 13, č. 4, s. 539 – 553.

Z HISTORIE EVROPSKÉ MEDICÍNSKÉ TERMINOLOGIE: ŘECKÉ ΚΟΛΟΝ, JEHO ETYMOLOGIE A DERIVÁTY¹

VÁCLAV BLAŽEK¹ – MARTINA ŠMEJKALOVÁ²

¹Ústav jazykovědy a baltistiky, Filozofická fakulta, Masarykova univerzita, Brno, Česko

²Katedra českého jazyka, Pedagogická fakulta, Univerzita Karlova, Praha, Česko

BLAŽEK, Václav – ŠMEJKALOVÁ, Martina: From the history of European medical terminology: the Greek word κόλον, its etymology and derivatives. *Jazykovedný časopis (Journal of Linguistics)*, 2025, Vol. 76, No. 2, pp. 412–432.

Abstract: The article introduces the Greek anatomical term κόλον in the context of its internal structure and etymological relations in other Indo-European branches. The studied anatomical terms are widely discussed from the point of view of their semantics. The second part follows the fate of Greek κόλον, when it was adapted into other languages, first into Latin. The Latin adaptation *colon*, reinterpreted as the indeclinable *n*-stem, became an integral part of medieval medicine terminology. From Latin, it was adapted into French, and later into English. Finally, the most frequent medicine term consisting of *colon* is discussed, namely, two variants of the compound *coloscopy* vs. *colonoscopy*.

Keywords: semantic field, etymology, adaptation, anatomical and medical terminology.

0. MOTIVACE

Fascinujícímu tématu řeckého anatomického termínu κόλον a jeho derivátům se věnují dvě naše přípravné studie. Jedna se zaměřuje na indoevropský kontext, včetně detailní diskuse z pohledu historické fonetiky, morfologie, sémantiky a etymologie (Blažek – Šmejkalová 2023). Druhá se koncentruje na uplatnění termínu a jeho derivátů v medicínské terminologii od antiky po současnost (Šmejkalová – Blažek 2023). Předkládaná studie nabízí propojení obou přístupů, kdy jsme se pokusili o syntézu etymologické a medicínské perspektivy, které jednotlivě dominovaly oběma předchozím již publikovaným článkům k tématu.

¹ Poděkování za podklady a analýzy: Tereza Valachovičová a Marek Bajger (studentské pomocné síly na KČJ PeDF UK, výskyt termínů *koloskopie/kolonoskopie* v uživatelské praxi).

Tato publikace vznikla s podporou Ministerstva školství, mládeže a tělovýchovy – Institucionální podpora dlouhodobého rozvoje výzkumných organizací Cooperatio HUM/Linguistics / Linguistics – Karlova univerzita, Pedagogická fakulta (2025).

1. DOKUMENTACE

První použití řeckého slova κόλον registrujeme v komedii *Jezdci*², kterou Aristofanés dokončil r. 424 př. Kr. Nejpodrobněji, v anatomickém kontextu, termín definuje Aristotelés v díle *Morfologie živočichů*³, které vzniklo kolem r. 350 př. Kr. Pro výklad lidské anatomie použil termín Níkandros z Kolofónu (3.–2. st. př. Kr.) ve svém díle *Alexifarmaka*⁴ o pokrmech, jidech a protijidech.

Ve složeninách se objevuje slovní kmen κολο-:

κολούλια n. pl. “plž (třída Gastropoda, doslova ‘břichonožci’) z kmene měkkýšů” [Xenokratés, 1. st. po Kr., apud Orobasius 2, 58, 79], kde druhou složku

² Aristofanés: *Jezdci* = ř. Ἰππῆς = lat. *Equites* (doplňujeme, že k použití střívek a dršťek jako tresta-jících nástrojů je zde vyzván Jelitář):

⁴⁵³παῖ' αὐτὸν ἀνδρείοτατα, καὶ

“Bij ho vši svou silou!

⁴⁵⁴γάστριζε καὶ τοῖς ἐντέροις

Udeř ho do břicha a svými střívky

⁴⁵⁵καὶ τοῖς κόλοις,

A svými **dršťkami!**

⁴⁵⁶χῶπως κολᾷ τὸν ἄνδρα.

Ztrestej toho muže oběma rukama!”

Edice F. W. Hall – W. M. Geldart (Aristophanes 1907)

Doslovný český překlad s přihlédnutím k volnému anglickému překladu Eugene O'Neill, Jr. (1938)

³ Aristotelés: *Morfologie živočichů* = ř. Περὶ ζῶων μορίων = lat. *De partibus animalium*:

3.14. Πᾶσι δὲ τοῖς μὴ εὐθυεντέροις προῖοῦσιν εὐρύτερον γίνεται τὸ μόριον τοῦτο, καὶ τὸ καλούμενον **κόλον** ἔχουσι, καὶ τοῦ ἐντέρου τυφλὸν τι καὶ ὀγκῶδες, εἴτ' ἐκ τούτου πάλιν στενώτερον καὶ εἰλιγμένον. Τὸ δὲ μετὰ τοῦτο εὐθὺ πρὸς τὴν ἔξοδον διατείνει τοῦ περιττώματος, καὶ τοῖς μὲν τοῦτο τὸ μόριον, ὁ καλούμενος ἄρχος, κνισσώδης ἐστὶ, τοῖς δ' ἀπίμελος. <https://el.wikisource.org/wiki/Περὶ_ζῶων_μορίων/3>

“Střevo, vyjma těch zvířat, kde je přímé, se neustále rozšiřuje, jak postupujeme dále od žaludku, a přichází k tomu, co se nazývá **kolon**, a ke slepému střevu. Poté se opět stává užším a složitějším. Pak pokračuje rovná část, která běží vpravo k otvoru. Tento otvor je znám jako konečník (anus) a u některých živočichů bývá obklopen tukem, u jiných nikoliv.”

Český překlad s přihlédnutím k anglickému překladu Williama Ogle (1912)

⁴ Níkandros: lat. *Alexipharmaca*

¹⁶τοῖο δὲ πάντα χαλινὰ καὶ οὐρανόεσσαν ὑπὴνιν

¹⁷οὐλὰ θ' ὑποστύφει χολόεν ποτόν, ἀμφὶ δὲ πρώτοις

¹⁸εἰλύεται στέρνοισι κακῇ ἀλάλγῃ βαρῦνον

¹⁹φῶτ' ἐπικαρδιόοντα· δὴ δ' ἐπιδάκνεται ἄκρον

²⁰νειαίρης, ἄκλειστον ἀειρόμενον στόμα γαστρός,

²¹τεύχεος ἦν κραδίην ἐπιδορπίου οἱ δὲ δοχαίην

²²κλείουσι στομάχοιο, πύλῃ δ' ἐπικέκλιται ἀρχαῖς

²³πρῶτα **κόλων** ὅθι πᾶσα βροτῶν ἄλις ἐμφέρεται δαῖς.

http://www.poesialatina.it/_ns/Greek/testi/Nicander/Alexipharmaca.html

“Píjákovy čelisti, patro úst a jeho dásně, to vše je staženo hořkým douškem, jak se tiskne o hořejšek hrudníku, drtíc se zlým dušením muže v mukách pálení žáhy. Hořejšek břicha ovládne bolest, narůstající, nezavřená ústa spodního žaludku, kterým někteří říkají „srdce“ trávicího traktu, jiní „příjemce“ žaludku – a brána je zavřena bezprostředně na počátku **střev**, kam je mužovo jídlo ve vším svém množství doneseno.”

Český překlad s přihlédnutím k anglickému překladu A. S. F. Gowa – A. F. Scholfielda (1953)

patrně tvoří substantivum označující “boty”, jež je známo z glosy ὀλίας· τοὺς καρπατινοὺς τόμους [Hesychios, c. 500 po Kr.];

κολοκορδόκολα f. „dršťky“ [Filodémos, *Anthologia Graeca* IV: epigram 10.103 s komickým obsahem], kde za první složkou zřejmě následuje komponenta χορδή f. “střevo, vnitřnosti” [Ferekratés, 5. st. př. Kr.] s iniciálou asimilovanou pod vlivem prvního a posledního κ, jež patří třetí komponentě identické první (Chantraine 1968, s. 557).

2. ETYMOLOGIE

Hlubším etymologickým výkladům se etymologické slovníky řečtiny spíše vyhýbají, a tak následující pasáž shrnuje a rozvíjí či zavrhuje dosud zformulované, leč nečetné etymologické návrhy a předkládá nová řešení.

2.1 Vnitřní řecká etymologie

2.1.1 Boisacq (1916, s. 488) spekoval o příbuznosti mezi κόλον a ř. κυλλός “deformovaný, zkřivený”, jež se poprvé objevuje ve složenině κυλλο-ποδίων “se zkřivenýma nohama, kulhavý” [*Il.*], a/nebo v Hesychiově glose κελλόν· στρεβλόν, πλάγιον “zkroucený, skloněný”. Podle Meiera-Brüggera (1990, s. 31) pochází adjektivum κυλλός z praformy **k^ul^hnó-*, zatímco Vine (1999, s. 566) rekonstruuje jeho východisko v podobě **k^uol[H]-iō-*. Chybí však jakákoliv opora pro předpoklad, že část střeva označovaná jako κόλον by byla “deformovaná, zkřivená, zkroucená”. A dále platí, že případné primární sloveso **k^uelh₁-* “pohybovat se, otáčet se” není slučitelné s glosou κελλόν, neboť bychom pak měli očekávat iniciální **τ-*, ale ani s dalšími příklady diskutovanými níže.

2.1.2 Lidén (1933, s. 23) identifikoval nadějně příbuzenstvo v rámci řeckého dialektového kontinua v kyperské glose καλίδια· έντερα. Κύπριοι [Hesychios, c. 500 po Kr.], tj. “vnitřnosti, střeva”. Týž sufix⁵ tvoří synonymum έντοσθίδια “útroby, nitro” oproti έντόσθια id. Obě slova zdokumentoval Aristotelés ve svém pojednání *Morfologie živočichů* [*De partibus animalium* 4.8–9]. Přidat lze též formu γαστρίδιον [Aristophanés], která je zdrobnělinou od γαστήρ “břicho” [*Odysea* +]. Formy κόλον a καλίδια jsou slučitelné na základě pravidel indoevropského ablautu: řecká polysloupnost καλV regulérně odráží prajazykovou sekvenci **k₁HV⁶*, zatímco slovo κόλον může být promítnuto do praformy **kolHo-*.

⁵ Suffixy na -iō- mohou odrážet starší *i*-kmeny (Chantraine 1933, s. 335–339; Schwyzler 1939, s. 464; Meier 1975, s. 13, § 3b; Balles 2008, s. 205, 215–216, 229, 233–234, 245), stejně dobře jako *o*-kmeny a konsonantické kmeny (Meier 1975, s. 67, § 44).

⁶ Sr. ř. πάλιν “zpět, opět, znovu” < **k^ul^hi-n* jako lokativ kořenového substantiva > dat. sg. πάλ-ι + postpozice -ν, utvořený od slovesa πέλομαι “pohnout se, stát se”, πολέω “chodit kolem, obíhat, otáčet se; objevit se” (Beekes 1988, s. 75; 2010, s. 1147, 1168–1169). Tento příklad též ilustruje postmykénskou změnu **k^u* > π- v daném kontextu.

2.2 Vnější příbuzenstvo

2.2.1 Pettersson (1921, s. 152–153) diskutoval možné vnější příbuzné výrazy. Z jeho příměrů se zdá být nejslibnější srovnání s arm. *k'atirt'*, *k'atert'*, s pozdější variantou *k'atird*, “κοιλία, intestinum”, tj. “tělní dutina”, konkrétně “hrudník s břišní dutinou”, dále “břicho, žaludek, střeva”, kde koncové *-rd* bývá vysvětlováno jako vliv arm. *leard* “játra” (Lidén 1933, s. 24). Olsen (1999, s. 192) vysvětluje *leard* ze syntagmatu **lisi (H)iekh₁rt* “tučná játra”, podobně promítá *neard* “šlachy” do praformy **sněh₁rt*, zatímco slovo *k'atird*, *k'atirt'* řadí k anatomickým termínům neznámého původu (op. cit., s. 942). Naproti tomu, Martirosyan (2010, s. 655–656) soudí, že forma *k'atird* je sekundární pro své pozdější zaznamenání, zatímco *k'atirt'* by mělo být odvozeno z **k'atit^h-rV*, kde **k'atit^h-* by mělo představovat adaptaci asyrského *kalītu* “ledvina” (< semit. **k^aaly-at-* id.; pouze v ostrovním jazyce soqotri duál *keloiti* je překládán jako “střeva” – viz Militarev – Kogan 2000, n. 156). Nicméně, kolísání *-d* ~ *-t'* může být vysvětleno též jako výsledek vnitroarménského procesu, který sám Martirosyan popsál jako progresivní asimilaci aspirace. Příkladem budiž takový pár jako *t'arp* “velká proutěná vrš na ryby” oproti *t'arb* “konstrukce z prutů, dřevěná mříž”, implikující asimilaci *t'...b* > *t'...p* (Martirosyan 2010, s. 281–282); podobně *t'lk'i* “javor” s variantou *t'lk(en)i* (op. cit., s. 294) nebo *c'ayt'* “stříkat, kropit” vs. *c'aytem* “postříkat, rozprašovat” (op. cit., s. 623). Kolísání vokálů *e* ~ *i* má také vnitřní arménské vysvětlení, sr. arm. *Movsisi* s variantou *Movsesi*, obě gen. z *Movsēs* (Schmitt 1981, s. 43). Jestliže vokál *i* byl primární jako v tvaru *Movsisi*, je možné odvodit *k'atird* z **k_lHērt*, kde pozdější finální *-d* vzniká sekundárně z **-t* analogicky pod vlivem takových anatomických termínů jako *leard* či *neard*. V tom případě by rekonstrukce měla být pouze **k_lHēr*. Alternativní řešení, ačkoliv pouze implicitně, nabídl Klingenschmitt (1982, s. 101), když odvodil arm. *saṭard* ~ *saṭart'* “větev s listy, ratolest” z **k_lH-d^hro-*. Přijmeme-li toto řešení, *k'atird* ~ *k'atirt'* by mělo být odvoditelné z **k_lHēd^hro-* nebo **kolHēd^hro-* (popř. **id^hro-*). Ještě další řešení vyplývá z pravidla, které zformulovala Olsen (1999, s. 774; též 206, 846), totiž **-h₁t-* > arm. *-t'*, jak ilustruje na příkladech jako *lowrt'* “azurový, nebesky modrý” < **k_luh₁tro-*; *kowt'k* “sklizeň” < **guh₁to-*; *šit'* “kapka” < **skeh₁to-*. Vyjdeme-li z tohoto pravidla, je možné analyzovat slovo *k'atirt'* jako složeninu sestávající ze složek **k'at^h* a **irt'* < **ētr^o* < **h₁eh₁tr^o*, odpovídající ř. ἥτρον “dutina břišní”, ἥτρον “srdce” [Il.+] (sr. Beekes 2010, s. 527). Přijmeme-li *k'atirt'* jako primární formu, varianta s finálním *-d* může být skutečně vysvětlena jako analogická změna pod vlivem *leard* “játra” a *neard* “šlachy”. Konečně, arm. *t* odráží skupiny **-lH-*, **-l(H)n-* nebo **-ls-*, srov. *owtiwł*, var. *owtił*, *owteł* “morek” < *owti* “cesta” ~ stprus. *aulis* “holeň” + **p_lh₁o-* = **“výplň kosti”* (Olsen 1999, s. 56–57), vedle arm. *hetowm* “nalévám, plním” < **pelh₁-nu-mi* (viz Martirosyan 2010, s. 403–404; Klingenschmitt 1982, s. 244–246); dále např. *atawri* “mlýn; žena meloucí zrna” ~ ř. ἀλετρις < **h₂lh₁-tri-* (Olsen 1999, s. 776; Martirosyan 2010, s. 31); *k'owł* “vlákno, nit” < **k^uōlh₁o-* ~ lat. *colus* “kolovrátek” (Olsen 1999, s. 780); *katin* “žalud” < **g^ulh₂én* vs. gen. **g^ulh₂nós*,

jenž pokračuje v ř. βάλανος “žalud” (Olsen 1999, s. 464); *atač* ‘em “modlím se” < **slh₂-ské/o-* ~ ř. ἰλάσκεσθαι (Olsen 1999, s. 776); *joť* “rákos, třtina” < **ǵ^holh₃o-* ~ lat. *helus* “bylina”, ř. χλωρός “zelený” (Olsen 1999, s. 781), aj. V každém případě arménská sekvence *k’at(i)*^o může představovat regulérní kontinuant ie. sekvence **k_lHV* či **kolHV*, tudíž přímo etymologicky odpovídat jak ř. καλῖδια, tak κόλον. Petersson (1921, s. 152) také přidal lit. *skilvỹs* “ptačí žaludek” a lot. *škilva* “slepičí žaludek”. Ale tato slova jsou transparentně utvořena ze sloves doložených v lit. *skilti*, lot. *škil̃t* “sekat, řezat” (Smoczyński 2018, s. 1198).

2.2.2 Mimo doposud uváděné příbuzenstvo v arménštině je třeba nově vzít v úvahu i slibný keltský materiál.

2.2.2.1 Prvním je brittonský anatomický termín **kalonā*⁷: střvelš. *callon*, pl. *cal(l)onneu* (13. stol.), velš. *calon*, pl. *calonnau* “srdce; ňadra, prsa; děloha, lůno; břicho, vnitřnosti, útroby; střed; duše, duch, mysl, intelekt, odvaha” (GPC); stkorn. *colon*, gl. ‘cor’, korn. *colon* m. “vnitřnosti”, f. “srdce”; stbret. *calonn*, střbret. *calon* (1499), bret. *kalon* f. “srdce; duše” (Deshayes 2003, s. 361).

2.2.2.2 V galštině lze potenciální příbuzenstvo hledat v toponymech jako *Calonia* [*Vita Domitiani: In praedio suo, quod dicebatur pridem Calonia, a fonte, qui Calonna vocabatur, trahens vocabulum*] a *Calonna*, dnes *Chalonnnes-sur-Loire* [*Magnobodi vita Maurilii: loco, qui nominatur Calonna super litus Ligeris*] (Holder 1896, c. 704), původně snad **kalonjā* f. sg. či n. pl. “centrální, střední”.

2.2.2.3 Tuto hypotézu podporují rétorománské substrátové lexémy, které již byly podezírány z keltského původu: *calùn* (Sutselva), *calung* (Surmeir), *chalu*n (Rumantsch Grischun) “bok; stehno, zadní strana stehna” (HwR, s. 144; Meyer-Lübke 1935, n. 1523). Významový posun lze vysvětlit tak, že jde o prostřední část těla. Sémantický posun vysvětluje fakt, že “bok” se nachází uprostřed těla. Souvislost mezi „středem“ a “srdcem” ilustrují např. česká slova *střed*, *středa* se svým praslovanským východiskem **serda*, jež je odvozeno z praslovanského slova pro “srdce” (Machek 1968/2010, s. 585). Obdobně např. v semitských jazycích: akkad. *giššu* “bok” vs. arab. *ǧawš* “prsa; střed”, *ǧaš* / *ǧāš* / *ǧāyīšat* “srdce; duše, duch” (Militarev – Kogan 2000, n. 97).

Brittonská praforma **kalonā* a její pravděpodobné kontinentální keltské příbuzenstvo **kalon^o* mohou být promítnuty do ie. praformy **k_lHon^o*, která je plně

⁷ Elsie (1979, s. 103), Pedersen (1909–1913, s. 147) a Campanile (1974, s. 29) zde viděli adaptaci zdroje stfr. *chaudun* “vepřové vnitřnosti určené k vaření”, stvalon. *caudain* id. (obě ze 13. stol. a později) apod., který měl odpovídat nedoloženému lat. **caldūmen*, jež bývá rekonstruováno podle románských kontinuantů a raných romanismů v západogermánských jazycích: ital. dial. *kaldom* (Bologna), *kaldume* (Benátky; arch.), plus ostrovní formy jako *quadumi* (Sicilie), *eskaldom* “frikasé” (Mallorca), jakož i rumun. *calmoniu* “jitnice s dršťkami”. Patří sem i předliterární starofrancouzská výpůjčka ve střhn., střhn. *kaldune* “dršťky”. Jestliže hypotetické latinské východisko **caldūmen* bylo utvořeno z adj. *calidus* “horký, teplý” (FEW 2, 78; Meyer-Lübke 1935, n. 1504), primární sémantika by byla “vařené prasečí vnitřnosti”, která je však velmi vzdálena od sémantiky brittonských slov s ústředním významem “srdce”, nemluvě o abstraktních metaforách jako “duše” či “odvaha”.

slučitelná s praformami řeckých a arménských anatomických termínů analyzovaných výše (§§ 2.1 a 2.2.1). Rekonstrukce **k_lHon^o* má svou obdobu např. v galském jménu měsíce *Samoni* (gen.) < **sm_hon^o* nebo osobním jménu *Banona* < **g^u_hH₂on^o* (viz Dočkalová – Blažek 2011, s. 442; de Bernardo Stempel 2023, s. 125). Sufixální rozšíření o **-n-* pravděpodobně představuje singulativ, který lze identifikovat v mnoha indoevropských anatomických termínech, včetně keltských: např. stír. *leccu*, *lecca* “horní čelist, tvář, strana”, gen. sg. *lecon*, *leacan*, *lecond* (Pronk 2015, s. 332–339; Brugmann 1906, s. 303). Sr. též germánský *n*-kmen **hertōn-* “srdce” (Kroonen 2013, s. 222) etc.

2.2.3 Toch. B *kele* m. se překládá jako (a) “pupek” a (b) “střed”, sr. následující textové ukázky (Adams 2013, s. 211):

(a1) *korne kelen=ārañcās paine tāntsi* “v hrdle, v pupku, do srdce, včetně nohou” [41b3/4^c]

(a2) *kätkre wartse kele* “hluboký, široký pupek” [73b2^c]

(b) *rīntse kelesa* “u středu (perlativ) města (gen.)” [244b4^c].

Toch. B *kele* bývá odvozováno z praformy **k^uolh₁o-*, která by měla být utvořena od slovesa **k^uelh₁-* “pohybovat se, otáčet se” (Lipp apud Rix 2001, s. 386–387), které je zdrojem takových jmenných derivátů jako stsl. *kolo*, gen. sg. *kolese* “kolo, kruh”, nom. pl. *kola* “vůz”, nebo stír. *cul* “vůz” < duál **k^uolh₁-oh₁* (Adams 2013, s. 211; Pinault 2008, s. 424; Hilmarsson 1996, s. 126). Pouze Adams (2013, s. 211) se pokouší vysvětlit sémantickou motivaci metaforickým posunem od primárního významu “náboj kola” k lidskému pupku. Termíny “pupek” a “náboj kola” bývají etymologicky spřízněny, sr. germ. **nablan-* m. “pupek” vs. **nabō-* f. “náboj kola” < **h₃nob^h-l-on-* & **h₃nob^h-eH₂-*, vedle **anban-* m. “břícho” < **h₃e/onb^h-on-* (Kroonen 2013, s. 24, 380–381); lot. *naba* f. < **h₃nob^h-eH₂-*, stprus. *nabis* “pupek” [EV 123],⁸ “náboj kola” [EV 297] < **h₃nob^hi-/-iō-* (Mažiulis 2013, s. 621–622); stír. *imbliu* “pupek” < **ambliiōn-* (Matasović 2009, s. 33) < **h₃ṇb^h-el-iōn-*; lat. *umbilicus* m. “pupek; střed” < **h₃ṇb^h-el-īko-*, vedle *umbō* m. “výčnělek ve středu štítu, hrbolek” < **h₃ṇb^h-on-*; ř. ὀμφαλός m. “pupek; výčnělek ve středu štítu” < **h₃ṇb^h-l-*; véd. *náb^hya-* n. “náboj kola”, *náb^hi-* f. “náboj kola; střed; pupek (těla nebo světa); původ, příbuznost, rodina” < **h₃nob^hi-* (se zdložením podle Brugmannova zákona) (Pokorny 1959, s. 314–315). Přidány by měly být příbuzné výrazy ze dvou větví: arm. *aniw* “kolo, osa vozu; kolo jako mučící nástroj” < **h₃nēb^ho-* (Olsen 1999, s. 23; Martirosyan 2010, s. 89–90), a jediný zástupce slovanské větve, č. *náboj* (Machek 1968/2010, s. 387: koncové *-oj* bylo modifikováno podle slova *náboj* “náboj střelné zbraně”) < **h₃nōb^h[i]-* se zdloženým **ō* v důsledku tzv. *vřdd^hi*-formace nebo díky přejetí z některého indoiránského zdroje, sr. véd. *náb^hi-* “náboj kola; střed; pupek”.

⁸ EV = Elbingský vokabulář, viz Mažiulis (2013).

Sémantický vztah termínů “náboj kola” a “pupek” je zjevný, ale navzdory četným sufixálním rozšířením slova pro “pupek” musí být tato viditelná část lidského těla předlohou pro označení technických artefaktů jako právě “náboj kola” (tak ještě Adams 1997, s. 391), a nikoliv naopak, neboť metaforické užití anatomických termínů pro názvy nejruznějších technických objektů je zcela přirozené, na rozdíl od opačného postupu. Proto je legitimní hledat alternativní etymologii toch. B *kele*. S ohledem na anatomické termíny analyzované výše (§§ 2.1.2, 2.2.1, 2.2.2) je lákavé přidat *kele* k nim.⁹ Slovo je regulérně odvoditelné z praformy **kolHo-*, podobně jako toch. B *seme* m. “naběračka” (Adams 2013, s. 765) z praformy **somHo-*, která je sama odvozena z ie. slovesa **semH-* “nabírat, nalít” > ř. ṣmḥ či ṣmḥ “vědro” (**smḥH-eh₂-*); lat. *sentīna* “dno člunu; voda a odpad nahromaděné na dně člunu”, lit. *sėmti*, *semiù* (var. *semù*), *sėmiau* “nabrat (do dlaní; sběračkou), vybírat tekutinu; zaplavit” (Zehnder – Kümmel apud Rix 2001, s. 531; Smoczyński 2018, s. 1154; Pokorný 1959, s. 901–902).

2.3 Vnitřní indoevropská etymologie

2.3.1 Jedno možné řešení předjímal již Petersson (1921, s. 152; viz § 2.2.1), který srovnával ř. *kólov* s lit. *skilvys* “ptačí žaludek”, lot. *šķilva*¹⁰ “slepičí žaludek” (přidejme ještě lit. *skilándis* “uzený žaludek prasete nebo ovce, naplněný tukem a krájeným masem”). Tyto baltské formy jsou odvozeny ze slovesa, které je doloženo v lit. *skilti*, lot. *šķīlt* “sekat, řezat” (Smoczyński 2018, s. 1197–1198). Mají bohaté příbuzenstvo, v první řadě v samotných baltských jazycích: lit. *skėlti*, *skeliù*, *skėliau* “rozštípat (dřevo), pokácet; udeřit, třísknout; vykřesat oheň”, lot. *šķēlt*, *šķēļu*, *šķēlu* “nasekat dřevo, rozštípat dřevo; vykřesat oheň”, a dále v ostatních ie. větvích: arm. *ec’el* “rozštěpil”; het. *iškallari* “rozřezává, rozštěpuje, krájí”; ř. σκάλλω “kopu; hledám, rozlišuji” (**sklH-je-*); střir. *scoiltim* “rozštěpuji”; gót. *skilja* “řezník”, stsev. *skilja* “rozštěpit; rozlišit”, aj. (Pokorný 1959, s. 923–926; Kümmel apud Rix 2001, s. 553; **skelH-*). V principu je tato “řeznická Wurzeletymologie” uplatnitelná, ale ze sémantického pohledu cítíme, že sotva vůči **všem** vnitřním orgánům zde diskutovaným.

2.3.2 S výjimkou “pupku”, který sem patří pouze v případě sémantického posunu “(část) břicha” → “pupek”, všechny ostatní anatomické termíny analyzované výše, tj. “střevo a jeho části”, “žaludek” nebo “srdce”, mají podobu vaku či pytle.

⁹ Sr. sémantický vztah mezi skt. *tundá-* n. “břicho” [Pāṇīni], pkt. *tuṇḍa-* id., kašmíri *ton* m. “břicho pod pupkem”, maráthi *tūd* “tučný, baculatý” vs. skt. *tunda-* m., *tund-* f. “pupek” [lex.], pandžábi *tuṇī* f., hindí *tūḍī* f. “pupek” (Turner 1966, n. 5858 a 5860). Podobnou sémantickou disperzi nacházíme u arm. *port* “pupek, břicho, prostředek” (Pokorný 1959, s. 96). Mimo indoevropské jazyky se obdobný sémantický rozptyl objevuje mj. v jazycích berberských: awjila *a-būt* “pupek”; awlemiden *tā-būt-ut*; semlal *a-buḍ* id. vs. ntifa *a-buḍ* “břicho” (Orel – Stolbova 1995, s. 85–86, n. 350, 352).

¹⁰ Stejný sufik tvoří lit. *pišvas* “žaludek, břicho” od slovesa *pilti* “naplnit” (Smoczyński 2018, s. 964).

A to může být důležitá stopa vedoucí ke společnému sémantickému jmenovateli a dále i k novému etymologickému řešení. Ve dvou indoevropských větvích nacházíme slibné kandidáty pro “protovak”:

2.3.2.1 Ř. κολέων “pochva meče” [Il.] < *κολεφόν, sr. mykénské mužské jméno *ko-re-wo* [KN Ln 1568.2b], odrážející *Kolewos* nebo *Kolewōn* (Aura Jorro 1985, s. 382; Bartoněk 2003, s. 405), odvoditelné z **kolh₁euo-*, plus κόλυθοι pl. “testiculi” [Arist.], odvoditelné z **kolHu-d^hro-*.

2.3.2.2 Lat. *culleus* “kožený vak” [Plautus+] (de Vaan 2008, s. 150), plus varianta *cullus* zaznamenaná Festem: *quoque masculine dixerunt. Est enim tormenti genus e corio* (viz Ernout – Meillet 2001, s. 155). Ačkoliv ř. κολέων a lat. *culleus* jsou podobné formou i významem, je obtížné demonstrovat jejich společný původ, jak konstatoval už Boisacq (1916, s. 484). Někteří lingvisté zde nacházejí důvod identifikovat je jako dvě nezávislé adaptace technického termínu neindoevropského původu (např. Ernout – Meillet 2001, s. 155). Analyzujeme tento termín ve světle historické fonetiky. Lat. -*ll-* může odrážet různé souhláskové skupiny. V tomto případě přichází v úvahu sekvence **KolHu-* jako v *sollus* “úplný” < **soluos* < **solHu-*, ale nikoliv **KļHu-*, jako v *salvus* “bezpečný” < **salauos* < **sļHu-* (Weiss 2009, s. 113, 124; Machajdíkova 2013, s. 26–42). Tento výběr implikuje kořenový vokál -*o-* v latině, odlišný od reálně zaznamenaného -*u-* v *culleus* (& *cullus* u Festa). Vokál -*o-* by mohl být zachován v prodloužené podobě v *cōlei*¹¹ pl. “testiculi” [Laberius+] s *ō* snad pod vlivem slova *cōlum* “síto, cedník” (Walde – Hofmann 1938, s. 244). Na druhé straně, forma s kořenovým vokálem -*u-* mohla být přejata z některého z osko-umberských jazyků, sr. protějšky lat. *sollus* “úplný” v osk. nom. pl. m. **sullus**, gen. pl. m. **sullum** (ale *solu* v latinské abecedě) “každý, všechen” (Untermann 2000, s. 714–715; de Vaan 2008, s. 572). Z tohoto řešení vyplývá existence předitalické praformy **kolHu-* & **kolHuējo-*, slučitelné s předeckou praformou **kolh₁euo-*.

2.3.2.3 Je legitimní se ptát, k jaké funkci tento hypotetický kořen sloužil. Co mají společného anatomické termíny připomínající vak jako “střevo a jeho části”, “žaludek” nebo “srdce”, plus “vak” samotný, který byl používán pro uchovávání tekutin? Odpovědí je, že byly naplněny svým obsahem. Tato idea vede ke stejné sémantické motivaci, formující lit. *piļvas* “žaludek, břicho”, které je transparentně odvozeno od slovesa *pīlti* “naplnit” (Smoczyński 2018, s. 964). Musíme se tedy ptát, zda je k dispozici kořen obdobné sémantiky, který by mohl být zdrojem vaku podobných anatomických termínů i slova “vak” samotného? Takový kandidát skutečně existuje: skt. *ā-kula-* “naplněný, plný; přetížený; dychtivě zaujatý” [MBh] < **kļHo-*, *kalila-* “plný, pokrytý” [MBh; BhP] (Monier-Williams 1993, s. 127, 262) < **kolHilo-*, a zřejmě též oset. iron. *kælyn*, digor. *kælyn* “téci; být stahován dolů; klopýtnout”;

¹¹ Cicero, *Fam.* 9.22.4, ‘*Testes*’ uerbum honestissimum in iudicio, alio loco non nimis; et honesti ‘*colei Lanuini*’, ‘*Cliternini*’ non honesti.

(caus./intr.) oset. iron. *kalyn*, digor. *kalun* “nalít, vylít; shodit na zem” (Abaev 1958, s. 569–570). Tyto tvary by měly být odděleny od véd. *kar-* “nalít, rozptýlit, rozsypat, rozházet, vrhat” [RV 1.32.13] (Monier-Williams 1993, s. 308; Mayrhofer 1986, s. 311, 323), mlav. *kar-* (caus.) “rozptýlit se, rozšiřovat se”, aj. (Cheung 2007, s. 239–240), které jsou spíše kontinuanty ie. kořene **kerH-* (Pokorný 1959, s. 933–935; Lipp apud Rix 2001, s. 353–355). Mimo indoiránské jazyky nacházíme slibné příbuzenstvo v germánské větvi: stsev. *hylja* “proudit, téci”; *hylr* m., gen. *hyljar* “louže; nahromadění vody” < germ. **hulwǵō*; sthn. *hul(i)wa* “louže, tůň, rybníček” < **hulwō(n)* (de Vries 1977, s. 274–275; Lloyd – Lühr 2009, c. 1209). Všechny tyto formy jsou odvoditelné z kořene **kelH-* s pravděpodobným významem “naplnit vodou”. Určení laryngály jako **h₁* podle ř. κολεόν “pochva meče” zůstává nejisté, neboť tento význam stojí přeci jen poněkud daleko od primární sémantiky “naplnit vodou”.

3. VNITŘNÍ STRUKTURA

3.1 Ř. κόλον a pravděpodobné příbuzenstvo z pohledu slootvorby:

kořen / extenze	<i>*kolH-</i>	<i>*k_lH-</i>
<i>*-o-</i>	ř. κόλον “část tlustého střeva” toch. B <i>kele</i> “pupek; střed”	ř. καλίδια “vnitřnosti, střeva” (před sufixálním rozšířením mohlo jít o <i>i-</i> nebo <i>o-</i> kmen)
<i>*-o-</i> + <i>*-n-</i>		britt. <i>*kalonā</i> “srdce” ?gal. <i>*kalonjā</i> “střed(ní)”
+ <i>*h₁eh₁tr^o</i>	?arm. <i>kʾatird/tʾ</i> “tělní dutina: hrudʾ s břišní dutinou; břicho, žaludek, střeva”	
<i>*-ērt(i)?</i>		?arm. <i>kʾatird/tʾ</i>
<i>*-ē-d^hro-</i> ?		?arm. <i>kʾatird/tʾ</i>
<i>*-u-d^hro-</i>	ř. κόλυθροι pl. “testiculi”	
<i>*-e_uo-</i>	prař. *κολεφόν “pochva meče”	
<i>*-e_io-</i>	lat. <i>cōleī</i> pl. “testiculi”	
<i>*-u_o- & *-u_ei_o-</i>	lat. <i>cullus</i> & <i>culleus</i> “kožený vak”	

3.1.1 Komentář k derivační morfologii

3.1.1.1 Ř. κόλον a toch. B *kele* jsou *o*-kmeny, ačkoliv se mohou lišit v akcentuaci a gramatickém rodu.

3.1.1.2 Britt. **kalonā* patrně představuje nazální extenzi *o*-kmenu, pravděpodobně se singulativní funkcí (§ 2.2.2.3).

3.1.1.3 Arm. *kʾatird/tʾ* poskytuje několik interpretací:

3.1.1.3.1 Složenina *kʾat-irtʾ* < **kolH-/k_lH-* + **h₁eh₁tr^o* “břišní dutina”.

3.1.1.3.2 Formace na **-ērt*, paralelní k *leard* a *neard* (§ 2.2.1).

3.1.1.3.3 Formace na **-ēr* s finální dentálou připojenou pod vlivem anatomických termínů *leard* a *neard*.

3.1.1.3.4 Formace na **-ē-d^hro-* je srovnatelná s řeckými formami jako μέλπηθρον “prostředek hraní, hračka” vs. μέλπω “oslat písní a tancem, zpívat”; στέργηθρον “kouzlo lásky” vs. στέργω “miluji, pocítuji lásku”; ἔλκηθρον “rukojeť pluhu” vs. ἔλκω “táhnou, vleku”, aj. (Chantraine 1933, s. 373). Zde je důležité zdůraznit, že týž sufix **-d^hro-* charakterizující nomina instrumenti tvoří i ř. κόλυθροι.

3.1.1.4 Ř. κόλυθροι představuje primárně *u*-kmen, rozšířený o sufix nomina instrumenti **-d^hro-*; sr. κόλυθρον “zralý fik” nebo σκόλυθρον “stolička” (Chantraine 1933, s. 373).

3.1.1.5 Denominální deriváty na **-e^ho-* jako prař. **κολεφόν* “pochva meče” bývají utvořeny z *u*-kmenů, jejichž stopy vidíme i ve slově κόλυθροι, ale také od jiných slovních kmenů, např. θυρεός “kámen položený proti dveřím tak, aby je držel zavřený” od slova θύρα “dveře” nebo πῖαλέος “bohatý” od πῖαλον n. “tuk” (Chantraine 1933, s. 51; Brugmann 1906, s. 204). Frisk (1973, s. 898) spekoval o odvození slova **κολεφόν* od slovesa καλύπτω “pokrývat, skrývat”. Z pohledu sémantiky jde o přijatelný, i když nikoliv jednoznačný významový vývoj. Vezmeme-li v úvahu pravděpodobné příbuzenstvo daného slovesa, konkrétně καλύβη “chatrč, chýše; stan” a καλυφή “ponořená země”, je možno souhlasit s Beekesem (2010, s. 628–629) v jeho závěru o jejich substrátovém původu. Frisk (1973, s. 769) srovnával ř. καλύπτω “pokrývat, skrývat” s lat. *occulere* “skrýt, ukrýt”, stir. *celim* “ukrývám”, kterému jsou dále příbuzné stang. *helan* “skrývat”, véd. *śárman-* “přístřešek, úkryt, skryš, útočiště”. Východiskem je pak kořen typu *aniṭ* **kel-* (Kümmel apud Rix 2001, s. 322–323; de Vaan 2008, s. 423–424). Opět, sémantika by byla v zásadě přijatelná, ale arm. *k’atird/t’* nemůže být utvořeno od jakéhokoli kořene s počátečním **k̥-* a současně anatomické termíny analyzované výše vyžadují rekonstrukci laryngály po **-l-*.

3.1.1.6 Formace na **-u^o-* jako v lat. *cullus* [Festus] (§ 2.3.2.2) pocházejí z *u*-kmenů. Lat. *culleus* představuje derivát na **-e^ho-* (Brugmann 1906, s. 198–199).

4. DALŠÍ VÝVOJ A UPLATNĚNÍ TERMÍNU ΚΟΛΟΝ V ANATOMICKÉ TERMINOLOGII LATINSKÉ I ŘECKÉ

4.1. Po Aristotelovi to byl zejména římský polyhistor Plinius (23–79 po Kr.), kdo shrnul stávající přírodovědné poznání a zasáhl tak i do anatomické terminologie, a to v kompendiu *Historia Naturalis*. Námi analyzované řecké slovo adaptoval v podobě *colon* do latiny. V jeho *Historii přírody* je najdeme na dvou místech.

Kniha 1 (fakticky jde o obsah k dalším knihám, zde k příslušné pasáži knihy 11): *haruspicum circa ea observationes et prodigia mira* (77) *praecordia. risus natura* (78) *de ventre. quibus nullus. quae sola vomant* (79) *lactes, hillae, alvus, colon. quare quaedam insatiabilia animalia* (80–83) *de omento. de splene. quibus animalium non sint de renibus ubi quaterni animalibus. quibus nulli pectus, costae vesica. qui-*

bus animalium non sit. ilia. de membranis (84–88) uterus. de locis. de volvis. de suum volva, sumine quae adipem, quae sebum habeant. de natura utriusque. quae non pinguescant de medullis. quibus non sint de ossibus. de spinis. quibus nec ossa nec spinae. cartilagines de nervis. quae sine nervis (89–92) arteriae, venae quae nec venas nec arterias habeant.

Vydal Karl Friedrich Theodor Mayhoff (1906).

“Pozorování těchto druhů haruspiky a úžasná znamení. (77) Bránice; podstata smíchu. (78) O žaludku a těch, kdo ho nemají; jediný druh, který zvrací. (79) Tenké střevo, lačník, žaludek, *colon* = tlusté střevo; proč někteří živočichové jsou tak žraví. (80–83) O bláně uzavírající střeva, o slezině, druhy bez sleziny. O ledvinách, kde žijí druhy se čtyřmi ledvinami a kde s žádnou. Hrud', žebra, močový měchýř a která zvířata jej nemají. Zvířecí střeva, blány. (84–88) Břicho – o částech, o děloze, o děloze sviní, cecíky. Které druhy mají tuk a které lůj; o podstatě každého; které druhy neztuční. O morku; druhy, které ho nemají. O kostech; o ostnech; o družích, které nemají ani kosti, ani ostny. Chrupavky. O šlachách; druhy bez šlach. (89–92) Tepny, žíly; druhy, které nemají ani žíly, ani tepny.”

Doslovný český překlad s přihlédnutím k volnému anglickému překladu H. Rackhama (1927).

Kniha 11.79/83

*media haec umbilico adnexa omnibus, in homine suillae infima parte similis, a graecis appellatur **colon**, ubi dolorum magna causa. angustissima canibus, qua de causa vehementi nisu nec sine cruciatu levant eam. insatiabilia animalium quibus a ventre protinus recto intestino transeunt cibi, ut lupis cervariis et inter aves mergis.*

Vydal Karl Friedrich Theodor Mayhoff (1906).

“U všech druhů existuje orgán, který se nachází uprostřed u pupku. U člověka v jeho dolní části připomíná břicho prasete. Řekové ho nazývají *colon*. Způsobuje velkou bolest. U psů je mimořádně úzký a z tohoto důvodu ji mohou zmírnit jen s velkým úsilím a nikoliv bez utrpení. Nejdravější živočichové jsou ti, kterým prochází strava přímo ze žaludku do střev. To je příklad rysa a u ptáků kormorána.”

Český překlad vznikl s přihlédnutím k volnějším anglickým překladům

J. Bostocka & H.T. Rileyho (1855) a H. Rackhama (1940).

4.2 Iulius Pollux (= Ἰούλιος Πολυδεύκης), řecky píšící autor ze 2. stol. po Kr., uvedl v život chybnou variantu κῶλον, a to ve svém spise *Onomasticon* [Ὀνομαστικόν 2.209]: λήγει δ' εἰς δύο τέλη, ὧν τὸ μὲν καλεῖται **κῶλον** καὶ κάτω κοιλία (ed. Wilhelm Dindorf 1824), s latinským překladem Rudolfa Walthera (Basel 1541): *Et in duos terminatur fines. Quorum alter, **alvus** vocatur, et venter inferior.* Změnu kořenového vokálu lze připsat vlivu ř. κῶλον “část těla”, zvláště “noha” (Chantraine 1968, s. 557), srov. κωλήν “stehno, noha”.

4.3 Serenus Sammonicus (zemřel 212 po Kr.), autor veršovaného pojednání *Liber medicinalis*, známého též jako *De medicina praecepta saluberrima*, který čítá 1115 hexametřů, slovo přizpůsobil latinské gramatice, když vytvořil ekvivalent řeckého neutra κῶλον (o chybné reinterpetaci původní formy κόλον viz § 4.2) v podobě lat. *cōlum*¹².

4.4 Na Serena navázal italský lexikograf Ambrogio Calepino (lat. Ambrosius Calepinus; c. 1440–1510), i když jeho adaptaci v podobě lat. *cōlum* nepřijal. Roku 1502 poprvé Calepino vydal svůj latinský slovník. Ve vydání z r. 1535 vysvětluje heslo **colon** slovy *crassius intestinum*, tj. “tlusté střevo”. Dále cituje Serenovy verše v podobě

*cum **colon**, inuisum morbi genus, intima carpit,
mande galeritam uolucrum quam nomine dicunt.*

K nim dodává: *Dicitur etiam **colon**, (κῶλον) membrum orationi.*

4.5 Isidor ze Sevilly (kol. r. 600 po Kr.) ve svých *Etymologiích* pro změnu navázal na Iulia Polluxe, když operoval pouze s řeckou formou i s chybně prodlouženým vokálem první slabiky [*Etymologiae* 4.7.38]: *Colica passio nomen sumpsit ab intestino, quem Graeci κῶλον appellant* “Nemoc kolika dostala své jméno podle střeva, které Řekové nazývají κῶλον”.

4.6 Kolem r. 1240 za svého pobytu v Magdeburgu Angličan Bartholomaeus Anglicus (1203–1272) sepsal latinské kompendium *De Proprietatibus Rerum* (*O vlastnostech věcí*). Zde ve 42. kapitole *De visceribus* (*O útrokách*) 5. knihy, která je věnována lidskému tělu, píše: *Tertium intestinum vocat a greck **colon*** (podle tisku z r. 1488). Je docela pravděpodobné, že termín i zmínku o jeho řeckém původu autor převzal od Isidora ze Sevilly (viz § 4.5), na kterého se odvolával v úvodu 42. kapitoly.

5. VYUŽITÍ TERMÍNU STŘEDOVĚKÝMI ŽIVÝMI JAZYKY

Ve 14. století opouští termín latinu jako univerzální metajazyk evropské vědy a proniká do soudobých moderních jazyků, ačkoliv pozice latiny bude ještě další půltisíciletí dominantní.

¹² XXX *Colo sedando*

⁵⁷⁴ *cum **colum** {rkp. AB colus}, inuisum morbi genus, intima carpit,*

⁵⁷⁵ *mande galeritam uolucrum quam nomine dicunt.*

⁵⁷⁶ *aut pavidū leporis madefacta coagula pota.*

⁵⁷⁷ *siue apio nepetas tereti cum mastice iunge*

⁵⁷⁸ *nec non et species ambas halantis anethi,*

⁵⁷⁹ *quarum decoctos patiens haurito liquores.*

5.1 Ve střední francouzštině je slovo *côlon* ‘partie du gross intestin, qui fait suite au caecum’ “část tlustého střeva, která tvoří pokračování {části zvané} caecum” doloženo od 14. st. (FEW 2, s. 920). Termín převzal z latiny a poprvé r. 1314 použil Henri de Mondeville (c. 1260–1320). Víme, že studoval medicínu na univerzitách v Montpellier a Paříži. Poté se vydal do Bologni, kde spolupracoval s nejlepším chirurgem své doby, Theodorikem Borgognonim. Vrací se do Francie, kde vyučuje anatomii a chirurgii na univerzitě v Montpellier (1301–1304), poté přichází do Paříže, kde se stává osobním lékařem králů Filipa Sličného a Ludvíka X., ale věnuje se i výuce na pařížských medicínských školách. Od r. 1306 až do své smrti píše zásadní dílo své doby, *La Chirurgie*, kde termín uplatnil v § 373: *Le .5. bouel est continué o cestui, et est apelé colon qui reçoit les fèces desnuees de toute chose profitable ...* “Pátá část střeva pokračuje k pasu a nazývá se *kolon*; přijímá výkaly ve velkém množství, čímž je především užitečná ...”. Žel, stihl dokončit jen dvě části z plánovaných pěti. Do dějin medicíny se Henri de Mondeville zapsal mj. prvními pitvami lidského těla a antiseptickými opatřeními při chirurgických úkonech.

5.2. V roce 1397/8 vznikl středoanglický překlad Bartholomaeova kompendia *De Proprietatibus Rerum*, jehož autorem byl John Trevisa. Klíčové slovo se objevuje v jeho překladu Bartholomaeo věty *Tertium intestinum vocat a greck colon* v následující podobě: *The thyrd grete gutte highte Colon, in the language of Grecke* “třetí {část} tlustého střeva se nazývá *kolon* v jazyce řeckém” (podle vydání z r. 1495; srov. 1582, Liber V, Cap. 42).

5.3 Jak už jsme uvedli dříve (Šmejkalová – Blažek 2023, s. 215), ve staročeském jazykovém materiálu se slovo *colon* vyskytuje u Klareta (v KlarGlosA v kapitole *De membris*; folio 194vb, verš 1313¹³; Klaret – Flajšhans 1926, s. 532), a to jako novotvar *žluček* ve významu *žlučník* (Hartmannová 2013, s. 129). Zjevně jde o změnu významu způsobenou zvukově blízkým ř. *χολή* “žluč”, dem. *χόλιον*, v přeneseném smyslu *χόλος* “žluč, žlučovité hněv”.

Klaretův *colon* se nalézá v řadě příkladů označujících prsty (narušuje ji?), což je zajímavé, neboť patrně nelze ani hypoteticky předpokládat souvislost mezi *colon* a označením prstu na ruce (srov. i studii A. M. Černé 1998). Klaret uvádí: *Palec (pollex), colon (žluček), ukazač (index), prstenec (annularis), ušinec (malik; auricu-*

¹³ Alena M. Černá. Zdroj: Elektronický slovník staré češtiny. Praha, oddělení vývoje jazyka Ústavu pro jazyk český AV ČR, v. v. i., 2006–, přístupné online: <http://vokabular.ujc.cas.cz> (verze dat 1.1.25, citován stav ze dne 6. 10. 2024). Staročeská textová banka [online]. Ústav pro jazyk český AV ČR, v. v. i., oddělení vývoje jazyka. Verze dat 1.1.15 [cit. 20. 6. 2020]. Dostupné z: <http://vokabular.ujc.cas.cz/banka.aspx?idz=STB>. Klaret; [Glosář]; konec 14. století; edice; Klaret a jeho družina. Sv. I. Slovníky veršované; Flajšhans, Václav; Praha; 1926; 14; KlarGlos; KlarGlosA; verš; slovník; ediční poznámka; Černá, Alena M. – Voleková, Kateřina.

laris), *mezě nec* (*medius* – vykládáno buď jako třetí prst, nebo jako čtvrtý prst¹⁴). Vysvětlením je chybná identifikace dvou řeckých kvasihomeonym, κόλον a κῶλον “část těla”, kterou uvedl v život Iulius Pollux (viz § 4.2).

Další výskyt nalézáme ve Veleslavínově spisu z roku 1586 *De homine et partibus humani corporis*, jenž je součástí kompendia *Nomeclator* (*Intestinum colon: veliké střevo / Weliké střewo / Großdarm*), který dokládá, že nejpозději v této době se *colon* = *střevo* stalo familiárním lékařským termínem. Dále už se situace opakuje ve stejném významu i u daleko mladšího slovníku Dobrovského z r. 1821 (*Grimmdarm, m. veliké střewo, colon*). Tyto poměrně chudé nálezy samozřejmě neznamenají, že by předtím vyšetření střev vůbec neexistovalo, a že by tedy neexistovaly slovníkové záznamy. Už Jungmann (a nikoliv jako první) popisoval střevo například takto: *střevo dvanácté (intestinum duodenum), střevo lačné (hladové, hladovité, prázdné, churavé, postní, tedy jejunum)*, akceptuje i náš termín *colon*, česky *denník* (Jungmann 1838, s. 346¹⁵). V další literatuře 19. století se většinou pracuje s termínem *colon* nebo *denní střevo*. Například *Slovník lékařské terminologie* z roku 1863 zachycuje výskyt termínu *Gich-Darm (colon): pakostniční, denní střevo* (s. 46). Podobně započal již dříve výraz *colon* do svého slovníku i známý obrozenec a lexikograf Josef Franta Šumavský (Šumavský 1844, s. 962). To už se ale více než výrazně vzdalujeme středověku.

6. ZROD SLOŽENINY KOLOSKOPIE S DUBLETOU KOLONOSKOPIE

6.1 Vraťme se ještě jednou k termínům, které byly impulsem našeho bádání. Utvářejícími složkami jsou řecká slova κόλον n. “tlusté střevo (od slepého střeva po rectum)”, podrobně analyzované výše, a σκοπία f. “místo určené pro pozorování; pozorovatelná; hlídka, stráž” [Homér+]. Slovo σκοπία spolu se substantivem σκοπός m. “slídil, špeh; cíl” [Homér+], slovesem σκοπέω “hledím” [Pindaros+] a mnoha dalšími deriváty představují regulérní apofonickou *o*-variantu od primárního kořene s kořenovým vokálem *e, jenž reprezentuje sloveso σκέπτομαι „pozoruji; pohlížím; ohledávám; rozvažuji; rozjímám“ [Homér+]. Další příbuzenstvo lze identifikovat v lat. *speciō, -ere* “vidět, pozorovat, sledovat, všimat si”, *haruspex* „věstec z vnitřností ovcí“, něm. *spähen* “(tajně) sledovat, špehovat”, alb. *pashë* “viděl jsem”, stind. *pásyati* “dívá se, pozoruje, všímá si, hlídá, střeží”, avest. *spasiia-* “objevit, pohlížet, vnímat” aj., vše od indoevropského kořene *spek-, který v řečtině prodělal metatezi *spek- > *skep- (Beekes 2010, s. 1347–1348; Chantraine 1970, s. 1014–1015; Pokorný 1959, s. 984; de Vaan 2008, s. 578–579).

¹⁴ Dostupné na: <https://vokabular.ujc.cas.cz/hledani.aspx>, dotaz *mezenec*.

¹⁵ Staročeské, srov. <https://vokabular.ujc.cas.cz/hledani.aspx>, dotaz *denník*. Jungmann výraz pravděpodobně pokládal za součást úzu.

6.2 Obě výše uvedené složeniny (§ 1.1), κολούλια a κολοκορδόκολα, dokumentují, že řečtina připouští pouze kompozitum *κολοσκοπία¹⁶, i když taková složenina v korpusu klasické, helénistické ani byzantské řečtiny není doložena.

6.3 Jedním z prvních¹⁷, kdo v anglické medicínské terminologii používal termíny *colonoscope* a *colonoscopy*, byl americký lékař z New Yorku, Heinrich Stern (1911, s. 42), jenž se ke svému autorství hlásil slovy: *I have called the instrument the colonoscope and I propose the name of colonoscopy for the inspection of the colon beyond the lowest portion of the sigmoid*. “Nazval jsem tento nástroj kolonoskop a navrhuji termín kolonoskopie pro prohlídku části tlustého střeva, zvaného *colon*”. Vnitřní struktura nové složeniny je zcela transparentní: (už v latině u Plinia nesklonné) *colon* + konektém *-o-* + *scopy*, tvořící i mnohé jiné lékařské termíny. Nicméně Stern již neříká, zda vyšel z tradice anglické medicínské terminologie, která užívá slovo *colon* od konce 14. stol., nebo jeho latinské předlohy *colon* z kompendia *De Proprietatibus Rerum* (1240), jehož autor Bartholomaeus Anglicus zjevně znal termín i jeho řecký původ od Isidora ze Sevilly z doby kolem r. 600 (*Etymologiae* 4.7.38). Ten se nepochybně inspiroval v *Dějínách přírody* (11.79), kde Plinius explicitně zmiňuje řecký původ slova. Sám Plinius převzal termín s největší pravděpodobností od Aristotela, který mu věnoval největší pozornost v anatomické perspektivě. Taková je tedy historická geneze termínu *kolonoskopie*.

Dodejme, že *Slovník Francouzské akademie* (*Dictionnaire de l'Académie française*, 9^e édition) zná pouze kratší variantu *coloscopie*¹⁸, která respektuje pravidla řecké gramatiky. Slovo se objevuje až ve 20. století.

Jak jsme rovněž již dříve zjistili, v české odborné literatuře se termíny *koloskopie* a *kolonoskopie* prosazují v souladu s rozvojem této zobrazovací metody až v 60. a 70. letech 20. století (Falt et al. 2015). A je pozoruhodné, že po krátkém iniciačním období (např. Tesař 1967) se již počínaje 70. lety 20. století úzus rozkolísává a již tehdy vedle sebe stojí varianty *koloskopie*/*kolonoskopie*. Uveďme ukázky příkladů. *Časopis lékařů českých* z roku 1969 uvádí termín *kolonoskopie*,¹⁹ týž zdroj

¹⁶ Obdobně jsou tvořena kompozita ζυγομαχία “svár” nebo ζυγοστασία “vážení”, kde první komponentou je neutrum ζυγόν “jho” [Homér+], též „rameno vah“ [Aischylos+]. Co se týče pozice přízvuku ve složenině s druhou komponentou tvořenou slovem σκοπία, srov. οἰωνοσκοπία „věštba na základě pozorování letu ptáků“ [Dionýsios z Halikarnassu, *Antiquitates Romanae* 3.47], kde první složkou je οἰωνός “velký pták, dravec” [Homér+].

¹⁷ Různé online slovníky se liší v datování prvního použití termínu v angličtině, aniž by uváděly konkrétní zdroje:

1902 – *colonoscopy*; 1896 – *procto-colonoscopy*; 1884 – *colonoscope*

<https://www.etymonline.com/word/colonoscopy#etymonline_v_28358>.

1900–1905 – *colonoscopy* <<https://www.dictionary.com/browse/colonoscopy>>.

1926 – *colonoscopy* <<https://www.merriam-webster.com/dictionary/colonoscopy>>.

¹⁸ Dostupné na: <https://www.dictionnaire-academie.fr/article/A9C3021>.

¹⁹ *Časopis lékařů českých* roč. 108, č. 10, 1969, s. 311.

v roce 1973 *koloskopie*,²⁰ odborná lékařská literatura z roku 1967 *koloskopie* (Tesař 1967, s. 56) a z roku 1973 *kolonoskopie* (Špaček 1973, s. 211). Někdy nebylo jednotné psaní dodrženo ani v rámci jedné publikace – tak např. Fučík (1972) uvádí na s. 241 variantu *koloskopie*, a v rejstříku na s. 396 nebo na s. 238 *kolonoskopie*, ve výkladu na s. 238 *kolonoskop*. V dokladech by bylo možno pokračovat.

Medicínská angličtina používá regulérní formace jako *colocentesis*, *colopathy*, *coloscopy*, *colostomy*, vedle *colonoscopy* a *colonopathy* (Dirckx 1997). Pak je zcela legitimní otázka, kde se vzaly formy se slabikou *no* jakoby navíc? Madsen (2015) s lehkou ironií konstatuje, že hypotetická řecká předloha této formy by musela obsahovat jako první komponentu slovo *κολωνός* “pahorek” [Hésiodos+], implikující složeninu **κολωνοσκόπια* s významem “pozorování pahorků” (tutéž úvahu vyslovil i anonymní autor wikipedického hesla *Colonoscopy*). Beran (2015) nabízí ještě bizarnější složeninu, kdy na místo první komponenty dosazuje latinské slovo *colōnus* “rolník”! Obě “řešení” jsou samozřejmě zcela mimoběžná. To, co vyhlíží jako terminologická chyba, má nicméně racionální vysvětlení, vzhledem k okolnostem šíření řecké medicínské terminologie do latiny a posléze do moderních evropských jazyků. Tvary s *-n-* se objevují díky tomu, že římský polyhistor Plinius (23–79 po Kr.) ve své monumentální encyklopedii *Historie přírody* adaptoval řecký anatomický termín *κόλον* v nesklonné podobě *colon*. Forma s *-n-* pak figuruje i v moderních složeninách typu *kolonoskopie*, v konkurenci s gramaticky jedinou správnou, i když méně frekventovanou formou *koloskopie*.

Zkratky: akkad. akkadský, alb. albánský, ang. anglický, arab. arabský, arm. arménský, av(est). avestský, bret. bretonský, britt. brittonský, č. český, digor. digorský, fr. francouzský, gal. galský, germ. (pra)germánský, gl. glosa, gót. gótský, het. hetitský, hn. hornoněmecký, ie. indoevropský, ind. indický, ir. irský, iron. ironský, ital. italský, korn. kornský, lat. latinský, lit. litevský, lot. lotyšský, ml- mlado-, něm. německý, niz. nizozemský, oset. osetský, osk. oskický, pkt. prákrt, rumun. rumunský, ř. řecký, sev. severský, skt. sanskrt, st- staro-, stř- středo-, stsl. staroslověnský, toch. tocharský, valon. valonský, véd. védský, velš. velšský.

Pozn.: Staroindické, řecké a latinské lexémy byly prověřovány ve standardních slovnících, jejichž autory jsou Monier-Williams (1899/1993), Liddell – Scott (1996) a Lewis – Short (1958). Odtud byly převzaty i standardní zkratky primárních textů, popř. jejich autorů.

Řecké, latinské, středofrancouzské a české prameny

ARISTOPHANES (1907): *Knights*. In: F. W. Hall – W. M. Geldart (eds.): *Aristophanes Comoediae*. Oxford: Clarendon Press.

²⁰ Op. cit., s. 112; 1973, s. 29.

ARISTOPHANES (1938): *Equites*. In: *The Complete Greek Drama*, Vol. 2, translated by Eugene O'Neill, Jr. New York: Random House.

ARISTOTLE (1912): *The Works of Aristotle*, Volume V: *De partibus animalium* translated by William Ogle; *De motu and de incessu animalium* translated by A. S. L. Farquharson; *De generatione animalium* translated by Arthur Platt. Oxford: Clarendon Press.

BARTHOLOMAEUS, Anglicus (1488): *De Proprietatibus Rerum*. Heidelberg: Knobloch.

CALEPINUS, Ambrosius (1535): *Dictionarium Latinae Linguae, Ambrosii Calepini: Et Aliorum hominum doctorum opera, ex optimorum scriptorum multa lectione accurata[ue] animaduersione in suum ordinem descriptum. De XII. item mensium Romanorum ratione, quomodo eis, praesertim in epistolis conscribendis, sit utendum*. Basel: Val.

DOBROVSKÝ, Josef (1821): *Deutsch-böhmisches Wörterbuch*. Praha. Edice © Mirek Čejka, Vojtěch Černý, Boris Lehečka, Zuzana Leštinová, Michaela Novotná, Martin Stluka, Iva Trhlíková, Robin Ujfalusi, 2008–2017. Dostupné na: <https://vokabular.ujc.cas.cz/nezapojene.aspx?idz=DDBW1821> [cit. 27-05-2025].

HESYCHIOS (1867): *Hesychii Alexandrini Lexicon*, curavit Mauricius Schmidt. Jena: Dufft.

ISIDORE of Seville (2006): *The Etymologies of Isidore of Seville*, translated by Stephen A. Barney, W. J. Lewis, J. A. Beach, Oliver Berghof, with the collaboration of Muriel Hall. Cambridge: University Press.

JUNG MANN, Josef (1838): *Slovník česko-německý, Díl IV. S–U*. Praha: Knížecí arcibiskupská knihtiskárna.

KLARET – FLAJŠHANS, Václav (1926): Sv. I, *Slovníky veršované*. Praha: Česká akademie věd a umění.

de MONDEVILLE, Henri (1897): *La chirurgie*, ed. A. Bos. Paris: Didot.

NICANDER (1953): *The Poems and Poetical Fragments*, edited and translated by A. S. F. Gow and A. F. Scholfield. Dostupné na: <http://www.attalus.org/poetry/nicander.html> [cit. 27-05-2025].

PLINIUS (1906): *Naturalis Historia*, ed. Karl Friedrich Theodor Mayhoff. Leipzig: Teubner.

PLINY the Elder (1855): *The Natural History*, translated by John Bostock & H. T. Riley. London: Taylor & Francis.

PLINY (1927): *Natural history* in ten volumes, Vol. I: Praefatio, libri I, II, with an English translation by H. Rackham. Cambridge (Mass.): Harvard University Press – London: Heinemann.

PLINY (1940): *Natural history* in ten volumes, Vol. III: Libri VIII–XI, with an English translation by H. Rackham. Cambridge (Mass.): Harvard University Press – London: Heinemann.

POLLUX, Julius (1824): *Onomasticon*, Ed. Wilhelm Dindorf. Leipzig: Kühn.

POLLUX, Julius (1541): *Onomasticon*. Rodolpho Gualthero Tigurino interprete. Basel: Winter.

SERENUS Q. Sammonicus (1916): *Quinti Sereni Liber Medicinalis*. Ed. Friedrich Vollmer. Leipzig: Teubner. Dostupné na: <https://latin.packhum.org/loc/1515/1/0#0> [cit. 27-05-2025].

ŠUMAVSKÝ, Josef Franta (1844): *Deutsch-böhmisches Wörterbuch*. Erster Band, A–J. Prag: Johann Spurný.

Z VELESĽAVĽNA, Daniel Adam (1586): *Nomenclator omnium rerum propria nomina tribus linguis – Latina, Boiémica et Germanica – explicata continens. Ex Hadriano Iunio Medico excerptus et pro usu scholarum Boiémicarum editus*. Praha: Daniel Adam z VelesĽavĽna 1586. Edice © Tilman Berger, Alena M. Ľerná, Ondřej Koupil, Petr Nejedlý, Štěpán Šimek, Miloslava Vajdlová, Petr Valenta, Jana Zdeňková, 2018. Dostupné na: <https://vokabular.ujc.cas.cz/nezapojene.aspx?idz=VelNomT> [cit. 27-05-2025].

Bibliografie

- ABAEV, Vasilij I. (1958): *Istoriko-ětimologičeskij slovar' osetinskogo jazyka*, I. (A–K'). Moskva, Leningrad: Izdatel'stvo Akademii nauk SSR.
- ADAMS, Douglas Q. (1997): 'Nave' and 'Navel'. In: J. P. Mallory – D. Q. Adams (eds.), *Encyclopedia of Indo-European Culture*, 391. London – Chicago: Fitzroy Dearborn Publishers.
- ADAMS, Douglas Q. (2013): *A Dictionary of Tocharian B* (Revised and Greatly Enlarged), Vol. 1–2. Amsterdam – New York: Rodopi.
- AURA JORRO, Francisco (1985): *Diccionario micénico*, Vol. I. Madrid: Consejo superior de investigaciones científicas – Instituto de filología.
- BALLES, Irene (2008): *Nominale Wortbildung des Indogermanischen in Grundzügen*, Band 1: *Latein, Altgriechisch*. Hamburg: Verlag Dr. Kovač.
- BARTONĚK, Antonín (2003): *Handbuch des mykenischen Griechisch*. Heidelberg: Winter.
- BEEKES, Robert (1988): Laryngeal development: A survey. In: A. Bammesberger (ed.): *Die Laryngalthorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter, s. 59–105.
- BEEKES, Robert, with the assistance of Lucien van Beek (2010): *Etymological Dictionary of Greek*, I–II. Leiden, Boston: Brill.
- BERAN, Aleš (2015): Koloskopie nebo kolonoskopie? In: P. Falt – O. Urban – P. Vitek et al.: *Koloskopie*. Praha: Grada Publishing, s. 27–28.
- de BERNARDO STEMPEL, Patrizia (2023): *The Accents of Celtic. New Light on the Older and Oldest Stages*. Heidelberg: Winter.
- BLAŽEK, Václav – ŠMEJKALOVÁ, Martina (2023): Mapping Indo-European anatomical terminology I: Greek κόλον “large intestine”, Armenian *k'atird & k'atirt* “cavity of the body”, Brittonic **kalonā* “heart”, Tocharian B *kele* “navel; center”, and the word family “navel” – “nave.” In: *Journal of Indo-European Studies*, roč. 51, č. 1–2, s. 34–54.
- BOISACQ, Émile (1916): *Dictionnaire étymologique de la langue grecque*. Heidelberg: Winter – Paris: Klincksieck.
- BRUGMANN, Karl (1906): *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, 2.1. Strassburg: Trübner.
- CAMPANILE, Enrico (1974): *Profilo etimologico del cornico antico*. Pisa: Pacini.
- ĽERNÁ, Alena (1998): Nejstarší české názvy prstů. In: *Naše řeč*, roč. 81, č. 4, s. 189–196. Dostupné na: <http://nase-rec.ujc.cas.cz/archiv.php?art=7458> [cit. 06-10-2024].
- CHANTRAINE, Pierre (1933): *La formation des noms en grec ancien*. Paris: Klincksieck.
- CHANTRAINE, Pierre (1968): *Dictionnaire étymologique de la langue grecque: Histoire des mots*. Paris: Klincksieck.

- CHEUNG, Johnny (2007): *Etymological Dictionary of the Iranian Verb*. Leiden – Boston: Brill.
- DESHAYES, Albert (2003): *Dictionnaire étymologique du breton*. Abri du Marin: Le Chasse-Marée.
- DIRCKX, John H. (ed.) (1997): *Stedman's concise medical dictionary for the health professions*. Baltimore: Williams & Wilkins.
- DOČKALOVÁ, Lenka – BLAŽEK, Václav (2011): Indo-European Year. In: *Journal of Indo-European Studies*, roč. 39, č. 3–4, s. 414–495.
- ELSIE, Robert W. (1979): *The Position of Brittonic – A Synchronic and Diachronic Analysis of Genetic Relationships in the Basic Vocabulary of Brittonic Celtic*. Bonn: Inaugural-Dissertation der Rheinischen Friedrich-Wilhelms-Universität zu Bonn.
- ERNOUT, Alfred – MEILLET, Antoine (2001): *Dictionnaire étymologique de la langue latine. Histoire des mots*. Ed. Jacques André. Paris: Klincksieck.
- FALT, Přemysl – URBAN, Ondřej – VÍTEK, Petr et al. (2015): *Koloskopie*. Praha: Grada Publishing.
- FEW = *Französisches etymologisches Wörterbuch – eine Darstellung des galloromanischen Sprachschatzes* (1928–2003), von Walther von Wartburg et al. Basel: Helbing & Lichtenhahn, od r. 1952 Zbinden – Pais: Champion – Genève: Slatkine. Dostupné na: <https://apps.atilf.fr/lecteurFEW/index.php/page/view> [cit. 27-05-2025].
- FRISK, Hjalmar (1973): *Griechisches etymologisches Wörterbuch*. Heidelberg: Winter.
- FUČÍK, Mojmír (1972): *Základní vyšetření ve vnitřním lékařství: učebnice pro lékařské fakulty*. Praha: Avicenum.
- GPC = *Geiriadur Prifysgol Cymru* (1967–2002; online 2014). Ed. R. J. Thomas. Cardiff: University of Wales Press / Board of Celtic studies. Dostupné na: <https://geiriadur.ac.uk/gpc/gpc.html> [cit. 27-05-2025].
- HARTMANNOVÁ, Soňa (2013): *Staročeské názvy částí lidského těla a orgánů, jejich původ a motivace pojmenování*. Diplomová práce. Brno: Masarykova univerzita, Filozofická fakulta. Dostupné na: <https://theses.cz/id/zvnvd8/> [cit. 27-05-2025].
- HILMARSSON, Jörundur (1996): *Materials for a Tocharian Historical and Etymological Dictionary*. Tocharian and Indo-European Studies, Supplementary Series. Reykjavik: Málvísindastofnun Háskóla Íslands.
- HOLDER, Alfred (1896): *Alt-Celtischer Sprachschatz*, I. Leipzig: Teubner.
- HwR = *Handwörterbuch des Rätoromanischen* (1994). Rut Bernardi – Alexi Decurtins – Wolfgang Eichenhofer – Ursina Saluz – Moritz Vögeli. Zürich: Offizin.
- CHANTRAINE, Pierre (1968–1970): *Dictionnaire étymologique de la langue grecque: Histoire des mots*. I–II. Paris: Klincksieck.
- Internetová jazyková příručka (2008–2022). Praha: Ústav pro jazyk český AV ČR. Dostupné na: <http://prirucka.ujc.cas.cz/> [cit. 06-02-2023].
- KLINGENSCHMITT, Gert (1982): *Das altarmenische Verbum*. Wiesbaden: Reichert.
- KROONEN, Guus (2013): *Etymological Dictionary of Proto-Germanic*. Leiden – Boston: Brill.
- LEWIS, Charlton T. – SHORT, Charles (1958): *A Latin Dictionary*. Oxford: Clarendon Press.
- LIDDELL, Henry George – SCOTT, Robert (1996): *A Greek-English Lexicon*, revised and augmented throughout by Sir Henry Stuart Jones. Oxford: Clarendon Press.
- LIDÉN, Evald (1933): Zur alten tieranatomischen Terminologie. In: *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen*, roč. 61, č. 1–2, s. 14–28.
- LLOYD, Albert L. – LÜHR, Rosemarie (with the collaboration of Gerlinde Kohlrusch, Maria Kozianka, Karen K. Purdy and Roland Schuhmann) (2009): *Etymologisches Wörterbuch des Althochdeutschen 4*. Göttingen: Vandenhoeck & Ruprecht.

- MADSEN, Eirik (2015): Koloskopi eller kolonoskopi? In: *Tidsskrift for den Norske Lægeforening*, roč. 135, č. 7, s. 663–663. DOI 10.4045/tidsskr.15.0159 [cit. 02-12-2023].
- MACHAJDÍKOVÁ, Barbora (2013): Sollum Osce totum et solidum significat. Úloha Festových glos v poznání latinskej a italickej lexiky a jeho prínos k problematike „Saussurovho efektu“. In: *Sambucus*, roč. 9, s. 26–42.
- MACHEK, Václav (1968/2010): *Etymologický slovník jazyka českého*. Praha: Academia / Nakladatelství Lidové noviny.
- MARTIROSYAN, Hrach (2010): *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden – Boston: Brill.
- MATASOVIĆ, Ranko (2009): *Etymological Dictionary of Proto-Celtic*. Leiden – Boston: Brill.
- MAYRHOFER, Manfred (1986): *Etymologisches Wörterbuch des Altindoarischen*, I. Heidelberg: Winter.
- MAŽIULIS, Vytautas (2013): *Prūsų kalbos etimologijos žodynas*. Vilnius: Mokslo ir enciklopedijų leidybos centras.
- MEIER, Michael (1975): -ið-. *Zur Geschichte eines griechischen Nominalsuffixes*. (Ergänzungshefte zur Zeitschrift für vergleichende Sprachforschung Nr. 23). Göttingen: Vandenhoeck & Ruprecht.
- MEIER-BRÜGGER, Michael (1990): Zu griechisch κολός. In: *Historische Sprachforschung / Historical Linguistics*, roč. 103, č. 1, s. 30–32.
- MEYER-LÜBKE, Wilhelm (1935): *Romanisches etymologisches Wörterbuch*. Heidelberg: Winter.
- MILITAREV, Alexander – KOGAN, Leonid (2000): *Semitic Etymological Dictionary: Anatomy of Man and Animals*. Münster: Ugarit-Verlag.
- MONIER-WILLIAMS, Monier (1899/1993): *A Sanskrit-English Dictionary*. Delhi: Motilal Banarsidass.
- OLSEN, Birgit A. (1999): *The Noun in Biblical Armenian. Origin and Word-Formation – with Special Emphasis on the Indo-European Heritage*. Berlin – New York: de Gruyter Mouton (Trends in Linguistics. Studies and Monographs 119).
- Online Etymology Dictionary: Colonoscopy*. Dostupné na: https://www.etymonline.com/word/colonoscopy#etymonline_v_28358 [cit. 23-02-2023].
- OREL, Vladimir E. – STOLBOVA, Olga V. (1995): *Hamito-Semitic Etymological Dictionary. Materials for a Reconstruction*. Leiden – New York – Köln: Brill.
- PEDERSEN, Holger (1909–1913): *Vergleichende Grammatik der keltischen Sprachen*, I–II. Göttingen: Vandenhoeck & Ruprecht.
- PETERSSON, Herbert (1921): *Studien über die indogermanische Heteroklisie*. Lund: Gleerup.
- PINAULT, Georges-Jean (2008): *Chrestomathie tokharienne. Textes et grammaire*. Leuven – Paris: Peeters.
- POKORNY, Julius (1959): *Indogermanisches etymologisches Wörterbuch*. Bern – München: Francke.
- PRONK, Tijmen (2015): Singulative *n*-stems in Indo-European. In: *Transactions of the Philological Society*, roč. 113, č. 3, s. 327–348.
- RIX, Helmut (ed.) (2001): *Lexikon der indogermanischen Verben*, by Martin Kümmel – Thomas Zehnder – Reiner Lipp – Brigitte Schirmer. Wiesbaden: Reichert.
- SCHMITT, Rüdiger (1981): *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: Institut für Sprachwissenschaft (Innsbrucker Beiträge zur Sprachwissenschaft 22).

- SCHWYZER, Eduard (1939): *Griechische Grammatik*, I. Band: *Allgemeiner Teil – Lautlehre – Wortbildung – Flexion*. München: Beck.
- SMOCZYŃSKI, Wojciech (2018): *Lithuanian Etymological Dictionary*. Eds. A. Holvoet – S. Young with the assistance of W. Browne. Berlin – Bern – Bruxelles – New York – Oxford – Warszawa – Wien: Peter Lang.
- STANĚK, Václav (1863): *Slovník lékařské terminologie*. Praha: Spolek českých lékařův.
- STERN, Heinrich (1911): Technic of Colonoscopy. In: *The Archives of Diagnosis*, roč. 4, s. 42–56.
- ŠMEJKALOVÁ, Martina – BLAŽEK, Václav (2023): Koloskopie, nebo kolonoskopie? In: *Časopis lékařů českých*, roč. 162, č. 5, s. 212–216.
- ŠPAČEK, Bohumil et al. (1973): *Speciální chirurgie*. Praha: Avicenum.
- TESAŘ, Otakar (1967): *Prudká krvácení do trávicí trubice*. Praha: Státní zdravotnické nakladatelství.
- TURNER, Ralph L. (1966): *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press.
- UNTERMANN, Jürgen (2000): *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.
- de VAAN, Michiel (2008): *Etymological Dictionary of Latin and the other Italic Languages*. Leiden – Boston: Brill.
- VINE, Brent (1999): On “Cowgill’s Law” in Greek. In: H. Eichner et al. (eds.): *Compositiones Indogermanicae in memoriam Jochem Schindler*. Prague: Enigma Corporation, s. 555–600.
- de VRIES, Jan (1977): *Altnordisches etymologisches Wörterbuch*. Leiden: Brill.
- WALDE, Alois – HOFMANN, Johann B. (1938): *Lateinisches etymologisches Wörterbuch*, I. Heidelberg: Winter.
- WEISS, Michael (2009): *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor – New York: Beech Steve Press.

TRAINING OF LARGE LANGUAGE MODEL MISTRAL ON SLOVAK LANGUAGE DATA

PETER BEDNÁR¹ – MAREK DOBEŠ² – RADOVAN GARABÍK³

¹Faculty of Electrical Engineering and Informatics of the Technical University of Košice, Košice, Slovakia

²Centre of Social and Psychological Sciences, Slovak Academy of Sciences, Košice, Slovakia

³L. Štúr Institute of Linguistics, Slovak Academy of Sciences, Bratislava, Slovakia

BEDNÁR, Peter – DOBEŠ, Marek – GARABÍK, Radovan: Training of large language model Mistral on Slovak language data. *Jazykovedný časopis (Journal of Linguistics)*, 2025, Vol. 76, No. 2, pp. 433–451.

Abstract: This study investigates the adaptation of the Mistral 7B large language model for the Slovak language, addressing the limited availability of high-quality open-source models for low-resource languages. While commercial models like GPT-4 and Claude exhibit strong Slovak proficiency, their proprietary nature restricts transparency and customization. To overcome this, we fine-tuned the open-weight Mistral 7B model using the Araneum Slovaccum VII Maximum corpus (5.3 billion tokens), creating a specialized Slovak variant, Mistral-SK-7b. The training, conducted on the Leonardo supercomputer (10,000 GPU hours), yielded significant improvements: the fine-tuned model generates grammatically correct and contextually coherent Slovak text, eliminating the errors (code-switching, repetition loops, and lexical interference from other languages) observed in the original Mistral-7B-v0.1. The resulting model, released under the Apache 2.0 license, provides a publicly accessible resource for Slovak NLP applications while preserving the base model's multilingual capabilities. Our work demonstrates the feasibility of adapting state-of-the-art LLMs for linguistically underrepresented languages and underscores the role of open models in promoting digital language preservation.

Keywords: large language models, Mistral, computational linguistics, Slovak language, Natural Language Processing, model fine-tuning

1. INTRODUCTION

1.1 Large Language Models in Natural Language Processing

Large language models (LLMs) are deep neural networks designed to learn and generate text in natural language. They achieve this through extensive training on vast amounts of textual data, allowing them to discern and create patterns that resemble human language. The training process involves analyzing numerous texts, from which the models build statistical relationships among letters, tokens, words, sentences, and larger text structures. As a result, LLMs are capable of producing coherent and contextually appropriate text.

The architecture of LLMs typically involves several layers of neural networks, each layer responsible for processing different aspects of the text. The initial layers might focus on identifying patterns at the letter or word level, while the deeper layers understand sentence structure and context.

One of the critical advancements in LLMs is the use of transformers, a type of neural network architecture that allows for better handling of long-range dependencies in text (Vaswani et al. 2017). Transformers have revolutionized the field by enabling models to consider the entire context of a sentence or paragraph, rather than processing text sequentially. This innovation has significantly improved the performance of LLMs in various natural language processing tasks, such as translation, summarization, and question answering (Devlin et al. 2019).

Training LLMs requires massive computational resources and access to extensive text corpora. For example, OpenAI's GPT-3 model was trained on hundreds of billions of words from diverse sources, including books, articles, and websites (Brown et al. 2020). The scale of data and computation involved in training these models allows them to capture a wide range of linguistic nuances and knowledge.

1.2 Accessibility and Limitations of Large Language Models

The most extensive language models currently available, such as GPT-4 and Claude, are primarily accessible only to the companies that develop them. These companies provide public access to their models under specific terms and conditions. For instance, OpenAI's GPT-4 is accessible through an API or a web based chat interface, but usage is governed by pricing, rate limits, and other restrictions set by OpenAI (OpenAI 2023). Similarly, Anthropic's Claude is another example of a powerful language model available to users under the conditions defined by its creators (Anthropic 2023).

One significant drawback of this proprietary model approach is the sudden switch to the lack of transparency. While some of the first generative LLMs were often rather transparent (including a description of the training data, training algorithms and parameters used for training), the landscape of the most performant LLMs changed to completely opaque. The broader community lacks detailed knowledge about the architecture, limitations, and training data of these models. This opacity can hinder academic research and independent verification of the models' capabilities and biases (Bommasani et al. 2021). Without access to the underlying details, it's challenging to fully understand how these models operate, their potential biases, or the robustness of their outputs. It is often not even possible to find out the proportion of multilingual data in the models, which makes it difficult to meaningfully compare performance across the models in less resourced languages (such as Slovak).

Because of these limitations, there is a growing movement towards the availability of open models. These open-source models can be downloaded,

modified, and used by anyone under creative licenses. This openness facilitates a more inclusive and collaborative approach to model development and usage, allowing for community-driven improvements and adaptations.

The availability of open models also democratizes access to advanced language processing capabilities, enabling smaller organizations and individual developers to leverage state-of-the-art tools without prohibitive costs. These open models can be fine-tuned for specific tasks, making them highly adaptable to various applications, from academic research to commercial use.

1.3 Multilingual Large Language Models and Slovak Language

While most language models are predominantly trained in English, there is a growing number of models that support multiple languages, including Slovak. Proprietary models such as GPT-4, Claude, and Gemini have shown good proficiency in Slovak due to extensive training on diverse multilingual datasets (OpenAI 2023; Anthropic 2023; Google 2024). These models benefit from the substantial resources and computational power available to their developers, enabling them to achieve high-quality outputs across various languages.

In addition to these closed models, there are also open-source models supporting Slovak, such as Gemma-3-4b-it (Google 2025) or Llama-3.3-70B-Instruct (Meta 2025). These open models, however, still face several challenges and exhibit notable deficiencies in their Slovak language capabilities. One significant factor contributing to this discrepancy is the limited availability of publicly accessible Slovak text data. The scarcity of high-quality, diverse training data in non-English languages, including Slovak, constrains the model's ability to learn and generalize effectively (Blasi et al. 2022).

Moreover, the development of open models often lacks the extensive fine-tuning and quality control processes that proprietary models undergo. Fine-tuning for specific languages, such as Slovak, requires significant time, expertise, and computational resources, which may not be as readily available in the open-source community (Touvron et al. 2023). As a result, the Slovak outputs from these open models may not yet match the quality of those produced by their closed counterparts.

To address these gaps, we present the following contributions:

- Slovak language adaptation: We fine-tune the Mistral 7B model – a state-of-the-art open-weight LLM – using the Araneum Slovacum VII Maximum corpus (5.3B tokens), creating Mistral-SK-7b, the first publicly available Slovak-tuned model of this scale.
- Performance validation: Our qualitative evaluation demonstrates that the fine-tuned model eliminates grammatical errors and code-switching artifacts prevalent in the base Mistral-7B-v0.1, achieving near-native fluency in Slovak text generation.
- Resource-efficient methodology: By leveraging the Leonardo supercomputer (10k GPU hours), we optimize the training process while preserving the mod-

el’s multilingual capabilities, providing a blueprint for adapting LLMs to other low-resource languages.

- Open-access release: The model is distributed under the Apache 2.0 license, enabling community use and further research in Slovak NLP applications.

2. RESOURCES

2.1 Decision on Model Training Strategy for Slovak Language

In our research, we faced a critical decision regarding the development of a Slovak language model: whether to train a native model from scratch or to fine-tune an existing model using Slovak text data. Training a model from the ground up demands an enormous amount of training data and computational resources. This process involves collecting and preprocessing a vast corpus of Slovak text, as well as dedicating significant time and computational power to the training process.

Given these substantial requirements, we opted to utilize an existing pre-trained model and fine-tune it with Slovak-specific data. This approach leverages the comprehensive training that the original model has already undergone, which includes learning general language patterns and structures (Radford et al. 2019). Fine-tuning enables us to adapt the model specifically to Slovak, using a smaller, more manageable dataset while still achieving high-quality results (Howard – Ruder 2018).

2.2 Mistral 7B Large Language Model

Mistral is a family of state-of-the-art language models in several sizes. The smallest version Mistral 7B (Mistral AI, 2023) has 7.3 billion parameters and is designed to offer high performance despite its relatively compact size. The model leverages advanced architectural features such as Grouped-Query Attention (GQA) and Sliding Window Attention (SWA) to enhance both its speed and ability to handle long sequences efficiently.

A key feature of Mistral 7B (Mistral 7B v0.1 that we use in our research) is its accessibility under an open-source Apache 2.0 license, which encourages broad usage and adaptation. This stands in contrast to many proprietary models that restrict usage through licensing terms. The open-weight nature of Mistral 7B means it can be fine-tuned for specific applications, providing flexibility for developers and researchers to tailor the model to their needs. On the other hand, the composition of training data (including the amount of texts in specific languages) has been deliberately kept undisclosed by the authors, thus limiting its transparency.

We decided to utilize the Mistral 7B model for several key reasons. Firstly, Mistral 7B is an open model, which allows us to fine-tune it with our selected data. The open-access nature of Mistral 7B, makes it an ideal candidate for customization and specific applications. This flexibility is crucial for adapting the model to handle Slovak text effectively.

Secondly, the model's parameter count is suitable for training with the computational resources and data we have available. Mistral 7B, with its 7.3 billion parameters, strikes a balance between complexity and manageability, enabling efficient training without requiring the extensive resources needed for much larger models like GPT-3, which has 175 billion parameters (Brown et al. 2020), and an efficient inference even on low end hardware.

Thirdly, Mistral 7B has been apparently trained only on sparse Slovak data. Starting with a model that lacks heavy prior exposure to Slovak allows us to monitor the training process closely and observe the development of structures responsible for understanding and generating Slovak text. This real-time insight into the model's learning process can be invaluable for refining and optimizing our fine-tuning approach (Radford et al. 2019).

Here are some key aspects of the model's architecture and parameters:

- **Transformer Architecture:** Mistral 7B is based on a decoder-only Transformer architecture. This structure is widely used in modern language models for its effectiveness in handling sequential data and generating coherent text.
- **Grouped-Query Attention (GQA):** This mechanism allows for faster inference and reduces memory usage by efficiently managing how queries, keys, and values are processed in the attention layers. It optimizes the attention calculation, making it suitable for large-scale models.
- **Sliding Window Attention (SWA):** SWA enables the model to handle longer context windows more effectively. It uses a sliding window approach to maintain a fixed-size cache, allowing the model to process sequences up to 128K tokens in length theoretically, although it typically uses an 8K context during training.
- **Byte-fallback BPE Tokenizer:** This tokenizer ensures that all characters are represented within the model's vocabulary, preventing out-of-vocabulary issues which are common in language models.
- **Model Parameters:** Mistral 7B contains approximately 7.3 billion parameters. It includes 32 attention heads, each with a dimension of 128, and a total hidden size of 4096. The model also uses rotary position embeddings to encode positional information effectively.

2.3 Training Data for Slovak Language Model

For training our Slovak language model, we are using the Araneum Slovacum VII Maximum web corpus, a state-of-the-art web corpus of Slovak language. The corpus forms a part of the ARANEA family of web corpora for two dozen languages, compiled by the Ľudovít Štúr Institute of Linguistics, Slovak Academy of Sciences, and the UNESCO Chair in Plurilingual and Multicultural Communication, Comenius University in Bratislava (Benko 2024). The size of the

Slovak corpus is approximately 5.3 billion tokens¹, 4.4 billion words, 280 million sentences, 106 million paragraphs, 14 million documents (i.e. web pages), offering a substantial volume of text necessary for effective model training, and at the time we started to train the model, it was the largest Slovak language corpus available to us for the training.

The corpus is built upon data that are being collected from 2011, with the latest collection in October 2023, thus offering reasonably up to date information, but also a vast amount of texts from recent history. The corpus is deduplicated on a document (i.e. webpage) and a paragraph level, and filtered for “reasonable quality” language (i.e. without texts lacking diacritics). Moreover, it is automatically lemmatized, POS tagged and morphosyntactically analyzed (features that are not used in LLM training).

Table 1 lists selected currently existing big or open-access corpora for Slovak. There is a significant overlap in the two web corpora in the table, the third one (ARANEUM + HPLT + FineWeb 0.1) is a union of existing three big Slovak web datasets, deduplicated on paragraph level and is currently the biggest existing general language corpus of Slovak (Garabík 2025; Webový korpus slovenčiny 2025).

<i>corpus</i>	<i>size</i> [million tokens]	<i>license</i>	<i>type</i>
Araneum Slovacom VII Maximum	5300	MIT/CC-0	web corpus
HPLT 2.0 Slovak Corpus v0.4	4090	CC-0	web corpus
ARANEUM + HPLT + FineWeb 0.1	104720	MIT/CC-0/ODC-by	web corpus
wiki-2019-08	50	CC-BY SA	Wikipedia & Necyklopédia
Corpus of Court Proceedings 3.0	11974	exempt from copyright	specialized
Corpus of Legal Texts 1.9	45	exempt from copyright	specialized
prim-11-public-all	1859	unavailable	national corpus

Table 1: List of selected big Slovak corpora. *Size* is in linguistic tokens (i.e. words + punctuation); *license* is the license of the package and annotation, not the content.

In addition to leveraging these existing corpora, we have initiated negotiations with various institutions in Slovakia that may have access to additional digitalized Slovak texts. These efforts aim to further expand our dataset, thereby enhancing the robustness and accuracy of our model. Access to a broader range of texts will be particularly valuable for fine-tuning and validating the model across diverse linguistic contexts and domains.

By combining publicly available linguistic resources with potential new sources of digitalized texts, we aim to develop a comprehensive and high-quality Slovak

¹ By a “token” here we mean either a word, number or a punctuation character, as usual in corpus linguistics. We feel the need to explicitly point this out, given a different usage of the term “token” in the context of LLMs.

language model. This approach not only maximizes the use of existing data but also opens opportunities for collaboration and further research in the field of computational linguistics.

2.4 Computational Resources for Model Training

To facilitate the training of our Slovak language model, we secured a grant from the National Supercomputing Centre Slovakia², which provides us with GPU computing access on the European supercomputer Leonardo. Leonardo is one of the most powerful supercomputers in Europe, ranking among the top in the Top500 list. It features NVIDIA Ampere A100 GPUs, each equipped with 64GB of high-bandwidth memory (HBM2e). This configuration is particularly advantageous for training large language models due to its high memory capacity and efficient data handling capabilities (HPC Cineca 2023).

Leonardo's architecture includes both Booster and Data-centric modules. The Booster module comprises BullSequana X2135 blades, each with a single Intel Xeon 8358 CPU and four NVIDIA A100 GPUs. This setup provides substantial computational power and memory, essential for handling extensive training datasets and complex model architectures. The Data-centric module, featuring Intel Sapphire Rapids CPUs, complements the Booster module by supporting a wide range of applications with high computational demands.

2.5 Virtual Environment and Dependencies

To run our Slovak language model effectively on the Leonardo supercomputer, we set up a specific environment to utilize the available computational resources efficiently. Here's an overview of the necessary environment and dependencies:

Module Loading:

We begin by loading essential modules that provide the necessary software environment for deep learning tasks. This includes a specific Python version and CUDA for GPU acceleration. The relevant modules are:

- `profile/deeplrn`: A module that provides an optimized environment for deep learning applications.
- `python/3.11.6--gcc--8.5.0`: This module loads Python version 3.11.6, compiled with GCC 8.5.0, ensuring compatibility with high-performance computing applications.
- `cuda`: CUDA is essential for leveraging the GPU capabilities of the Leonardo supercomputer, which uses NVIDIA A100 GPUs with 64GB of HBM2e memory.

² <https://nsccl.sk>

Python Virtual Environment:

To manage dependencies, we create a Python virtual environment. This isolated environment helps to avoid conflicts between different package versions and system-wide Python packages.

Package Installation:

Within the virtual environment, several key Python packages are installed to support the model training and execution:

- wheel and setuptools: These are fundamental packages for building and distributing Python packages.
- torch and accelerate: PyTorch is a popular deep learning framework, and Accelerate helps in optimizing the training process across multiple GPUs.
- datasets: This library provides easy-to-use tools for loading and preprocessing datasets.
- transformers: Essential for working with transformer-based models, such as our language model.
- flash-attn: Optimized attention mechanisms for faster model training.
- sentencepiece: A tokenizer that handles the text preprocessing needed for our language model.
- bitsandbytes: Useful for optimizing memory usage during model training.
- tensorboardX: Allows us to visualize and track the training process using TensorBoard.

This setup ensures that our environment is optimized for the high computational demands of training large language models, leveraging the full capabilities of the Leonardo supercomputer.

2.6 Memory Management

The computational power of a supercomputer enables handling models that are both computationally and memory intensive. This capability is essential for running advanced language models, but it also complicates the programming process. Several key considerations must be addressed to optimize the model's performance on such a powerful system.

Firstly, decisions must be made to ensure the model can run in parallel across multiple cores. This parallelization is crucial to leverage the supercomputer's full potential and reduce training times significantly. Techniques such as data parallelism and model parallelism are employed to distribute the workload effectively across the available cores. Data parallelism involves splitting the dataset into smaller chunks, which are then processed simultaneously by different cores. Model parallelism, on the other hand, involves splitting the model itself so that different parts of the model are processed by different cores (Dean et al. 2012).

Secondly, efficient communication between cores is paramount. The interconnects used in supercomputers like Leonardo, such as NVIDIA Mellanox HDR 200 Gb/s InfiniBand, provide high bandwidth and low latency, which are critical for minimizing the communication overhead between cores. Effective communication strategies ensure that data transfer between cores does not become a bottleneck, thereby maintaining high computational efficiency.

Lastly, maximizing the memory usage is essential to handle large models. The GPUs in Leonardo, such as the NVIDIA A100 with 64GB of HBM2e memory, allow for large portions of the model to be loaded into memory at once. This capability reduces the need for frequent memory swaps, which can significantly slow down the processing. Techniques like gradient checkpointing can also be used to save memory during training by storing only a subset of intermediate states and recomputing them as needed during the backward pass (Chen et al. 2016).

2.7 Training Process

The training program for the Mistral 7B model is designed to efficiently utilize available computational resources and optimize the training process. Here is an overview of the training script we use.

Environment Setup:

We start by setting up the working directories and loading the necessary modules and datasets. The `WORK_DIR`, `OUTPUT_DIR`, `DATASET_DIR`, and `BASE_MODEL_DIR` are defined to organize where the datasets and models are stored and where the outputs will be saved.

Model and Tokenizer Initialization:

The `train` function begins by loading the pre-trained Mistral 7B model using the `AutoModelForCausalLM` class from the Hugging Face library. This model is configured to use `bfloat16` precision and optimized for low CPU memory usage. The `AutoTokenizer` is also loaded to handle tokenization of the text data. The original Mistral tokenizer was used.

Data Preparation:

The tokenizer's padding token is set to the end-of-sequence token to handle padding efficiently. The training and validation datasets are packed using the `pack` function to ensure that the context length is fully utilized with fragments from one or more documents. This approach is customary in training large language models to maximize the usage of the context window.

Training Configuration:

Our script sets several key training parameters:

- **Batch Size and Gradient Accumulation:** The per-device training batch size is set to 2, with gradient accumulation steps set to 8. This means that gradients are accumulated over 8 batches before performing a weight update. With 1 node and 4 GPUs, the effective mini-batch size was $2 \times 8 \times 4 = 64$.

- **Training Steps:** The total number of training steps is calculated based on the available GPUs, batch size, gradient accumulation steps, and context length.
- **Save and Evaluation Steps:** The model is configured to save checkpoints and perform evaluations at intervals determined by dividing the total training steps.

Training Arguments:

The `TrainingArguments` class is used to specify various training parameters such as the optimizer (`adamw_bnb_8bit`), learning rate, learning rate scheduler (cosine), warmup steps, evaluation strategy, and others. These settings are crucial for managing the training process, ensuring that it runs smoothly and efficiently.

Trainer Initialization and Training:

The `Trainer` class is instantiated with the model, training arguments, tokenizer, and datasets. The `DataCollatorForLanguageModeling` is used to handle data collation without masking (as it is not required for causal language modeling).

The training process is started with the `trainer.train` method, which optionally resumes from the last checkpoint if available.

The training ran successfully, with loss diminishing as seen in Figure 1. The total training time was approximately 10 thousand GPU hours.

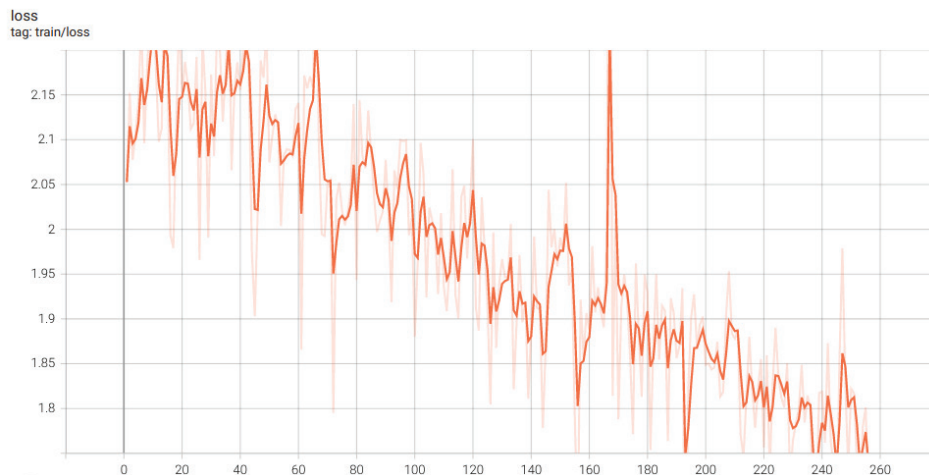


Figure 1: Loss function across the training. X-axis: epochs, Y-axis: loss value.

The unusual spike in training loss around Epoch 160 likely stems from a combination of unclipped gradients and optimizer dynamics. As gradient clipping was not used, large gradients could have caused abrupt updates to the model's weights, temporarily destabilizing training. This effect might have been amplified by the 8-bit

Adam optimizer (adamw_bnb_8bit), which – while memory-efficient – can sometimes exhibit higher variance in gradient updates. Additionally, the spike could reflect a batch of unusually noisy or challenging samples, creating a brief divergence before the model recovered. Notably, the loss quickly resumed its downward trend, suggesting the overall training process remained robust.

3. EVALUATION OF THE MODEL

To assess the performance of large language models (LLMs), various metrics have been developed or adapted from existing machine translation metrics. These metrics can broadly be classified into qualitative and quantitative categories. Quantitative metrics, such as perplexity, BLEU (Bilingual Evaluation Understudy; Papineni et al. 2002) and ROUGE (Recall-Oriented Understudy for Gisting Evaluation; Lin, Och, 2004), provide numerical evaluations of language model outputs based on their accuracy, fluency, and relevance to a reference text. On the other hand, qualitative metrics involve human evaluations that focus on aspects like coherence, creativity, and context appropriateness. While these metrics have been instrumental in evaluating LLMs, they face significant limitations when applied to languages other than English. Most metrics, particularly quantitative ones, rely on linguistic features or reference datasets that are heavily biased towards English, making it challenging to accurately gauge the performance of LLMs in non-English contexts (Bender et al. 2021; Ruder et al. 2021). Moreover, qualitative evaluations often require native speakers and culturally relevant contexts, which are not always readily available or well-defined for many languages (Blasi et al. 2022). Often, existing evaluation datasets are just machine translated from English.

We have not performed any thorough quantitative evaluation (apart from perplexity) of our model, because of lack of any relevant Slovak language metrics. Such a battery of general tests is being prepared by a wide community of researchers, and we plan to evaluate and compare our model, once the test suite is ready. We performed simple tests evaluating the percentage of out-of-lexicon words (based on comprehensive lexicographic and morphological databases), but our model already saturated the accuracy, being virtually indistinguishable from corpus samples of natural language.

Based on our (admittedly limited) experiments with the model, we conclude that it retains its original capability of in-context learning. For example, when presented with a prompt describing a bilingual dictionary and short bilingual list of Slovak-English (or English-Slovak) word pairs, the model outputs the correct equivalent of the following word, followed by a surprisingly accurate bilingual glossary. However, the results for other languages apart from English are very inaccurate. The model also lost the ability to freely switch the output language to

that of the prompt and keeps switching to Slovak, after a few words in the language of the prompt.

For comparison, we include several examples of text generation of the original Mistral-7B-v0.1 model and our mistral-sk-7b model, with the prompt marked with underline.

We used greedy search for these examples, which picks a token with the highest probability as the next token, and no repetition penalty (reflected in the first and last output of the Mistral-7B-v0.1 model).

The output from the original Mistral-7B-v0.1, we mark grammatical and lexical (but not factual errors nor texts in English) errors with text *in italics*:

Kde bolo, tam bolo, bol raz jeden kráľ. A ten kráľ mal tri dcéry.³ A dcéry *maly tri synov. A synov maly tri dcéry. maly tri synov. A synov maly tri dcéry. maly tri synov. A synov maly tri dcéry. maly tri synov. A synov maly tri dcéry.*

Veľké jazykové modely sú novým trendom v umelej inteligencii. Sú zaujímavé najmä⁴
pre to, že sú vytvorené na základe veľkých množín textov, ktoré sa vyskytujú v rôznych jazykových korpusoch.

Na základe dostupných informácií vieme, že Ľudovít Štúr sa narodil⁵ 28. 10. 1815 v obci Hlboké nad Váhom.

Ľudovít Štúr

Ľudovít Štúr was born on 28 October 1815 in the village of Hlboké nad Váhom. He was the son of a Lutheran pastor, Ján Štúr, and his wife, Anna, née Kollár.

Sociálna psychológia je pododbor psychológie, ktorý sa zaoberá⁶ vzťahmi medzi *ludmi*. Sociálna psychológia sa *zabýva* vzťahmi medzi *jednotlivci*, skupinami, organizáciami a širším spoločenským kontextom.

Slnko je v znamení⁷ raka

The Sun is in the sign of Cancer

The Sun is in the sign of Cancer.

The Sun is in the sign of Cancer.

The output from mistral-sk-7b (remarkably, there are no grammatical errors in this output):

Kde bolo, tam bolo, bol raz jeden kráľ. A ten kráľ mal tri dcéry. Najstaršia bola princezná, ktorá sa volala Zlatovláska. Druhá bola princezná, ktorá sa volala Ružovláska. A tretia bola princezná, ktorá sa volala Blondáta.

Veľké jazykové modely sú novým trendom v umelej inteligencii. Sú zaujímavé najmä
preto, že dokážu spracovať veľké množstvo dát a vytvoriť z nich zmysluplné výstupy.

³ Once upon a time, there was a king. And that king had three daughters.

⁴ Large language models are a new trend in artificial intelligence. They are interesting chiefly

⁵ According to available information we know that Ľudovít Štúr was born

⁶ Social psychology is a subfield of psychology dealing (with)

⁷ The Sun is in the sign (of)

Veľké jazykové modely sú založené na princípe transformácie prirodzeného jazyka na abstraktný jazyk, ktorý je ľahšie spracovateľný počítačom.

Na základe dostupných informácií vieme, že Ľudovít Štúr sa narodil 28. októbra 1815 v Uhrovci. V roku 1830 sa stal študentom evanjelického lýcea v Bratislave. V roku 1833 sa stal študentom evanjelického kolégia v Prešove.

Sociálna psychológia je pododbor psychológie, ktorý sa zaoberá skúmaním vzťahov medzi ľuďmi. Jej predmetom je skúmanie vzťahov medzi ľuďmi, ich správanie, vzájomné ovplyvňovanie, vzťahy medzi ľuďmi a ich prostredím.

Slnko je v znamení Vodnára a je v trigóne s Jupiterom v znamení Rýb a v kvadráte s Uránom v znamení Barana. Merkúr je v znamení Vodnára a je v trigóne s Venušou v znamení Rýb a v kvadráte s Plutom v znamení Kozorožca.

The texts generated by the original Mistral-7B-v0.1 are riddled with numerous grammatical and stylistic errors. They also frequently incorporate words from Czech, sometimes also from Croatian or Slovene⁸, and often switch to English, or get stuck in repetition loops. Almost all of these issues are also visible in our examples above. This aligns with previous observations that untrained models often lack coherence and consistency in language-specific outputs (Ruder et al. 2021). In contrast, the second text, produced by the trained model, does not contain any mistakes in the provided example (which was not specifically selected). In fact, during our work with the model, we noticed grammatical mistakes are very rare and not obvious (a kind of mistakes a native speaker would make as well).

This demonstrates the positive impact of targeted training on language models' ability to generate more linguistically accurate and culturally appropriate content (Blasi et al. 2022).

4. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Conclusions

Large language models can be trained for various purposes, including text generation, question answering, and text error checking, among others. These models leverage vast amounts of data to understand and produce human-like text, adapting to different tasks based on their training objectives and data.

⁸ After Czech, Polish and Sorbian(s) are the nearest literary languages to Slovak. However, Sorbian is used very infrequently and we would be surprised if any significant amount of Sorbian texts made it into the original training data. On the other hand, Polish is very well represented on the Internet, but the orthography is significantly distinct from Slovak, which leaves Slovene and Croatian (and its sister Latin script languages) as the languages with somewhat compatible orthography for the model to draw from. We observed this behaviour (mixing or switching to a mixture of Slovene and Croatian) with several other available LLMs with low support for Slovak.

In the first phase of our project, we decided to update an existing model for generating Slovak text. Given the nature of our training data, focusing on text generation is the most straightforward approach. Training for text generation involves feeding the model extensive datasets of Slovak text, allowing it to learn linguistic patterns, structures, and vocabulary specific to the language. This foundational step ensures that the model can produce coherent and contextually relevant Slovak text.

Text generation is a crucial capability, serving as the backbone for many applications, including automated content creation, conversational agents, and language translation (Brown et al. 2020). By starting with this fundamental task, we can establish a robust baseline for our model, making it easier to expand its capabilities in the future.

The research presented in the article highlights the significant disparity between commercial and open-source language models in handling the Slovak language. While proprietary models like GPT-4 and Claude have demonstrated strong capabilities in generating and understanding Slovak and other low-resource languages due to extensive multilingual training datasets, open-source models such as Gemma and LLaMA3 lag behind in their understanding and production of the language (Dargis et al. 2024, Costa-Jussa et al. 2022). This gap is primarily due to the limited availability of high-quality Slovak text data and the lack of extensive fine-tuning in open models.

To address these challenges, we decided to fine-tune the existing Mistral 7B model with Slovak text data. This approach offers several advantages:

- **Open Accessibility:** By using an open-source model, we ensure that the resulting Slovak language model is publicly accessible, promoting further research and application development.
- **Resource Efficiency:** Fine-tuning an existing model is more resource-efficient compared to training a new model from scratch, leveraging pre-existing training on general language patterns.
- **Monitoring and Insights:** Training the model on Slovak data from scratch provides valuable insights into the formation of Slovak linguistic structures within the neural network, contributing to our understanding of language processing.

Our initial focus was on text generation due to the straightforward nature of the available training data. This foundational capability is essential for various applications, including automated content creation and conversational agents. The model is available under Apache 2.0 license at <https://www.juls.savba.sk/llm.html>.

Unresolved Questions

Several unresolved questions and challenges emerged from our research:

1. **Data Scarcity:** The limited availability of high-quality, diverse Slovak text data remains a significant barrier. There is a need for more comprehensive corpora to improve model training and performance.
2. **Fine-Tuning Strategies:** Determining the most effective strategies for fine-tuning models on low-resource languages like Slovak. This includes identifying the optimal balance between data volume and model performance.
3. **Model Evaluation:** Establishing robust evaluation metrics and benchmarks for Slovak language models to ensure consistent and accurate performance assessment.

Future Research Directions

Future research should focus on the following areas to further advance the development and application of Slovak language models:

1. **Data Collection and Curation:** Collaborating with institutions to gather and curate more extensive and diverse Slovak text corpora. This will provide a richer training dataset, enhancing model performance and generalization.
2. **Multilingual Training Techniques:** Exploring advanced multilingual training techniques that can leverage cross-linguistic data to improve performance on low-resource languages. In particular, extensive fine-tuning an existing multilingual model with additional monolingual data might diminish or destroy capabilities in other languages.
3. **Contextual Understanding:** Developing models with enhanced capabilities for contextual understanding and reasoning in Slovak. This includes training on specialized tasks such as question answering and dialogue systems.
4. **Community Engagement:** Encouraging the open-source community to contribute to the development and fine-tuning of Slovak language models. This collaborative approach can drive innovation and improve availability of models.
5. **Evaluation of Generalization Capabilities:** Future work will be to assess whether fine-tuned Slovak models (like Mistral-SK 7B) retain their inherent zero-shot and few-shot learning abilities after task-specific adaptation. Key steps include designing standardized benchmarks for Slovak to test multi-task performance (e.g., translation, sentiment analysis) without task-specific training, quantifying trade-offs between specialization (e.g., NER accuracy) and generalizability (e.g., few-shot prompt adaptability), and comparing fine-tuned models with their base versions to identify how much linguistic flexibility is lost during domain adaptation. This is important for applications requiring dynamic task switching, such as virtual assistants or educational tools. Once a training database for Slovak question answering becomes available, we plan to extend the model's functionality to include this task. Question answering requires the model to understand and generate precise

responses based on the given queries, which involves a more complex understanding of context and information retrieval.

- 6. Comparative Human Evaluation of Slovak Language Models:** So far, we performed only very limited qualitative tests of the model. Although we already prepared a set of prompts in Slovak specific to Slovak cultural environment, a thorough evaluation of several Open Source models (including ours) is yet to be performed. The evaluation criteria could include fluency, grammatical correctness, cultural appropriateness, and task adherence. While resource-intensive, this comparative analysis would provide valuable insights into the relative strengths of different approaches for Slovak language processing and establish clearer benchmarks for model quality in this linguistic context.

Once a training database for Slovak question answering becomes available, we plan to extend the model's functionality to include this task. Question answering requires the model to understand and generate precise responses based on the given queries, which involves a more complex understanding of context and information retrieval (Rajpurkar et al. 2016). This extension will enhance the model's versatility, enabling it to handle a wider range of applications, such as educational tools and customer support systems.

In summary, while significant progress has been made in adapting large language models for the Slovak language, continued efforts are needed to address data limitations and improve model performance. The insights gained from this research underscore the importance of developing accessible and high-quality language models as cultural artifacts that preserve and advance linguistic heritage.

Acknowledgments

Part of the Research results was obtained using the high performance computing resources operated by CINECA and awarded within the the National Leonardo access call 2023 by the Centre of Operations, Slovak Academy of Sciences and the Slovak National Supercomputing centre.

This project was supported by the Slovak Research and Development Agency under the Contract no. APVV-22-0414.

This work was supported by the Slovak Research and Development Agency under the Contract no. APVV-22-0370.

This work was supported by DiusAI, a. s.

References

- Anthropic (2023): *Claude*. Available at: <https://www.anthropic.com/>. [accessed 2025-07-24]
BENKO, Vladimír (2024): The Aranea Corpora Family: Ten+ Years of Processing Web-Crawled Data. In: E. Nöth – A. Horák – P. Sojka (Eds.): *Text, Speech, and Dialogue*. TSD 2024

- Lecture Notes in Artificial Intelligence, Vol. 15048, Heidelberg: Springer, pp. 55–70. DOI: https://doi.org/10.1007/978-3-031-70563-2_5.
- BENDER, Emily M. – GEBRU, Timnit – MCMILLAN-MAJOR, Angelina – SHMITCHELL, Shmargaret (2021): On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York: Association for Computing Machinery, pp. 610–623.
- BLASI, Damian – ANASTASOPOULOS, Antonios – NEUBIG, Graham (2022): Systematic Inequalities in Language Technology Performance across the World’s Languages. In: S. Muresan – P. Nakov – A. Villavicencio (Eds.): *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. Dublin: Association for Computational Linguistics, pp. 5486–5505.
- BOMMASANI, Rishi – HUDSON, Drew A. – CARD, Dallas – DURMUS, Esin – SRINIVASAN, Krishnan et al. (2021): On the Opportunities and Risks of Foundation Models. In: *arXiv preprint arXiv:2108.07258*.
- BROWN, Tom B. – MANN, Benjamin – RYDER, Nick – SUBBIAH, Mellanie – KAPLAN, Jared D. – DHARIWAL, Prafulla – ... – AMODEI, Dario (2020): Language models are few-shot learners. In: *Advances in Neural Information Processing Systems*, Vol. 33, pp. 1877–1901.
- CHEN, Tianqi – XU, Bing – ZHANG, Chiyuan – GUESTRIN, Carlos (2016): Training Deep Nets with Sublinear Memory Cost. In: *arXiv preprint arXiv:1604.06174*.
- COSTA-JUSSÀ, Marta R. – CROSS, James – ÇELEBI, Onur – ELBAYAD, Maha – HEAFIELD, Kenneth – HEFFERNAN, Kevin – ... – NLLB Team (2022): No language left behind: Scaling human-centered machine translation. In: *arXiv preprint arXiv:2207.04672*.
- DARĞIS, Roberts – BĀRZDINS, Guntis – SKADĪNA, Inguna – SAULĪTE, Baiba – GRŪŽĪTIS, Normunds (2024): Evaluating open-source LLMs in low-resource languages: Insights from Latvian high school exams. In: M. Härmäläinen – E. Öhman – S. Miyagawa – K. Alnajjar – Y. Bizzoni (Eds.): *Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities*, Miami: Association for Computational Linguistics, pp. 289–293.
- DEAN, Jeffrey – CORRADO, Greg – MONGA, Rajat – CHEN, Kai – DEVIN, Matthieu – MAO, Mark, et al. (2012): Large Scale Distributed Deep Networks. In: *Advances in Neural Information Processing Systems*, Vol. 25.
- DEVLIN, Jacob – CHANG, Ming-Wei – LEE, Kenton – TOUTANOVA, Kristina (2019): Bert: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies*, Vol. 1 (long and short papers), pp. 4171–4186.
- GARABÍK, Radovan (2025): Webový korpus slovenčiny ARANEUM + HPLT + FineWeb2. In: *Kultúra slova*, Vol. 59, No. 5, pp. 292–296.
- Google (2024): *Gemini*. Available at: <https://gemini.google.com/>. [accessed 2025-07-24]
- Google (2025): *Gemma 3*. Available at: <https://huggingface.co/google/gemma-3-4b-it/>. [accessed 2025-07-24]
- HPC Cineca (2023): *Leonardo HPC System*. Available at: <https://leonardo-supercomputer.cineca.eu/>. [accessed 2025-07-24]
- HOWARD, Jeremy – RUDER, Sebastian (2018): Universal Language Model Fine-tuning for Text Classification. In: *arXiv preprint arXiv:1801.06146*.
- LIN, Chin-Yew – OCH, Franz Josef (2004): Looking for a few good metrics: ROUGE and its evaluation. In: *Ntcir workshop*, pp. 1–8.

- Meta (2025): *Llama 3*. Available at: https://www.llama.com/docs/model-cards-and-prompt-formats/llama3_3/. [accessed 2025-07-24]
- Mistral AI (2023): *Mistral 7B v0.1*. Available at: <https://huggingface.co/mistralai/Mistral-7B-v0.1/>. [accessed 2025-07-24]
- OpenAI. (2023): *GPT-4*. Available at: <https://platform.openai.com/docs/models/overview>. [accessed 2025-07-24]
- PAPINENI, Kishore – ROUKOS, Salim – WARD, Todd – ZHU, Wei-Jing (2002): Bleu: a method for automatic evaluation of machine translation. In: P. Isabelle – E. Charniak – D. Lin (Eds.): *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*. Philadelphia, Pennsylvania: Association for Computational Linguistics, pp. 311–318.
- RADFORD, Alec – WU, Jeffrey – CHILD, Rewon – LUAN, David – AMODEI, Dario – SUTSKEVER, Ilya (2019): Language Models are Unsupervised Multitask Learners. In: *OpenAI Blog*. Available at: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf. [accessed 2025-07-24]
- RAJPURKAR, Pranav – ZHANG, Jian – LOPYREV, Konstantin – LIANG, Percy (2016): SquAD: 100,000+ Questions for Machine Comprehension of Text. In: *arXiv preprint arXiv:1606.05250*.
- RUDER, Sebastian – PETERS, Matthew E. – SWAYAMDIPTA, Swabha – WOLF, Thomas (2021): Transfer Learning in Natural Language Processing. In: K. Toutanova – A. Rumshisky – L. Zettlemoyer et al. (Eds.): *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, pp. 765–780.
- TOUVRON, Hugo – LAVRIL, Thibaut – IZACARD, Gautier – MARTINET, Xavier – LACHAUX, Marie-Anne – LACROIX, Timothée et al. (2023): LLaMA: Open and Efficient Foundation Language Models. In: *arXiv preprint arXiv:2302.13971*.
- VASWANI, Ashish – SHAZEER, Noam – PARMAR, Niki – USZKOREIT, Jakob – JONES, Llion – GOMEZ, Aidan N. – KAISER, Lukasz – POLOSUKHIN, Illia (2017): Attention is All you Need. In: *Advances in Neural Information Processing Systems*, Vol. 30, pp. 5998–6008.
- Wobový korpus slovenčiny ARANEUM + HPLT + FineWeb* (2025). Available at: <https://www.juls.savba.sk/webskahfcorp.html>. [accessed 2025-07-24]

APPENDIX

Slovak support in selected LLMs. If there are multiple versions, we include the version that first supported Slovak, as well as the immediately preceding version (if applicable), and the version with the smallest number of parameters that supports Slovak, as well as the version with strictly smaller number of parameters. The judgement was necessarily subjective, although the level of support was clearly delineated, with no borderline cases.

The table reflects the situation at the time of submitting the final version of the article, the situation at the start of our work was different and much less favourable for Slovak – of the models mentioned in the table, only Llama-2, Phi-1.5, gemma-7b and Mistral-7B-v0.1 were available.

<i>model</i>	<i>provider</i>	<i>supports Slovak</i>
Llama-2-70b-hf	Meta	no (cs)
Llama-3.1-8B	Meta	yes
Mistral-7B-v0.1	Mistral AI	(very) badly
gemma-7b	Google	badly
gemma-2-2b-it	Google	badly
gemma-2-2b	Google	yes
gemma-3-1b-pt	Google	yes
Phi-1.5	Microsoft	no
Phi-3-Small-8K-Instruct	Microsoft	badly
Qwen3-4B	Alibaba Cloud	badly
Qwen3-8B	Alibaba Cloud	yes

Table 2: Slovak support in various smaller Open Source LLMs.

Legend:

- yes – Slovak is fully supported, lexical or grammatical errors only sporadic
- no – Slovak is not supported, model either outputs text in other languages or the output is garbled
- badly – the output is intelligible, clearly in Slovak, but with many grammatical and orthographic mistakes
- (cs) – the output is in Czech, despite the input prompt being in Slovak

SLOVENSKÁ PRÁVNÁ DOKUMENTÁCIA PRE UTEČENCOV A ŽIADATEĽOV O AZYL A OTÁZKA JEJ ZROZUMITEĽNOSTI V PREKLADOCH DO ARABČINY¹

ZUZANA GAŽÁKOVÁ²

Filozofická fakulta Univerzity Komenského v Bratislave, Bratislava, Slovensko

GAŽÁKOVÁ, Zuzana: Slovak legal documentation for refugees and asylum seekers and the question of its comprehensibility in Arabic translations. *Jazykovedný časopis (Journal of Linguistics)*, 2025, Vol. 75, No. 2, pp. 452–467.

Abstract: Creating comprehensible documents for asylum seekers from Arabic-speaking regions presents a challenge due to differences in legal cultures, institutionally established norms, and vastly different social and cultural backgrounds. Such materials should consider the recipients' level of understanding within their social context, cultural conventions, and the diverse connotations of terms used. This article examines selected features of original Slovak legal documentation on migration and asylum seekers translated into Arabic, alongside its updated and revised version developed within the framework of the RESPLAIN project. It also explores ways to simplify Arabic legal language for recipients from diglossic Arabic-speaking environments arriving in Slovakia, where understanding complex legal texts in Modern Standard Arabic (MSA) cannot be taken for granted.

Keywords: Modern Standard Arabic, Educated Spoken Arabic, diglossia, spoken dialects, legal terminology, comprehensibility.

1. ÚVOD

K problematike tvorby a aktualizácie dokumentov v arabskom jazyku, ktoré sa týkajú žiadateľov o azyl a ochrany utečencov, sme sa dostali prostredníctvom projektu RESPLAIN – *Research and Response: Plain Language in Administrative Communication with Refugees* [Výskum a odpoveď. Zrozumiteľný jazyk v administratívnej komunikácii s utečencami]. Jeho hlavným cieľom bolo zlepšiť prístup k informáciám a zdokonaľiť komunikáciu so žiadateľmi o azyl a utečencami, ktorí využívajú služby poskytované verejnými subjektami na území Slovenska. Tento cieľ

¹ Tento článok vznikol v rámci projektu RESPLAIN *Research and Response: Plain Language in Administrative Communication with Refugees*, ktorý sa realizoval v spolupráci s Filozofickou fakultou Univerzity Komenského v Bratislave, Migračným úradom SR, Jazykovedným ústavom Ľudovíta Štúra SAV, Ministerstvom vnútra SR, UNHCR a Ligou za ľudské práva, č. zmluvy SVK01/2023/0000000051/000. JÚEŠ SAV sa podieľal na rámcovom vypracovaní skrátených znení *Poučenia* (pozri pozn. č. 3) a táborových poriadkov. Tieto texty boli upravované a aktualizované aj v slovenskej verzii.

² Za doplnenia a pripomienky k jednotlivým častiam textu srdečne ďakujem Mgr. Bronislave Chocholovej, PhD., z Jazykovedného ústavu Ľudovíta Štúra SAV a doc. Mgr. Helene Tužinskej, PhD., z Filozofickej fakulty UK.

sme sa snažili dosiahnuť aj prostredníctvom aktualizácie a harmonizácie pôvodnej dokumentácie Ministerstva vnútra SR, ktorá sa týkala žiadateľov o azyl a ochrany utečencov, s terminológiou Európskej únie a jej následného prekladu do anglického, francúzskeho, ukrajinského, ruského, tureckého, arabského, perzského a bengálskeho jazyka, ako aj do jazykov paštú, hindu, urdu a dárí. Náplňou tohto príspevku je zdokumentovať a odôvodniť hlavné zmeny a úpravy arabskej verzie tejto dokumentácie, čo by mohlo byť užitočné pri budúcich úpravách.

Počas trvania projektu RESPLAIN došlo k aktualizácii pôvodnej verzie a k vytvoreniu zjednodušenej verzie rozsiahleho materiálu Ministerstva vnútra SR s názvom *Poučenie žiadateľa o udelenie azylu o právach a povinnostiach* (ďalej *Poučenie*),³ ako aj táborových poriadkov jednotlivých záchytných a pobytových táborov. Nemáme informácie, kedy bola pôvodná verzia *Poučenia* preložená do arabčiny, ale predpokladáme, že zrejme v 90. rokoch, v období po osamostatnení Slovenskej republiky. Je možné, že text bol následne upravovaný viacerými prekladateľmi, čo zrejme viedlo k použitiu rozličnej terminológie na označenie rovnakých javov. V priebehu riešenia projektu sa zmenila časť legislatívy a vznikla aj možnosť výraznejšie skrátiť arabskú verziu textu *Poučenia* z pôvodných 32 strán (!) na 9. Týmto zmenám bolo možné prispôbiť aj vizuálnu stránku, čím sa darilo naplniť cieľ projektu – priblížiť a sprehľadniť obsah dokumentácie pre potreby žiadateľov o azyl s obmedzenými možnosťami porozumenia právneho textu.⁴ Situáciu však komplikovala výrazná odlišnosť jazykových a kultúrnych kontextov (napr. medzi Slovenskom a arabskými krajinami) a diglosná povaha arabčiny (pozri ďalej).

Chceli by sme tiež konštatovať, že týmto príspevkom nemáme v úmysle vrhať nepriaznivé svetlo na pôvodný stav tejto dokumentácie, nakoľko v minulosti s určitou plnila svoj účel. Okrem nižšie uvedených nedostatkov sa v arabských materiáloch objavovala istá neobratnosť, ktorú možno pripísať nedostatočne rozvinutým metodikám prekladu právnych textov do arabského jazyka v našom prostredí v tých časoch.⁵ Vzhľadom na úplnú neznalosť niektorých jazykových mutácií (arabčina, perzština, paštu a pod.) si zrejme jednotlivé inštitúcie ani neuvedomovali možnú neaktuálnosť príslušných textov.

Celková potreba aktualizácie vyplynula z pretrvávajúcej migračnej krízy, z nevyhnutnosti terminologicky zosúladiť texty s dokumentáciou Európskej únie, ako aj pod vplyvom viacerých zmien vo vnútroštátnej legislatíve. Prepracovanie materiálov však nebolo ľahké aj vzhľadom na časovú náročnosť a pre nedostatok odborníkov v určitých jazykových špecializáciách.

³ Pozri: *Poučenie žiadateľa o udelenie azylu o právach a povinnostiach*. 8.11. 2025 [online].

⁴ V časovom strese a v situácii ohrozenia žiadateľa o azyl neraz potvrdili svojimi podpismi, že týmto dokumentom porozumeli, aj keď miera porozumenia nebola overená (Tužinská 2020, s. 139 – 152).

⁵ Pôvodná arabská verzia *Poučenia* mala v záhlaví ako podnadpis po arabsky uvedené „slovenský jazyk“. Dlhو sme nerozumeli prečo, a len potom sme si všimli, že slovenská verzia má v záhlaví uvedené „slovenčina“.

2. ŠPECIFIKÁ ARABSKY HOVORIACEHO PROSTREDIA

Výsledky prezentované v tejto štúdii sa opierajú o aktualizáciu arabskej verzie dokumentácie určenej arabsky hovoriacim žiadateľom o azyl a na ňu nadväzujúce školenia tlmočníkov, ktoré prebehli v rokoch 2023 – 2024. Jazykové a kultúrne špecifiká tu zohrávali dôležitú úlohu, preto považujeme za potrebné na úvod stručne objasniť základné charakteristiky arabsky hovoriaceho prostredia. Náplňou tohto príspevku však nie je podrobný opis arabskej diglosnej situácie ani dôkladnejšie vyhodnotenie výskumu v tejto oblasti. Preto v prehľade uvádzame aspoň bibliografické odkazy na relevantné práce k tejto problematike.

Moderná spisovná arabčina (Modern Standard Arabic, MSA, *al-lughā 'l-'arabīya 'l-fuṣṣḥā*) je oficiálnym alebo jedným z oficiálnych jazykov v geograficky rozsiahlom pásme, ktoré pokrýva viac ako 20 krajín. Môžeme ich ohraničiť na západe Atlantickým oceánom, na severe Stredozemným morom, na juhu Saharskou púšťou a na východe Arabským polostrovom a Irakom. Práve z tohto geografického pásma prichádza do Európy veľké množstvo utečencov.

Jazykovú situáciu tohto regiónu komplikuje viacero skutočností a primárne hlavne fakt, že tento významný a bohatý jazyk⁷ (MSA) nie je materinským a ani primárnym komunikačným jazykom jeho používateľov. Panuje tu totiž špecifická situácia známa ako tzv. *diglosia* (podľa Ferguson 1959) (arab. *izdiwādž lughawī, izdiwādžīja*), čiže existencia dvoch základných jazykových foriem, ktoré sa navzájom dopĺňajú.⁸ Zjednodušene by sme mohli tento lingvistický jav opísať takto: popri kodifikovanej, gramaticky zložitejšej a nadradenej variete MSA tu existujú subštandardné hovorové jazyky (označované aj ako dialekty), ktoré sa používajú v bežnej komunikácii (arab. *'ammīja, lahdža, lughā dāridžīja*).

Moderná spisovná arabčina (MSA) sa ako *lingua franca* používa v celom arabskom svete a uplatňuje sa najmä vo formálnej písomnej a ústnej komunikácii, ako sú medzinárodné konferencie,⁹ arabské spravodajské stanice, právne dokumenty, denná tlač, učebnice alebo dabing zahraničných seriálov, ktoré sú následne distribuované do rôznych arabských krajín. MSA si stále zachováva dominantnú pozíciu aj v oblasti umeleckej literatúry, hoci je zrejmé, že jazykové preferencie a praktiky autorov

⁶ Pri prepise arabských mien do latinky používame zjednodušenú vedeckú transkripciu podľa Gažáková – Pauliny 2013, s. 3.

⁷ Odhliadnuc od historického kontextu a lexikálnej bohatosti, arabčina je nositeľom posolstva svätého textu – Koránu – a zosobňuje tak prepojenie s duchovným svetom moslimov.

⁸ Viacerí jazykovedci oprávnenne nesúhlasia s prísnou dvojvariantnou dichotómiou diglosie. Za mnohých spomenieme napr. prácu E. Badawiho (1973), ktorý identifikoval pri štúdiu sociolingvistickej situácie v Egypte až päť úrovní (*mustawjāt*) arabského jazyka. Vo všeobecnosti však platí konsenzus, že *al-fuṣṣḥā* a hovorové jazyky predstavujú hraničné body variačného kontinua. Pozri aj Suleiman 2006, s. 173 – 178; u nás Drozdík 2006; Brieška 2009a, 2009b.

⁹ Aj v tomto prípade arabský rečník spravidla svoj prejav číta, pretože má problém v MSA spontánne prednášať bez zachádzania do hovorového jazyka.

sa postupne menia.¹⁰ Ľudia si MSA osvojujú najmä prostredníctvom školského vzdelávania, keďže ju presadzujú vzdelávacie a štátne inštitúcie ako univerzity (tu ale napr. s výnimkou Egypta), vedecké akadémie arabského jazyka, ministerstvá, kazatelia v mešitách a pod. Vzdelaní Arabi sú tak spravidla s MSA v každodennom kontakte, väčšina z nich v nej pravidelne číta a rozumie jej. Keď sa však pokúšajú v MSA spontánne komunikovať, v prevažnej väčšine tento štandard nedosahujú, aj keď si to mnohí o sebe myslia (Gažáková – Pauliny 2013, s. 127 – 130; Drozdík 2006, s. 67). Nezriedka dochádza k tomu, že aj rečníci v televíznom programe začínajú diskutovať v MSA, no postupne zachádzajú do svojho hovorového jazyka. Podobný jav možno pozorovať aj u profesorov na prednáškach.¹¹

Ľudia bez vyššieho vzdelania sú v každodennom živote s MSA v obmedzenejšom kontakte a ovládajú ju spravidla pasívne. Na bežnú komunikáciu v arabskom prostredí používajú takmer výlučne regionálny hovorový jazyk. Hovorové jazyky však po stáročia podliehajú citeľným inojazyčným vplyvom, a to najmä francúzštiny, španielčiny a berberských jazykov v oblasti arabského Maghribu či angličtiny a francúzštiny v oblasti arabského Mašriku. Napríklad v Egypte v minulosti zohral pri formovaní hovorových jazykov dôležitú úlohu silný vplyv osmanskej turečtiny. Aj preto sú medzi subštandardnými hovorovými jazykmi zásadné rozdiely najmä pokiaľ prebieha komunikácia medzi aktérmi z geograficky vzdialených oblastí arabského sveta a ich vzájomné porozumenie môže byť, bez prechodu do vyššej variety, výrazne obmedzené.¹² Subštandardné hovorové jazyky však nemajú žiaden oficiálny status a ich písomná podoba nie je ustálená, keďže sa v nich až donedávna písomne

¹⁰ V 20. storočí sa využitie hovorového jazyka v arabskej literatúre obmedzovalo predovšetkým na sociolingvistickú charakteristiku postáv, dialógy v dramatických dielach a na oblasti humoru či satiry, pričom v poézii sa objavovala sporadicky, najmä v Egypte.

¹¹ Osobné skúsenosti autorky z pobytov v Egypte a Alžírsku.

¹² Za najtradičnejšie sa pokladá rozdelenie arabských hovorových jazykov na geografickom základe. Tu rozlišujeme dve základné skupiny: západné (*maghribí*) a východné (*mašriqí*) dialekty, ktoré sa líšia na všetkých lingvistických úrovniach. V rámci tohto delenia rozpoznávame niekoľko dialektových okruhov:

1. Veľkú skupinu tvoria *severoafrikké (maghribí)* dialekty, kde zaraďujeme hovorové jazyky Líbye, Alžírsku, Tunisku, Maroka, dialekt Mauretánie, Mali a Západnej Sahary (tzv. *ħassánija*), ale aj maltský jazyk a zaniknuté dialekty Andalúzie a Sicílie.
2. Východná (*mašriqí*) dialektová skupina sa začína *egyptskými dialektmi*, kam patria hovorové jazyky Sudánu, Čadu, Eritrey, Džibutska a Somálska.
3. *Sýrsko-palestínske* dialekty zasahujú okrem Sýrie a Palestíny aj do Jordánska a Libanonu.
4. *Dialekty Arabského polostrova* zahŕňajú väčšinu Arabského polostrova (oblasti Nadždu, al-Ĥidžázu, Jemenu, Kuvajtu, Ománu, Spojených arabských emirátov ako aj Bahrajn a Katar) vrátane beduínskych dialektov. Napriek tomu, že táto oblasť predstavuje kolísku arabskej civilizácie, hovorové jazyky tohto regiónu patria k najmenej preskúmaným.
5. *Irackými* (alebo aj tzv. mezopotámskymi) *dialektmi* sa rozpráva okrem Iraku a východnej Sýrie aj v dnešnom Iráne (najmä v provincii Ĥúzistán). Podľa Gažáková – Pauliny 2013, s. 131 – 132.

komunikovalo len v obmedzenej miere. Až s nástupom internetu, SMS a sociálnych sietí začali Arabi každodennú hovorovú komunikáciu bežne zapisovať. Tieto zápisy vznikali spontánne, mimo akýchkoľvek normotvorných mechanizmov, s množstvom odlišných niekedy až idiosynkratických pravopisných konvencií.¹³ Hovorové jazyky začali následne masívne prenikať do ľudovej kultúry, seriálov, populárnych piesní či reklamnej sféry a úspešne sa presadzujú aj v umeleckej literatúre (najmä u mladej generácie autorov). Stredové variety sa súhrnne označujú ako tzv. zmiešaná alebo stredná arabčina, ktorá je v ústnom prejave často odborne nazývaná aj prestížna orálna arabčina¹⁴ (POA) (Drozdík 2006, s. 64), aj keď jej koncept je ťažko uchopiteľný.¹⁵ Zmiešané varianty arabčiny, ktoré z MSA preberajú terminológiu, pričom si zachovávajú hovorový ráz, začali prevládať vo virtuálnom priestore (blogy, online diskusie a fóra), kde sa zložité pravidlá MSA nevnímajú ako záväzné.

V bežnej komunikácii sú teda hovorové jazyky charakteristické značnou variabilitou a aj vzájomnou nezrozumiteľnosťou, najmä pokiaľ ich používatelia pochádzajú z uzavretých komunít či zo vzdialených častí arabského sveta (Mauretánia, Alžírsko, Jemen, Irak). Aj preto mnohí bádatelia a arabskí intelektuáli prirovnávajú diglosné arabské prostredie ku kríze (*azma*), veľkému problému (*qadīja*) alebo konfliktu (*širāʿ*) (Boussofara-Omar 2006, s. 629). Vo všeobecnosti platí, že čím je kontext odbornejší, tým viac sa v komunikácii využíva slovná zásoba a prvky MSA. Zmena však väčšinou neprebíha v podobe náhleho prepnutia kódov (*code switching*), ale skôr zvýšením výskytu prvkov opačnej variety (Versteegh 1997, s. 194). Aj preto zohráva vzdelanie uchádzačov o azyl a ich ochota spolupracovať dôležitú úlohu pri miere porozumenia daných materiálov.

Z našich skúseností vyplýva, že sa tieto skutočnosti neberú do úvahy v dostatočnej miere pri práci so žiadateľmi o azyl z arabského prostredia. Rozdiely medzi modernou spisovnou arabčinou (MSA) a hovorovými jazykmi sú často chápané zjednodušene, prípadne prirovnávané k rozdielom medzi spisovnou slovenčinou a nárečiami Slovenska, táto problematika je však zložitejšia. Znalosť MSA v oblasti práva sa na druhej strane u bežných ľudí preceňuje. V budúcich školeniach je preto nevyhnutné venovať čas dôkladnej príprave materiálov na tlmočenie a prispôbiť ich tak viac potrebám účastníkov z arabsky hovoriacich krajín, ako aj nižšie uvede-
ným špecifikám.

¹³ Sem patrí napríklad používanie latinky, integrácia anglických lexikálnych prvkov a využívanie číslíc na grafické zaznačenie foném, pre ktoré latinka nemá ekvivalent (napr. pre spoluhlásky *ʾajn*, *há*, *qáfa* i.).

¹⁴ *Educated Spoken Arabic* (ESA) (Mitchell 1986); alebo napr. *lughat al-muthaqqafīn* [jazyk intelektuálov] (Badawi 1973).

¹⁵ Everyone claims to believe that Educated Spoken Arabic is rule-governed, but none seems to be able to come up with the rules. [Každý tvrdí, že prestížna orálna arabčina sa riadi určitými pravidlami, no zdá sa, že nikto nedokáže tieto pravidlá stanoviť.] (Parkinson 2003, s. 29).

3. PÔVODNÁ A AKTUALIZOVANÁ VERZIA ARABSKEJ DOKUMENTÁCIE

Tento rozsiahlejší všeobecný úvod je dôležitý pre lepšie pochopenie východiskovej situácie arabsky hovoriacich žiadateľov o azyl v SR a ich potrieb. Často im chýba vyššie vzdelanie, nezriedka pochádzajú z ekonomicky znevýhodneného prostredia, pričom nemusia mať skúsenosti s niektorými základnými zariadeniami a moderným technickým vybavením. Nasvedčujú tomu aj vybrané pokyny v jednotlivých táborových poriadkoch, napr. táborový poriadok pre pobytový tábor v Opatovskej Novej Vsi, bod. 16 „Kuchyňa“: *„Prosíme, aby ste používali rýchlovarnú elektrickú kanvicu výlučne na zohriatie vody, ak do nej dáte mlieko, zničí sa.“*

Od týchto skutočností sa ďalej odvíja aj ich miera (ne)pochopenia právnych textov, pretože sú písané v stylisticky zložitom jazyku, ktorý sa nepoužíva v bežnej komunikácii a ktorý nie je ani ich materinským jazykom. Pre lepšie porozumenie týchto dokumentov (pozri nižšie) nemusí byť dostatočná ani prítomnosť kvalifikovaného tlmočníka.¹⁶ Tlmočníci, hoci sú väčšinou arabského pôvodu, môžu pochádzať vo vzťahu k žiadateľovi z geograficky vzdialeného regiónu a následne sa vo všeobecnosti snažia komunikovať takou hovorovou varietou arabského jazyka, ktorá je bližšie k spisovnej norme a bez lokálnej terminológie (tzv. prestížna orálna arabčina POA),¹⁷ pokiaľ sú toho schopní. Ich úlohou je tak preklenúť kultúrnu priepasť medzi príslušnou arabskou krajinou a slovenským inštitucionálnym prostredím, čo v praxi okrem iného znamená aj prerozprávanie komplexných právnych formulácií do pre prijímateľa zrozumiteľného hovorového či menej formalizovaného variantu, hoci prerozprávanie už nemusí byť úplne verné pôvodine. Kultúru tu pritom treba chápať nie v užšom zmysle pokročilého intelektuálneho rozvoja človeka, ktorý sa odráža v umení, ale v širšom antropologickom zmysle všetkých spoločensky podmienených aspektov ľudského života. Súvisí teda so všeobecnými poznatkami, ktoré zvyčajne zahŕňajú prácu politických inštitúcií, vzdelanie, históriu a aktuálne dianie. Problémom pre prekladateľa sa stáva dodržiavanie noriem kultúrnej vzdialených prostredí, resp. rozhodovanie, ktoré normy sa majú v konkrétnych situáciách uprednostniť (Schäffner – Adab 1995, s. 329).

Materiály, s ktorými sú žiadatelia o azyl v SR konfrontovaní, zahŕňajú najmä spomínané rozsiahle *Poučenie* Ministerstva vnútra SR, s ktorým sa žiadateľ oboznamuje na začiatku azylového konania, a jednotlivé táborové poriadky. Je pochopiteľné, že tieto materiály musia spĺňať zásadné právne kritériá a že každé slovo takéhoto textu musí mať presne vymedzený význam. Chceli by sme tiež dodať, že sa nebudeme dôkladnejšie zaoberať zjavnými gramatickými chybami v pôvodnom znení, kto-

¹⁶ Od tlmočníkov sa okrem jazykových kompetencií a bohatých skúseností neraz očakáva právna, sociálna, kultúrna či dokonca psychologická expertíza, ktorou nedisponujú (viac Tužinská 2023).

¹⁷ Zväčša sú k dispozícii tzv. komunitní tlmočníci bez dostatočnej praktickej znalosti etického kódexu (viac pozri Šveda – Tužinská 2021).

ré sa tu príležitostne objavujú, napr. *فستعين المحكمة وصياً أو ولياً*, správne by malo byť *فستعين المحكمة وصياً أو ولياً*.¹⁸

Plné znenie hutného právneho textu *Poučenia* obsahuje komplikované slovenské súvetia s pomerne náročnou terminológiou, ktoré sú následne, vzhľadom na právnu povahu textu, prekladané s filologickou dôkladnosťou do MSA. Text je prirodzene orientovaný na slovenský právny systém,¹⁹ s ktorým arabskí žiadatelia o azyl nemajú praktickú skúsenosť. Výsledná arabská verzia tak stále zostáva – aj napriek štylistickej aktualizácii – značne komplikovaná. Spôsobuje to nielen náročná predloha, ale aj samotný arabský právny jazyk, ktorý je charakteristický množstvom pasívnych slovesných tvarov a častým použitím tzv. *mašdarov*,²⁰ najmä vo viacnásobných genitívnych spojeniach (napr. *تمديد فترة قرار البت بالجوء* – „predĺženie lehoty na vydanie rozhodnutia o žiadosti o azyl“), pretože práve tento gramatický tvar vyjadruje univerzálnosť, neosobnosť, neutralitu a obsahovú presnosť potrebnú v právnej oblasti.

K nedostatočnej zrozumiteľnosti MSA v oficiálnych právnych kontextoch prispieva aj samotný charakter arabského písma (tzv. *scriptio defectiva*), ktoré sa vyznačuje celkovou absenciou vyznačovania krátkych samohlások a v niektorých prípadoch aj dlhých. Táto črta sa vo všeobecnosti pociťuje aj ako jedna z hlavných prekážok v boji s negramotnosťou (Drozdík 2006, s. 24). **Vokalizácia a diakritické prvky nesú zásadnú gramaticko-lexikálnu záťaž**, no v štandardnom zápise sa **systematicky vynechávajú**. To vedie k viacerým **možným interpretáciám** jedného grafického tvaru, čo možno vyriešiť len kontextovo. Písaný text sa tak nedá „rovno čítať“, treba mu najskôr porozumieť, čo znamená nevyhnutú predprípravu bez ohľadu na profesiu a úroveň vzdelania. Právny text charakterizovaný nadmerne dlhými a syntakticky preťaženými vetnými konštrukciami je tak vo všeobecnosti interpretácie náročný, najmä ak sa dokonca predpokladá (a vyplýva to aj z obsahu dokumentu), že žiadateľ o azyl je maloletý, čiže ešte nedovršil osemnásť rokov.²¹

Na ilustráciu si uveďme niekoľko príkladov z tohto materiálu (*Poučenie*, s. 7, bod. 12):

Lehotu na rozhodnutie môže opakovane, najviac o 9 mesiacov predĺžiť nadriadený vedúci zamestnanec zamestnanca migračného úradu, ktorý vo veci koná, ak si

¹⁸ Slov. „...súd vymenuje poručníka alebo opatrovníka.“ V arabskom znení chýba na konci posledného slova tzv. *alif tanwín*. Ide o bežnú chybu z nepozornosti.

¹⁹ Bližšie o prekladoch právnych textov a ich orientácii na právne systémy rôznych kultúr pozri Madsen 1997, s. 291 – 299.

²⁰ V arabskej gramatike je *mašdar* podstatné meno, ktoré označuje *zdrojovú formu slovesa*. V európskych jazykoch sa často prekladá ako „slovesné podstatné meno“ alebo „verboid“. Ide však o oveľa bohatší koncept, ktorý v arabčine plní špecifické syntaktické funkcie a môže byť základom tvorby ďalších derivátov (viac pozri Ditters 2008, s. 164 – 167).

²¹ Pozri napr. *Poučenie*, s. 15, bod 9: „Ak ste umiestnený v zariadení sociálnoprávnej ochrany detí a sociálnej kurately (v centre pre deti a rodiny), ste povinný dostať sa do pobytového tábora do troch dní od nadobudnutia plnoletosti. Uvedené neplatí, ak ste požiadalí o zotrvanie v tomto zariadení.“

rozhodovanie o žiadosti o udelenie azylu vyžaduje posúdenie zložitých skutkových alebo právnych otázok alebo veľký počet žiadateľov podalo súčasne žiadosť o udelenie azylu a je veľmi zložité rozhodnúť o nej do 6 mesiacov od začatia konania alebo nie je možné rozhodnúť do 6 mesiacov od začatia konania z dôvodu, že žiadateľ nespôlupracuje alebo iným spôsobom sťažuje posúdenie žiadosti o udelenie azylu.

يجوز تمديد فترة قرار البت باللجوء، بعد أقصى لمدة 9 أشهر، من قبل المسؤول الأعلى في مكتب الهجرة، والذي يقوم بتقييم الحالة. إذا كان القرار بشأن طلب اللجوء يتطلب تقييماً لقضايا معقدة من الوقائع أو القانون أو عدداً كبيراً من المتقدمين قدموا طلب اللجوء في نفس الوقت، فمن الصعب جداً اتخاذ قرار في غضون 6 أشهر من بدء الإجراءات. كما لا يمكن اتخاذ قرار بشأن اللجوء في غضون 6 أشهر من بدء الإجراءات بسبب عدم تعاون مقدم الطلب أو أنه بشكل آخر يعرقل تقييم طلب اللجوء .

Snád' ešte nezrozumiteľnejší je v arabčine bod 27 *Poučenia*, ktorý sa týka závažnej otázky – podmienok, za akých je dovolené žiadateľovi o azyl pracovať. Kým tu slovenský originál rešpektuje bežnú legislatívnu hierarchiu (hlavná norma, výnimky, doplnenia), arabský text vyznieva neprirodzene a je syntakticky preťažený. Dôvodom je nielen vysoká koncentrácia abstraktnej právnej terminológie, ale aj doslovný preklad podmienkovej vety, ktorá využíva síce gramaticky správnu, ale ťažkopádnu zápornú konštrukciu. (*Poučenie*, s. 11, bod 27):

Na území Slovenskej republiky môžete pracovať po šiestich mesiacoch od začatia konania o udelenie azylu okrem prípadu, ak:

a) podanie správnej žaloby proti rozhodnutiu migračného úradu nemá odkladný účinok a správny súd nerozhodol o priznaní odkladného účinku, alebo

b) Najvyšší súd SR nerozhodol o priznaní odkladného účinku kasačnej sťažnosti proti rozhodnutiu správneho súdu.

يمكنك العمل في أراضي الجمهورية السلوفاكية بعد ستة أشهر من بداية إجراءات اللجوء إلا إذا:
أ) تم تقديم دعوى إدارية ضد قرار مجلس الهجرة ليس له أثر رجعي ولم تصدر المحكمة الإدارية حكماً بشأن منح الأثر الرجعي، أو
ب) لم تصدر المحكمة العليا للجمهورية السلوفاكية حكماً بشأن الاعتراف بالأثر الرجعي لشكوى النقص ضد قرار صادر عن محكمة إدارية

Značná nezrozumiteľnosť dokumentácie tak môže viesť k nepochopeniu pokynov, čo môže byť kvalifikované ako nespôlupráca žiadateľa o azyl s autoritami prijímajúcej krajiny. V krajnom prípade to môže mať vážne právne následky, preto je dôležité podporiť spätnú väzbu týkajúcu sa kvality porozumenia.

Dalším zásadným problémom pôvodnej dokumentácie bola terminológia, ktorú bolo treba aktualizovať a spresniť najmä prostredníctvom eliminácie často využívannej arabskej synonymie a s dôrazom na významovú jednoznačnosť. Ako príklad vyberáme dva kľúčové termíny pre každého žiadateľa o azyl: „pobyťový tábor“ a „záchranný tábor“. V pôvodnej arabskej dokumentácii už samotné slovo „tábor“ priebež-

ne variovalo medzi termínmi *mu'askar* a *muchajjam*, pričom *mu'askar* sa spravidla používa na označenie organizovaného dlhodobého fungujúceho hromadného ubytovacieho zariadenia pôvodne vojenského typu (arab. *'askarī* – ‚vojak‘) (al-Mundžid 1997, s. 505). Časom sa začal termín používať aj v iných kontextoch, napr. „trénin-gový tábor karate“ (arab. *mu'askar tadrib li 'l-kárátéh*) alebo „západný“, resp. „vý-chodný blok“ (arab. *mu'askar gharbí*, resp. *šarqí*) a pod., kde nadobudol nový alebo prenesený význam. Výraz *muchajjam* sa primárne používa na označenie dočasného stanového tábora (porov. arab. *chajma* – ‚stan‘), no vyskytuje sa aj v spojeniach typu „utečenecký tábor“ (arab. *muchajjam al-ládži'ín*), kde poukazuje na prechodný pro-vizórny charakter ubytovacieho zariadenia. Aj z tohto dôvodu sme sa pri unifikácii terminológie priklonili k termínu *mu'askar*, keďže označuje všeobecný typ tábora a zároveň najadekvátnejšie korešponduje s anglickou verziou textov, v ktorej sa kon-sistentne používa výraz *camp*.

Vážnejší problém nastal pri termíne „záchytný tábor“, ktorý bol v pôvodnej verzii prekladaný ako *mu'askar i' tiqál*. Ten síce môže označovať aj záchytný tábor, čiže zariadenie, ktorého klient je dočasne obmedzený na slobode²² (arab. *i' taqala* – ‚zadržať, internovať, uväzniť‘). Zároveň sa však bežne používa na označenie inter-načného tábora – často zriadeného z politických alebo vojenských dôvodov a spoje-ného s neľudskými podmienkami – alebo dokonca koncentračného tábora (arab. *mu-'askarát al-i' tiqál an-názija* – ‚nacistické koncentračné tábory‘).²³ Známy arabsko-anglický slovník H. Wehra (Wehr 1980, s. 630) ponúka výlučne túto alternatívu. Termín sme konzultovali aj s niekoľkými rodenými hovoriacimi, ktorí potvrdili jeho výraznú negatívnu konotáciu a uviedli, že takéto označenie azylového zariadenia im evokuje obraz „miesta niekde uprostred púšte, kam človek ide a už sa nevráti“.²⁴

Je teda pravdepodobné, že použitie tohto termínu mohlo v minulosti vyvolať negatívne emócie a nedôveru arabských žiadateľov o azyl voči našim prijímajúcim azylovým zariadeniam, čo mohlo prispieť k budovaniu nedôvery či k nespolupráci s úradmi.

Negatívne konotácie termínu sa pri tvorbe pôvodnej verzie s veľkou pravdepodo-bnosťou nebrali do úvahy. Pri hľadaní vhodnejšej alternatívy sme zvažovali aj ter-mín *mu'askar ihtidžáz*, ktorý sa tiež objavoval v pôvodnej verzii. Tento termín by mohol do istej miery vystihovať význam slovenského slova „záchytný“ (arab. *ih-tadžaza* – ‚zadržať‘) a má formálne neutrálnejší charakter, avšak napokon sme ho ne-

²² „Záchytný tábor je účelové zariadenie Ministerstva vnútra SR, do ktorého sú žiadatelia umiestnení po začatí azylového konania aj z dôvodu prijímacej zdravotnej karantény. Počas celého obdobia trvania zdravotných vyšetrení, ktoré trvajú spravidla dva až tri týždne, nie je dovolené opustiť toto zariadenie.“ (Ekonomická encyklopédia) 1. 12. 2024 [online].

²³ Napr.: *Júri 2024 tuwadždžih al-anžár naḥwa našab mu'askar i' tiqál nází*. [Euro 2024 upriamuje pozornosť na pamätník nacistického koncentračného tábora.] 25. 6. 2024 [online].

²⁴ Na základe konzultácie s rodeným hovoriacim arabčiny Michaelom Moawedom Bastom 5. 7. 2024.

zvolili. V minulosti sa totiž tiež používal na označovanie internačných táborov na severe Afriky, v období francúzskej koloniálnej správy, a preto v sebe nesie historické a politicky zaťažené asociácie, ktorým sme sa rozhodli vyhnúť.²⁵

Nakoniec sme prijali riešenie, ktoré sa opiera o terminologické postupy Európskej únie. V jej dokumentácii sa anglický termín *reception camp* prekladá ako *mu‘askar al-istiqbál* – ‚prijímací tábor‘, ktorý už názvom naznačuje, že ide o zariadenie, prostredníctvom ktorého je žiadateľ o azyl prijímaný na našom území. V slovenských materiáloch sa však naďalej používa pojem „záchytný tábor“,²⁶ čo môže viesť k nejasnostiam pri spätnom preklade z arabčiny do slovenčiny. Variabilitu terminologických riešení v rámci EÚ dokladá aj nemecká prax, kde sa záchytný tábor po arabsky označuje ako *markaz quḍúm al-ládži‘ín*²⁷ – „centrum príchodu migrantov“. Podobný posun sledujeme aj v Taliansku, kde sa tiež slovo „tábor“ nahrádza neutrálnejším termínom „centrum“²⁸. Je pravdepodobné, že rovnakým smerom sa bude uberať aj slovenská terminológia, keďže aj samotný výraz „tábor“ nesie určitú negatívnu konotáciu.

V pôvodnej verzii *Poučenia* sme ďalej identifikovali viacero variácií pomenovania totožných javov či určitých štylizácií s použitím rôznych synonym. Takáto nejednotnosť sa môže javiť ako marginálna, avšak v ojedinelých prípadoch môže viesť k nejasnostiam. Tu môžeme spomenúť niekoľko zdanlivo menej závažných, ale dôležitých označení, ktorých zjednotenie uľahčuje prehľadnosť textu, napr. označenie Slovenskej republiky ako *al-Džumhúrija as-slúfákija* (aj *džumhúrija Slúfákija*), označenie rozhodnutia Migračného úradu ako *qarár* alebo aj *batt*, ako aj vyššie spomínané označovania jednotlivých typov táborov, napr. záchytný tábor bol v rámci textu a dokonca v rámci tej istej vety označovaný ako *mu‘askar i‘tiqál* (ale aj *mu‘askar iḥtidžáz*) a pobytový ako *mu‘askar iqáma* (ale aj *muchajjam iqáma* či *mu‘askar sakanī*), čím sa navodzoval dojem, že text odkazuje na viacero odlišných zaříadení, napr. *Poučenie*, str. 5, bod 1:

بعد تقديم طلب اللجوء الخاص بك ، سيتم وضعك في معسكر اعتقال ، حيث سيتم تقييد حرية حركتك في منطقة معسكر الاحتجاز ، حتى يقرر طبيبك أن حالتك الطبية تسمح بنقلك إلى معسكر سكني

Po podaní žiadosti o azyl budete umiestnený do záchytného tábora (mu‘askar i‘tiqál), kde bude vaša sloboda pohybu obmedzená na priestor tohto záchytného zariadenia (mu‘askar al-iḥtidžáz), až kým váš lekár nerozhodne, že váš zdravotný stav dovoľuje presun do ubytovacieho tábora.

²⁵ *Mu‘askarát al-iḥtidžáz wa ‘l-‘amal fī šimāl Ifríqijá*. [Záchytné a pracovné tábory na severe Afriky.] 20. 11. 2024 [online].

²⁶ Zariadenia migračného úradu. 25. 6. 2024 [online].

²⁷ *Marákiz quḍúm al-ládži‘ín wa marákiz al-quḍúm, ittichádh al-qarár wa ‘t-tarhíl*. [Centrá príchodu utečencov a centrá príchodu, rozhodovacie konanie, vyhostenie.] 20. 11. 2024 [online].

²⁸ *Ḥaqquka fī ‘l-baqá‘ fī markaz al-istiqbál*. [Tvoje právo zostať v prijímacom centre.] 10. 7. 2024 [online].

Spomedzi synonymických štylistických fráz spomeňme synonymické spojenia *anta mulzam* – ‚si povinný‘ ale aj *anta muṭálib* – ‚očkáva sa od teba‘, ktoré bolo potrebné v texte unifikovať. Ako ďalší príklad možno uviesť pomenovanie dokladu o oprávnení na poskytovanie zdravotnej starostlivosti, ktoré malo v Poučení pôvodne dve podoby: *ithbát li ʿt-tarchiṣ li taqdīm ar-riʿája aṣ-ṣiḥḥíja* – ‚potvrdenie o nároku na zdravotnú starostlivosť‘, ako aj *ithbát istiḥqáqika li ʿr-riʿája ʿṣ-ṣiḥḥíja* – ‚doklad o vašom nároku na zdravotnú starostlivosť‘. Obe podoby sú síce z jazykového hľadiska akceptovateľné, avšak použitie dvoch odlišných označení pre ten istý dokument nepôsobí konzistentne.

Medzi dôležité úpravy celej dokumentácie patrilo aj doplnenie aktuálnych adries a kontaktných údajov na rôzne slovenské úrady a inštitúcie, s ktorými môže žiadateľ o azyl prísť do styku. V pôvodných materiáloch boli tieto údaje uvedené takmer výlučne v arabčine, napr. Úrad na ochranu osobných údajov (مكتب حماية البيانات الشخصية) na ulici Hraničná 12 (هرانيتشنا 12), čo mohlo spôsobovať problémy, pokiaľ by mal žiadateľ o azyl skutočný záujem takúto inštitúciu vyhľadať, resp. kontaktovať hoci aj s pomocou nearabsky hovoriaceho človeka. Už samotné slovo „hraničná“ je pre Araba ťažko vysloviteľné.

Zároveň je potrebné skonštatovať, že pôvodná aj aktualizovaná verzia dokumentácie je napísaná takmer výlučne v mužskom rode, aj keď sa medzi žiadateľmi o azyl objavujú prirodzene aj ženy. V arabčine sa mužský rodový tvar používa ako predvolený štandard vo väčšine kontextov vrátane právnych a administratívnych dokumentov. Je to výsledok štrukturálnej charakteristiky arabského jazyka, ktorý má síce precízne vymedzené tvary pre ženský rod dokonca vo väčšej miere, ako je to v slovenskom jazyku,²⁹ avšak s prevládajúcimi mužskými formami, keď sa odkazuje na zmiešané skupiny alebo všeobecné témy. Napriek početným diskusiám a návrhom na používanie rodovo neutrálnejšieho jazyka najmä v súvislosti s oficiálnymi dokumentmi³⁰ táto problematika vyvoláva množstvo niekedy až úsmevných nejasností³¹ a aj my sme pri aktualizácii materiálov radšej uprednost-

²⁹ Popri samostatných osobných zámenách v 3. osobe jednotného a množného čísla (*huwa* – ‚on‘ pre muža a *hija* – ‚ona‘ pre ženu; *hum* – ‚oni‘ pre mužov a *hunna* – ‚ony‘ pre ženy), má arabský jazyk samostatné osobné zámená aj pre 2. osobu jednotného a množného čísla (*anti* – ‚ty‘ pre ženu a *anta* – ‚ty‘ pre muža; *antum* – ‚vy‘ pre mužov a *antunna* – ‚vy‘ pre ženy). Rozdiely v jednotnom čísle sa strácajú, ak nie sú v texte vyznačené samohlásky. Tieto odlišnosti sa následne prejavujú aj v koncovkách slovies, ktoré sa líšia a prispôsobujú ženským alebo mužským tvarom. Podobne je to aj s prívlastňovacími zámenami.

³⁰ Pozri napr. *Gender Sensitive Guidelines* pre United Nations Economic and Social Commission for Wester Asia s návrhmi pre arabský jazyk 28. 11. 2024 [online], s. 6 – 8.

³¹ Ak si odmyslíme hlboko zakorenené tradičné konvencie arabského písomníctva, stačí uviesť často citovaný príklad arabského slova *náʿib* – ‚poslanec‘, ktoré má aj podobu v ženskom rode *náʿiba* – ‚poslankyňa‘. Slovo *náʿiba* je však aj označením pre ‚pohromu‘ či ‚katastrofu‘, a tak sa môže používať v ironickom alebo zámerne nejednoznačnom zmysle (Al-Haq Al-Abed – Al-Hayek Salah 2012, s. 250 – 251).

nili tradičný „úzus“ t. j. príklon k spoločenskej norme príjemcu v duchu zaužívaných kultúrnych konvencií, podľa ktorých sa zmiešaný kolektív označuje mužským rodом. Toto rozhodnutie sme prijali aj z praktických dôvodov, nakoľko vyznačovanie ženských tvarov a koncoviek v zátvorkách nielen pri osobných zámenách ale následne po tom aj napr. pri slovesných tvaroch by ešte viac sťažilo prehľadnosť už aj tak informačne a graficky náročného textu. Tam, kde to bolo možné, sme využili konštrukcie, ktoré sa vzhľadom na absenciu samohlások v arabskom texte môžu čítať podľa potreby v mužskom či ženskom tvare, napr. *ladajka(-ki) haqq fi...* – ‚ty muž/ty žena máš právo na...‘. Keď sa texty týkali výlučne žien (napr. pasáže súvisiace s tehotenstvom žiadateľiek o azyl, pôrodom a pod.), tak sme dôsledne doplnili vokalizáciu. Tá slúžila na jasné rozlíšenie ženských zámen (*anti* – ‚ty‘ pre ženu), spolu s príslušnými ženskými tvarmi sloves, aby bol obsah jednoznačný.

Ešte treba dodať, že arabské právne texty nepoužívajú vykanie v klasickom zmysle, aké poznáme v európskych jazykoch. Preto aj my v slovenskej dokumentácii používame priame vyjadrovanie pripomínajúce tykanie, ktoré však nie je považované za nezdvorilé, pričom sa berie ohľad na kontext a neutrálny jazyk, aby text pôsobil univerzálne a profesionálne.

4. ZJEDNODUŠENÉ VERZIE POUČENIA A TÁBOROVÝCH PORIADKOV

Na základe uvedených skutočností v rámci projektu RESPLAIN tím právnych a jazykových expertov z Migračného úradu MV SR, Ligy za ľudské práva, Jazykovedného ústavu Ľ. Štúra SAV a Filozofickej fakulty UK zostavil zjednodušené (plain) verzie *Poučenia* a táborových poriadkov, ktoré zároveň obsahujú všetky potrebné informácie a ktoré sa následne prekladali do všetkých jazykov.

Zjednodušenie dokumentov aj v arabskej verzii bolo vypracované na základe odporúčaní *Európskych pravidiel tvorby ľahko čitateľných a ľahko zrozumiteľných informácií* (2017). Naším cieľom bolo nájsť stratégiu, ktorá zabezpečí právnu presnosť a zároveň zrozumiteľnosť informácií tak, aby aj ľudia, ktorí z rôznych dôvodov nerozumejú ich rozšírenej verzii (z dôvodu nižšieho vzdelania, stresu a pod.) dôkladne pochopili svoje práva a povinnosti.

Za týmto účelom boli rozsiahle komplikované súvetia a podmienkové vety nahradzané krátkymi vetami v heslovitej podobe. Právnický jazyk bol po konzultáciách s právnikmi zredukovaný a niektoré menej závažné informácie boli vypúšťané. Abstraktná terminológia bola nahradená, pokiaľ to bolo možné, každodennou slovnou zásobou. Výsledkom bolo sprehľadnenie určitých inštitucionálnych procesov do postupných a prehľadných „krokov“.

Jednoduchosť a presnosť vyjadrovania sa docielila aj prostredníctvom eliminácie synonym a zdĺhavých štylizácií v MSA. Používateľskú prístupnosť materiálu

sme sa snažili docieľiť aj využitím neosobných konštrukcií (*jurdžá anna* – ,vyžaduje sa‘, ,je želateľné‘ alebo *jadžúzu/lá jadžúzu* – ,dovoľuje sa‘/,nedovoľuje sa‘), vypúšťaním modálnych slovies vyjadrujúcich zákaz a povinnosť a ich nahrádzaním opisnými tvarmi vyjadrujúcimi nevyhnutnosť (*min ađ-đarúrí* – ,je potrebné‘) či slovesami možnosti vyjadrujúcimi povolenie, možnosť alebo nemožnosť vykonania určitej činnosti.

Imperatív sme používali výlučne tam, kde bola pri pokynoch a inštrukciách kľúčová stručnosť a jasnosť, napr. „*Neopúšťaj tábor bez povolenia na odchod.*“³² Vo všeobecnosti sme sa však vyhýbali autoritatívnemu tónu a uprednostňovali sme formulácie, ktoré dokumentom dodávajú skôr podporujúci a rešpektujúci charakter. Ten sme sa snažili docieľiť aj úvodnými osobnejšími formuláciami táborových poriadkov, ktoré začínajú slovami: „*teraz sa nachádzaš v pobytovom tábore v meste Rohovce. Tu ti bude poskytnutá zdravotná a sociálna starostlivosť, jedlo, ubytovanie, základné hygienické potreby, ďalšie nevyhnutné veci a vreckové.*“³³

Zjednodušenie formulácií oproti pôvodnému *Poučeniu* môžeme pozorovať aj na nasledujúcej ukážke (porovnaj s neskráteným vyššie citovaným paragrafom *Poučenia*, s. 7, bod. 12), pričom sme z textu vypúšťali dôvody mimoriadnych prípadov:

سيقرر مكتب الهجرة بشأن طلبك للجوء في غضون 6 أشهر، يجوز تمديد فترة قرار بالجوء في حالات استثنائية بحد أقصى لمدة 18 شهراً

Úrad pre migráciu rozhodne o vašej žiadosti o azyl do 6 mesiacov. V mimoriadnych prípadoch môže byť lehota na rozhodnutie o azyle predĺžená najviac o 18 mesiacov.

Vyššia zrozumiteľnosť textu bola zabezpečená aj tým, že sme nahradili pôvodné usporiadanie textu novým vizuálnym formátom s osobitnými podnadpismi, do ktorého boli zapracované farebné grafické prvky s inštruktážnymi ikonami. Zároveň bolo rozmiestnenie textu citlivo prispôbené orientácii sprava doľava so zvýraznením kľúčových bodov pomocou tučného písma a odrážok. Tieto úpravy napomáhajú čitateľovi rýchlejšie sa orientovať v texte a zlepšujú celkový používateľský komfort, keďže používateľ má tak lepší prehľad o časovej následnosti jednotlivých postupov a úkonov.

³² *Poučenie žiadateľa o udelenie azylu o právach a povinnostiach* (arabská mutácia, zjednodušená verzia), [online], s 1.

³³ Porovnaj s pôvodným právne ladeným strohým znením: „*Pobytový tábor Rohovce. Všeobecné podmienky: 1 – Vnútorný poriadok utečeneckého tábora (ďalej len „tábor“) určuje podmienky pobytu žiadateľov o azyl a oboznamuje ich s ich právami a povinnosťami, ako aj s pravidlami zachovania bezpečnosti, čistoty a poriadku v tábore.*“ Zbierka interných predpisov Migračného úradu Ministerstva vnútra Slovenskej republiky.



أنت الآن في معسكر الاستقبال في مدينة هومني Humenné



هنا سيتم فحصك من قبل الطبيب.

ستبقى هنا من 3 إلى 4 أسابيع، في حال الضرورة لمدة أطول (إذا كانت حالتك الصحية سيئة)، وبعد إجراءات الدخول سيتم نقلك إلى معسكر الإقامة.



هنا سيتم إصدار بطاقة طالب اللجوء الخاصة بك.

يمكنك أن تستخدمها كوثيقة هويتك إذا كنت قد وصلت إلى الجمهورية السوفاكية مع جواز سفرك



إذا فقدتها يرجى الإبلاغ عن ذلك في المعسكر



انتباه! إذا كنت قد وصلت بدون جواز سفرك فبطاقة طالب اللجوء لا تصلح كوثيقة هويتك

مع بطاقة طالب اللجوء لديك الحق في الرعاية الصحية العاجلة، وهي ضرورية عند كل زيارة للطبيب

سيقدم لك مكتب الهجرة مزيداً من المعلومات



هنا يؤمن لك الطعام، السكن، مستلزمات النظافة الأساسية، وأشياء

ضرورية أخرى ومصرف الجيب



لا تغادر المعسكر بدون تصريح الخروج.

يمكنك الحصول على تصريح الخروج مباشرة في المعسكر وهذا بعد اجتياز الفحص الطبي بنتيجة إيجابية. تصريح الخروج من نوعين:

• تصريح الخروج لمدة أقصاها 24 ساعة إذا لم تكن قد أتممت مقابلتك للدخول.

5. ZÁVER

Z uvedených skutočností vyplýva, že pri vypracovaní právnej dokumentácie v arabčine, ktorá sa týka azylu, zohrávajú kľúčovú úlohu odborné kompetencie a kultúrna rozhladenosť zamestnancov zapojených inštitúcií, tlmočníkov a prekladateľov (pozri aj Tužinská 2023). Považujeme zároveň za nevyhnutné primerane zohľadňovať aj situáciu diglosného arabsky hovoriaceho prostredia, citlivo vyberať terminológiu a vyhýbať sa zložitým právnym štylizáciám s dôrazom na stručnosť, priamočiarosť a jednoznačnosť. Okrem podpory komunikačných zručností zúčastnených aktérov odporúčame zohľadňovať aj etnokultúrne aspekty arabsko-islamskej spoločnosti a reflektovať komunikačné stereotypy.

Arabská dokumentácia vypracovaná v rámci projektu RESPLAIN spĺňa aktuálne obsahové aj grafické kritériá. Dospelo sa k nim vďaka dôkladnému prepracovaniu, terminologickému zjednoteniu a štylistickému zjednodušeniu, ktoré významnou mierou znižujú riziko nesprávnych výkladov či chýb pri aplikácii právnych princípov. Tento typ

medzisektorovej odbornej spolupráce je zároveň príkladom prínosu humanitných vied pre administratívu štátnej správy – od návrhu projektu, zohľadnenia aktuálnych potrieb cez realizáciu aplikovaného výskumu až po tvorbu konkrétnych výstupov.

Bibliografia

AL-HAQ AL-ABED, Fawwaz – AL-HAYEK SALAH, Reema (2012): Gender in Arabic. From a Language Planning Perspective. In: *Dacoromania*, serie nuová, roč. XVII, č. 2, s. 245 – 264.

BADAWI, Elsaïd (1973): *Mustawaját al-‘arabíja al-mu‘ášira fí Mišr*. [Úrovne arabského jazyka v Egypte]. Al-Qáhira: Dár al-Ma‘árif. 223 s.

BOUSSOFARA-OMAR, Naima (2006): Diglossia. In: K. Versteegh et al (eds.): *The Encyclopedia of Arabic Language and Linguistics*. Zv. I. Leiden: Brill, s. 629 – 637.

BRIEŠKA, Marek (2009a): Genéza subštandardnej orálnej nehovorovej arabčiny. In: *Studia Orientalia Slovaca*, roč. 8, s. 271 – 292.

BRIEŠKA, Marek (2009b): Arabská dvojvariantná komunikácia v prácach akadémie arabského jazyka – prvá dekáda 21. storočia. In: *Jazykovedný časopis*, roč. 60, č. 1, s. 75 – 86.

DITTERS, Everhard (2008): Mašdar. In: K. Versteegh et al (eds.): *The Encyclopedia of Arabic Language and Linguistics*. Zv. III. Leiden: Brill, s. 164 – 169.

DROZDÍK, Ladislav (2006): Arabčina a diglosívna komunikácia v súčasnom arabskom svete. In: L. Drozdík – V. Krupa – A. Rácová (eds.): *Jazyková politika a problémy komunikácie*. Bratislava: Zing Print, s. 10 – 80.

FERGUSON, Charles (1959): Diglossia. In: *Word*, roč. 15, s. 325 – 340.

GAŽÁKOVÁ, Zuzana – PAULINY, Ján (2013): *Arabský jazyk a kultúra*. Bratislava: Univerzita Komenského. 230 s.

MADSEN, Dorte (1997): A Model for Translation of legal texts. In: Z. Jettmarova – K. Kaindl – M. Snell-Hornby (eds.): *Translation as Intercultural Communication: Selected Papers from the EST Congress, Prague 1995*. Amsterdam: John Benjamins Publishing Company, s. 291 – 299.

MITCHELL, Terence Frederick (1986): What is Educated Spoken Arabic? In: *Journal for the Sociology of Language*, roč. 61, s. 7 – 32.

Al-Mundžid fí ‘l-lughā wa ‘l-a‘lām. [Slovník jazyka a významných osobností] (1997) Bajrút: Dár al-Mašriq. 714 s.

PARKINSON, Dilworth B. (2003): Verbal features in oral fušḥā in Cairo. In: *International Journal of the Sociology of Language*, roč. 163, s. 27 – 41.

SCHÄFFNER, Christina – ADAB, Beverly (1997): Translation as intercultural communication – Contact as conflict. In: Z. Jettmarova – K. Kaindl – M. Snell-Hornby (eds.): *Translation as Intercultural Communication: Selected Papers from the EST Congress, Prague 1995*. Amsterdam: John Benjamins Publishing Company, s. 325 – 337.

SULEIMAN, Yasir (2006): Arabiyya. In: K. Versteegh et al (eds.): *The Encyclopedia of Arabic Language and Linguistics*. Zv. I. Leiden: Brill, s. 173 – 178.

ŠVEDA, Pavol – TUŽINSKÁ, Helena (2021): Public Service Interpreting in the Context of Social and Political Tensions. In: P. Šveda et al. (eds.): *Changing paradigms and approaches in interpreter training. Perspectives from Central Europe*. London: Routledge, s. 25 – 42.

Táborové poriadky pre záchytný tábor v Humennom a pobytové tábory v Rohovciach a Opatovskej Novej Vsi. Zbierka interných predpisov Migračného úradu Ministerstva vnútra Slovenskej republiky.

TUŽINSKÁ, Helena (2020): *Medzi riadkami. Etnografia tlmočenia azylových súdnych pojednávaní*. Bratislava: AKAMedia. 360 s.

TUŽINSKÁ, Helena (2023): Interpreters as cultural experts? Asylum practices in Slovakia. In: *Jindal Global Law Review*, Roč. 14, č. 2, s. 241 – 257.

VERSTEEGH, Kees (1997): *The Arabic language*. Edinburgh: Edinburgh University Press. 288 s.

WEHR, Hans (1980): *Dictionary of Modern Written Arabic*. Beirut: Librairie du Liban. 1130 s.

Internetové a iné zdroje

Európske pravidlá tvorby ľahko čitateľných a ľahko zrozumiteľných informácií. Dostupné na: https://www.inclusion-europe.eu/wp-content/uploads/2017/06/SK_Information_for_all.pdf [cit. 30. 11. 2024].

Ekonomická encyklopédia. Dostupné na: <https://www.ekonomicka.sk/zachytny-tabor/> [cit. 1. 12. 2024].

Gender Sensitive Guidelines. United Nations Economic and Social Commission for Western Asia. Dostupné na: https://www.unescwa.org/sites/default/files/services/doc/guidelines_gender-sensitive_language_e-a.pdf [cit. 28. 11. 2024].

Haqquka fi 'l-baqá' fi markaz al-istiqbál [Tvoje právo zostať v prijímacom centre]. Dostupné na: <https://italy.refugee.info/ar/articles/5388911887383> [cit. 10. 7. 2024].

Júriú 2024 tuwadždžih al-anžár naḥwa našab mu'askar i' tiqál názi [Euro 2024 upriamuje pozornosť na pamätník nacistického koncentračného tábora]. Al-Džazíra, 25/6/2024. Dostupné na: <https://www.aljazeera.net/sport/2024/6/25/buchenwald-concentration-camp> [cit. 28. 11. 2024].

Marákiz quḍúm al-ládži'in wa marákiz al-quḍúm, ittichádh al-qarár wa 'l-tarḥíl [Centrá príchodu utečencov a centrá príchodu, rozhodovacie konanie, vyhostenie]. Bundesamt für Migration und Flüchtlinge. Dostupné na: <https://www.bamf.de/AR/Themen/AsylFluechtlingsschutz/Ankunftszentren/ankunftszentren-node.html> [cit. 20. 11. 2024].

Mu'askarát al-iḥtidžáz wa 'l-'amal fi šamál Ifriqijá [Záchytné a pracovné tábory na severe Afriky]. United States Holocaust Memorial Museum. Dostupné na: <https://encyclopedia.ushmm.org/content/ar/article/labor-and-internment-camps-in-north-africa> [cit. 20. 11. 2024].

Poučenie žiadateľa o udelenie azylu o právach a povinnostiach. Dokument Ministerstva vnútra Slovenskej republiky, dokumenty na stiahnutie, slovenská verzia, 17 s. Dostupné na: <https://www.minv.sk/?poucenie-ziadateľa-o-udelenie-azylu-o-pravach-a-povinnostiach> [cit. 8. 11. 2025].

Poučenie žiadateľa o udelenie azylu o právach a povinnostiach. Dokument Ministerstva vnútra Slovenskej republiky, dokumenty na stiahnutie, arabská verzia, 9 s. Dostupné na: <https://www.minv.sk/?poucenie-ziadateľa-o-udelenie-azylu-o-pravach-a-povinnostiach> [cit. 8. 11. 2025].

Zariadenia migračného úradu. 25/6/2024. Dostupné na: <https://www.minv.sk/?zariadenia-migracneho-uradu-1> [cit. 14. 11. 2024].

Zbierka interných predpisov Migračného úradu Ministerstva vnútra Slovenskej republiky.

PERSISTING GRAMMATICAL ERRORS OF MACHINE TRANSLATION¹

KATARÍNA WELNITZOVÁ¹ – DAŠA MUNKOVÁ²

¹University of Ss. Cyril and Methodius in Trnava, Slovakia (corresponding author)

²Constantine the Philosopher University in Nitra, Slovakia

WELNITZOVÁ, Katarína – MUNKOVÁ, Daša: Persisting grammatical errors of machine translation. *Jazykovedný časopis (Journal of Linguistics)*, 2025, Vol. 76, No. 2, pp. 468–492.

Abstract: The paper investigates the quality of machine translation (MT) and traces its development through two main approaches – Statistical Machine Translation (SMT) and Neural Machine Translation (NMT) – by comparing English-to-Slovak outputs produced by Google Translate. The aim of the paper is to evaluate the quality of MT outputs from the point of view of two typologically different languages – English, a predominantly analytic language, and Slovak, a primarily synthetic language – using a sample of newspaper texts, which are often translated by machine due to their wide vocabulary and varied subject matter. The research results indicate that NMT (obtained in 2023), compared to its predecessor SMT (obtained in 2017), has significantly improved in almost all framework categories. The NMT output is much more fluent, sounding more natural and comprehensible. In contrast, shortcomings can be found in the category of syntactic-semantic correlativeness and lexical semantics. In such cases, neural MT may struggle to select the appropriate fit-in-context meaning; moreover, these lexemes can further shift the meaning of the entire sentence, clause, or even utterance.

Keywords: machine translation quality assessment, Slovak, English, statistical MT, neural MT.

1. INTRODUCTION

As part of technological advancements, significant progress is being observed in the translation industry as well. This includes, among other developments, significant improvements to one of the most widely used and well-known machine translation systems – Google Translate (GT). Its quality, efficiency, and popularity are continually improving, and it is becoming popular among millions of users worldwide due to its free access and availability (Wang et al. 2022, p. 143; Melby 2020, p. 422). Koponen (2016, p. 131) claims that developments have shifted GT from a peripheral position to a more central role in the translation industry. Machine translation (MT) is designed for both common language users who do not have a high proficiency in the target language and professional human translators using

¹ This work was supported by the Slovak Research and Development Agency under the Contract no. APVV-23-0554.

MT in their computer-assisted (CAT) tools (e.g., TRADOS) to foster translation. In terms of the translation process, MT demonstrates a considerably greater improvement in productivity compared to human translation (HT). This has been empirically demonstrated in many cases over the last decade, relying on phrase-based and rule-based paradigms of MT across various text types, including technical documents (Plitt – Masselot 2010) and news (Martín – Serra 2014).

The translation industry is still evolving, and as advancements continue, different translators' competencies are required. In these terms, Bednářová-Gibová et al. (2024, p. 104) point at the fact that technological competence, subsuming the knowledge of IT applications, CAT tools, essentials of MT translation, data literacy and workflow management tools, sits at the heart of the revised 2022 EMT framework. They describe a translator as an augmented translator whose work is significantly aided and enhanced by technology.

1.1 Research objective

The primary goal of this research is to assess the MT quality and evaluate its advancements. The most frequent and significant errors in statistical machine translation (SMT) and neural machine translation (NMT) outputs will be identified, classified, and analysed. Moreover, the improvements and shortcomings of NMT compared to SMT will be highlighted. It will be compared whether a statistically significant difference exists between the two approaches used in automated translation.

Various grammar categories, focusing on predicativeness, syntactic-semantic correlativeness, and lexical semantics, which have a crucial impact on the comprehension of MT output, will also be discussed. This study builds upon research initiated by Welnitzová and Munková (2021, p. 90), which found that the most frequent errors leading to misunderstandings are related to predicativeness, nominal and verbal morpho-syntax, word order, and lexical adequacy.

1.2 Machine translation and its background

MT is defined as a fully automated process that transforms a text from one language into another without human intervention (Quah 2006, p. 2). Since this process is fully automated, the output often requires correction or revision (post-editing) by a human. The goal of post-editing is to ensure both the accuracy of the MT output (so that the target text remains faithful to the source text) and the fluency of the text.

Melby (2020, p. 419) defines the most common MT paradigms – SMT and NMT – as follows: SMT works with extensive bilingual corpora. It is based on the acquisition of a translation model and the decoding of sentences from the source text to find an adequate translation in the target text. Its aim is to create a system that would match sentences from the source text with the sentences from the target text. It creates target text using language and translation models and statistical probability.

However, the disadvantage of SMT is its limited training (translation is “trained” on one type of corpus) and equivalence (since alignment of words can be unreliable) (Munková 2017, p. 21).

NMT has appeared as a new paradigm in MT, and has been shown to improve the translation quality, regardless of the language pair (Toral – Sánchez-Cartagena 2017). It seems that translations produced by NMT are much more fluent (Bentivogli et al. 2016) compared to those derived by phase-based SMT, and that NMT does not lead to literal translations, as it was in the case of phase-based SMT. Melby also states that nowadays sufficient training data for a viable NMT system are available only for about twenty out of more than 4,000 languages in the world (2020, p. 420). The move to NMT is the most significant change in the MT approach. NMT, in principle, processes large amounts of data, translates a sentence as a whole, and therefore, in most cases, NMT output is much clearer than phrase-based SMT output. Although NMT provides a reader with seemingly more comprehensible and flawless outputs, the quality and comprehensibility of NMT output are questionable.

NMT has been brought to the fore as it has shown significant results in translations from English into French (Luong et al. 2015) and from English into German (Jean et al. 2015). It is based on neural networks, trained end-to-end, with a small memory track and its ability to generalise long sequences of words. It uses deep machine learning represented by neural networks (Bessenyei 2017), using algorithms which enable it to learn and consequently decide about translation solutions. It can process source segments (one segment usually corresponding to one sentence) and transform them into target segments, considering whole sentences, not just phrases.

Moorkens and Lewis (2020, p. 474) state that NMT is regarded as a form of weak artificial intelligence, as it determines the next transformation to perform autonomously rather than executing explicit instructions from the user. NMT requires vast quantities of human-created data for training. An example can be found at language data brokers who sell language data for MT learning system training and media stories of a boom in language data (Diño 2018).

Caffrey and Valentini (2020, pp. 142–143) claim that NMT is based on neural network models which learn from previously translated texts. NMT output produces a more natural word order than SMT. When comparing the systems, NMT frequently outperform SMT in automatic evaluations, but human evaluations are less definite (Castilho et al. 2017). In the comparison of output translated from English to Japanese using SMT and NMT, a human evaluator preferred the NMT output to the SMT output. The difference was not significant (4%), and in the overall evaluation, it can be stated that both systems (NMT and SMT) were comparable for the majority of sentences.

NMT fluency has greatly improved, albeit sometimes at the expense of adequacy. In such cases, NMT offers surprising translation solutions – it often creates sentences which seem to be fluent, but are, on the contrary, not adequate in meaning (cf. Munková 2017, p. 23).

1.3 Typological characteristics of Slovak and English in newspaper style

When analysing errors in MT outputs, it is essential to evaluate them in the context of the specific style and type of text, as well as in relation to the source and target languages. Our research examines MT outputs of newspaper texts translated from English into Slovak using two MT approaches – SMT and NMT. To fully understand the errors that occur in MT outputs, it is crucial to consider the linguistic principles of both languages within the given stylistic contexts.

Slovak is primarily classified as a synthetic language, whereas English is considered an analytic language (Dolník 2013, p. 87). According to Vaňko and Auxová (2015, p. 24), in the analytic type, grammatical meaning is expressed analytically – that is, through separate words, one of which carries lexical meaning and the other (auxiliary) conveys grammatical meaning. Analytic languages do not formally distinguish between nominative and accusative cases, and their word order is typically rigid.

In contrast, Slovak features synthetic morphology (Ondruš – Sabol 1984, p. 186). This type is manifested by the richness of forms of open class words, a marked gender differentiation of forms, the expression of a complex of grammatical meanings (e.g., gender, number, and case) by a single formal element within a word (e.g., the morpheme *-u* in the form *žen-u* (the form *woman* in the accusative case) reflects feminine gender, singular number, and the accusative case), the occurrence of synonymy and homonymy of case suffixes, and the differentiation of forms by changes at the end of the morphological base.

Based on the theoretical background and previously mentioned studies, we can state that the main differences between English and Slovak (and thus the challenges in both human and machine translation) lie in almost all examined grammatical categories. In the category of ‘*predicativeness*’, it is agreement between the subject and verb in person, number, and gender. In the category of ‘*syntactic-semantic correlation*’, the key aspects are ‘*nominal morpho-syntax*’ and ‘*verbal morpho-syntax*’ (mainly in noun phrases and verb phrases, which follow similar rules as subject-verb agreement in ‘*predicativeness*’), as well as word order, which is fixed in English but loose in Slovak. Word order in Slovak mostly applies the sentence pattern S – V – O. On the other hand, it is not grammatically relevant, and it does not determine the grammatical (syntactic) function of a word in a sentence. For example, the word order in the sentence *Evu ľúbi Peter.* (*Peter loves Eve.*) can be changed into *Peter ľúbi Evu.* (*Peter loves Eve.*) without any changes in the syntactic functions of nouns *Peter* and *Eva*. Fixed word order in Slovak is required within noun phrases (particularly in pre-modifiers and post-modifiers) and complements (Vaňko 2015, pp. 80–81). Moreover, serious issues in machine translation can arise because, unlike in English, the subject in Slovak can be expressed both explicitly and implicitly (e.g., ENG: *She said.* SVK: *Ona povedal-a./Povedal-a.*). The pronoun *ona* (*she*) can be omitted, and cohesion in the text is conveyed through the suffix *-a* in the verb

povedal-a (said, referring to the feminine gender). Noun phrases functioning as subjects and objects (mainly expressed by noun-adjective or noun-noun combinations) are structurally determined by the gender of the head noun. Inflectional endings reflect the grammatical case of an adjective or noun, thereby determine the role of the word or phrase (the subject is in the nominative case, while the object is primarily in the accusative case). Further discrepancies between English and Slovak can be found in the category of ‘lexical semantics’, as English has significantly more polysemantic and homonymous words than Slovak (see Ondruš – Sabol 1984, pp. 228–229). To ensure translation adequacy in such cases is challenging, even for a human translator.

Table 1: Examples illustrating structural divergence in Slovak and English (Welnitzová 2024, p. 27).

Slovak (SVK)	English (ENG)
<i>chodím (chod-im)</i>	<i>I go</i>
<i>študenta (študent-a)</i>	<i>of the student</i>
<i>nepôjdu (ne-pôjd-u)</i>	<i>they will not go</i>

The analytic form *I go* (2 words in English) is expressed by the form *chodím* (1 word in Slovak), just as the structure *of the student* (3 words in English) has the equivalent *študenta* (1 word in Slovak) (Dolník 2013, p. 88). The structure *they will not go* (4 words in English) is translated as *nepôjdu* (1 word in Slovak). It is evident that grammatical categories are expressed in different morphemes, structures, and the number of words.

Regarding the English–Slovak language pair, it is necessary to add that the number of verb tenses in the given grammar systems is different: Slovak has three grammaticalised verb tenses, while English has six simple temporal verb forms, which can also be combined with progressive forms (e.g., *I work, I am working*; see Dušková 2012). The ante-preterit is also recognised within the Slovak tense system (Oravec et al. 1984, p. 148); however, it is regarded as a marginal stylistic phenomenon, primarily used as a stylistic device (Vajičková 2023, p. 7) and not typically employed in newspaper style (Welnitzová et al. 2020, p. 166).

Table 2: Examples illustrating tense systems of Slovak and English.

Slovak (SVK)	English (ENG)
Present <i>hliadk-u-jú</i>	present simple <i>they patrol</i>
Past <i>hliadk-ova-l-i</i>	past simple <i>they patroll-ed</i>
Future <i>budú hliadk-ova-t'</i>	simple future <i>they will patrol</i>
	present perfect <i>they have patroll-ed</i>
	past perfect <i>they had patroll-ed</i>
	future perfect <i>they will have patroll-ed</i>

In synthetic languages, common phenomena such as synonymy and homonymy of case endings are typical. In English, the grammatical categories of gender, number, and case are morphologically marked, though not by means of declension as in the Slovak language.

In addition to the aforementioned issues, which reflect differences between English and Slovak related to grammatical and lexical equivalence, there are also challenges concerning pragmatic and textual equivalence. These include the translation of names, idioms, or geographically marked expressions (see Kvetko 2021, p. 27).

Unlike the grammar of English and Slovak, the register of the newspaper style is similar in both contexts. According to Mistrík (1989, p. 460), the newspaper style is characterised by a high frequency of nouns, numbers, abbreviations, names, and symbols, while verbs occur less frequently. He notes that within verb categories, aspect (with a perfective-to-progressive ratio of 1:2), person (with the third-person singular being the most commonly used), and tense (predominantly past tense) are particularly significant. Additionally, syntactic and grammatical constructions in this style are relatively flexible. According to Findra (2013, pp. 262–270), the primary objective of newspaper texts is to convey information to the reader; consequently, clarity and comprehensibility are essential. The newspaper style frequently employs both simple and complex sentence structures. Its lexis is relatively diverse, depending on the genre and domain. From the morphological point of view, it is characterised by predominantly nominal structures, often using personal names, surnames, geographic names, and culturally specific references (realia). Biber and Conrad (2009, p. 109) define *newspaper style* (also referred to as *newspaper writing*) in English as a style characterised by a written register, emphasising its informative function. The primary goal of this style is to report and describe events rather than interpret them. In terms of nominal features, various noun forms – including nouns, nouns used as pre-modifiers or post-modifiers, and noun phrases – are common. Regarding verbs, newspaper style primarily uses the present simple and past tense to narrate sequences of events.

The newspaper style has been labelled in various ways, such as *publicistic style*, with more recent designations including *media communication sphere* (Slančová et al. 2022, p. 300) or *the sphere of media communication* (Hlavatá et al. 2019, p. 44). However, as the term *newspaper style* is present in both Slovak and English linguistic contexts, it will be consistently applied in this study.

2. MATERIAL AND METHODS

At the beginning of research (in 2017), we created a corpus consisting of 59 newspaper articles (3,376 segments / 54,442 words) from the British online newspaper *The Guardian*. The data were pre-processed through tokenisation and

segmentation before being translated by statistical GT (in 2017) and later by neural GT (in 2023). The MT outputs, along with their original texts, were then imported into the virtual environment OSTEPERE – a system for translation, post-editing, and MT evaluation (Munková – Munk 2016; Benko – Munková 2016) – in which the texts were post-edited by professional translators.

The identification and classification of errors in the SMT and NMT outputs were performed by two Slovak language experts, who categorised and evaluated the MT errors. The MT output analysis was aligned with general error typology frameworks (e.g., Font Llitjos et al. 2005, Vilar et al. 2006, and Lommel 2018). Due to discrepancies between English and Slovak, as well as the character of texts examined, we focused on a more detailed assessment of morpho-syntactic, syntactic-semantic relations and lexical semantics – language categories with a significant number of errors in Slovak, as designed by Vaňko (2017, pp. 83–100). In this context, four key grammatical realms were examined: **1. *Predicativity and Modal and communication sentence framework*** (whether the main sentence elements – the subject and predicate – are correctly identified in the source segment and accurately transferred into the target segment by MT and whether the modal framework in the source segment is identified and accurately transferred into the modal framework of the target segment by MT), **2. *Syntactic-semantic correlativeness*** (the correctness of expressing semantic-syntactic relationships between content words within phrases and sentences, as well as the grammatical means of their realisation [e.g., agreement, pre-modification, post-modification] is examined), **3. *Compound/complex sentences*** (whether the semantic relationships between sentences in the source segment are correctly transposed into the target segment – for example, the correct use of conjunctions, time shifts, and the transformation of compound/complex sentences from the source segment into compound/complex sentences in the target segment), and **4. *Lexical semantics*** (is closely linked to the categories of syntactic-semantic correlativeness and compound/complex sentences; an arrangement of words into grammatically correct phrases and sentences is essential; otherwise, the segment may become difficult or even impossible to understand). After identifying and classifying the MT errors according to this framework, we calculated their frequencies in the analysed newspaper texts. This analysis led to the formulation of the following research hypothesis:

H0: *There is no statistically significant difference between statistical and neural MT.*

3. RESULTS

The following section presents the research results, highlighting the occurrence of errors in specific language categories within both SMT and NMT outputs. Categories without errors are not included in the results.

Based on the results of Levene’s Test for Homogeneity of Variances (Table 3), the error frequencies in SMT and NMT exhibit approximately equal variability. Therefore, Univariate tests for repeated measures can be used for the statistical verification of our hypothesis.

Table 3. Levene’s Test for Homogeneity of Variances.

	MS Effect	MS Error	F	p
SMT	285.798	201.310	1.420	0.2540
NMT	27.757	20.321	1.366	0.2697

Based on the results of the test, we reject the null hypothesis, suggesting a statistically significant difference between SMT and NMT ($p = 0.002118$).

Specifically, the analysis reveals a statistically significant difference in error frequency between SMT and NMT, in favour of NMT (Table 4). On average, fewer errors were identified in NMT (5.526) compared to SMT (18.533).

Table 4. Multiple comparisons: MT.

MT	Mean	1	2
NMT	5.526	****	
SMT	18.553		****

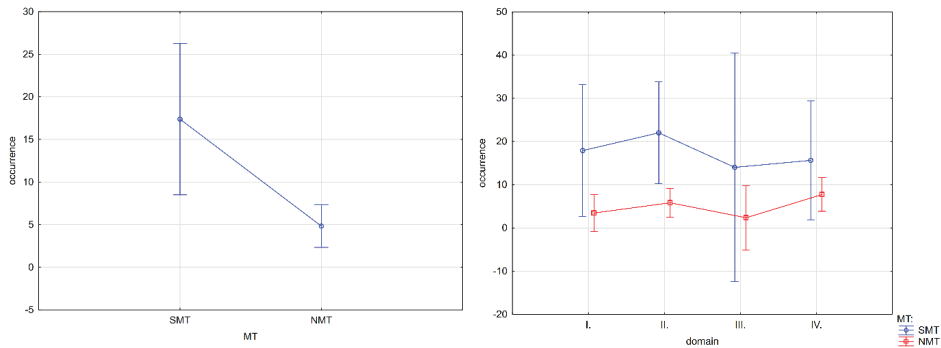


Figure 1: Error frequency for SMT and NMT a) overall b) by categories (domain).

Subsequently, we aimed to examine which specific error categories showed significant differences between SMT and NMT.

3.1 Predicativeness and Modal and communication sentence framework

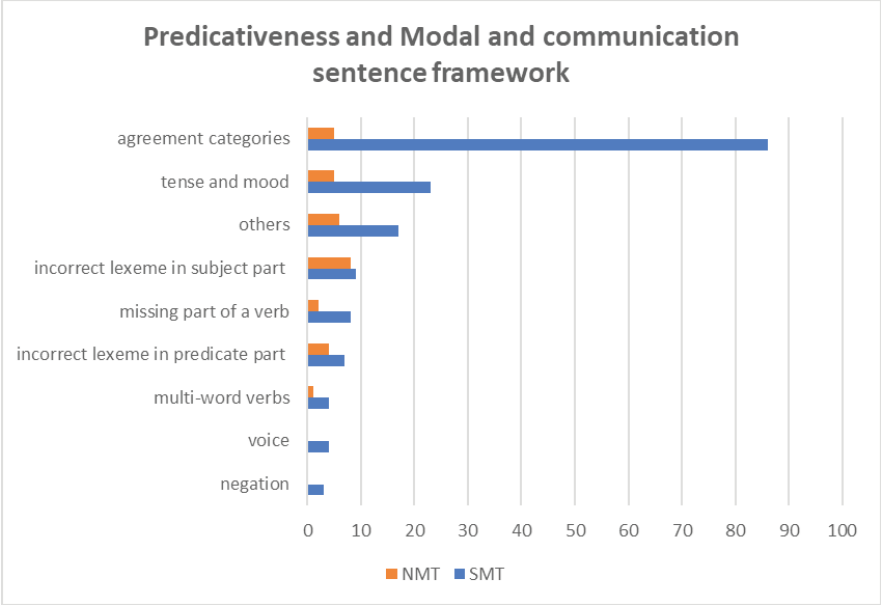


Figure 2: The results in Predicativeness and Modal and communication sentence framework for SMT and NMT.

Figure 2 shows the results in the category of *Predicativeness*, covering ‘*predicative categories*’ (‘*infinitive form of a verb*’, ‘*missing part of a verb*’, ‘*voice*’, ‘*multi-word verbs*’, ‘*incorrect lexeme in predicate part*’), ‘*agreement categories*’ (‘*agreement in person, number, gender*’, ‘*subject and its issues*’, ‘*incorrect lexeme in subject part*’), ‘*others*’ (the errors which were not classified due to lacking subcategories); and in the item of *Modal and communication sentence framework*, for both SMT and NMT outputs. The most frequent errors were identified in the category of ‘*agreement in person*’, ‘*number*’, and ‘*gender*’ (SMT 86, NMT 5), ‘*tense*’ and ‘*mood*’ (SMT 23, NMT 5), and ‘*other issues*’ (SMT 17, NMT 6). The latter category primarily involves the English form *’s*, which can indicate either verb contraction or possessive case. These errors were, in fact, quite frequent. Errors in other categories included the following: ‘*incorrect lexeme in subject part*’ (SMT 9, NMT 8), ‘*missing part of a verb*’ (SMT 8, NMT 2), ‘*incorrect lexeme in predicate part*’ (SMT 7, NMT 4), ‘*multi-word verbs*’ (SMT 4, NMT 1), ‘*voice (active/passive)*’ (SMT 4, NMT 0).

The highest number of errors in SMT was observed in the categories of ‘*agreement in person*’, ‘*number*’, and ‘*gender*’ (Figure 2). This aligns with the typological differences between English and Slovak, as well as the nature of newspaper

texts. In such texts, nominal features – nouns in various forms (e.g., nouns, nouns used as pre-modifiers or post-modifiers, nouns within noun phrases, and noun phrases embedded in prepositional phrases) – are more prevalent. Consequently, the frequent occurrence of these structures and the related challenges (e.g., noun inflection, noun-adjective agreement, and noun-verb agreement) were anticipated.

Example (1) illustrates the improvement of machine translation in the ‘category of agreement in gender’. In SMT, for example, the structure *seasonal, casual nature of the work was often problematic* was translated as *sezónne, príležitostná povaha práce bolo často problematické*. Observing the affixes in the word, it can be concluded that the word *sezónn-e* refers to the neuter gender, the noun phrase *príležitostn-á povaha* refers to the feminine gender, the verb structure *bolo problematické* to the neuter gender (the correct structure is *sezónny, neplánovaný charakter práce bol často problematický*, with all these words referring to the same gender, as shown below, which displays the human translation (HT). We can say that the level of agreement in the given SMT subject-verb structure is relatively low. On the contrary, NMT suggests the translation *sezónny, príležitostný charakter práce bol často problematický* which is correct. All adjectives (*sezónny, príležitostný, problematický*) correspond with the noun (*charakter*) and the auxiliary verb (*bol*) in the ‘category of gender’ (as well as in ‘number’ and ‘case’). We can conclude that the improvement in the categories of ‘agreement in person’, ‘number’, and ‘gender’ is significant, and that the NMT output is much more comprehensible than the SMT output:

(1)

ST: They acknowledged that the **seasonal, casual nature** of the work **was** often **problematic** for people with families who live here permanently, trying to pay mortgages.

HT: Priznali, že **sezónny, neplánovaný charakter práce bol** často **problematický** pre ľudí s rodinami, ktorí tu žijú trvale a snažia sa splácať hypotéky.

SMT: Uznali, že **sezónne, príležitostná povaha práce bolo** často **problematické** pre ľudí s rodinami, ktorí tu žijú trvalo, snažia sa platiť hypotéky.

NMT: Uznali, že **sezónny, príležitostný charakter práce bol** často **problematický** pre ľudí s rodinami, ktorí tu žijú trvale a snažia sa platiť hypotéky.

In the following part, examples of additional frequently occurring errors within the subcategories of *Predicateness*, as shown in the graph (Figure 2), will be presented. Due to space limitations, errors will be highlighted in bold, without detailed commentaries. It is important to note that some segments may contain other errors beyond those emphasised; however, all errors were thoroughly analysed and assigned to their respective categories. The segments are organised as follows: ST (source text), HT (human translation), SMT (statistical machine translation), and NMT (neural machine translation). In some cases, the error appears in either the SMT or NMT output, or both. Each SMT and NMT segment is marked as either correct or incorrect and can be compared with the HT version for reference.

(2) Example of an error in the category of ‘tense’:

ST: *Revenue **has been falling** too, with \$4.8bn pulled in last quarter.*

HT: *Aj tržby **klesajú**, v poslednom štvrtroku predstavovali 4,8 miliárd dolárov.*

SMT: *Výnosy **klesá** príliš, s 4,8 mld \$ vytiahol v poslednom štvrtroku. (correct)*

NMT: *Príjmy tiež **klesali**, keď v poslednom štvrtroku pritiahli 4,8 miliárd dolárov. (incorrect)*

(3) Example of an error in the category of ‘mood’:

ST: *For more evidence of that theory, **look** to the Mac.*

HT: *Viac dôkazov o tejto teórii **nájdete** na počítači Mac. (From the point of view of context, comprehensibility, and fluency, it can be considered a correct translation.)*

SMT: *Ďalší dôkaz o tom teoreticky **vyzerat'** na Mac. (incorrect)*

NMT: *Pre viac dôkazov o tejto teórii **sledujte** Mac. (correct)*

(4) Example of an error in the category of ‘other’ errors:

ST: *The American's manager, Mark Steinberg, delivered added detail on another tale of woe.*

HT: *Američanov manažér Mark Steinberg priniesol ďalšie podrobnosti k ďalšej nešťastnej udalosti.*

SMT: *Americký manažér, Mark Steinberg, vydal ďalšie podrobnosti o ďalšom príbehu beda. (incorrect)*

NMT: *Američanov manažér Mark Steinberg dodal ďalšie detaily ohľadom zlých správ. (correct)*

(5) Example of an error in the category of ‘incorrect lexeme in the subject part’:

ST: *For one thing, **revenue and unit sales** just keep rising – though from what, and to what, we don't know.*

HT: *Na jednej strane to, že **príjem a jednotkový predaj** stále narastá - hoci nevieme povedať, z čoho a do čoho.*

SMT: *Pre jednu vec, **príjmov a jednotkové predajné** jednoducho stále rastie - aj keď z toho, čo a čo nevieme. (incorrect)*

NMT: *Jednak to, že **tržby a jednotkový predaj** stále rastú - aj keď z čoho a do čoho nevieme. (correct)*

(6) Example of errors in the category of ‘incorrect lexeme in a verb’:

ST: *In other provinces that have adopted western diets you see pretty young girls but when they smile they have rotten teeth, because the sugar **has broken down** their teeth.*

HT: *V iných provinciách, ktoré si osvojili západnú stravu, vidíte pekné mladé dievčatá, ale keď sa usmievajú, majú pokazené zuby, pretože cukor im **pokazil** zuby.*

SMT: *V iných provincií, ktoré prijali západné stravy vidíte celkom mladé dievčatá, ale keď sa usmievajú majú pokazené zuby, pretože cukor **sa pokazil** zuby. (incorrect)*

NMT: *V iných provinciách, ktoré si osvojili západnú stravu, vidíte pekné mladé dievčatá, ale keď sa usmievajú, majú zhnité zuby, pretože cukor **si zlomil** zuby. (incorrect)*

(7) Example of an error in the category of ‘multi-word verbs’:

ST: *Tiger Woods **pulls out** of Dubai Desert Classic with back injury*

HT: *Tiger Woods **odstupuje** pre zranenie chrbta zo súťaže Dubai Desert*

SMT: *Tiger Woods **vytiahne** z Dubai Desert Classic s bolesťami chrbta (incorrect)*

NMT: *Tiger Woods **sa sťahuje** z Dubajskej púšte Classic so zranením chrbta (correct)*

Figure 2 also includes the results of the *Modal and communication sentence framework*: In the ‘category of negation’, 3 errors were recorded in SMT, while no errors were found in NMT.

(8) Example for the category of ‘negation’:

ST: Cadot said **no** explosives **had been found** in the man’s bag and there **was** “**no threat**”.

HT: Cadot povedal, že v pánskej taške **neboli** nájdené **žiadne** výbušniny a že „**neexistuje** **žiadna hrozba**“.

SMT: Cadot uviedol žiadne výbušniny boli nájdené v mužovej sáčku a tam bolo „**žiadna hrozba**“. (incorrect)

NMT: Cadot povedal, že v pánskej taške **neboli** nájdené **žiadne** výbušniny a že „**neexistujú** **žiadne hrozby**“. (correct)

3.2 Syntactic-semantic correlativeness

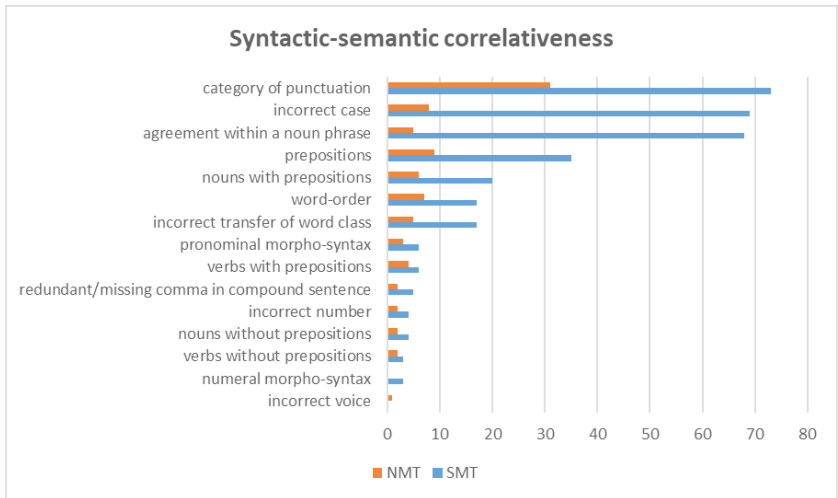


Figure 3: Results in Syntactic-semantic correlativeness for SMT and NMT.

Figure 3 shows the results in the item of *Syntactic-semantic correlativeness*, covering ‘nominal morpho-syntax’, ‘pronominal morpho-syntax’, ‘verbal morpho-syntax’, ‘word order’, and ‘other issues’.

SMT numbered the most errors (Figure 3) in the ‘category of punctuation’ (SMT 73, NMT 31), ‘incorrect case’ (SMT 69, NMT 8), ‘agreement within a noun phrase’ (SMT 68, NMT 5). ‘Other issues’ in machine translation outputs can be seen in the ‘category of prepositions’ (SMT 35, NMT 9), ‘nouns with prepositions’ (SMT 20, NMT 6), ‘word order’ (SMT 17, NMT 7) and ‘incorrect transfer of a word class’ (SMT 17, NMT 5). The categories with the occurrence of errors below 10 are: ‘verbs with prepositions’ (SMT 6, NMT 4), ‘pronominal morpho-syntax’ (SMT 6, NMT 3),

‘*redundant/missing comma in compound sentence*’ (SMT 5, NMT 2), ‘*nouns without prepositions*’ (SMT 4, NMT 2), ‘*incorrect number*’ (SMT 4, NMT 2), ‘*verbs without prepositions*’ (SMT 3, NMT 2), ‘*numeral morpho-syntax*’ (SMT 3, NMT 0), and ‘*incorrect voice*’ (SMT 0, NMT 1).

Figure 3 reflects a significant decrease in errors in the category of ‘*agreement within a noun phrase*’, ‘*incorrect case*’, ‘*prepositions*’, and ‘*punctuation*’; then in ‘*prepositional phrase*’, ‘*word order*’, and ‘*incorrect transfer of a word class*’. The issue of *Syntactic-semantic correlativeness* is partially connected with the *Predicativeness* in the sense that the category of ‘*agreement within a noun phrase*’ in *Syntactic-semantic correlativeness* is connected with the category of ‘*agreement in person*’, ‘*number*’, ‘*gender*’ in *Predicativeness*. In both cases, nouns are paired with adjectives or other nouns to ensure agreement within a noun phrase (e.g., *English teacher/anglický učiteľ*) or to form the genitive construction (e.g., *a teacher of English/učiteľ angličtiny*). In both cases, the agreement within noun phrases must be considered. Clearly, this issue relates to other problems recorded in the graph (Figure 3), such as incorrect case or word order.

The use of commas and quotation marks – subject to different norms in English and Slovak – is an error-prone issue within the category of ‘*punctuation*’. In Slovak, direct speech is marked by a pair of low-high inverted commas (either single or double) (,,...“ or ,,...‘). In contrast, English uses common high-high marks (aligning with the top of capital letters) (‘...’ or “...”).

The improvement in machine translation between the SMT and NMT is evident even in this category. As shown in example (9), the quotation marks in the NMT output are used correctly:

(9) Example of an error in the category of ‘*punctuation/quotation marks*’:

ST: *The recent reading was described by some experts as “unimaginable”.*

HT: *Nedávne meranie niektorí odborníci označili za „nepredstaviteľné“.*

SMT: *Nedávne čítanie, popísaný niektorými expertmi ako "nepredstaviteľné". (incorrect)*

NMT: *Nedávne čítanie niektorí odborníci označili za „nepredstaviteľné“. (correct)*

Other punctuation errors occurred quite frequently following adjuncts in their initial position in sentences.

(10) Example of an error in the category of ‘*punctuation/commas*’:

ST: *At meeting in Bedford, local agricultural business leaders, residents and politicians spoke to home affairs committee MPs about immigration*

HT: *Na stretnutí v Bedforde hovorili lídri miestnych poľnohospodárskych podnikov, obyvatelia a politici s poslancami Výboru pre vnútorné záležitosti o imigrácii*

SMT: *Na stretnutí v Bedfordu, miestne poľnohospodárske Business Leaders, obyvatelia a politici prehovoril k poslancom vnútorné veci o imigrácii (incorrect)*

NMT: *Na stretnutí v Bedforde hovorili vedúci miestnych poľnohospodárskych podnikov, obyvatelia a politici s poslancami Výboru pre vnútorné záležitosti o imigrácii* (correct)

(11) Example of an error in the category of ‘incorrect case’:

ST: *The average selling price of iPads in the last quarter was the same as it was a year ago.*

HT: *Priemerná predajná cena iPadov v poslednom štvrtroku bola rovnaká ako pred rokom.*

SMT: *Priemerná predajná cena iPady v poslednom štvrtroku bol rovnaký ako to bolo pred rokom.* (incorrect)

NMT: *Priemerná predajná cena iPadov v poslednom štvrtroku bola rovnaká ako pred rokom.* (correct)

Numerous errors occurred in the category ‘agreement within a noun phrase’, like in the structure ... (have) so much local food (grown organically). In SMT output, the structure which is equivalent to the noun phrase ... (máme) *tol’ko miestne potraviny* (pestované organicky) does not take into account the fact that in Slovak, the expression *tol’ko* is followed by the genitive case rather than the nominative or accusative. NMT output used the structure with the correct case: ... (máme) *tol’ko* (ekologicky pestovaných) *potravín*, thus the output is comprehensible:

(12)

ST: *It is easy to boil noodles or rice, but they have almost no nutritional value and there is no need to eat imported food when we have so much local food grown organically on our islands.*

HT: *Je ľahké variť rezance alebo ryžu, ale nemajú takmer žiadnu výživovú hodnotu a nie je nutné jesť dovážané potraviny, keď máme na našich ostrovoch **tol’ko** ekologicky pestovaných **potravín**.*

SMT: *Je ľahké variť rezance alebo ryža, ale nemajú takmer žiadnu výživnú hodnotu a nie je nutné jesť dovážané potraviny, keď máme **tol’ko miestne potraviny** pestované organicky na našich ostrovoch.* (incorrect)

NMT: *Je ľahké variť rezance alebo ryžu, ale nemajú takmer žiadnu výživnú hodnotu a dovážané jedlo nie je potrebné jesť, keď máme na našich ostrovoch **tol’ko miestnych** ekologicky pestovaných **potravín**.* (correct)

(13) Example of an error in the category of ‘prepositions’:

ST: *He talked to Matthew and Danny, feels awful, and he feels terrible for the tournament.*

HT: *Rozprával sa s Matthewom a Dannym, cíti sa hrozne a cíti sa tak pre turnaj.*

SMT: *Hovoril Matúša a Danny, cíti strašné, a cíti hrozné na turnaj.* (incorrect)

NMT: *Hovoril s Matthewom a Dannym, je to strašné a pre turnaj sa cíti strašne.* (correct)

(14) Example of errors in the category of ‘nouns with prepositions’:

ST: *If the labour market was to tighten up we would struggle to fill those seasonal jobs with UK nationals.*

HT: *Ak by sa mal trh práce sprísniť, snažili by sme sa zaplniť tieto sezónne pracovné miesta občanmi Spojeného kráľovstva.*

SMT: *Ak sa na trhu práce mala sprísniť by sme sa snaží vyplniť tieto sezónne pracovné miesta s britských štátnych príslušníkov.* (incorrect when comparing SMT with HT, correct when accepting a possible more word-for-word translation *britskými štátnymi príslušníkmi*)

NMT: *Keby sa trh práce mal sprísniť, usilovali by sme sa obslúžiť tieto sezónne pracovné miesta u občanov Spojeného kráľovstva. (incorrect)*

(15) Example of errors in the category of ‘word order’:

ST: *There is no need to eat **imported food** when we have so much local food grown organically on our islands.*

HT: *Nie je nutné jesť **dovážané potraviny**, keď máme **na našich ostrovoch** toľko ekologicky pestovaných potravín.*

SMT: *Nie je nutné jesť **dovážané potraviny**, keď máme toľko miestne potraviny pestované organicky **na našich ostrovoch**. (imported food – correct, on our islands – incorrect)*

NMT: ***Dovážané jedlo** nie je potrebné jesť, keď máme **na našich ostrovoch** toľko ekologicky pestovaných potravín. (imported food – incorrect, on our islands – correct)*

(16) Example of an error in the category of ‘incorrect transfer of word class’:

ST: *Yvette Cooper, the committee’s **chair**, said she wanted to encourage people to talk frankly about immigration.*

HT: ***Predsedička** výboru Yvette Cooperová uviedla, že chce povzbudiť ľudí k tomu, aby o imigrácii hovorili na rovinu.*

SMT: *Yvette Cooper, **predseda** výboru, povedal, že chce, aby sa ľudia otvorene hovorili o imigrácii. (incorrect)*

NMT: ***Predsedička** výboru Yvette Cooperová povedala, že chce povzbudiť ľudí, aby úprimne hovorili o imigrácii. (correct)*

3.3 Compound/complex sentences

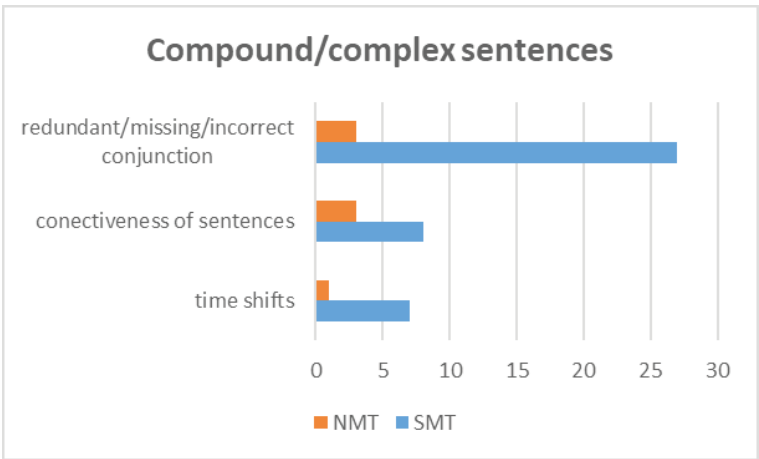


Figure 4: Results of Compound/complex sentences for SMT and NMT.

In the category of *Compound/complex sentences*, the results are like those in the previous category, with more errors recorded in SMT outputs than in NMT outputs (Figure 4). In the category of ‘*redundant/missing/incorrect conjunction*’

SMT scored 27 errors and NMT 3 errors, in the category of ‘connectiveness of sentences’, SMT numbered 8 errors and NMT 3 errors, and in the category of ‘time shifts’ in SMT 7 errors and in NMT only 1 error was identified (Figure 4). Overall, there were a few errors in the category of *Compound/complex sentences*, except in the subcategory of ‘*redundant/missing/incorrect conjunction*’.

One common issue arises from English prepositions, which often have multiple meanings and, consequently, several possible translations into Slovak. For example, the preposition *for* can be translated as *pre*, *ku*, *na*, *po*, or *za*, depending on the context. Selecting the correct equivalent from these options poses a significant challenge for machine translation.

Another issue involves the use of *that* in nominal clauses (which substitute an object or complement), functioning as a complementizer. In informal English, *that* is often omitted, resulting in a *zero that*-clause. Since machine translation tends to follow a more literal rather than free translation approach, if the source text omits *that*, the MT output does as well. This omission can disrupt the structure of the target segment and reduce the overall comprehensibility of the translation.

This issue is illustrated in example (17). Unlike SMT, NMT successfully recognised the omitted relativizer *that* (the category of ‘*redundant/missing/incorrect conjunction*’) and transferred the source sentence correctly:

(17)

ST: *Michel Cadot, the Paris police prefect, said the man had headed towards soldiers “armed with a machete”.*

HT: *Prefekt parížskej polície Michel Cadot povedal, že muž smeroval k vojakom „ozbrojený mačetou“.*

SMT: *Michel Cadot, Parížska polícia prefekt, povedal muž smeroval k vojakom „vyzbrojených mačetou“.* (incorrect)

NMT: *Prefekt parížskej polície Michel Cadot povedal, že muž smeruje k vojakom „ozbrojeným mačetou“.* (correct)

Obviously, other errors concerning different categories are also found here – for example, the use of the verb form *had headed*, which is correctly translated in SMT (*smeroval*) but not in NMT (*smeruje*), or the phrase *armed with a machete*, which is incorrectly rendered in both SMT and NMT (the correct form is *ozbrojený mačetou*). However, as explained above, the examples presented focus primarily on the highlighted category, which in this case is ‘*redundant/missing/incorrect conjunctions*’.

(18) Example of errors in the category of ‘connectiveness of sentences’:

ST: *It hopes individuals will come to the regional meetings to give their views on what approach the government should take to different kinds of migration – skilled to unskilled, students to refugees – and on what steps can be taken to manage the impact of migration in communities.*

- HT: *Dúfa, že jednotlivci prídu na regionálne stretnutia, aby vyjadrili svoje názory **na to, aký** prístup by mala vláda zaujať k rôznym druhom migrácie – od kvalifikovaných k nekvalifikovaným, od študentov k utečencom – **a na to, aké kroky môžu podniknúť** na zvládnutie dopadu migrácie na spoločnosť.*
- SMT: *Dúfa, že jednotlivci budú prichádzať do krajských stretnutí, aby vyjadrili svoj názor **na to, čo sa** blíži vláda by mala trvať na rôzne druhy migrácie – kvalifikovaných aby nekvalifikovaní, študentov, aby utečencov – **a aké kroky možno urobiť** v rámci konania o vplyve migrácie v komunitách. (**na to, čo** – incorrect connection followed by an incomprehensible structure (the issue is also related to the polysemous word what), **a aké kroky možno urobiť** – correct connection)*
- NMT: *Dúfa, že jednotlivci prídu na regionálne stretnutia, aby predniesli svoje názory **na to, aký** prístup by mala vláda zaujať k rôznym druhom migrácie – kvalifikovaným a nekvalifikovaným, študentom pre utečencov – **a na aké kroky sa môžu podniknúť** na zvládnutie dopadu migrácie na komunitu. (**na to, aký** – correct connection, **a na aké kroky možno urobiť** – correct connection, but followed by an incomprehensible part)*

(19) Example of errors in the category of ‘time shifts’:

- ST: *The French prime minister, Bernard Cazeneuve, **said it appeared** to be an “attack of terrorist nature”.*
- HT: *Francúzsky premiér Bernard Cazeneuve **povedal, že to vyzerá** na „teroristický útok“.*
- SMT: *Francúzsky premiér Bernard Cazeneuve, **že to vyzeralo** ako „útok teroristickej povahy“.* (incorrect)
- NMT: *Francúzsky premiér Bernard Cazeneuve **uviedol, že sa javil** ako „útok teroristickej povahy“.* (incorrect)

3.4 Lexical semantics

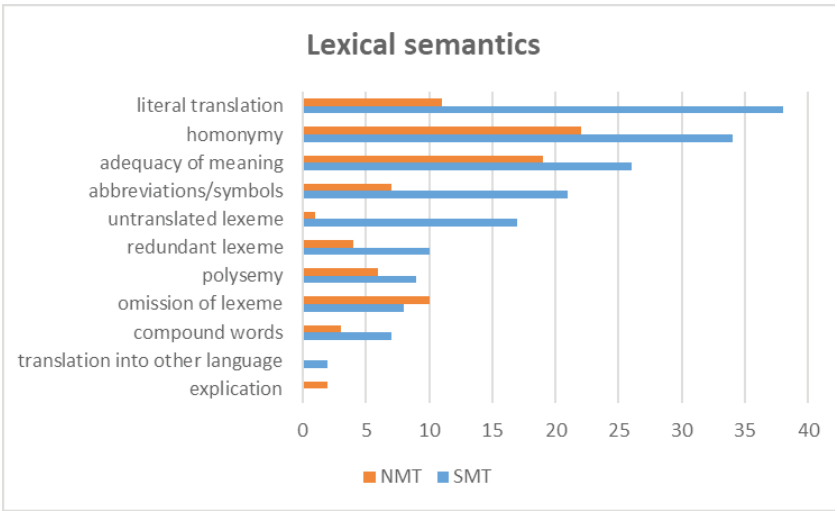


Figure 5: The results in Lexical semantics for SMT and NMT.

In Figure 5, the results from the category *Lexical semantics* for both SMT and NMT are summarized. The highest number of errors was recorded in the following categories: ‘*literal translation*’ (word-for-word translation) (SMT 38, NMT 11), ‘*homonymy*’ (SMT 34, NMT 22), ‘*adequacy of meaning*’ (SMT 26, NMT 19), ‘*abbreviations/symbols*’ (SMT 21, NMT 7), ‘*untranslated lexeme*’ (SMT 17, NMT 1) and ‘*redundant lexeme*’ (SMT 10, NMT 4). In the categories of ‘*literal translation*’ and ‘*untranslated lexeme*’, there is the most significant improvement in NMT, compared to SMT. The other categories reflecting the positive development of machine translation were the following: ‘*polysemy*’ (SMT 9, NMT 6), ‘*omission of lexeme*’ (SMT 8, NMT 10), ‘*compound words*’ (SMT 7, NMT 3), ‘*translation into other language*’ (SMT 2, NMT 0), and ‘*explication*’ (SMT 0, NMT 2) were the categories with a slight increase in errors.

The results in Figure 5 indicate that the most frequent errors in machine translation outputs related to lexical semantics are those connected to ‘*literal translation*’, ‘*homonymy*’, and ‘*adequacy of meaning*’. These errors stem from the typological differences between English and Slovak. Homonymy is characteristic for languages with a high number of non-derived words (e.g., English), whereas languages with a rich system of word derivation tend to have fewer homonymous words (e.g., Slovak) (see Ondruš – Sabol 1984, pp. 228–229). As a result, both machine and human translators often struggle to find appropriate equivalents for homonymous words, sometimes leading to inadequate translation solutions.

(20) Example of an error in the category of ‘*literal translation*’:

ST: *Still, serious doubts over the 41-year-old's longevity naturally remain.*

HT: *Stále sa však objavujú vážne obavy nad zotrvaním 41-ročného športovca.*

SMT: *Napriek tomu vážne pochybnosti nad 41-ročného dlhovekosti prirodzene zostávajú. (incorrect)*

NMT: *Stále však pretrvávajú vážne pochybnosti o dlhovekosti 41 rokov. (correct from the point of view of literal translation, but the expression o dlhovekosti 41 rokov has lower comprehensibility. The problem also concerns the meaning of the word longevity and the missing reference to its bearer – namely, the sportsman.)*

(21) Example of errors in the category of ‘*homonymy*’:

ST: *May is expected to stay only for the morning session and working lunch.*

HT: *Očakáva sa, že Mayová zostane iba na ranné zasadnutie a pracovný obed.*

SMT: *Môže sa očakáva, že bude len na dopoludňajšie a pracovným obed. (incorrect)*

NMT: *Očakáva sa, že máj zostane iba na ranné zasadnutie a pracovný obed. (incorrect)*

(22) Example of an error in the category of ‘*adequacy of meaning*’:

ST: *One soldier was slightly wounded, and another soldier fired back five shots.*

HT: *Jeden vojak bol ľahko zranený a ďalší vojak naspäť vystrelil päť rán.*

SMT: *Jeden vojak bol ľahko zranený a ďalší vojak vypálil päť striel. (correct)*

NMT: *Jeden vojak bol ľahko zranený a ďalší vojak vystrelil päť záberov. (incorrect)*

(23) Example of errors in the category of ‘abbreviations’:

- ST: *The estimated cost of decommissioning the plant and decontaminating the surrounding area had risen to 21.5**tn** yen (£150**bn**), nearly double an estimate released in 2016.*
- HT: *Odhadované náklady na vyradenie elektrárne z prevádzky a na dekontamináciu okolia sa zvýšili na 21,5 **biliónov** jenov (150 **miliárd** GBP), čo je takmer dvojnásobok odhadu zverejneného v roku 2016.*
- SMT: *Odhadované náklady na odstavenie elektrárne a dekontamináciu okolia stúplo na 21.5**tn** jenov (150 **miliárd** Sk), takmer dvojnásobným odhadu vydaná v roku 2016. (tn – tn untranslated abbreviation, bn – miliárd **correct**)*
- NMT: *Odhadované náklady na vyradenie elektrárne z prevádzky a na dekontamináciu okolia sa zvýšili na 21,5 **miliárd** jenov (150 **miliárd** GBP), čo je takmer dvojnásobok odhadu zverejneného v roku 2016. (tn – miliárd is **incorrect**, bn – miliárd is **correct**)*

(24) Example of an error in the category of ‘untranslated lexeme’:

- ST: *A **spokeswoman** for the Louvre said the museum was “closed for the moment”.*
- HT: ***Hovorkyňa** Louvru uviedla, že múzeum je momentálne zatvorené.*
- SMT: ***Spokewoman** pre Louvre povedal, že múzeum bolo „uzavretý pre túto chvíľu“. (**incorrect**)*
- NMT: ***Hovorkyňa** Louvru uviedla, že múzeum bolo „momentálne uzavreté“. (**correct**)*

(25) Example of errors in the category of ‘polysemy’:

- ST: *The **shooting** comes with France on its highest state of alert with thousands of troops patrolling the streets following a string of attacks in the last few years.*
- HT: ***Streľba** prichádza v čase, keď je Francúzsko v stave najvyššej pohotovosti, pričom po sérii útokov v posledných rokoch hliadkujú v uliciach tisíce vojakov.*
- SMT: ***Natáčanie** je dodávaný s Francúzskom na jeho najvyššom stave pohotovosti s tisíckami vojakov hliadkujúci v uliciach po reťazec útokov v posledných niekoľkých rokoch. (**incorrect**)*
- NMT: ***Natáčanie** prichádza s Francúzskom na najvyššom stupni pohotovosti, keď tisíce vojakov hliadkujú po uliciach po niekoľkých útokoch v posledných rokoch. (**incorrect**)*

(26) Example of an error in the category of ‘omission of lexeme’:

- ST: *Then there’s the red-headed **stepchild** of the Apple revenue streams.*
- HT: *Potom je tu ešte „ryšavé“ **nevlastné dieťa** zdrojov príjmov spoločnosti Apple.*
- SMT: *Potom je tu hrdzavý nevlastní zdroje príjmov Apple. (**incorrect**)*
- NMT: *Potom je tu **nevlastná dcéra** tokov výnosov spoločnosti Apple. (**correct**, maybe more adequate translation would be *nevlastné dieťa*)*

(27) Example of an error in the category of ‘compound words’:

- ST: *Paris **anti-terrorist** police are investigating after a man carrying a rucksack wounded one soldier with a knife.*
- HT: *Parížska **protiteroristická** polícia vyšetroje prípad, keď muž s batohom zranil vojaka nožom.*
- SMT: *Paríž **anti-teroristickej** polícia vyšetroje po tom, čo muž niesol batoh zranené jedného vojaka s nožom. (**incorrect**)*
- NMT: *Parížska **protiteroristická** polícia vyšetroje, keď muž, ktorý nosil ruksak, zranil jedného vojaka nožom. (**correct**)*

4. DISCUSSION AND CONCLUSIONS

Our goal was to analyse the developments and improvements of NMT compared to SMT by identifying and classifying the differences in their outputs. We focused on typological differences between the two languages – English, primarily an analytic language, and Slovak, primarily a synthetic language – within newspaper texts.

To evaluate these differences, we derived from Vaňko's (2017, pp. 83–100) categorical framework for evaluating MT errors, which provides a more detailed analysis of morpho-syntactic and syntactic-semantic relations, the areas where significant errors occur in the Slovak language. The quality of SMT and NMT outputs was analysed using a sample of newspaper articles from the British online newspaper *The Guardian*. Newspaper texts are frequently translated by MT systems due to their broad vocabulary and diverse range of topics.

Based on the results, NMT performed significantly better in terms of translation fluency, adequacy, and accuracy. Its output was more natural, clear, and comprehensible compared to SMT. However, NMT is not flawless – it still has certain shortcomings that require correction by a human post-editor. Numerous errors occurred in the category of *Syntactic-semantic correlativeness* (namely in the category of 'punctuation', 'incorrect case', 'agreement within a noun phrase' and in the category of *Lexical semantics* ('literal translation', 'homonymy', 'adequacy of meaning', 'abbreviations/symbols', and 'untranslated lexeme'). Surprisingly, the category of 'omission of lexeme' recorded more errors in NMT output than in SMT output. This finding supports previous research (see Bentivogli et al. 2016; Munková 2017, p. 23), which indicates that NMT outputs tend to be more fluent than those produced by phrase-based SMT. While NMT reduces literal translations and improves fluency, this often comes at the expense of adequacy. NMT sometimes produces unexpected translation solutions, generating sentences that appear fluent but lack semantic accuracy. These results are aligned with the findings of Petráš and Munková (2023, p. 87), which suggest that present NMT is more successful in translating longer and more complex syntactic structures. However, while the translation may appear fluent, its adequacy remains questionable (see Welnitzová 2023).

The detected errors in *Syntactic-semantic correlativeness* and *Lexical semantics* correspond to the typological differences between English and Slovak, as well as the nature of the examined texts. As newspaper texts are characterised by a high occurrence of nouns in various forms (e.g., nouns as pre-modifiers, post-modifiers, and within noun phrases), issues related to nouns are both frequent and significant. The category of nominalisation, then prepositional phrases after nouns, and attributive adjectives are common features of newspaper style, further contributing to translation challenges. Based on this, noun-related issues are both frequent and significant in terms of comprehensibility.

Despite persisting errors – primarily in *Syntactic-semantic correlativeness* and *Lexical semantics* – NMT has shown significant improvements. Unlike SMT, which uses a more word-by-word approach, NMT translates a sentence as a whole, leading to more fluent and comprehensible translations. However, challenges remain at the lexical level, particularly due to rich homonymy in English and the issues in adequacy. Other notable issues include literal translation and omission of lexemes, which continue to affect translation accuracy. These challenges present potential areas for future research and further advancements in machine translation.

In this context, Pondelíková and Luprichová (2024, p. 198) claim that AI-powered tools are not only transforming the way translations are performed but also reshaping the broader field of language and literacy. Moreover, the use of AI tools can significantly contribute to the improvement of users' linguistic and cognitive abilities, suggesting that AI can play an important role not only in education but also in translation practice. Absolon (2024, p. 9) predicts the decline in translation and localisation firm revenue losses due to the impact of AI, particularly the rise of large language models, which extend beyond translation services to other linguistic domains such as terminology management, research, and more.

References

- ABSOLON, Jakub (2024): Education for the Needs of the Post-Localisation Era: Transforming the Translation Industry in the AI Age. In: *Acta Anglica Tyrnaviensia*, Vol. 2, pp. 9–14.
- BEDNÁROVÁ-GIBOVÁ, Klaudia – GAVUROVÁ, Miroslava – GREŠTY, Jonathan (2024): *A Propaedeutics of Translation Studies*. Prešov: Vydavateľstvo Prešovskej univerzity. 194 p.
- BENKO, Ľubomír – MUNKOVÁ, Daša (2016): Application of POS Tagging in Machine Translation Evaluation. In: *DIVAI 2016: 11th International Scientific Conference on Distance Learning in Applied Informatics*. May 2–4, 2016, Štúrovo, Slovakia. Prague: Wolters Kluwer, pp. 471–489.
- BENTIVOGLI, Luisa – BISAZZA, Arianna – CETTOLO, Mauro – FEDERICO, Marcello (2016): Neural versus phrase-based machine translation quality: a case study. In: J. Su – K. Duh – X. Carreras (eds.): *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Austin: Association for Computational Linguistics, pp. 257–267.
- BESSENYEI, György (2017): Neural Machine Translation: The Rising Star. Memsource. Available at: www.memsource.com/blog/2024/09/19/neural-machine-translation-the-rising-star/?utm_source=mailchimp&utm_medium=email&utm_content=blog_article.
- BIBER, Douglas – CONRAD, Susan (2009): *Register, Genre, and Style*. Cambridge: Cambridge University Press. 344 p.

CAFFREY, Colm – VALENTINI, Cristina (2020): Technology applications at WIPO. In: M. O'Hagan (ed.): *The Routledge Handbook of Translation and Technology*. New York: Routledge – Taylor and Francis Group, pp. 127–147.

CASTILHO, Sheila – MOORKENS, Joss – GASPERI, Frederica – CALIXTO, Iacer – TINSELY, John – WAY, Andy (2017): Is neural machine translation the new state of the art? In: *The Prague Bulletin of Mathematical Linguistics*, Vol. 108, No. 1, pp. 109–120.

DIÑO, Gino (2018): *Korean Voice Assistant Highlights Tech's Insatiable Hunger for Language Data*. Slator. Available at: <https://slator.com/technology/korean-voice-assistant-highlights-techs-insatiable-hunger-for-language-data/>

DOLNÍK, Juraj (2013): *Všeobecná jazykoveda. Opis a vysvetľovanie jazyka*. Bratislava: Veda, vydavateľstvo SAV. 375 p.

DUŠKOVÁ, Libuše (2012): *Mluvnice súčasné angličtiny na pozadí češtiny*. Praha: Academia. 676 p.

FINDRA, Ján (2013): *Štylistika súčasnej slovenčiny*. Martin: Osveta. 320 p.

FONT LLITJOS, Ariadna – CARBONELL, G. Jaime – LAVIE, Alon (2005): A Framework for Interactive and Automatic Refinement of Transfer-based Machine Translation. In: *Proceedings of the 10th EAMT Conference: Practical applications of machine translation*. Budapest, Hungary: European Association for MT, pp. 87–96.

HLAVATÁ, Renáta – KRALČÁK, Ľubomír – GLOVNÁ, Juraj – RENDÁROVÁ, Mária – DUDOVÁ, Katarína (2019): *Jazyk ako text (vysokoškolský učebný materiál z praktickej štylistiky)*. Nitra: UKF. 174 p.

JEAN, Sébastien – FIRAT, Orhan – CHO, Kyunghyun – MEMISEVIC, Roland – BENGIO, Yoshua (2015): Montreal Neural Machine Translation Systems for WMT'15. In: O. Bojar – R. Chatterjee – Ch. Federmann – B. Haddow – Ch. Hokamp – M. Huck – V. Logacheva – P. Pecina (eds.): *Proceedings of the 10th Workshop on Statistical Machine Translation*. September 17–18, 2015 Lisboa. Portugal, Lisbon: Association for Computational Linguistics, pp. 134–140.

KOPONEN, Maarit (2016): Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. In: *The Journal of Specialised Translation*, Vol. 25, pp. 131–148.

KVETKO, Pavol (2021): *Translation Studies*. Trnava: UCM. 170 p.

LOMMEL, Arle (2018): Metrics for Translation Quality Assessment: A Case for Standardising Error Typologies. In: J. Moorkens – S. Castilho – F. Gaspari – S. Doherty (eds): *Translation Quality Assessment. Machine Translation: Technologies and Applications*. Springer: Cham, pp. 109–127.

LUONG, Thang – SUTSKEVER, Ilya – LE, Quoc – VINYALS, Oriol – ZAREMBA, Wojciech (2015): Addressing the rare word problem in neural machine translation. In: Ch. Zong – M. Strube (eds.): *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*. Beijing, China: Association for CL, Vol. 1, pp. 11–19.

MARTÍN, Juan Alberto Alonso – SERRA, Anna Civil (2014): Integration of a machine translation system into the editorial process flow of a daily newspaper. In: *Procesamiento del Lenguaje Natural*, Vol. 53, pp. 193–196.

MELBY, K. Alan (2020): Future of machine translation. Musings on Weaver's memo. In: M. O'Hagan (ed.): *The Routledge Handbook of Translation and Technology*. New York: Routledge – Taylor and Francis Group, pp. 419–436.

MISTRÍK, Jozef (1989): *Štylistika*. Bratislava: SPN.

MOORKENS, Joss – LEWIS, Dave (2020): Copyright and re-use of data. In: M. O'Hagan (ed.): *The Routledge Handbook of Translation and Technology*. New York: Routledge – Taylor and Francis Group, pp. 469–481.

MUNKOVÁ, Daša – MUNK, Michal (2016): *Evalvácia strojového prekladu*. Nitra: UKF. 173 p.

MUNKOVÁ, Daša (2017): Prekladateľské technológie a translatológia. In: D. Munková – J. Vaňko – J. Absolon – T. Bánik – Ľ. Benko – R. Machová – M. Munk – P. Petráš – K. Welnitzová (eds): *Mýliť sa je ľudské (ale aj strojové)*. Nitra: UKF, pp. 15–36.

ONDRUŠ, Šimon – SABOL, Ján (1984): *Úvod do štúdia jazykov*. Bratislava: SPN. 344 p.

ORAVEC, Ján – BAJZÍKOVÁ, Eugénia – FURDÍK, Juraj (1984): *Súčasný slovenský spisovný jazyk. Morfológia*. Bratislava: SPN. 232 p.

PETRÁŠ, Patrik – MUNKOVÁ, Daša (2023): Strojový preklad založený na neurónových sieťach – sľubná cesta prekladu z analytických jazykov do flektívnej slovenčiny? In: *Slovenská reč*, Vol. 88, No. 1, pp. 74–89.

PLITT, Mirko – MASSELOT, Francois (2010): A productivity test of statistical machine translation post-editing in a typical localisation context. In: *The Prague Bulletin of Mathematical Linguistics*, Vol. 93, pp. 7–16.

PONDELÍKOVÁ, Ivana – LUPRICHOVÁ, Jana (2024): AI-assisted enhancing of gender awareness through reading comprehension in history and literature courses of anglophone cultures. In: G. D. Sampso – D. Ifenthaler – P. Isaias (eds): *CELDA 2024: 21st International Conference on Cognition and Exploratory Learning in Digital Age: Proceedings*, pp. 198–208.

QUAH, Chiew Kin (2006): *Translation and Technology*. London: Palgrave Macmillan. 256 p.

SLANČOVÁ, Daniela – SPÁČILOVÁ, Stanislava (2022): Štylistické výskumy v časopise Slovenská reč. In: *Slovenská reč*, Vol. 87, No. 3, pp. 283–306.

TORAL, Antonio – SÁNCHEZ-CARTAGENA, Víctor (2017): A multifaceted evaluation of neural versus phrase-based machine translation for 9 language directions. In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Vol. 1, pp. 1063–1073.

VAJČKOVÁ, Mária (2023): Plusquamperfekt/Antepräteritum im Deutschen und im Slowakischen als Kohäsions- und Kohärenzmittel. (Pluperfect/Antepräterite in German and Slovak as a Means of Cohesion and Coherence). In: *Philologia*, Vol. 33, No. 2, pp. 7–22.

VAŇKO, Juraj – AUXOVÁ, Darina (2015): *Morfológia slovenského jazyka*. Nitra: UKF. 158 p.

VAŇKO, Juraj (2015): *Syntax slovenského jazyka*. Nitra: UKF. 150 p.

VAŇKO, Juraj (2017): Kategoriálny rámec pre analýzu chýb strojového prekladu. In: D. Munková – J. Vaňko – J. Absolon – T. Bánik – Ľ. Benko – R. Machová – M. Munk – P. Petráš – K. Welnitzová (eds): *Mýliť sa je ľudské (ale aj strojové)*. Nitra: UKF, pp. 83–100.

VILAR, David – XU, Jia – D'HARO, Luis Fernando – NEY, Hermann (2006): Error Analysis of Statistical Machine Translation Output, In: *Proceedings of LREC2006*. May 22–28, 2006, Genoa, Italy. Genoa: European Language Resources Association, pp. 697–702.

WANG, Haifeng – WU, Hua – HE Zhongjun – HUANG, Liang – CHURCH Ward, Kenneth (2022): Progress in Machine Translation. In: *Engineering*, Vol. 18, pp. 143–153.

WELNITZOVÁ, Katarína – MUNKOVÁ, Daša (2021): Sentence-structure errors of machine translation into Slovak. In: *Topics in Linguistics*, Vol. 22, No. 1, pp. 78–92.

WELNITZOVÁ, Katarína – MUNKOVÁ, Daša – WREDE, Oľga (2020): Chybovosť v predikativnosti a kvalita strojového prekladu. In: *Jazyk a kultúra*, Vol. 11, No. 41–42, pp. 160–172.

WELNITZOVÁ, Katarína (2023): *Chybovosť strojového prekladu*. Praha: Verbum. 130 p.

WELNITZOVÁ, Katarína (2024): Machine Translation and its Limits. In: *Acta Anglica Tyrnaviensia*, Vol. 1, No. 2, pp. 22–33.

R e s u m é

PRETRVÁVAJÚCE GRAMATICKÉ CHYBY STROJOVÉHO PREKLADU

Článok sa zaoberá problematikou gramatických chýb vo výstupoch strojového prekladu (SP) a hodnotí kvalitu dvoch prístupov: štatistického SP (SMT – Statistical Machine Translation) a neurónového SP (NMT – Neural Machine Translation). Výskum porovnáva kvalitu výstupov strojových prekladov v smere angličtina – slovenčina a zameriava sa na rozdiely medzi danými typologicky rozdielnymi jazykmi (angličtina klasifikovaná zväčša ako analytický jazyk a slovenčina ako syntetický jazyk).

Kvalitu SP sme hodnotili na vzorke publicistických textov vyexcerpovaných z britského denníka *The Guardian*. Texty sme dali preložiť strojovému prekladaču Google Translate pri dvoch rozdielnych prístupoch (SMT a NMT), ktoré následne posteditovali dvaja ľudskí prekladatelia v prostredí OSTEPERE (systém na hodnotenie, preklad a post-editáciu SP). Výstupy strojového prekladu analyzovali odborníci na slovenský jazyk, ktorí identifikovali a klasifikovali chyby v jednotlivých gramatických kategóriách.

Výsledky výskumu ukázali, že neurónový strojový preklad (NMT) možno považovať za oveľa kvalitnejší ako štatistický strojový preklad (SMT) v takmer všetkých hodnotených kategóriách. Preklad prostredníctvom NMT je plynulejší, prirodzenejší a zrozumiteľnejší. Pri preklade strojový prekladač pracuje s vetou ako s celkom a nie ako s jej časťami (frázami), ako to bolo v prípade SMT. Najvýraznejšie zlepšenie sme zaznamenali v kategórii predikativnosti, v ktorej NMT dosahuje výrazne lepšie výsledky napr. v zhode medzi podmetom a prísudkom, a to v osobe, čísle a rode. V kategórii predikativnosti sa znížila početnosť chýb aj pri nekorektnom používaní času a rodu.

Ďalšou významnou kategóriou, v ktorej NMT preukázal zlepšenie, je syntakticko-sémantická korelácia. V tejto oblasti sa znížil počet chýb v interpunkcii, zhode a väzbe a v neposlednom rade aj slovoslede. Typologické rozdiely medzi angličtinou a slovenčinou – najmä v oblasti syntaktickej flexibility a morfológických zmien – predstavujú pre strojový preklad výraznú výzvu. Anglický jazyk ako analytický jazyk definuje zväčša pevný slovosled, no pre slovenský jazyk je typický slovosled voľnejší. NMT dokáže tieto rozdielne charakteristiky

lepšie uplatňovať, a preto výstupy strojového prekladu pôsobia zrozumiteľnejšie a prirodzenejšie.

Problémy SP pretrvávajú v oblasti lexikálnej sémantiky, najmä pokiaľ ide o polysémiu, homonymiu a adekvátny transfer významu slova. Výsledky výskumu ukázali, že hoci sa pri NMT v porovnaní so SMT znížil počet tzv. doslovných prekladov, NMT nie je spoľahlivý pri výbere lexém v komplexnejších vetných štruktúrach, čo môže negatívne ovplyvniť celkový význam vety ako celku. Daný fakt potvrdzuje potrebu post-editácie výstupu SP ľudským prekladateľom alebo posteditorom, ktorý zabezpečí korekciu chýb a nepresností.

Napriek tomu, že neurónový strojový preklad (NMT) predstavuje v porovnaní so svojím predchodcom štatistickým strojovým prekladom (SMT) výrazné zlepšenie, gramatické chyby pretrvávajú. Výsledky výskumu zároveň potvrdzujú, že technologický pokrok v oblasti strojového prekladu neustále zlepšuje kvalitu strojového prekladu, a tým sa postupne zmenšuje rozdiel medzi kvalitou strojového a ľudského prekladu. Odborníci na SP vkladajú nádeje do vývoja hĺbkového učenia a umelej inteligencie, ktoré by mohli odstrániť aj v súčasnosti pretrvávajúce problémy a zvýšiť spoľahlivosť prekladov aj pri jazykovo rozdielnych jazykoch, akými sú angličtina a slovenčina.

ROZHLADY

PURISM AND NATIONALISM IN LANGUAGE NORMS: A COMPARATIVE ANALYSIS OF THE NORWEGIAN AND SLOVAK SPELLING REFORM ADOPTED DURING WORLD WAR II

TOMÁŠ BRATINA

Faculty of Arts, Masaryk University, Brno, Czech Republic

BRATINA, Tomáš: Purism and nationalism in language norms: A comparative analysis of the Norwegian and Slovak spelling reform adopted during World War II. *Jazykovedný časopis (Journal of Linguistics)*, 2025, Vol. 76, No. 2, pp. 493–508.

Abstract: The linguistic situation in interwar Norway and Czechoslovakia had several parallels. The official languages in both countries were represented by two written standards: Norwegian in the form of Bokmål and Nynorsk, and, in Czechoslovakia, Czech and Slovak – separate but closely related languages that were officially referred to as the Czechoslovak language. At the same time, both countries adopted spelling norms during this period to gradually unify the language variants. In Norway, this process took the form of a mutual approximation of Bokmål and Nynorsk, while in Czechoslovakia, it primarily involved bringing Slovak closer to Czech. Both the Norwegian spelling reform of 1938 and the Rules of Slovak Orthography of 1931 caused controversy. Moreover, during the Second World War, both Norway and the newly established Slovak Republic – at that time Nazi Germany's allies – introduced revised spelling rules: the Norwegian Spelling Reform of 1941 and the Rules of Slovak Orthography of 1940. This article focuses on these wartime reforms and explores the impact of political changes on language planning. Through comparative analysis, it seeks to identify the elements, approaches, or ideas that may connect the new rules despite the different linguistic natures of the analysed languages.

Keywords: orthography, spelling reform, purism, nationalism, language policy, Slovak, Norwegian

1. INTRODUCTION

Interwar Norway and Czechoslovakia were connected not only by diplomatic, trade, and cultural contacts¹ but also by notable parallels in their linguistic situations. An official language with two varieties was used in both countries. In Norway, it was the Norwegian language with two written standards, Bokmål 'Book language' and

¹ An important figure who linked Norway and Czechoslovakia was Bjørnstjerne Bjørnson, winner of the Nobel Prize in Literature. In 1907, Bjørnson spoke out strongly against the new Hungarian school law, which significantly restricted the teaching of Slovak pupils in their mother tongue. As a sign of gratitude, a street in Bratislava was named after Bjørnson in 1930.

Nynorsk ‘New Norwegian’ (until 1929, these standards were called Riksmål ‘National language’ and Landsmål ‘Country language’, respectively). In Czechoslovakia, the so-called Czechoslovak language, a political construct that had two versions, Czech and Slovak, was used at the official level. Both countries found themselves in this situation for different reasons.

In Norway, this situation arose as early as in the Middle Ages. As a result of the plague epidemic in the mid-14th century, Norway lost a significant part of its clergy and nobility, groups that were carriers of written culture. Thus weakened, Norway was unable to resist the influence of the Swedish and later the Danish language, not only culturally but also politically. Furthermore, Old Norse (Norwegian *norront*, less commonly *gammelnorsk*) gradually disappeared as a written language. During the 19th century, in the era of national revival, impulses for the development of a distinct Norwegian written language began to appear. This effort resulted in two forms of the Norwegian language: Landsmål, which was based on Norwegian dialects, and Riksmål, which implemented Norwegian elements into Danish. However, neither form could secure sufficient support; therefore, in 1885, the Storting, the Norwegian parliament, decided to adopt a compromise proposal. The decision, known as *jamstillingsvedtaket* (lit. ‘the equality decision’), equalised both forms and is still in force today.

In Czechoslovakia, the language situation had a political character. The concept of the newly established state was based on the idea of Czechoslovakism, and it stemmed from the notion of a Czechoslovak nation-state. Furthermore, the state-forming Czechoslovak nation became significantly predominant only in this conception in a country where more than three million Germans, over half a million Hungarians, and Rusyns, Jews, Poles, and Roma lived. The constitution of the new state established that the state’s official language would be the Czechoslovak language, used in two variants, namely Czech and Slovak. However, the Czechoslovak language did not exist (Švagrovský 2006, p. 333). This concept was based on the need for an internally strong state (Kováč 2011, p. 181).

The coexistence of two forms of the official language ultimately led to attempts to bring them closer together in both countries. In Norway, these efforts had a mutual character, meaning that both forms were supposed to gradually converge by incorporating elements from the opposite form or by eliminating some of their characteristic features. On the other hand, in Czechoslovakia, it was primarily Slovak that was supposed to approach Czech.² The reforms that sought to realise this goal – the Norwegian spelling reform of 1917 (*Den nye rettskrivning: Regler og Ordliste* 1918), but especially of 1938 (*Ny rettskrivning 1938: Bokmål. Regler og ordliste*

² For completeness, it should be added that Slovakisms also penetrated into Czech (Nábělková 2017); however, compared to the changes that occurred in Slovak, it was disproportionate. Therefore, this paper focuses exclusively on Slovak.

1938; *Ny rettskriving 1938: Nynorsk. Reglar og ordliste* 1938), and the Rules of Slovak Orthography of 1931 (*Pravidlá slovenského pravopisu s abecedným pravopisným slovníkom* 1931) – caused controversy.

The new, revised orthographic norms of both Slovak and Norwegian were finally adopted during the Second World War, at a time when governments sympathetic to Nazi Germany seized power in both countries, namely Norway and the newly established Slovak Republic. In Norway, Vidkun Quisling and his fascist party *Nasjonal Samling* (lit. ‘National Gathering’) ruled, while in Slovakia, Jozef Tiso and the far-right party *Hlinkova slovenská ľudová strana* (lit. ‘Hlinka’s Slovak People’s Party’) was in power.

These parallels make it possible to compare these “wartime” reforms, namely the Norwegian spelling reform of 1941 (*Ny rettskrivning: Regler for bokmål. Regler for nynorsk. Fellesregler* 1941) and the Rules of Slovak Orthography of 1940 (*Pravidlá slovenského pravopisu s pravopisným slovníkom* 1940).³ In this paper, I aim to consider whether common ideas, approaches, or goals may connect these reforms at a higher level despite their different natures: Norwegian as a North Germanic language and Slovak as a West Slavic language.

I will outline a comparative analysis of the reforms from two perspectives. First, I will focus on the processes involved in shaping the new rules, identify the individual responsible for their preparation, and explore the declared goal. Second, I will attend to the content of the reforms themselves. In this regard, I am interested in the intention behind the implemented changes. I will then compare the findings and identify the similarities and differences between the analysed reforms. I will also reflect upon the impact of the reforms on the subsequent development of the language.

2. THE NORWEGIAN SPELLING REFORM OF 1941

2.1 Background

In 1885, the Norwegian parliament, the Storting, adopted a resolution on the equalisation of the two written standards of Norwegian, namely *Landsmål* – which was based on dialects and, since 1929, has been called *Nynorsk* – and *Riksmål* – which implemented Norwegian elements into Danish and, since 1929, has been called *Bokmål*. Over time, however, there were debates regarding whether it would be more practical to bring both varieties closer together so that they would merge into one in the future. The spelling reform of 1938 was significant, yet controversial, in this regard. A part of society viewed the reform critically. When Quisling’s government came to power in occupied Norway, therefore, the government decided not to accept the rules

³ In line with their titles, these normative manuals are referred to here as spelling or orthographic rules. However, it should be noted that their scope extended beyond orthography, since the proposed and adopted changes also affected other levels of language, especially morphology and lexicon.

and to prepare a new spelling in an attempt to gain public sympathy (Tjelle 1994, p. 26).

Quisling's propaganda derogatorily referred to the 1938 spelling as *det kohtske knot*, loosely translated as 'Koht's mishmash', referring to Halvdan Koht, the pre-war Norwegian Minister of Foreign Affairs who was the driving force behind the spelling changes. Additionally, the propaganda depicted the spelling as a manifestation of the "dictatorship of the proletariat", as it included *folkemålsformer*, forms characteristic of dialects, in the rules (Torp – Dahl – Lundeby 1991, p. 131).⁴ However, in general, the occupation government was not sceptical about the gradual convergence of the written standards, Bokmål and Nynorsk; rather, it objected to the way this idea was realised in 1938. Eventually, the merger of the forms was part of Quisling's party's programme. Their language paragraph stated that "the natural merging of two language forms into one written language is promoted, but without coercion" (Lunde 1942, p. 172). The merger of language forms was also in line with the party's ideological goal, namely to unite and harmonise society as a whole, including cultural life (Tjelle 1994, p. 102).

However, it is necessary to note that *Nasjonal Samling*, as a small nationalist party, did not focus significantly on the language issue.⁵ Even the intense debate about the future of the Norwegian language, which took place in society in the 1930s, did not gain substantial space in the party newspaper *Fritt Folk* 'Free People' (Tjelle 1994, p. 102). However, the disinterest in language issues did not apply to all party members. An important figure in this regard was the Minister of Culture in the Quisling government, Gulbrand Lunde, a linguistic idealist responsible for drafting the new spelling reform.

When Lunde presented the new spelling rules at a press conference on 26 November 1941, he argued with a historical-national ideology (Tjelle 1994, p. 106). The spelling reform aimed to create *høynorsk høvisk mål*, loosely translated as 'a noble Norwegian language'.⁶ In other words, instead of convergence through the implementation of dialectal elements, fusion was to be based on a cultivated language. In practice, this idea meant that if there were doubts about a word, the focus should be on Old Norse to the greatest extent possible. This principle affected

⁴ For example, in Bokmål, the past tense forms *glidde* 'slid', *klyvde* 'climbed', and *nyste* 'sneezed' were considered vulgar and unrefined. These forms, which followed the standardised past tense endings *-dde*, *-de*, and *-te*, stood in contrast to the traditional irregular forms *gled*, *kløv* and *nøs*.

⁵ The fact that the party did not take a significant interest in the language before the war may be evidenced by the ambiguity and vagueness of the paragraph. Tjelle (1994, p. 102) highlights that the sentence contradicts itself in its essence. As long as there are people who oppose the convergence of forms, it is impossible to promote natural merging without coercion.

⁶ This concept is not easy to define, as it is not known what exactly Lunde meant by *høynorsk* (lit. 'High Norwegian'). Tjelle (1994, p. 113) believes that the term referred to the language used on formal occasions in Norway in the Middle Ages, later replaced by Danish.

Bokmål to a much greater extent, as Nynorsk, which was based on Norwegian dialects, already contained Old Norse elements before the reform.⁷ Nynorsk, thus, functioned de facto as a *reservoir* of cultivated words that Bokmål could adopt (Tjelle 1994, p. 113; Vikør 2022).

It was not easy to appoint a commission to prepare the spelling changes. Most of the experts who were approached refused the offer; therefore, the reform was ultimately decided by people without the necessary expertise and education. In addition to Minister Gulbrand Lunde, who had the final say, the commission included Rolf R. Nygaard, a university student of the Norwegian language; Johan Fr. Voss, a teacher; Sigvat Heggstad, a civil servant; and Johannes K. Norvik. The latter was the only one to link the preparation of the rules with *the new era* and called for strengthening relations with the German language (Tjelle 1994, p. 133).⁸ The commission, lacking proper linguistic authority, prepared the new rules rapidly and hurriedly. The first edition was published in the autumn of 1941, a year and a half after Quisling's government took office.

The new spelling rules began to be used in newspapers, official communication, and translation literature in mid-1942 (Ramsfjell – Vinje 1978, p. 47).⁹ The reform was also supposed to be introduced in schools, but this intention remained unrealised for several reasons. Apart from the shortage of paper, there was also insufficient support and sabotage from teachers. Moreover, there remained a considerable number of recently printed textbooks with the spelling of 1938 in storage. Therefore, publishers lacked motivation and economic resources to reprint books (Vikør 2022). In summary, most people came into contact with the new rules through the press.

2.2 Content

The spelling reform of 1941 was intended to be another step towards unifying the two written standards of the Norwegian language. However, in practice, there was no significant progress in this direction, and Guttu (2017) has even argued that the merger slowed down.

The lack of progress was due to the inconsistency and incoherence of the changes, especially on the Bokmål side. This irregularity is most visible in morphology. The reform, in line with convergence efforts, increased the number of feminine nouns that end in *-a* in the singular definite form, as in Nynorsk. However, in connection with verbs, the reform went in the opposite direction. Regular verbs,

⁷ For example, unlike Bokmål, Nynorsk has retained the Old Norse system of conjugating irregular verbs (Nesse 2013, p. 51). The present tense is formed with one syllable, not two as is common in Bokmål. For example, compare *eg skriv* (Nynorsk) and *jeg skriver* (Bokmål), in English *I write*.

⁸ For example, Norvik advocated against the past tense ending of weak verbs *-et*, and (unsuccessfully) proposed that the ending *-te*, used in German, should be adopted (Tjelle 1994, p. 124).

⁹ However, the obligation did not apply to Norwegian writers. They could write as they pleased, and many boycotted the reform (Otnes – Aamotsbakken 2006, p. 164).

which, according to the spelling rules of 1938, could end in the past tense either in *-a* (a form typical of Nynorsk) or *-et*, have exclusively had the ending *-et* since 1941.¹⁰

Greater regularity and more consistent efforts towards unification were found in phonology. The reform built on previous rules and, at the same time, codified new forms in Bokmål inspired by Old Norse, such as the diphthongs *au* [æu], *øy* [œy], and *ei* [æi], as well as forms like *gammal* ['gam:al] 'old' (instead of *gammel* ['gam:(e)l]) and *fram* [fram:] 'forward' (instead of *frem* [frem:]). Aligning with this principle, forms with *o* instead of *u* were also introduced, for example, in the word *no* [nu:] 'now' (instead of *nu* [nu:]), and, nowadays, *nå* [no:]). This change is often cited as the most characteristic feature of the spelling reform of 1941 (cf. Vinje 1978; Torp – Dahl – Lundebø 1991; Nesse 2013).¹¹

Nynorsk, rather than merging with Bokmål, was supposed to return to its traditional forms, which were abolished in 1938. These forms included, for instance, the ending *-i* in the singular definite form of feminine nouns (e.g. *bygdi* 'the village') and in the plural definite form of neuter nouns (e.g. *ordi* 'the words'). The abolition of these forms, which had a strong tradition in literary and religious texts, was considered by Lunde to be almost a crime (Tjelle 1994, p. 110).

As mentioned above, Nynorsk was closer to the ideal of "*høynorsk*", as it contained several Old Norse elements. Therefore, it is unsurprising that Bokmål was primarily supposed to approach this concept, although the extent to which the new rules succeeded in this regard may be debated.

A comparison of the reforms of 1938 and 1941 on an ideological level reveals at least a difference in argumentation. Instead of the social and democratic argumentation that characterised the previous reform of 1938, Quisling's spelling rules apply a more national-historical conception, although inconsistently. In some places, aesthetic principles take precedence, for example, when the ending *-a* in the plural definite form of masculine nouns (e.g. *gutta* 'the boys' or *hesta* 'the horses') was not retained in Bokmål, unlike in Nynorsk, due to its "vulgar nature" (Tjelle 1994, p. 122).¹² Therefore, in connection with the rules of 1941, it is not possible to speak of a unified and coherent language ideology (Tjelle 1994, p. 134).

¹⁰ Lunde realised that the form with *-a* had a future, but, at the same time, he believed that its implementation was too premature. The form with *-et* was, therefore, intended to be transitional until the next spelling reform was implemented (Tjelle 1994, p. 111).

¹¹ The form *no* was never used in Bokmål before 1941, and it is necessary to note that it was not used after 1945, when the reform became invalid. Therefore, if the form *no* appears in a text written in Bokmål, there is a high probability that the document dates from the Second World War period (Torp – Dahl – Lundebø 1991, p. 131).

¹² Rolf R. Nygaard, a member of the commission that drafted the spelling reform of 1941, described the ending *-a* in the masculine definite form as "one of the worst vulgarisms in the language" and, in a letter to the Ministry of Culture and Enlightenment, insisted that this form must be excluded from the written language at all costs, warning that its retention would compromise the purity of the language (Tjelle 1994, p. 122).

3. THE RULES OF SLOVAK ORTHOGRAPHY OF 1940

3.1 Background

After the First World War, the standard Slovak language form known as the *martinský úzus* (lit. ‘Martin usage’, named after the city of Turčiansky svätý Martin) persisted in official communication, in literature, and in other styles of public communication (Krajčovič – Žigo 2004, p. 182).¹³ This form was shaped under the influence of the cultural centre in Turčiansky svätý Martin at the end of the 19th century. It represented a relatively stabilised form of standard Slovak (Krajčovič – Žigo 2011, p. 211). However, after the establishment of Czechoslovakia in 1918, this language form came under pressure. The insufficient number of Slovak-speaking intelligentsia capable of leading state administration or education led to this role being predominantly assumed by Czech people. In this context, Slovak was enriched with new necessary terms from the Czech language, but, at the same time, the norm of the language began to fluctuate due to the unregulated adoption of Czech words. The infiltration of the Slovak language by Czech elements affected all linguistic levels, including phonology, morphology, syntax, and, especially, vocabulary (Kačala – Krajčovič 2011, p. 124). Therefore, the language situation in Slovakia began to resemble linguistic chaos (Kačala – Krajčovič 2011, p. 124).¹⁴

Furthermore, the concept of the newly established state was based on the idea of Czechoslovakism, presenting a unified Czechoslovak nation and language, used in two varieties, namely Czech and Slovak.¹⁵ In this spirit, Slovak was to be reformed to get closer to Czech (Kopecká 2019, p. 79) and, potentially, merge with it in the future (Hrancová 2019, p. 78). The first step towards realising the thesis of a unified Czechoslovak language was the reform named the Rules of Slovak Orthography (*Pravidlá slovenského pravopisu s abecedným pravopisným slovníkom* 1931), published in 1931 under the leadership of the Czech linguist Václav Vážný. The Rules aimed to achieve Czechoslovak linguistic unity by “suppressing unique features of standard Slovak and incorporating Bohemisms into the standard norm” (Ružička 1970, p. 62).

¹³ Turčiansky svätý Martin (today’s name Martin) was the centre of Slovak culture and the seat of several editorial offices. According to Jóna (1973, p. 25), the *martinský úzus* was based on the speech of the inhabitants of Martin in the 1860s and 1870s.

¹⁴ It should be noted that interaction between Slovak and Czech predates the establishment of Czechoslovakia. Various forms of linguistic blending existed, including the so-called “slovakised Czech” (see Kesselová – Slančová 2010). In addition, Slovak scholars frequently used biblical Czech as a literary language in the 18th and 19th centuries. However, the creation of Czechoslovakia in 1918 marked a turning point, as this interaction was redefined by the ideology of Czechoslovakism and by institutional efforts to bring Slovak closer to Czech, as discussed in the main text.

¹⁵ Some later textbooks also define the standard language in this way. For instance, Stanislav (1938, p. 177) in his *Československá mluvnica* ‘Czechoslovak grammar’ stated that “by the Czechoslovak language, we understand a distinct linguistic group within the Western Slavic languages, encompassing all dialects from Cheb to Uzhhorod, with two written standards – Czech and Slovak”.

After the Rules were issued, the Slovak public was divided into two camps. While the Czechoslovak centralist press, authorities, and schools accepted them as the norm, the autonomist press and the *Matica slovenská* (the oldest Slovak national, cultural, and scientific organisation) did not accept most of the rules in their publications (Hrancová 2019, p. 80). A year later, at the general assembly of the *Matica slovenská*, the Rules of 1931 were rejected precisely due to their convergence tendencies towards the Czech language. Furthermore, the authorities ordered the preparation of a completely new norm, “which would primarily meet national-representative, cultural-representative, and national-integrative criteria and be a sufficient reflection of the distinctive features between Slovak and Czech” (Žigo 1997, p. 183).

In response to this situation, *Matica slovenská* launched *Slovenská reč* ‘Slovak speech’ in 1932 – the first Slovak linguistically oriented journal. The journal actively promoted the principles of the *martinský úzus* and aimed to stabilise the standard by advocating domestic forms and opposing excessive Czech influence (Krajčovič 2004, p. 325).¹⁶ Its chief editor was Henrich Bartek, representative of the Language Department of *Matica slovenská* and a supporter of the movement for the distinctiveness and purity of standard Slovak. Bartek not only shaped the editorial direction of the journal but also made the main contribution to the nearly seven-year preparation and overall implementation of the new proposal of the Rules.

The proposal followed the tradition of the *martinský úzus*, and linguistic correctness was sought both in earlier historical stages of Slovak and in dialects, which were considered a legitimate internal source for enriching the language’s vocabulary, particularly as an alternative to borrowing foreign words (Peciar 1950, p. 258). In addition, the proposal reflected the perspective of the Prague linguistic circle, which emphasised a synchronic approach to the written variety of the national language and the stability and functionality of its linguistic means (Kopecká 2019, p. 93).

The new Rules of Slovak Orthography proposal was submitted by the *Matica slovenská* to the Minister of Education on 22 March 1939, eight days after the establishment of the independent Slovak Republic. Despite expectations that the approval process would be spontaneous in the new political and linguistic situation, the proposal was not approved by the ministry and was sent back for revision. A university commission composed of philologists and historians working at the Slovak University (originally and currently Comenius University) classified the

¹⁶ In the interwar period, *Slovenská reč* published both analytical studies and articles on various aspects of the Slovak language. However, a substantial portion of its content consisted of short purist contributions intended to stabilise the standard language. According to Jánsky (1940), nearly half of the 3,134 such contributions published during the journal’s first seven years focused on expressions of Czech origin (46.5%), while considerably fewer addressed German loanwords (8.2%) and words of Hungarian origin (3%).

proposal as overly purist and insufficiently scientifically prepared (Krajčovič – Žigo 2011, p. 230). The subjects of dispute were lexical borrowings from Czech, as well as the spelling changes concerning the unified writing of the verb ending *-li* in the past tense¹⁷; the phonetic spelling of the prefix *s-/z-/zo-*; and the binding of the prepositions *z, zo* (of, from) with the genitive and *s, so* (with) with the instrumental (Kopecká 2019, p. 96). The revised Rules were published in 1940.

3.2 Content

The Rules of Slovak Orthography of 1940 respected Bartek's proposal to eliminate most of the Bohemisms. However, the 1940's spelling reform refused to make some Bartek's adjustments that would have significantly transformed Slovak spelling, potentially moving it further away from Czech. For example, the previously mentioned unified writing of the suffix *-li* in the plural form of the past tense or the writing of the prefixes *s-*, *z-*, *so-*, and *zo-* following pronunciation rather than the principle of meaning (direction from top to bottom, from one place to another, etc.) were among the adjustments that were ultimately rejected. This stance was justified, among other reasons, by Škultéty's views on the Slavic context of Slovak spelling (Jóna 1963, p. 248).

The revised Rules, thus, represented a compromise: While a radical spelling reform proposed in 1939 was rejected, as were the *extreme* demands of anti-Czech purism¹⁸ (Blanár – Jóna – Ružička 1974, p. 209), most of the demands of the autonomists were accepted. For example, Czech doublets were removed, with a few exceptions, and the introduction to the Rules (*Pravidlá slovenského pravopisu s pravopisným slovníkom* 1940, p. 3) stated that "the main aim was to eliminate various foreign elements from Slovak spelling that had entered it due to undesirable extralinguistic influences". In connection with the Rules of 1940, this revision is, therefore, more of a return from a unifying, Czech-oriented position to a position represented by the previous Martin period (Jarošová 2012, p. 257) than an attempt at a more significant transformation of the Slovak language.

The spelling reform also eliminated the split in language practice of the 1930s, when one part of the public respected the 1931 Rules in communication and the other part adhered to the recommendations of the *Matica slovenská*.

¹⁷ Until then, a distinction was made between forms *-li* for the masculine animate plural and *-ly* for the masculine inanimate, feminine, and neuter plural; for example, *muži spievali*, 'the men were singing', but *ženy spievaly*, 'the women were singing'.

¹⁸ Ondrejovič (2003, p. 11) and Krajčovič – Žigo (2011, p. 223) describe cases of extreme purism as instances where the decisive, and sometimes the only argument for rejecting a certain term, was its Czech origin or its existence in the Czech language. Kopecká (2019, p. 81) cites *mrazivo* as an example proposed by purists to be used instead of the (now widespread) word *zmrzlina* (in English *ice cream*), borrowed from Czech.

4. THE NORWEGIAN AND SLOVAK SPELLING REFORMS IN COMPARISON

4.1 What do the reforms have in common?

Several parallels can be observed between the Norwegian spelling reform of 1938 and the Rules of Slovak Orthography of 1931. As mentioned earlier, these rules aimed to merge language varieties into one. On the Norwegian side, the goal was to bring Bokmål and Nynorsk closer to each other, while on the Czechoslovak side, the goal was to bring Slovak closer to Czech. However, both reforms encountered resistance, and the question of preparing new rules gradually became more pressing. In both cases, this pressure ultimately led to the adoption of revised spelling reforms during the Second World War: in Slovakia in 1940 and in Norway in 1941.

The aims of the wartime reforms reveal that they are connected by the effort to establish or return to what is “native” and “authentic” for the language, depending on how the authors of the reforms themselves understood and defined this phenomenon. In the case of the Norwegian language, it was a return to forms typical of Old Norse. In Slovak, this effort involved the removal of “various foreign elements” (*Pravidlá slovenského pravopisu s pravopisným slovníkom* 1940, p. 3), meaning Bohemisms, and a return to usage that prevailed before intensified language contact with Czech and the implementation of the idea of Czechoslovakism.

Considering the period in which the new spelling rules were adopted, the question arises as to the extent to which these efforts can be attributed to political changes. Both the *Nasjonal Samling*, which came to power in Norway, and Lunde, who was responsible for the spelling reform, looked back at the Old Norse period as a great national epoch (Tjelle 1994, p. 99).¹⁹ The national-historical ideology that prevailed in the creation of the new spelling represents a change from the social and democratic approach that characterised the 1938 reform. In this sense, there is room to admit that the nationalist ideology of the new regime influenced the goal of the reform. A similar development is observed on the Slovak side as well. The ideology of the newly formed Slovak Republic, which emphasised national – that is, Slovak – identity, aligned with the effort to return from the Czech-oriented position to the position represented by the previous Martin codification (Jarošová 2012, p. 257).

While the political environment may have influenced language strategies in both cases, crucially, it was not the sole determining factor. Slovakia provides an example. The commission responsible for deciding on the new spelling had an even more puristic alternative on the table, developed during the First Czechoslovak Republic (i. e., the Bartek’s proposal from 1939), which they rejected. Apart from concerns about “insufficient scientific preparation” (Krajčovič – Žigo 2011, p. 230),

¹⁹ For example, within the *Nasjonal Samling* party, the Old Norse phrase *heil og sæl* (lit. ‘happy and healthy’) was commonly used as a greeting (Hamre 2019, p. 45).

personal animosities between the proponent, Henrich Bartek, and the evaluator, Ján Stanislav (Švagrovský – Žigo 1994, p. 99), as well as the effort to preserve the Slavic continuum (Jóna 1963, p. 248), also played a role. Factors other than national-historical criteria also entered into the process of creating the Norwegian reform. An example is the aesthetic factor that prevailed in Nynorsk regarding the past tense forms of the auxiliary verbs *kunne* ‘could’, *skulle* ‘should’, and *ville* ‘would/wanted to’. These verbs were proposed to be written with *d* – *kunde*, *skulde*, and *vilde*, as they appeared more dignified in this form (Tjelle 1994, p. 111). These facts indicate that, alongside political changes, a number of other factors influenced the development of new spelling rules. Notably, this is the second point where the analysed reforms intersect.

4.2 What makes the reforms different?

Aside from the parallels between the analysed spelling reforms, their differences are also noteworthy. There were differences on several levels, from the goals of the reform through their preparation and implementation to their impact on further development.

As I hinted at in the previous section, despite the common effort to reconnect with what is “native” and “authentic” in both languages, the new spelling rules differed in the question of the gradual convergence of standard forms. This goal was understandably irrelevant in Slovak due to new circumstances because the political concept of a unified Czechoslovak language lost its relevance with the establishment of the independent Slovak Republic. Nevertheless, the Norwegian reform declared an interest in continuing efforts to merge Bokmål and Nynorsk, although it can be debated as to what extent this goal was achieved. At the very least, in the declarative sense, this divergence may be seen as a notable contradiction.

It is also important to reiterate that the starting points for preparing new rules were substantially different. While work on the Slovak reform began as early as during the First Czechoslovak Republic in the 1930s, the Norwegian reform was exclusively developed in the atmosphere of the World War and haste. The first edition was published in the autumn of 1941, a year and a half after Quisling’s government came to power. Perhaps it is precisely from here, in combination with an unprofessional commission, that the inconsistency and incoherence of the Norwegian reform arise. To clarify, most of the work in preparing the new rules was done by a 22-year-old Norwegian language student, Rolf R. Nygaard, who had no prior experience with this type of work (Tjelle 1994, p. 131). In contrast, in Slovakia, a commission comprising philologists and historians working at the Slovak University (formerly and currently Comenius University) was formed to assess the proposal of Bartek. In other words, unlike Norway, the Slovak commission did not have to deal with questions regarding its authority and quality.

Lastly, it is crucial to highlight how the new spelling reforms were received by the broader public. The Norwegian reform faced sabotage, not only from teachers but also from publishers. Similarly, in the Slovak Republic, some people disagreed with the spelling reform of 1940, such as Henrich Bartek, the author of the unsuccessful proposal from 1939. However, the intensity of their disagreement did not reach the same levels as in Norway.

The combination of these points may partly explain why the reforms fundamentally diverge in terms of their impact on further language development.

4.3 Impact on further language development

After the Second World War, Quisling's spelling was abolished, and Norway reverted to the rules of 1938. Otnes and Aamotsbakken (2006, p. 164) stated that the 1941 reform left no traces, and Tjelle (1994, p. 69) added that nothing would have changed even if experts had been involved in its preparation. The biggest challenge with the spelling was the lack of legitimacy as a product of the *Nasjonal Samling* (Tjelle 1994, p. 69).

In contrast, the Slovak spelling reform adopted in 1940 continued to serve as a norm even after the Second World War and the restoration of Czechoslovakia until 1953. The decision to replace the spelling of 1940 was not driven by its origin during the Slovak Republic (1939–1945), but rather by its practical inefficiency (Krajčovič – Žigo 2011, p. 227). The reason was the complex spelling principle, which relied on various semantic criteria and served as a compromise between Bartek's initial proposal and the demands of the university commission, such as direction from top to bottom, from one place to another, etc. (Švagrovský – Žigo 1994, p. 103). In line with structuralist methodology, there was a need to evaluate the existing spelling norm and prepare a new codification, which was published in 1953.

In the linguistic journal *Naše řeč* 'Our speech', Czech philologist Váhala (1954, p. 34) stated that the spelling of 1940 was marked by "reactionary bourgeois nationalism", manifesting in language policy as "chauvinistically nationalist purism". These words, however, must be understood in the context of the regime at that time. While it is undeniable that the Rules of 1940 eliminated most of the Bohemisms, as I have pointed out, not all purist demands were incorporated. Paradoxically, in this regard, the new reform of 1953 (*Pravidlá slovenského pravopisu s pravopisným a gramatickým slovníkom* 1953) was more consistent, and it almost entirely respected Bartek's proposal from 1939 (Švagrovský – Žigo 1994, p. 103).²⁰ However, for ideological reasons, this step could not be publicly declared (Kopecká 2019, p. 99).

²⁰ More precisely, the 1953 reform adopted Bartek's proposed orthographic changes regarding the past tense suffix *-li* and the use of *s-/so-/z-/zo-*. However, his purist stance concerning lexicon, especially the rejection of Bohemisms, was not fully incorporated.

In terms of long-term impact, while the Norwegian spelling reform of 1941 represents a dead end in the development of Norwegian, Kačala and Krajčovič (2011, p. 167) describe the period after 1940 as “the period of the most intensive development and the greatest boom” in the history of standard Slovak.

Ultimately, the differing trajectories of the two wartime reforms seem to hinge not only on their linguistic features or their degree of alignment with natural language usage, but also on their symbolic legitimacy within the socio-political context in which they were implemented. The Slovak reform of 1940, although adopted during an authoritarian regime, was perceived as a continuation of earlier national codification efforts initiated in the interwar period. It was based on work carried out by established linguistic institutions such as the *Matica slovenská* and represented a compromise, whereas a more radical reform of spelling had been rejected. In contrast, the Norwegian reform of 1941 was closely associated with the collaborationist Quisling regime. Its lack of pre-war legitimacy and politically character contributed to its eventual rejection, not solely on linguistic grounds, but as part of a broader act of symbolic distancing in the post-war period.²¹

5. CONCLUSION

Despite the similarities between the Norwegian spelling reform of 1941 and the Rules of Slovak Orthography of 1940, the differences appear to be more prominent. The distinctions are related not only to the preparation, implementation, content, and impact of the reforms but also to the number of publications on the reforms available. While there are dozens of sources discussing the Slovak reform (including the unapproved proposal from 1939, cf. Kačala – Krajčovič 2011; Krajčovič – Žigo 2011; Kopecká 2019), there are significantly fewer such sources on the Norwegian side. Even when attention is given to the reform, it often consists of scant information (cf. Vinje 1978; Otnes – Aamotsbakken 2006; Nesse 2013). An exception is Arne Tjelle’s thesis (1994), which comprehensively captures the reform. This lack of attention creates opportunities for further exploration, particularly on the Norwegian side.

I believe two factors significantly influence this situation. First, as mentioned earlier, the Norwegian reform ceased to be effective after the end of the Second World War. Second, the reform was a product of the Quisling government. Therefore, there might have been a strategy not to pay excessive attention to the reform. The situation in Slovakia was different. The reform remained in effect until 1953, shaping the language for a longer period, and its validity, as well as its preparation, was not exclusively tied to the Slovak Republic (1939–1945). Additionally, the perception of

²¹ This was not the only case where authoritative language planning met with resistance in Norway. A later example is the *Samnorsk* policy – a post-war attempt to merge Bokmål and Nynorsk into a unified written standard – which also faced considerable opposition and was officially abandoned in 2002.

the existence of this period is not uniformly negative in society. According to a 2024 survey (Tomečková 2024), slightly less than a third of the Slovak population holds a negative view of the wartime Slovak state.

Another interesting discrepancy I noticed while studying the materials is the characterisation of the analysed reforms in publications. While the Norwegian spelling reform of 1941 is also described with expressive terms such as “Quisling’s” (cf. Vinje 1978) and “Nazi” (cf. Otnes – Aamotsbakken 2006; Ramsfjell – Vinje 1978) or regarded as a curiosity in the development of Norwegian (Tjelle 1994, p. 5), I have not observed this phenomenon on the Slovak side.²² A comparative analysis of the two ‘wartime’ reforms has, therefore, uncovered interesting aspects to explore in future research.

References

- BLANÁR, Vincent – JÓNA, Eugen – RUŽIČKA, Jozef (1974): *Dejiny spisovnej slovenčiny 2*. Bratislava: Slovenské pedagogické nakladateľstvo. 254 p.
- Den nye rettskrivning: Regler og ordlister (1918). Kristiania: Det Mallingske Bogtrykkeri. 97 p.
- GUTTU, Tor (2017): Quisling-regimet og samnorsk. In: *Riksmålsforbundet*. Available at: https://www.riksmalsforbundet.no/qa_faqs/quisling-regimet-samnorsk/ [cit 27-05-2025].
- HAMRE, Martin K. (2019): Norwegian Fascism in a Transnational Perspective: The Influence of German National Socialism and Italian Fascism on the Nasjonal Samling, 1933–1936. In: *Fascism: Journal of Comparative Fascist Studies*, Vol. 8, No. 1, pp. 36–60. DOI 10.1163/22116257-00801003. [cit 27-05-2025].
- HRANCOVÁ, Hana (2019): Postoje Anny Horákovej-Gašparíkovej v kontexte problematiky sporu o Pravidlá slovenského pravopisu z roku 1931. In: *Slavica litteraria*, Vol. 22, No. 1, pp. 77–88. DOI 10.5817/SL2019-1-8. [cit 27-05-2025].
- JAROŠOVÁ, Alexandra (2012): Bohemizmy a kodifikačné „kyvadlo“. In: A. Bohunická (ed.): *Jazykoveda v pohybe*. Bratislava: Univerzita Komenského, pp. 255–263.
- JÁNSKY, Ladislav M. (1940): Slovenský jazykový purizmus vo svetle štatistiky. In: *Linguistica Slovaca* 6. Bratislava: Erudita Societas Slovaca – Slovenská učená spoločnosť, pp. 5–14.
- JÓNA, Eugen (1963): Práca Matice slovenskej v oblasti jazykovedy a spisovného jazyka po roku 1919. In: J. Mésároš – M. Kropilák (eds.): *Matica slovenská v slovenských dejinách*. Bratislava: Vydavateľstvo Slovenskej akadémie vied, pp. 240–249.
- JÓNA, Eugen (1973): Teória slovenského spisovného jazyka a jeho kodifikácia na rozhraní 19. a 20. storočia (1875–1918). In: V. Betáková (ed.): *Slovenčina na rozhraní 19. a 20. storočia*. Bratislava: Slovenské pedagogické nakladateľstvo, pp. 15–45.
- KAČALA, Ján – KRAJČOVIČ, Rudolf (2011): *Prehľad dejín spisovnej slovenčiny*. Martin: Matica slovenská. 233 p.
- KESSELOVÁ, Jana – SLANČOVÁ, Daniela (2010): Slovak. In: *Revue belge de philologie et d’histoire*, Vol. 88, No. 3, pp. 873–896. DOI 10.3406/rbph.2010.7807. [cit 27-05-2025].

²² It is worth noting, however, that although the 1940 Slovak reform is not explicitly framed in these ideological terms, its purist tendencies were increasingly problematised after the war (cf. Peciar 1950; Váhala 1954). As Múcsková (2017, p. 42) points out, even today, attempts to discuss linguistic purism in a neutral, descriptive manner – without evaluating it as either beneficial or harmful to the standard language – remain problematic.

- KOPECKÁ, Martina (2019): *Kodifikačné tendencie v Pravidlách slovenského pravopisu z rokov 1931 – 2013. Jazykovedné štúdie 34*. Bratislava: Veda, vydavateľstvo SAV. 378 p.
- KOVÁČ, Dušan (2011): *Dějiny Slovenska*. Praha: Nakladatelství Lidové noviny. 434 p.
- KRAJČOVIČ, Rudolf (2004) Z galérie osobností v dejinách spisovnej slovenčiny (XII.) Medzivojnové obdobie. In: *Kultúra slova*, Vol. 38, No. 6, pp. 321–329.
- KRAJČOVIČ, Rudolf – ŽIGO, Pavol (2004): *Príručka k dejinám spisovnej slovenčiny*. Bratislava: Univerzita Komenského. 188 p.
- KRAJČOVIČ, Rudolf – ŽIGO, Pavol (2011): *Dejiny spisovnej slovenčiny*. Bratislava: Univerzita Komenského. 249 p.
- LUNDE, Gulbrand (1942): *Kampen for Norge II: Foredrag og artikler 1940–1941*. Oslo: Nasjonal Samlings Rikstrykkeri. 259 p.
- MÚCSKOVÁ, Gabriela (2017): O slovenskom purizme a anti-purizme v kontexte jazykových ideológií. In: I. Lanstyák – G. Múcsková – J. Tancer (eds.): *Jazyky a jazykové ideológie v kontexte viacjazyčnosti na Slovensku*. Bratislava: Univerzita Komenského, pp. 39–79.
- NÁBĚLKOVÁ, Mira (2017): Slovakismy v češtině. In: P. Karlík – M. Nekula – J. Pleskalová (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*. Available at: https://www.czechency.org/slovník/SLOVAKISMY_V_ČEŠTINĚ [cit 27-05-2025].
- NESSE, Agnete (2013): *Innføring i norsk språkhistorie*. Oslo: Cappelen Damm AS. 232 p.
- Ny rettskrivning 1938: Bokmål. Regler og ordliste (1938)*. Oslo: Olaf Norlis Forlag. 87 p.
- Ny rettskriving 1938: Nynorsk. Reglar og ordliste (1938)*. Oslo: Olaf Norlis Forlag. 45 p.
- Ny rettskrivning 1941: Regler for bokmål. Regler for nynorsk. Fellesregler (1941)*. Oslo: Blix Forlag. 52 p.
- ONDREJOVIČ, Slavomír (2003): Sedem decéníí časopisu Slovenská reč. In: S. Ondrejovič (ed.): *Slovenská reč 1932 – 2002*. Bratislava: Veda, vydavateľstvo SAV, pp. 9–34.
- OTNES, Hildegunn – AAMOTSBAKKEN, Bente (2006): *Tekst i tid og rom: Norsk språkhistorie*. Oslo: Det Norske Samlaget. 277 p.
- PECIAR, Štefan (1950): Jazykový purizmus a apretorská prax. In: *Slovenská reč*, Vol. 15, No. 9–10, pp. 257–273.
- Pravidlá slovenského pravopisu s abecedným pravopisným slovníkom (1931)*. Turčiansky Sv. Martin: Matica slovenská. 357 p.
- Pravidlá slovenského pravopisu s pravopisným slovníkom (1940)*. Turčiansky Sv. Martin: Matica Slovenská. 475 p.
- Pravidlá slovenského pravopisu s pravopisným a gramatickým slovníkom (1953)*. Bratislava: Vydavateľstvo Slovenskej akadémie vied. 408 p.
- RAMSFJELL, Berit S. – VINJE, Finn-Erik (1978): *Språkkunnskap*. Oslo: Aschehoug. 146 p.
- RUŽIČKA, Ján (1970): *Spisovná slovenčina v Československu*. Bratislava: Vydavateľstvo Slovenskej akadémie vied. 250 p.
- STANISLAV, Ján (1938): *Československá mluvnica*. Praha – Prešov: Československá grafická únia. 235 p.
- ŠVAGROVSKÝ, Štefan (2006): Jazykovedné dielo S. Czambly v osídľach politiky. In: *Slovenská reč*, Vol. 71, No. 6, pp. 330–339.
- ŠVAGROVSKÝ, Štefan – ŽIGO, Pavol (1994): O pravopisnej koncepcii neschválených Pravidiel slovenského pravopisu z r. 1939. In: *Slovenská reč*, Vol. 59, No. 2, pp. 98–103.
- TJELLE, Arne (1994): *Rettskrivinga av 1941: Bakgrunn, politisk spel og ideologisk analyse*. Bergen: Nordisk institutt, Universitetet i Bergen. 174 p.
- TOMEČKOVÁ, Nicol (2024): Len tretina populácie má negatívny postoj k Slovenskému štátu a deportáciám Židov. In: *RTVS Správy*. Available at: <https://spravy.stvr.sk/2024/03/len>

- tretina-populacie-ma-negativny-postoj-k-slovenskemu-stat-u-a-deportaciám-zidov/ [cit 27-05-2025].
- TORP, Arne – DAHL, Berit H. – LUNDEBY, Ingard (1991): *Språklinjer: Språkhistorie for den videregående skolen*. Oslo: Norsk Undervisningsforlag. 208 p.
- VÁHALA, František (1954): Pravidla slovenského pravopisu. In: *Naše řeč*, Vol. 37, No. 1–2, pp. 33–39.
- VIKØR, Lars S. (2022): Rettskrivingsreforma av 1941. In: *Store norske leksikon*. Available at: https://snl.no/Rettskrivingsreforma_av_1941 [cit 27-05-2025].
- VINJE, Finn-Erik (1978): *Et språk i utvikling: Noen hovedlinjer i norsk språkhistorie fra reformasjonen til våre dager*. Oslo: Aschehoug. 423 p.
- ŽIGO, Pavol (1997): Henrich Bartek a vývin slovenskej spisovnej normy v medzivojnovom období. In: J. Mlacek (ed.): *Studia Academica Slovaca 26, Prednášky 33. letného seminára slovenského jazyka a kultúry*. Bratislava: Stimul – Centrum informatiky a vzdelávania FF UK 1997, pp. 180–188.

Resumé

PURIZMUS A NACIONALIZMUS V JAZYKOVEJ NORME: KOMPARATÍVNA ANALÝZA PRAVOPISNÝCH REFORIEM NÓRČINY A SLOVENČINY PRIJATÝCH POČAS 2. SVETOVEJ VOJNY

Jazyková situácia v medzivojnovom Nórsku a Československu vykazuje viacero paralel. Úradné jazyky v oboch krajinách existovali v dvoch variantoch: nórčina vo forme bokmål a nynorsk a československý jazyk v českom a slovenskom znení. Zároveň obe krajiny prijali v tomto období pravopisné normy s cieľom postupného zjednotenia variantov jazyka do jedného. V Nórsku mal tento proces podobu vzájomného sa približovania foriem bokmål a nynorsk, zatiaľ čo v Československu išlo predovšetkým o priblíženie sa slovenčiny češtine. Obe normy – nórska z roku 1938 a slovenská z roku 1931 – vyvolali kontroverzie, a v oboch prípadoch sa počas druhej svetovej vojny, v čase, keď Nórsko i novovzniknutý Slovenský štát boli spojencami nacistického Nemecka, prijali nové pravidlá: nórska pravopisná reforma z roku 1941 a *Pravidlá slovenského pravopisu* z roku 1940. Tento článok sa zameriava na uvedené reformy prijaté počas druhej svetovej vojny a skúma, aký vplyv mali politické zmeny na jazykové plánovanie. Prostredníctvom komparatívnej analýzy sa snaží identifikovať prvky, východiská či idey, ktoré spájajú nové pravidlá napriek odlišnej lingvistickej povahe analyzovaných jazykov.

FONOLOGICKÉ PROCESY V PRÚDE POSUNKOVANIA

ROMAN VOJTECHOVSKÝ

Pedagogická fakulta Trnavskej univerzity v Trnave, Trnava, Slovensko

VOJTECHOVSKÝ, Roman: Phonological processes in the stream of signing. *Jazykovedný časopis (Journal of Linguistics)*, 2025, Vol. 76, No. 2, pp. 509–533.

Abstract: Segments in spoken and sign language interaction do not produce individual units in isolation but in a continuous stream of expression. When using language for communication, the forms of segments can be simplified to make communication more efficient for both communicators and communicants. This communication can result in changes in the way formal units of language are formed. It follows that, as in spoken languages, there are considerable differences in how sign languages are articulated. The differences may manifest in lexis and syntax and in phonetic differences related to social and linguistic factors. In this paper, we focus on the linguistic factors causing differences in the articulation of signs, specifically on the phonological processes within which different adaptations in the forms of signs may occur. For this reason, we will approach several phonological processes in the stream of signing occurring in different national sign languages compared to Slovak sign language and their characteristic features. Understanding phonological processes in the realization of signing can enrich the field of phonetics and phonology in general.

Keywords: sign language, sign parameters, phonological processes, assimilation, reduction, elision, epenthesis, metathesis

1. NA ÚVOD O ARTIKULÁTOROCH

V porovnaní s fonológiou hovoreného jazyka je fonológia posunkového jazyka mladá veda a začala sa rozvíjať až v 60. rokoch 20. storočia. Na prvý pohľad najvýraznejším rozdielom medzi hovorenými, teda zvukovými, a posunkovými jazykmi v ľudskej komunikácii je **rozdiel v modalite**. Slová sa v hovorených jazykoch artikulujú pohybmi rečových orgánov na tvorenie zvukov ľudskej reči a vnímajú sa sluchom (**vokálno-auditívna modalita**), zatiaľ čo posunky sa artikulujú pohybmi hornej časti tela vrátane rúk a percipujú sa zrakom (**motoricko-vizuálna modalita**). Opticky vnímateľné pohyblivé časti tela – ruky, mimika, hlava a trup – sú **artikulátormi posunkového jazyka**. V hovorenom jazyku sa samotný jazykový signál prenáša vokálne od komunikátora ku komunikantovi formou zvukových vln a v posunkovom nevodí – motoricky prostredníctvom **svetelných, optických vln** (porov. van der Kooij – Crasborn 2016; Anderson a kol. 2022).

Pre posunkové jazyky existujú dve hlavné kategórie artikulátorov, ktoré sa spoločne podieľajú na vyjadrení formy a významu posunkov. **Manuálne artikulá-**

tory, tvoriace paže, ruky a prsty, sú hlavnými artikulátormi, ktoré sa uplatňujú pri posunkovaní. Pohybujú sa pomocou kĺbov hornej končatiny. Súčasne s manuálnymi artikulátormi sa na artikulovanie posunkového prejavu používa horná časť tela, a to trup, hlava a výrazy tváre. Tie sa nazývajú **nemanuálne artikulátory**. Ide o komplexný súbor rôznorodých pohybov, ako je napríklad zmena zrakového kontaktu, zužovanie a rozširovanie viečok, zdvíhanie a zvrátenie obočí, nakláňanie a rotácia trupu, predklon a záklon hlavy, nafúknutie a vtiahnutie líc, zaoblenie a rozťahnutie pier, odhalenie zubov a pod. (Anderson a kol. 2022). Zdá sa, že na rozdiel od hovorených jazykov majú posunkové jazyky viacero aktívnych artikulátorov. Okrem úst, ktoré sa občas považujú za nezávislý artikulátor, sa na artikulácii zúčastňujú obidva manuálne artikulátory, ktoré sa môžu pohybovať v závislosti od seba (napr. kým sa jedna ruka drží v posunkovom priestore, aby odkazovala na referent, druhá ruka môže pokračovať v posunkovaní; van der Kooij – Crasborn 2016). Posunkovanie sa produkuje vo vymedzenom, tzv. **posunkovom** (artikulačnom) **priestore**, ktorý má podobu trojdimenzionálneho priestoru. Vo vertikálnej rovine sa rozprestiera približne tesne nad hlavou až po boky, v horizontálnej od jedného lakťa až po druhý lakeť a v sagitálnej (predozadnej) rovine od roviny hrude až po končeky prstov. Centrálnym poľom posunkového priestoru je oblasť od úst až po oblasť tesne pod hrudníkom a ohniskom pohľadu sú ústa. Samostatné manuálne posunky sa vnímajú periférne.

Základnou samostatnou významovou jednotkou posunkového jazyka je **posunok**, ktorý chápeme ako ekvivalent slova v hovorenom jazyku. Podobne ako v hovorenom jazyku i v posunkovom jazyku sa vyskytujú výstavbové prvky, ktoré sa často označujú ako **parametre** (porov. van der Kooij – Crasborn 2016). Posunok reprezentuje „komplex parametrov, ktorý vytvára lexikálnu jednotku posunkového jazyka“ (Csonka 1995, s. 4). Jednotlivé manuálne posunky pozostávajú zo štyroch súčasne artikulovaných formálnych parametrov: **tvár ruky** (ako vyzerá určitý tvar ruky s jedným alebo viacerými prstami v danej polohe pri artikulácii – narovnaný, zatvorený, zakrivený a ohnutý), **miesto artikulácie** (kde sa daný posunok artikuluje v priestore pred telom, na tele či v jeho blízkosti – na hrudi, bruchu, ramene, predlaktí, hlave, tvári a pod.), **pohyb** (ako sa realizuje dráhový pohyb ruky či rúk z jednej lokácie na druhú a/alebo vnútorný pohyb prstov) a **orientácia ruky** (kam smeruje dlaň a prsty), pričom každý posunok utvára jedinečnú špecifickú kombináciu týchto prvkov. Uvedieme niekoľko príkladov slovenských posunkov, ktoré sa odlišujú v jednom parametri. Posunky HNEDÝ a RUŽOVÝ sa realizujú na zhodnom mieste – na líci, vykonávajú sa totožným – opakovaným kruhovým pohybom. Orientácia dlane dvoch posunkov je identicky nasmerovaná k telu, konkrétne k lícu, avšak medzi nimi je rozdiel v tvare ruky: kým prvý posunok ilustruje zovretú ruku s voľným palcom, druhý zase zovretú ruku so vztýčeným prostredníkom a ukazovákom pri sebe.



Obr. č. 1 – 2: Model rozpoznávania posunkov HNEDÝ a RUŽOVÝ.

Naproti tomu posunky RUŽOVÝ a PIATOK majú zhodný tvar ruky, identické miesto artikulácie na líci, takú istú orientáciu dlane smerujúcu k lícu, ale odlišný pohyb ruky. Pri posunku RUŽOVÝ sa vykonáva opakovaný kruhový pohyb (ide o trenie) a posunok PIATOK sa realizuje opakovaným priamočiarym pohybom k lícu. Ďalším príkladom sú posunky POZNAŤ a PIATOK, ktoré sa produkujú na odlišných miestach, ale ostatné formálne parametre sú zhodné. Posunok POZNAŤ sa artikuluje na spánkovej časti čela a PIATOK na líci. Iné príklady posunkov s odlišnou orientáciou ruky sú posunky PIATOK a ZÁCHOD. Kým orientácia dlane posunku PIATOK smeruje k lícu, orientácia dlane posunku ZÁCHOD je mierená smerom von doľava (u praváka). Ukazuje sa teda, že slovenskí posunkujúci vedia rozlišovať manuálne parametre posunkov, pričom aj minimálne rozdiely v jednom parametri majú dištinktívnu platnosť. Všetky simultánne artikulované parametre posunku sú obligatórne, nemožno vylúčiť ani jeden z nich a zároveň ani nie je možné produkovať jednotlivé parametre sekvenčne, teda lineárne. Niektorí autori (napr. Sandler – Lillo-Martin 2006, cit. podľa Johnston – Schembri 2007) tvrdia, že všetky parametre posunku môžu byť dôkazom sekvenčného usporiadania posunku, napr. parameter pohybu často zahŕňa postupnosť pohybov od jedného tvaru ruky k druhému alebo z jedného miesta artikulácie na druhé. Dokumentuje to slovesný posunok SVIETIŤ, keď tvar ruky v počiatočnej fáze má podobu takmer zovretej päste a postupne ju roztvára. Konštatovanie, že slová v hovorenom jazyku nevznikajú náhodnou kombináciou zvukov, ale sú vytvorené z obmedzeného súboru zvukov, platí aj pre posunkový jazyk. Niektorí lingvisti (Sandler 1989, cit. podľa Quer a kol. 2017; van der Kooij 2002) zahŕňajú tvar aj orientáciu ruky pod pojem **konfigurácia ruky**.

Na portáli *Posunky.sk: Prekladový slovník slovenského jazyka – slovenského posunkového jazyka* (Vojtechovský 2018 – 2024) uvádzame všetky realizácie parametrov posunkov používané v slovenskom posunkovom jazyku. Konkrétne ide o 61 tvarov ruky, 21 miest artikulácie a 26 pohybov ruky, príp. rúk. V súčasnosti pre aktu-

álne výskumy vo fonetike a fonológii slovenského posunkového jazyka nemožno s istotou zhrnúť, z koľkých foném s fonologickou platnosťou inventár fonologického systému slovenského posunkového jazyka pozostáva.

Typológia klasifikuje posunky na **jednoručné**, ktoré sú vytvárané jednou, a to dominantnou rukou, a posunky **dvojručné**. Na ich artikulovaní sa môžu podieľať obidve ruky aktívne alebo jedna ruka aktívne a druhá pasívne. Ruka v pasívnej pozícii vždy vytvára miesto pre aktívnu ruku. Dvojručné posunky sa delia na **symetrické posunky**, charakterizované zhodným tvarom oboch rúk, a **asymetrické posunky**, počas ktorých je každá ruka v inom tvare. Pri dvojručných symetrických posunkoch sa môže líšiť súhra pohybu medzi dvoma aktívnymi rukami, teda pohyb môže byť synchrónny alebo alternatívny.

So štyrmi manuálnymi parametrami sa súčasne artikuluje aj relevantný piaty parameter, ktorý utvárajú **nemanuálne prostriedky**. Sú nimi pohyb hlavy a hornej časti tela a výraz tváre (mimiku) vrátane pohybu úst. Pohyb úst na fonologickej úrovni zohráva dôležitú úlohu pri rozlišovaní manuálnych homónym, ktoré sú založené na identickej podobe posunku s odlišným, nesúvisiacim významom (napr. MÁJ – BISKUP) a manuálnych polysémantov – posunky majú viac navzájom súvisiacich významov (napr. STROMČEK – VIANOCE). Bezhlasné pohyby úst sa častejšie klasifikujú do dvoch komponentov: **hovorené komponenty** naznačujúce súvisť s konkrétnym hovoreným jazykom (napr. zo slovenského – *auto*, *mama*, *bába*) a **orálne komponenty** prezentujúce nesúvisť s daným hovoreným jazykom (napr. zovreté pery, našpúlené pery, vibrujúce pery a pod.). Posunky podobne ako slová nie sú teda holistické gestá bez vnútornej štruktúry (porov. Sandler 2012).

Okrem jednotiek podobných segmentom sa potvrdzuje, že v posunkových jazykoch existujú aj **slabiky**. Odborníci na fonológiu posunkových jazykov sa zhodujú v tom, že za jadro slabiky sa pokladá pohyb (Emmorey 2007; Sandler 2012). W. Sandlerová (2006, s. 206) definuje slabiku ako „jeden pohyb, a to buď: a) po dráhe, b) vnútorne prostredníctvom zmeny tvaru alebo orientácie ruky, alebo c) oba súčasne.“

Podľa D. Brentariovvej (2012) vo všetkých jazykoch existujú **obmedzenia** týkajúce sa spôsobov, akými sa môžu jednotlivé parametre navzájom kombinovať. Obmedzenia v posunkovom jazyku sa môžu týkať tvaru ruky s vybraným prstom/vybranými prstami, vnútorného pohybu a štruktúry slabík. Lingvisti (napr. van der Kooij – Crasborn 2016) poukázali na to, že anatómia nášho tela obmedzuje možnosti produkcie posunkov, napríklad extrémne polohy kĺbov, ktoré sú namáhavé pre svaly a šľachy, sa v posunkových jazykoch vyskytujú zriedka. Z podobného dôvodu sa zdá, že posunkové jazyky na celom svete viac využívajú tvary rúk, pri ktorých je vzpriamený len ukazovák alebo malíček, ako tie tvary rúk, pri ktorých je vzpriamený len prostredník alebo prstenník. E. van der Kooij a O. Crasborn (2016) poukázali aj na obmedzenia v produkcii dvojručných posunkov, ktoré sú platné pre všetky posunkové jazyky. Prvou je tzv. **podmienka symetrie**, ktorá sa uplatňuje, keď sa obe ruky pohybujú, majú rovnaký tvar a vykonávajú synchrónny alebo alternatívny pohyb.

Druhou je **podmienka dominancie**, v ktorej jedna aktívna ruka vykonáva pohyb, zatiaľ čo druhá pasívna ruka sa stáva miestom artikulácie pre aktívnu ruku. Avšak počet tvarov pasívnej ruky, tzv. **nepríznakových tvarov**, je výrazne obmedzený. Podľa portálu *Posunky.sk: Prekladový slovník slovenského jazyka – slovenského posunkového jazyka* (Vojtechovský 2018 – 2024) sa vyskytuje sedem nepríznakových tvarov pasívnej ruky.

Každý posunok má východiskovú a záverečnú podobu. Posunkujúci artikulujú posunkami tak, že prirodzene a plynulo prechádzajú z jedného posunkú do druhého. Medzi nimi nerobia dlhšie pauzy, ktoré by sťažovali komunikáciu. Pohyb, ktorý je foneticky potrebný na presun ruky/rúk z koncového bodu jedného posunkú na začiatok ďalšieho posunkú a ktorý nie je súčasťou lexikálnej špecifikácie žiadneho z dvoch susedných posunkov, sa nazýva **prechodný pohyb** (Quer a kol. 2017).

Predpokladá sa, že posunkujúci má vo svojom mentálnom lexikóne uloženú jednu formu každého posunkú. Táto forma sa nazýva **citátová forma**, či **fonologická forma** alebo **základná forma**. Realizácia posunkú je **fonetická forma** a môže byť zakaždým iná. Posunky sa do lexikónov vkladajú vo svojej citátovej forme (Twilhaar – van den Bogaerde 2016).

2. FONETIKA A FONOLÓGIA POSUNKOVÉHO JAZYKA

Mnohé štúdie (napr. Sandler 2006, van der Kooij – Crasborn 2016, Quer a kol. 2017) prezentovali poznatok, že fonológia posunkového jazyka má veľa spoločného s fonológiou hovoreného jazyka vrátane existencie bezpríznakových fonologických prvkov a tried prvkov, fonologických obmedzení, pravidiel asimilácie a ďalších znakov. W. Sandlerová (2006, s. 185) uvádza, že „fonetika týchto dvoch jazykových modalít je čiastočne formovaná artikulačnými prostriedkami s veľmi odlišnými anatomickými a fyziologickými charakteristikami. Z týchto dvoch dôvodov ponúka posunkový jazyk zaujímavý pohľad na vzťah medzi fonetikou a fonológiou.“ Americký lingvista W. Stokoe v prvej lingvistickej a zásadnej štúdii z roku 1960 zaoberajúcej sa analýzou stavebných prvkov amerického posunkového jazyka uprednostňoval špecifické termíny pre jednotlivé modalitty, a to *cherológia* a *cheréma* (z gr. *cheir* – ‚ruka‘), ktoré zodpovedajú termínom *fonológia* a *fonéma*. V súčasnosti viacerí lingvisti posunkového jazyka (napr. Parkhurst 2008; van der Kooij – Crasborn 2016; Anderson a kol. 2022) bežne používajú termíny *fonológia* a *fonéma* bez ohľadu na modalitu.

V lingvistike sa od seba funkčne rozlišujú fonetika a fonológia. **Fonetika** skúma artikuláciu a vnímanie reči a posunkov, a preto sa zameriava na modálny rozdiel medzi vizuálno-priestorovými a auditívno-orálnymi jazykmi. V oboch typoch jazykov sa v komunikačnom reťazci rozlišuje medzi artikuláciou, signálom a vnímaním tohto signálu (Crasborn – van der Kooij 2016). **Fonológia** skúma prvky používané na tvorbu slov a posunkov a ich možné kombinácie. Vo fonológii hovorených jazy-

kov sa analyzujú zvuky konkrétneho jazyka a to, ako sa tieto zvuky spájajú do slabík a slov vo vzťahu k významu slov (van der Kooij – Crasborn 2016). Rovnako sa analyzujú časti tela príslušného posunkového jazyka vo fonológii. Fonetika vo výskume posunkového jazyka sa na jednej strane zaoberá vizuálnou percepciou posunkových jazykov a na druhej strane fyziologickými vlastnosťami, ktoré zohrávajú úlohu pri produkcii posunkových jazykov. Pohyby, ktoré sú výsledkom aktivácie kĺbov a svalov, sa môžu opísať na troch úrovniach: a) ktoré svaly sú zapojené v ktorom posunkú a ako sa realizujú, b) ktorý kĺb sa zapája, a c) ktoré časti tela sú aktívne (Twilhaar – van den Bogaerde 2016). Vizuálna fonetika skúma, ako percipient vníma a spracúva vizuálne signály, pričom sa zameriava na vzťah medzi fyziologickými vlastnosťami posunkového prejavu a ich porozumením v posunkovom priestore. Naopak, fonológia posunkových jazykov sa zameriava na základné prvky, z ktorých sa posunok skladá. Skúma ich realizácie, ktoré môžu meniť význam, a pravidlá, podľa ktorých sa tieto prvky kombinujú v kontexte iných posunkov.

3. FONOLOGICKÉ PROCESY VYSKYTUJÚCE SA V POSUNKOVÝCH JAZYKOCH

Fonetika posunkového jazyka skúma to, ako sa posunok artikuluje a ako sa perceptive vníma. S. Parkhurst (2008, s. 37) tvrdí, že „artikulácia a percepcia sú riadené dvoma všeobecnými silami, ktoré sú často vo vzájomnom konflikte.“ Autor (tamže) to vysvetľuje tak, že ak chceme, aby naša komunikácia bola čo najjednoduchšia na produkciu, máme tendenciu sa vyhýbať zložitým tvarom rúk, orientáciám alebo miestam artikulácie. Preto sa môže stať, že odlišné tvary rúk z dvoch posunkov sa môžu spojiť a vytvoriť jeden tvar ruky, ktorý už zahŕňa oba posunký (príklad asimilácie) tak, aby sa posunok ľahšie artikuloval. Niektoré dvojručné posunký sa môžu stať jednoručnými. Ale ak chceme zreteľne komunikovať, každý posunok by sa mal odlišovať od všetkých ostatných posunkov. Posunký rovnako ako hovorené slová sú pri rýchlej artikulácii ovplyvnené podobnými fonologickými procesmi (Liddell – Johnson 1989, cit. podľa Dikyuva a kol. 2017, s. 96).

Fonologický proces je podľa J. Quera a kol. (2017, s. 53) „výsledkom uplatnenia súboru obmedzení alebo pravidiel manipulujúcich s fonologickým tvarom (základných) vstupných foriem s cieľom dostať určité výstupné formy.“ Autor (tamže) dodáva, že vďaka fonologickým procesom môže dôjsť k jednoduchosti artikulácie na jednej strane a jednoduchosti vnímania na druhej strane. Ďalším typom fonologického procesu, ktorý ovplyvňuje lexiku, je adaptácia formy výpožičiek z iného (posunkového) jazyka. Z toho vyplýva, že pri použití posunkov vo vete môže mať vplyv forma predchádzajúceho alebo nasledujúceho posunkú, tento jav je známy ako **koartikulácia**. Niekedy koartikulácia môže viesť k čiastočnej zmene fonologických prvkov posunkú. Teda fonológia posunkového jazyka neskúma len inventár parametrov posunkov, ktoré rozlišujú význam, ale jej cieľom je aj opísať modifikácie

parametrov posunkov, ku ktorým dochádza v prúde posunkovania, keď sa artikulo-
vané posunky spájajú vo výpovediach, v texte (porov. van der Kooij – Crasborn
2016).

Skôr ako sa zameriame na opis fonologických procesov v procese posunkova-
nia, je vhodné priblížiť, ktoré fonologické procesy uvádzajú lingvisti posunkových
jazykov vo svojich štúdiách, keďže viacerí z nich ich neklasifikujú zhodne. Niektorí
z nich neuviedli termíny – buď ich nepoznali, alebo jednotlivé procesy nepovažovali
za svojbytné fonologické procesy. Odborné termíny uvedené v tabuľke sú vysvetle-
né nižšie.

	Johnston – Schembri (2007)	Twilhaar – van den Bogaerde (2016)	Baker a kol. (2016)	Quer a kol. (2017)	Dikyuva a kol. (2017)	Klomp (2021)
Asimilácia tvaru ruky	áno	áno	áno	áno	áno	áno
Asimilácia miesta artikul.	áno	áno	–	–	áno	áno
Asimilácia pohybu	–	áno	áno	–	–	–
Redukcia pohybu	áno	áno	–	áno	–	áno
Redukcia druhej ruky	–	–	áno	–	–	–
Extenzia pohybu	–	–	–	áno	–	áno
Elízia segmentu držania	áno	–	–	–	–	–
Elízia segmentu pohybu	áno	–	–	áno	áno	áno
Elízia druhej ruky	–	áno	áno	áno	áno	áno
Epentéza pohybu	áno	–	–	áno	áno	áno
Metatéza	–	–	–	áno	–	áno
Naturalizácia	–	–	–	áno	–	áno
Koalescencia	–	–	–	áno	–	áno
Perseverácia nedominantnej ruky	áno	–	–	–	–	–
Anticipácia nedominantnej ruky	áno	–	–	–	–	–

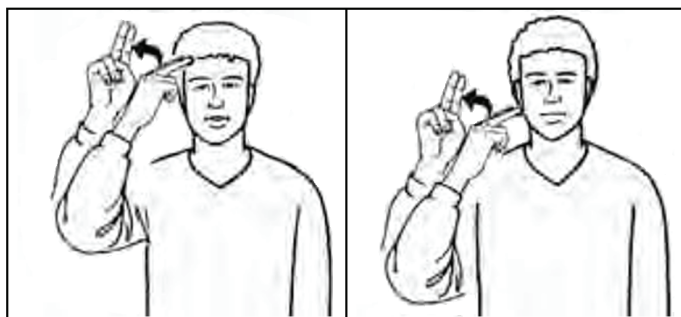
Tab. 1: Prehľad fonologických procesov vyskytujúcich sa v štúdiách lingvistov zameraných na
posunkové jazyky.

K uvedeným javom by sme doplnili ďalšie fonologické procesy, ktoré uviedla
vo svojej práci D. Lachmanová (2016) a ktoré sa vyskytujú aj v slovenskom posun-
kovom jazyku. Sú to **redukcia miesta artikulácie** a **redukcia tvaru ruky**.

3.1 ASIMILÁCIA

Jedným z fonologických procesov je **asimilácia**, ktorá je v posunkových jazykoch pomerne rozšíreným javom (porov. Quer a kol. 2017). Asimiláciu chápeme ako „fonologický proces, ktorý umožňuje, aby jedna alebo viacero vlastností fonémy nadobudlo rovnakú hodnotu ako iná fonéma v rámci určitej oblasti. Výsledkom je, že forma, ktorá prechádza fonologickou zmenou, sa viac podobá blízkej forme“ (Quer a kol. 2017, s. 53). Autori ďalej tvrdili, že asimilácia môže byť buď čiastočná, v ktorej sa kopírujú len niektoré prvky – parametre, alebo úplná, v ktorej sa kopírujú všetky prvky – parametre. Pokiaľ ide o zdroj asimilácie, rozlišujeme: a) **regresívnu asimiláciu**, ak je zdrojom asimilácie nasledujúca forma, teda keď sa v posunkovom jazyku fonologický prvok stotožňuje s prvkom nasledujúceho posunkú, b) **progresívnu asimiláciu**, ak je zdrojom asimilácie predchádzajúca forma, čiže keď je v posunkovom jazyku fonologický prvok nasledujúceho posunkú ovplyvnený predchádzajúcim posunkom, a c) **obojsmernú asimiláciu**, ak je zdrojom asimilácie predchádzajúca i nasledujúca forma potrebná na vytvorenie príslušného kontextu. Príklady týchto typov asimilácie uvádzame nižšie.

Asimilácia môže byť zameraná na ktorýkoľvek z fonologických parametrov posunkú. **Asimilácia miesta artikulácie** sa vyznačuje v tom procese, v ktorom sa miesto artikulácie jedného posunkú asimiluje do miesta artikulácie iného posunkú, ktorý sa artikuluje sekvenčne. Napríklad vo výpovedi austrálskeho posunkového jazyka IX₁ MENO S-A-M, v ktorom posunok MENO je artikulovaný nižšie v posunkovom priestore pred pravou stranou tváre, pretože predchádzajúci deiktický, teda indexový posunok IX₁ znamenajúci *ja* je vykonaný na nižšom mieste artikulácie (progresívna asimilácia). Inak je citátová forma posunkú MENO artikulovaná na čele (Johnston – Schembri 2007, s. 116).



Obr. 3: Citátová forma MENO a znížená forma posunkú (vpravo) (Zdroj: Johnston – Schembri 2007, s. 116).

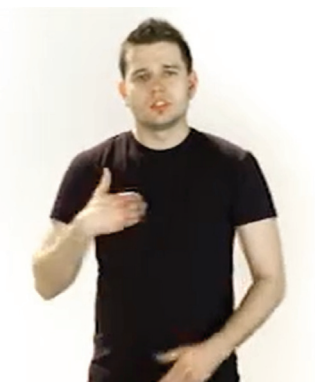
Ďalší príklad asimilácie je z tureckého posunkového jazyka, keď sa posunok pre vetnú negáciu NIE artikuluje hneď za posunkom slovesa VEDIET, preberá

miesto artikulácie predchádzajúceho slovesa (progressívna asimilácia). Citátová forma záporného posunku NIE je artikulovaná ipsilaterálne (artikulujúca na bočnej strane dominantnej ruky) vedľa hlavy. Asimilácia miesta artikulácie posunku VEDIEŤ však vedie len k artikulácii negácie prítomnej na brade, kde je artikulovaný posunok VEDIEŤ (Dikyuva a kol. 2017, s. 98).



Obr. 4: Asimilácia miesta artikulácie zloženého posunku VEDIEŤ^NIE (Zdroj: Dikyuva a kol. 2017, s. 98).

K **asimilácii tvaru ruky** dochádza počas skladania dvoch posunkov majúcich rôzne tvary rúk. Pri tejto asimilácii sa tvar ruky stáva podobnejší tvaru ruky posunku, ktorý mu vo vete predchádza, alebo po ňom nasleduje (Johnston – Schembri 2007). Napríklad v deiktickom posunku IX₁ (= JA) sa ilustruje zovretá ruka so vztýčeným palcom a druhom austrálskom posunku NEVEDIEŤ sa znázorňuje plochá ruka so vztýčeným palcom. Lenže v konštrukcii IX₁ NEVEDIEŤ sa deiktický tvar ruky asimiluje k tvaru ruky druhého posunku (regresívna asimilácia), teda deiktický tvar ruky znázorňuje už zhodný tvar ruky ovplyvnený nasledujúcim posunkom (tamže). Podobný príklad je aj v slovenskom posunkovom jazyku, v ktorom sa tvar ruky prvého posunku stáva podobnejším tvaru ruky druhého posunku, ktorý mu vo výpovedi predchádza: IX₁ VELMI-POTREBOVAŤ POSTEĽ, namiesto deiktického posunku IX₁ sa objaví ohnutá plochá ruka so vztýčeným palcom, keďže sa daný tvar ruky bude využívať v nasledujúcom posunku VELMI-POTREBOVAŤ.



Obr. 5: Asimilovaný tvar ruky deiktického posunkú (Zdroj: autor).

Z týchto príkladov je možné vyvodit', že tvar ruky deiktických posunkov sa častejšie asimiluje s daným tvarom ruky nasledujúcich posunkov, pretože deiktické posunky majú vysokú frekvenciu použitia v texte (porov. Lachmanová 2016, s. 57).

Asimilácia tvaru ruky sa môže realizovať aj pri dvojručných asymetrických posunkoch v rámci jednej lexikálnej jednotky (Liddell – Johnson 1989, cit. podľa Lachmanová 2016, s. 54). Prejavuje sa tým, že pasívna ruka sa svojím tvarom asimiluje k aktívnej ruke. Uvedieme príklad slovenského posunkú OBJEDNAŤ, v ktorom pasívna ruka citátovej formy znázorňuje zovretú päšť, avšak v prúde posunkovania sa pasívna ruka prispôsobí tvaru aktívnej ruky – zovretá päšť so vztyčeným ukazovákom.



Obr. 6 – 7: Pasívna ruka v citátovej forme (vľavo), asimilovaná pasívna ruka (vpravo)
(Zdroj: autor).

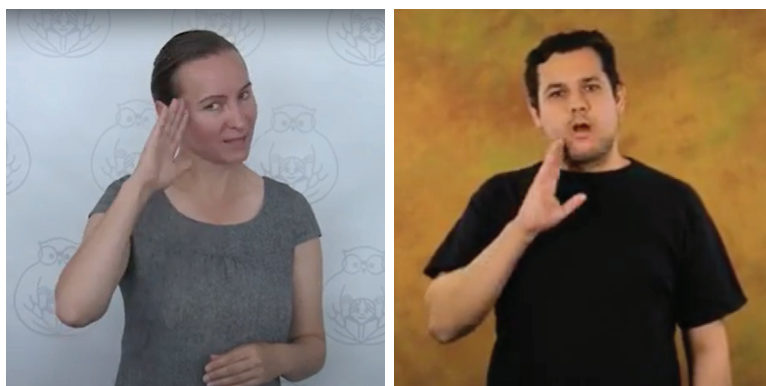
Asimilácia sa môže prejavit' pri **zмене počtu rúk** zúčastňujúcich sa na artikulácii. Konkrétne ide o zmenu jednoručných posunkov v okolí dvojručných posunkov

na dvojručnú formu posunku a o asimiláciu dvojručných posunkov na jednoručné posunky v okolí jednoručných posunkov (Liddell – Johnson 1989, cit. podľa Lachmanová 2016, s. 59).

Realizuje sa aj z pragmatického hľadiska pri expresívnych výpovediach, intenzifikácii niektorých posunkov, pri vyjadrení kriku či hnevu alebo v situácii, keď sa vo výpovedi môže artikulovať dvojručný posunok namiesto jednoručného, ktorý vykazuje citátovú podobu posunku. Druhá aktívna ruka asimiluje k prvej aktívnej ruke pôvodného jednoručného posunku. Môže ísť napríklad o slovenské posunky ako MUSIEŤ, POVINNOSŤ, SLABÝ.

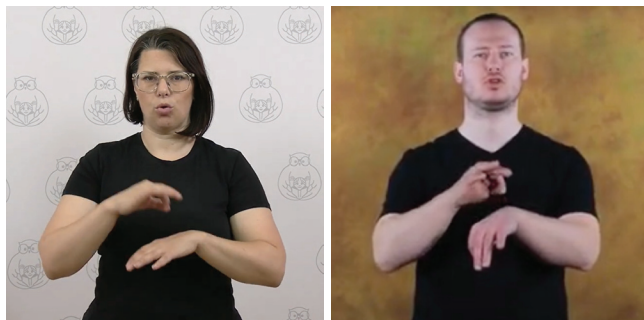
3.2 REDUKCIA

Druhým fonologickým procesom je **redukcia**, ktorá môže byť spôsobená týmito parametrami posunku: miesto artikulácie, tvar ruky a pohyb (porov. Lachmanová 2016). Redukcia znamená, že „v artikulácii sa stráca lexikálne špecifikovaná fonologická informácia“ (van der Kooij – Crasborn 2016, s. 271). Výsledkom **redukcie miesta artikulácie** sú posunky, ktoré sa v citátovej podobe artikulujú na svojich definovaných miestach artikulácie a v prúde posunkovania u nich dochádza k zmene parametra miesta artikulácie v podobe posunu smerom do centra posunkového priestoru (Lachmanová 2016, s. 46 – 47). Môžeme uviesť príklady v slovenskom posunkovom jazyku, v ktorom posunky v oblasti hlavy bývajú v prúde posunkovania umiestnené nižšie: POZNAŤ – z laterálnej strany čela je umiestnený na horný okraj líca, NEVEDIEŤ (konkrétne gramatický variant posunku NEVIEM) – z laterálnej strany čela je totožne umiestnený na horný okraj líca, VERIŤ – zo spánkových častí hlavy je premiestnený na miesto pred hlavou, ZABUDNÚŤ – z čela je posunutý na miesto pred hlavou.



Obr. 8 – 9: Citátová forma NEVEDIEŤ (vľavo; zdroj: Vojtechovský 2018 – 2024), redukovaná forma NEVEDIEŤ (vpravo; zdroj: Liddell 2018).

K redukcii tvaru ruky dochádza v takom prípade, keď „v priebehu artikulácie posunku nedôjde k plnej artikulácii hodnoty parametra tvaru ruky“ (Lachmanová 2016, s. 48). Daná redukcia sa môže prejavovať v dvoch situáciách. Jednou z nich je redukcia tvaru pasívnej ruky pri dvojručných posunkoch, v ktorých sa pasívna ruka nedrží v takom tvare, ako by sa mala, ale skôr ochabuje. Napríklad slovenský dvojručný asymetrický posunok LEKÁR, pri ktorom by pasívna ruka tvoriaca plochú ruku so vztýčeným palcom mala stáť vzpriamene, ale redukuje sa – prsty smerujú nadol k zemi.



Obr. 10 – 11: Citátová forma LEKÁR (vľavo; zdroj: Vojtechovský 2018 – 2024) a redukovaný tvar pasívnej ruky (vpravo; zdroj: Lydell 2018).

Druhou z nich je redukcia tvaru ruky pri jednoručných a dvojručných posunkoch. Uvedieme dva príklady slovenských posunkov: a) DIEVČA – namiesto plochej ruky so vztýčeným palcom sa redukuje na zovretú ruku so vztýčeným ukazovákom a prostredníkom pri sebe a palcom, b) DECEMBER – namiesto mierne zakrivenej roztiahnutej ruky (ako mierne skrčená číslica 5) sa redukuje na zovretú ruku s mierne zakriveným ukazovákom a prostredníkom od seba a palcom (ako mierne pokrčená číslica 3).



Obr. 12 – 13: Citátová forma DECEMBER (vľavo; zdroj: autor) a redukovaný tvar ruky (vpravo; zdroj: autor).

Výsledkom **redukcie pohybu** sú posunky s menším pohybom než je reálny pohyb v citátovej forme. K takejto modifikácii pohybu môže dôjsť dvomi spôsobmi: buď sa redukcia vykoná v tom istom kĺbe, v ktorom sa vykonáva pohyb citátovej formy, alebo sa pohyb vykonáva v inom kĺbe (Quer a kol. 2017, s. 56). Názor- ne uvedieme príklad talianskeho i slovenského posunku VOLEJBAL. Citátová podoba VOLEJBAL sa realizuje opakovaným pohybom v lakt'ovom kĺbe a redu- kovaná forma sa vykonáva opakovaným pohybom vo vretenozápästnom kĺbe. Príkladmi zo slovenských posunkov je posunok POLÍCIA – v citátovej forme ho znázorňuje dráhový pohyb, ktorý sa začína dotykom na kontralaterálnej strane hrudníka a končí sa na ipsilaterálnej strane boku a v redukovanej forme sa vztýče- ný palec trvalo umiestňuje na stred hrudníka a vnútorným pohybom zápästia uka- zovák smeruje oblúkovým pohybom zhora nadol – z kontralaterálnej strany na ip- silaterálnu stranu. Citátová podoba PLÁVAŤ sa realizuje opakovaným pohybom v lakt'ovom kĺbe a redukovaná forma sa vykonáva opakovaným pohybom vo vre- tennozápästnom kĺbe.



Obr. 14 – 15: Citátová forma POLÍCIA (vľavo; zdroj: autor) a redukovaný dráhový pohyb ruky (vpravo; zdroj: autor).

3.3 EXTENZIA

Redukcia pohybu sa môže vyskytovať aj v takých prípadoch, keď je komuni- kant tak blízko, že mu môže stačiť malý pohyb komunikátora vykonaný na zápästí. Naopak, vo väčšej vzdialenosti, napr. 30 metrov, je potrebný výraznejší pohyb, aby sa zabezpečilo rozpoznanie posunkov. Ide o **extenziu pohybu** (rozšírenie pohybu), výsledkom tohto procesu sú posunky s väčším pohybom, než je vykázaný pohyb v citátovej forme (Quer a kol. 2017, s. 56). Všeobecnejšie povedané, pohybové mo- difikácie redukcie a extenzie sa bežne vyskytujú v spôsobe šepotu či kriku. Podobne sa v situácii šepotu môže objaviť výrazná redukcia miesta artikulácie s redukciou pohybu a zachovaným tvarom ruky/rúk, napr. pri tzv. obscénnych posunkoch, ako sú BLBEC, KRAVA, PAKO. Príslušné posunky sa realizujú v oblasti hlavy, avšak pri

tajnom šepote sa obscénne posunky posúvajú z vyššej pozície posunkového priestoru smerom nadol a artikulujú sa na nižšej pozícii posunkového priestoru.

3.4 ELÍZIA

Elízia sa chápe ako „jav, keď sa vynecháva jazykový prvok“ (Twilhaar – van den Bogaerde 2016, s. 53). **Elízia druhej ruky** sa vzťahuje na dvojručné posunky, keď sa jedna ruka artikulačne vynecháva, a tým činom sa zapája len druhá, a to dominantná ruka; tento proces fonologického vypustenia sa v niektorých odborných štúdiách označuje ako *Weak Hand Drop* alebo len *Weak Drop* (porov. Sandler 2012; van der Kooij – Crasborn 2016), čo v preklade môže znamenať **slabý pokles ruky**. Napríklad turecký posunok PLAT sa artikuluje pozdĺžnym kontaktným pohybom aktívnej ruky po pasívnej ruke. V dôsledku elízie sa úplne vynecháva pasívna ruka, ktorá je miestom artikulácie pre aktívnu ruku, a daný posunok sa artikuluje len dominantnou rukou bez zmeny iných parametrov, t. j. miesta artikulácie, tvaru ruky, pohybu a orientácie. Autori A. Bakerová a kol. (2016) vo svojej štúdii zmienili termín **redukcia druhej ruky** namiesto elízie druhej ruky.



Obr. 16 – 17: Elízia druhej ruky (Zdroj: Dikyuva a kol. 2017, s. 100).

V posunkovej komunikácii sa vyskytujú prípady, keď dvojručné posunky menia svoju podobu na jednoručné: a) posunkujúci má zaneprázdnenú ruku inou činnosťou, napr. držaním tašky, volantú vozidla, držadla v dopravnom prostriedku, b) posunkujúci súbežne artikuluje dva posunky a každá ruka referuje k inému posunkú (simultánna konštrukcia), napr. PO-PRVÉ (ľavá ruka) – ČÍTAŤ (pravá ruka), PO-DRUHÉ (ľavá ruka) – PÍSAŤ (pravá ruka), PO-TRETIE (ľavá ruka) – VYJADRIŤ-SA (pravá ruka), c) posunkujúci má dva voľné artikulátory k artikulácii, a predsa

posunkuje jednou rukou, čo môže byť ovplyvnené mimojazykovými faktormi: momentálne psychické rozpoloženie posunkujúceho (únava, emócie), postoj posunkujúceho ku komunikačnému aktu a snaha o nenápadnosť či o šepot (Battison 1978, cit. podľa Lachmanová 2016).

3.5 METATÉZA

Piatym fonologickým procesom je **metatéza**. „Je to proces, pri ktorom sa mení poradie foném v slove/posunkú. V posunkových jazykoch môže tento proces ovplyvniť posunky zahŕňajúce zmenu miesta artikulácie v dôsledku pohybu“ (Quer a kol. 2017, s. 59). Príkladom na ňu je posunok NEPOČUJÚCI v americkom posunkovom jazyku (Lucas a kol. 2001, cit. podľa Quer a kol. 2017). V citátovej forme sa tento posunok so zovretou päťou a vztýčeným ukazovákom artikuluje oblúkovým pohybom z miesta ucha na ipsilaterálnu stranu úst. V podobe metatézy je poradie oboch kontaktných bodov obrátené, čiže ukazovák sa najprv dotkne ipsilaterálnej časti úst a potom ucha. Aj v slovenskom posunkovom jazyku v prúde posunkovania môže dochádzať k prešmyknutiu poradia segmentov miesta artikulácie, napr. SÁM, NORMÁLNY, v ktorom posunkujúci pravák využíva citátové podoby z kontralaterálnej strany hrudníka oblúkovým pohybom na ipsilaterálnu stranu hrudníka, v procese metatézy sa zmenia miesta artikulácie a s tým súvisí zmena smeru pohybu. Autori A. Bakerová a kol. (2016) a J. N. Twilhaar – B. van den Bogaerde (2016) vo svojich štúdiách spomínali termín **asimilácia pohybu** namiesto metatézy.



Obr. 18 – 19: Citátová forma SÁM (vľavo; zdroj: autor) a prešmyknutá forma posunkú (vpravo; zdroj: autor).

Autori vo svojich štúdiách uviedli metatézu vzťahujúcu sa len na zmenu poradia segmentov – miest artikulácie v posunkoch súvisiacich s pohybom. Fonolo-

gické javy v slovenskom posunkovom jazyku potvrdzujú závery D. Lachmanovej (2016), podľa ktorej by teoreticky bolo možné uvažovať, že proces metatézy v rámci jedného posunkku môže prebiehať vo všetkých parametroch, ktoré sú obsadené minimálne dvoma hodnotami. V prúde posunkovania potom dochádza k zámene poradia či smeru artikulácie hodnôt. Lachmanová (tamže, s. 53) vo svojej práci spomína **metatézu tvaru ruky** (ide o segmentálnu zmenu poradia z jednotlivých tvarov rúk na druhý tvar v rámci jednej lexikálnej jednotky) a **zámenu dominantnej ruky** (týka sa výmeny dominantnej ruky, teda namiesto pravej ruky ľavá). Naopak, prieskum slovenského posunkového jazyka potvrdil, že niektorí slovenskí posunkujúci využívajú **metatézu pohybu**, ak sa kruhové pohyby rúk vykonávajú obrátene, a **metatézu orientácie ruky**, ak orientácia ruky smeruje iným smerom oproti citátovej podobe posunkku. Uvedieme štyri príklady. Citátová forma HLADAŤ, pri ktorej sa opakovaný kruhový pohyb dvoch rúk vo vertikálnej rovine začína najprv nadol, následne nahor. V podobe metatézy je opakovaný kruhový pohyb dvoch rúk obrátený, začínajúci nahor, následne nadol. Druhý príklad citátovej podoby OPAKOVÁŤ, pri ktorej sa opakovaný kruhový pohyb dvoch rúk v sagitálnej rovine začína nahor a dopredu od tela a vo forme metatézy je opakovaný kruhový pohyb dvoch rúk opačný, teda začína nahor a dozadu k telu. Ďalšia forma UVIDIETĚ v zmysle „ešte to nie je isté“ sa realizuje opakovaným priamočiarym pohybom k telu v sagitálnej rovine a orientácia dlane smeruje k nosu a v podobe metatézy sa forma posunkku zmení orientáciou v uhle o 90 stupňov a je smerovaná doľava. Taký identický prípad sa vyskytuje aj pri posunkku POZNAŤ, avšak sa realizuje na spánkovej časti čela.



Obr. 20 – 21: Citátová forma HLADAŤ (vľavo; zdroj: autor) a prešmyknutá forma posunkku (vpravo; zdroj: autor).



Obr. 22 – 23: Štylistická forma UVIDIEŤ (vľavo; zdroj: autor) a prešmyknutá forma posunku (vpravo; zdroj: autor).

3.6 NATURALIZÁCIA

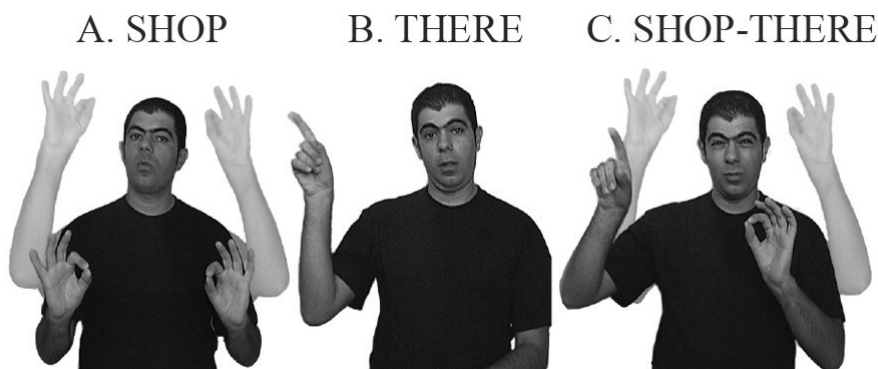
Šiestym fonologickým procesom je **naturalizácia** (nativizácia), ktorú chápeme takto: „ak fonéma vypožičaného posunku nepatrí do fonematického inventára cieľového jazyka, adaptačné procesy môžu zmeniť pôvodný tvar posunku s cieľom splniť požiadavky na správnu tvorbu“ (Quer a kol. 2017, s. 59). Príkladom adaptácie v talianskom posunkovom jazyku je posunok WORKSHOP, ktorý je prevzatý z amerického posunkového jazyka. V posunku východiskového jazyka sa používa tvar ruky znázornený ako zovretá päť so vztýčeným ukazovákom, prostredníkom a prstenníkom od seba. Aby bol daný posunok kompatibilný s fonológiou talianskeho posunkového jazyka, posunkujúci produkujú WORKSHOP s prirodzeným tvarom ruky – rozťahnutá ruka so založeným palcom (ako číslica 4). Príkladom prispôsobenia sa v slovenskom posunkovom jazyku je prevzatý bieloruský posunok MINSK, ktorý ilustruje zovretú päť so vztýčeným ukazovákom, prostredníkom a prstenníkom pri sebe, dotýkajúc sa pritom ipsilaterálnej časti prstov. Avšak daný tvar ruky je upravený na tvar ruky znázorňujúci plochú ruku so založeným palcom, čiže na artikulácii sa zúčastňuje aj malíček.



Obr. 24 – 25: Bieloruský posunok MINSK (vľavo; zdroj: autor) a slovenský posunok MINSK (vpravo; zdroj: autor).

3.7 KOALESCENCIA

Ďalším fonologickým procesom je **koalescencia**, ktorú podľa J. Quera a kol. (2017, s. 55) chápeme ako „splnutie dvoch fonetických jednotiek do jednej“. Príkladom je klitizácia zámen v izraelskom posunkovom jazyku, v ktorom sa aktívna ruka symetrického dvojručného posunkú môže stať nositeľom deiktického posunkú (Sandler 1999, cit. podľa Quer a kol. 2017). Autori Quer a kol. (2017, s. 803) tvrdia, že „koalescencia sa vzťahuje na špeciálny typ klitizácie; najčastejšie ide o klitizáciu deiktického posunkú k predchádzajúcemu symetrickému dvojručnému posunkú, takže vzniká jedno prozodické slovo.“ Termínom klitizácia sa v lingvistike posunkových jazykov označuje splnutie dvoch (alebo viacerých) slov v syntaktickom reťazci (Quer a kol. 2017, s. 63). Iná definícia klitizácie podľa J. Quera a kol. (2017, s. 803) ju vysvetľuje ako „proces, pri ktorom sa funkčný prvok fonologicky pripája k lexikálnemu prvku tak, že vzniká jedno prozodické slovo (napr. anglické *can't* a francúzske *j'aime*); funkčný prvok sa označuje ako klitika“.



Obr. 26: Koalescencia dvoch fonetických jednotiek do jednej (Sandler 2006, s. 198).

V citátovej podobe sa izraelský posunok OBCHOD artikuluje ako symetrický dvojručný posunok. Daný posunok spĺňa fonologickú požiadavku na vznik koalescencie. Obrázok na pravej strane demonštruje, ako sa realizuje spojenie, teda počas prechodu z jedného posunkú na druhý, v ktorom jedna aktívna ruka dvojručného posunkú mení svoj tvar na deiktický posunok určujúci miesto TAM, čím sa vytvára spojenie OBCHOD-TAM.

3.8 PERSEVERÁCIA A ANTICIPÁCIA

V posunkových jazykoch môžeme identifikovať ďalšie fonologické procesy, a to **perseveráciu** a **anticipáciu nedominantnej ruky**. V posunkovej komunikácii sa realizujú jednoručné a dvojručné posunky. Keď po dvojručnom posunku nasleduje jednoručný posunok, často sa druhá nedominantná ruka posunku drží na tom istom mieste po predchádzajúcom dvojručnom posunku namiesto toho, aby sa vrátila do pokojovej polohy nižšie pred telom, teda pred bruchom. Tento jav sa nazýva **perseverácia**. Naopak, keď po jednoručnom posunku nasleduje dvojručný posunok, často sa druhá nedominantná ruka, ktorá sa drží v pokojovej polohe pred telom, pripravuje na dvojručný posunok, ktorý bude nasledovať (Johnston – Schembri 2007).

3.9 EPENTÉZA A ELÍZIA

V lingvistike posunkového jazyka sa termínom **epentéza** chápe vkladanie fonematického prvku s cieľom upraviť nevhodne utvorené slabičné štruktúry (Quer a kol. 2017, s. 60), teda pridávať pohyb. Z toho vyplýva, že všetky posunky by mali obsahovať jasne viditeľný pohyb. Typy epentetického pohybu sú: opakovaný pohyb a priamočiary pohyb. Príkladom epentézy pohybu je taliansky posunok HLAVA, ktorého citátová forma neobsahuje žiadnu pohybovú zložku, len držanie ukazováka smerujúceho na laterálnu časť čela. Avšak kvôli požiadavkám na manuálnu sonoritu správne tvorených slabík sa v tomto posunkovom jazyku zavádza epentetický krátky opakovaný pohyb smerujúci na laterálnu časť čela (Geraci 2009, cit. podľa Quer a kol. 2017). Epentézu používa aj časť slovenských posunkujúcich, keď namiesto dvojslabičného posunku BANÁN použijú trojslabičný posunok, ktorým sa motorickou ikonickosťou znázorňuje trojnásobné šúpanie banána zhora nadol.

Ďalší prípad **pohybovej epentézy sui generis** predstavuje prechod medzi posunkami, pri ktorých sa artikulácia nepreruší. V podstate spočíva v premiestnení ruky/rúk z ukončenej fázy konečného miesta artikulácie posunku do počiatočného miesta nasledujúceho posunku. Počas tohto prechodu ruky zvyčajne menia tvar a orientáciu na posunok, ktorý sa má artikulovať (Quer a kol. 2017). V tejto súvislosti sa realizuje **elízia segmentu držania** ako ďalšieho segmentu v posunku. Posunky, ktoré sa obvykle vytvárajú s držaním ako súčasť ich citátových foriem, sa môžu v súvislej produkcii viacerých posunkov posunkovať inak, takto sa môže realizovať elízia segmentu držania medzi posunkami a, naopak, pribudne epentetický segment pohybu. Napríklad austrálsky posunok NEVEDIEŤ sa tento posunok v citátovej forme so segmentmi vykazuje ako HMM (poznámka: *H* – hold – držanie, *M* – movement – pohyb). Naproti tomu ten istý posunok v posunkovanej výpovedi nemusí vykazovať žiadne držanie. Tri posunky IX₁, NEVEDIEŤ a KDE by sa vytvárali ako súvislý prúd pohybov, pričom medzi posunkami by sa pridávali prechodné pohyby ako súčasť procesu pohybovej epentézy. Na obrázku je znázornený príklad v zjednodušenom zápise pohybu a držania (Johnston – Schembri 2007, s. 115).

Sign	PRO-I		NOT-KNOW		WHERE
Citation form	M	H	H	M	H
Movement epenthesis	M	H	M	H	M
Hold deletion	M	M	M	M	M

Obr. 27: Reprezentácia pohybu a držania pri elízii držania a epentéze pohybu (Jonhston – Schembri 2007, s. 115).

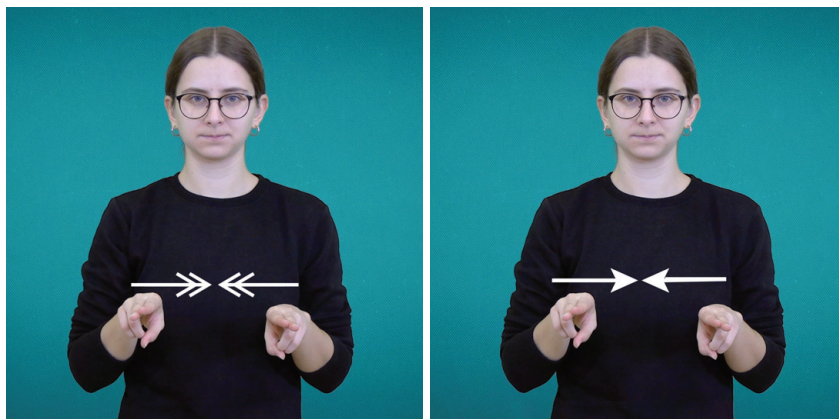
V slovenskom posunkovom jazyku sa tento jav realizuje napr. vo výpovedi zahŕňajúcej dva slovenské posunky MYŠ a KLOPAŤ (*Myš klope.*), pri ktorých sa v oboch v začiatkovej a konečnej fáze citátových podôb vykazuje segment držania. Pri posunkovom prejave sa plynule prechádza z jedného na druhý, preto sa medzi posunkami nevykazujú segmenty držania a namiesto nich sa vkladá segment prechodného pohybu.



Obr. 28: Epentetický pohyb medzi dvoma posunkami MYŠ a KLOPAŤ (Zdroj: Snepeda 2015 – 2024).

Elízia segmentu pohybu je fonologický proces, pri ktorom sa vypúšťa segment pohybu v posunku, aby sa posunok mohol realizovať rýchlejšie a s menšou manuálnou aktivitou (porov. Dikyuva a kol. 2017, s. 99). Elízia pohybu sa často vyskytuje pri opakovaných pohyboch posunkov a vedie k produkcii posunkov s jedno-

duchým pohybom (Johnston – Schembri 2007, s. 115). V lexike slovenského posunkového jazyka sa realizuje napríklad v posunku BRAT, keď sa citátová forma vyznačuje opakovaným priamočiarym pohybom a pri prejave sa namiesto opakovaného môže vyskytovať jednoduchý priamočiary pohyb.



Obr. 29 – 30: Citátová forma posunku BRAT a upravená forma posunku BRAT (Zdroj: Vojtechovský a kol. 2024).

3.10 KOARTIKULÁCIA

Medzi fonologické procesy v posunkovom jazyku sa zaraďuje aj **koartikulácia**, definovaná podľa J. N. Twilhaara a B. van den Bogaerdeovej (2016, s. 31) ako „fonetická variácia artikulácie ovplyvnená jazykovým faktorom slov alebo posunkov v jeho bezprostrednom okolí.“ Tí istí autori (2016, s. 146) uviedli ďalšiu definíciu: „Koartikulácia je forma fonetickej variácie, ktorá zahŕňa miernu úpravu artikulácie“ a exemplifikovali ju na príklade holandského posunku DOBRÝ, ktorý môže mať rôzne umiestnenie v závislosti od posunkov, ktoré mu predchádzajú alebo po ňom nasledujú. Posunok sa môže produkovať vyššie, ako je určený v citátovej forme, keď sa kombinuje s výrazom VEČER, ako je napríklad fráza DOBRÝ VEČER. Pozícia začiatočného posunku VEČER v posunkovom priestore je umiestnená vyššie ako pozícia posunku DOBRÝ, preto sa pozícia daného posunku presúva na vyššiu pozíciu, na ktorej sa artikuluje VEČER. Ide o príklad regresívnej koartikulácie, keďže sa prispôsobuje nasledujúcemu posunkovi.

Podobným spôsobom sa v slovenskom posunkovom jazyku realizujú dve slovenské výpovede zložené z troch posunkov, v ktorých sa miesto artikulácie mení vplyvom miest artikulácie okolitých posunkov a pripodobní sa k miestam artikulácie predchádzajúcich a nasledujúcich posunkov. V prvej výpovedi KUCHÁR VEDIEŤ₃ VARIEŤ (*Kuchár vie varieť.*), v ktorej KUCHÁR a VARIEŤ sú artikulované pred telom a VEDIEŤ₃ v oblasti hlavy, sa vplyvom dvoch susediacich posunkov posúva miesto

artikulácie posunkú VEDIĚ₃, a to vo vertikálnej rovine výrazne nižšie. Ide o obojsmernú koartikuláciu. V druhej výpovedi MATKA VARĚ OBED (*Matka varĚ obed.*) sú posunky MATKA a OBED realizované v časti hlavy a VARĚ pred telom – vplyvom dvoch susediacich posunkov sa miesto artikulácie posunkú VARĚ posúva vo vertikálnej rovine vyššie.

Uvedení autori priblížili jestvujúci rozdiel medzi asimiláciou a koartikuláciou v tom zmysle, že „existuje možnosť, že asimiláciu možno v určitých prípadoch považovať za koartikuláciu. Vo všeobecnosti sa však koartikulácia vzťahuje na malé prispôbenie sa v artikulácii, keď napríklad realizácia miesta artikulácie alebo tvaru ruky pre posunok vykazuje miernu fonetickú odchýlku od citátovej formy posunkú. Tieto realizácie sú vždy v tej istej fonéme ako v citátovej forme posunkú. Pri asimilácii je vždy fonologická adaptácia, pri ktorej realizácia parametra vedie k inej fonéme ako v citátovej forme posunkú.“

Môžeme teda zhrnúť, že koartikulácia je malé prispôbenie (fonetická variácia) a asimilácia je fonologická adaptácia, ktorá vedie k inej fonéme.

4. ZÁVER

Posunkové jazyky rovnako ako všetky orálno-auditívne jazyky majú fonetickú a fonologickú rovinu. Tieto roviny tvorí systém diskretných formálnych prvkov, ktoré sa špecifickými spôsobmi kombinujú pri vytváraní jednotiek s lexikálnym významom. Na základnej úrovni sa skladajú z výstavbových prvkov posunkového jazyka, ako sú tvar ruky a jej orientácia, miesto artikulácie, pohyb a nemanuálne prostriedky. Tieto parametre posunkú predstavujú v slovenskom posunkovom jazyku minimálne dištinktívne jednotky s fonologickou platnosťou. Sú to stavebné prvky posunkového kódu, ktoré fungujú analogicky ako fonémy v orálno-auditívnych jazykoch: samy o sebe nenesú lexikálny význam, ale ich zámena alebo zmena v rámci posunkú vedie k zmene lexikálneho významu, ako to ilustrujú minimálne páry uvedené vyššie. V tejto súvislosti je esenciálne aj v posunkovom jazyku vnímať a rozlišovať vzťah fonetiky a fonológie. Fonológia vyberá a klasifikuje tie motoricko-vizuálne realizácie, ktoré majú v posunkovom kóde dištinktívnu funkciu. Fonetika, naopak, popisuje celý rozsah fyzických a artikulačných realizácií vrátane nedištinktívnych variácií.

Venovali sme sa predovšetkým fonologickej adaptácii, teda formám posunkov, ktoré sa môžu meniť v dôsledku fonologických procesov, ako sú asimilácia, redukcia, elízia, metatéza, epentéza, extenzia, naturalizácia, koalescencia, perseverácia a anticipácia. Zdokumentovali sme, že v slovenskom posunkovom jazyku sa nachádzajú takmer všetky tieto fonologické procesy, a to konkrétne: asimilácia tvaru ruky, redukcia pohybu, miesta artikulácie a tvaru ruky, extenzia pohybu, elízia segmentu držania, segmentu pohybu a druhej ruky (redukcia druhej ruky), epentéza pohybu, metatéza miesta artikulácie, pohybu a orientácie ruky, naturalizácia, koalescencia, perseverácia nedominantnej ruky a anticipácia nedominantnej ruky. K týmto svoj-

bytným fonologickým procesom možno priradiť ďalšie procesy, ktoré sme zistili prostredníctvom prieskumnej sondy, konkrétne redukciu miesta artikulácie, redukciu tvaru ruky, metatézu pohybu a metatézu orientácie ruky. Fonologické zmeny sa prejavujú v rámci lexikálnych jednotiek – posunkov. Z týchto procesov sme priblížili aj koartikuláciu – formu fonetickej variácie, ktorá sa v slovenskom posunkovom jazyku vyskytuje častejšie než asimilácia miesta artikulácie. Poznanie týchto fonologických procesov je dôležité v pragmatickej komunikácii v posunkovom jazyku, keďže v intrakultúrnej komunikácii posunkujúcich sa často vyskytuje tendencia k zjednodušenej artikulácii posunkov alebo k (častočnej) zmene artikulácie posunkov, ktoré sa realizujú inak ako ich citátová podoba. Menšia posunková artikulácia alebo zmenená artikulácia stoja menej úsilia manuálnej aktivity, sú prejavom ekonomizácie posunkového prejavu, najmä ak komunikant má z kontextu dostatok informácií na rozpoznanie posunkov, teda redukovaná alebo zmenená artikulácia nespôsobuje v komunikácii entropiu. Pozorovanie prejavu komunikátora, ktorý komunikuje v posunkovom jazyku, predpokladá od príjemcu schopnosť pozorovať celkový prejav artikulovaných pohybov zahŕňajúcich ruky, paže, hlavu, telo a tvár.

Bibliografia

- ANDERSON, Catherine – BJORKMAN, Bronwyn – DENIS, Derek – DONER, Julianne – GRANT, Margaret – SANDERS, Nathan – TANIGUCHI, Ai (2022): *Essentials of Linguistics*. 2. vyd. Hamilton: eCampusOntario. 884 s.
- BAKER, Anne. – VAN DEN BOGAERDE, Beppie – PFAU, Roland – SCHERMER, Trude (2016): *The Linguistics of Sign Languages : An Introduction*. Amsterdam – Philadelphia: John Benjamins Publishing Company. 378 s.
- BATTISON, Robbin (1978): *Lexical borrowings in American Sign Language*. Silver Spring: Linkstok Press. 240 s.
- BRENTARI, Diane (2012): Phonology. In: R. Pfau – M. Steinbach – B. Woll (eds.): *Sign Language: An International Handbook*. Berlín: Mouton de Gruyter, s. 21 – 54.
- CRASBORN, Onno – van der KOOIJ, Els (2016): Phonetics. In: A. Baker – B. van den Bogaerde – R. Pfau – T. Schermer (eds.): *The Linguistics of Sign Languages: An Introduction*. Amsterdam – Philadelphia: John Benjamins Publishing Company, s. 229 – 250.
- CSONKA, Štefan (1995): Osvojovanie posunkovej reči počujúcimi. In: *Efeta – otvor sa*, roč. 5, č. 3, s. 2 – 5.
- DIKYUVA, Hasan – MAKAROGLU, Bahtiyar – ARIK, Engin (2017): *Turkish Sign Language Grammar*. Ankara: The Republic of Turkish Ministry of Family and Social Policies. 318 s.
- EMMOREY, Karen (2007). The psycholinguistics of signed and spoken languages: How biology affects processing. In: M. G. Gaskell (ed.): *The Oxford Handbook of Psycholinguistics*. Oxford University Press. Chapter 43, s. 703 – 722. Dostupné na: www.researchgate.net/publication/283518841_The_psycholinguistics_of_signed_and_spoken_languages_How_biology_affects_processing. [cit. 26-10-2024].

FENLON, Jordan – CORMIER, Kearsy – BRENTARI, Diane (2018): *Sign language phonology*. Dostupné na: www.researchgate.net/publication/326293270_Sign_language_phonology. [cit. 26-10-2024].

GERACI, Carlo (2009): Epenthesis in Italian Sign Language. In: *Sign Language & Linguistics*, roč. 12, č. 1, s. 3 – 51.

JOHNSTON, Trevor – SCHEMBRI, Adam (2007): *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press. 324 s.

KLOMP, Ulrike (2021): *A descriptive grammar of Sign Language of the Netherlands*. Amsterdam: Universiteit van Amsterdam. 407 s.

LACHMANOVÁ, Denisa (2016): *Změny ve výšce realizace znaků českého znakového jazyka vlivem koartikulace*. [diplomová práce]. Praha: Filozofická fakulta Univerzity Karlovy. 146 s.

LIDDELL, Scott K. – JOHNSON, Robert, E. (1989): American Sign Language: The Phonological Base. In: *Sign Language Studies*, roč. 18, č. 64, s. 195 – 277.

LUCAS, Ceil – BAYLEY, Robert – VALLI, Clayton (2001): *Sociolinguistic Variation in American Sign Language*. Washington: Gallaudet University Press. 256 s.

LYDELL, Thomas a kol. (2018): *Spreadthesign*. Örebro: European Sign Language Center. Dostupné na: www.spreadthesign.com/sk.sk/search/. [cit. 26-10-2024].

PARKHURST, Stephen (2008): Phonology, the ebb and flow of two great forces. In: H. van der Hulst – R. Channon (eds.): *The Phonetics and Phonology of Sign Languages First SignTyp Conference*. Storrs: University of Connecticut, s. 36 – 37. Dostupné na: www.academia.edu/1151797/The_Phonetics_and_Phonology_of_Sign_Languages_First_SignTyp_Conference_June_26_28_2008. [cit. 26-10-2024].

PARKHURST, Stephen (2008): The role of the input structure in sign language phonology. In: H. van der Hulst – R. Channon (eds.): *The Phonetics and Phonology of Sign Languages First SignTyp Conference*. Storrs: University of Connecticut, s. 38. Dostupné na: www.academia.edu/1151797/The_Phonetics_and_Phonology_of_Sign_Languages_First_SignTyp_Conference_June_26_28_2008. [cit. 26-10-2024].

QUER, Josep – CECCHETTO, Carlo – DONATI, Caterina – GERACI, Carlo – KELEPIR, Meltem – PFAU, Roland – STEINBACH, Markus (2017): *SignGram Blueprint: A Guide to Sign Language Grammar Writing*. Berlín – Boston: Walter de Gruyter BmbH. 896 s.

SANDLER, Wendy (1989): *Phonological representation of the sign. Linearity and non-linearity in American Sign Language*. Dordrecht: Foris. 228 s.

SANDLER, Wendy (1999): *Cliticization and prosodic words in a sign language*. In: T. Hall – U. Kleinhenz (eds.): *Studies on the phonological word*. Amsterdam: John Benjamins, s. 223 – 254.

SANDLER, Wendy (2006): Phonology, phonetics and the nondominant hand. In: L. Goldstein – D. H. Whalen – C. Best (eds.): *Laboratory Phonology 8*. Berlín: De Gruyter Mouton, s. 185 – 212. Dostupné na: https://signlab.haifa.ac.il/wp-content/uploads/2013/12/Phonetics_Phon_and_Nondominant_hand1.pdf. [cit. 26-10-2024].

SANDLER, Wendy (2011): The Phonology of Movement in Sign Language. In: M. Oostendorp – C. J. Ewen – E. Hume – K. Rice (eds.): *The Blackwell Companion to Phonology*, s. 1 – 27. Dostupné na: <https://doi.org/10.1002/9781444335262.wbctp0024>. [cit. 26-10-2024].

SANDLER, Wendy (2012): The Phonological Organization of Sign Languages. In: *Language and Linguistics Compass*. roč. 6, č. 3, s. 162 – 182. Dostupné na: <https://pubmed.ncbi.nlm.nih.gov/23539295/>. [cit. 26-10-2024].

SANDLER, Wendy – LILLO-MARTIN, Diane (2006): *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press. 570 s.

Snepeda (2015 – 2024): *Domček, domček, kto v tebe býva?* Bratislava: Snepeda. Dostupné na: www.snepeda.sk/Domcek-domcek-kto-v-tebe-byva.html. [cit. 26-10-2024].

STOKOE, William C. (1960): Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. In: *Studies in Linguistics*. Occasional papers, č. 8, s. 1 – 78.

TWILHAAR, Jan Nijen – VAN DEN BOGAERDE, Beppie (2016): *Concise Lexicon for Sign Linguistics*. Amsterdam – Philadelphia: John Benjamins Publishing Company. 244 s.

VAN DER KOOIJ, Els (2002). *Phonological Categories in Sign Language of the Netherlands: The role of Phonetic Implementation and Iconicity*. Utrecht: LOT. 327 s.

VAN DER KOOIJ, Els – CRASBORN, Onno (2016): Phonology. In: A. Baker – B. van den Bogaerde – R. Pfau – T. Schermer (eds.): *The Linguistics of Sign Languages: An Introduction*. Amsterdam – Philadelphia: John Benjamins Publishing Company, s. 251 – 278.

VOJTECHOVSKÝ, Roman. (ed.) (2018 – 2024): *Posunky.sk: Prekladový slovník slovenského jazyka – slovenského posunkového jazyka*. Bratislava: Snepeda. Dostupné na: www.posunky.sk. [cit. 26-10-2024].

VOJTECHOVSKÝ, Roman – MAJERČÁKOVÁ, Diana – HEFTY, Michal – VOJTECHOVSKÁ, Veronika – CEHLÁRIK, Jaroslav – HEFTY, Angela (2024): *Slovenský posunkový jazyk pre študentov – úplných začiatčníkov I.: Ja a moja rodina*. Trnava: Pedagogická fakulta Trnavskej univerzity. 242 s. Dostupné na: <https://pdfweb.truni.sk/veda-vyskum?e-kniznica#online;online=posunkov>. [cit. 26-10-2024].

RECENZIE

PUCHOVSKÁ, Zuzana: *SLOVESNÝ PRÍTOMNÝ A MINULÝ ČAS Z FRANCÚZSKO-SLOVENSKEJ KONTRASTÍVNEJ PERSPEKTÍVY*. Bratislava: Univerzita Komenského v Bratislave, 2024, 264 s. ISBN: 978-80-223-5796-8.

V čase globalizácie, v období multikulturalizmu a rapídneho rozmachu informačných technológií, ktoré sprostredkujú takmer okamžitý a nepretržitý prístup k informáciám, nastali v mnohých oblastiach duševnej činnosti ľudí výrazné zmeny, spôsobené okrem iného narastajúcim diapazónom dostupných nástrojov a informačných systémov. Medzi intelektuálne činnosti zasiahnuté technologickým pokrokom s určitosťou patrí aj osvojovanie si cudzieho jazyka, jeho aktívne využívanie v komunikácii a práca s cudzojazyčným textom. Zmeny sa odohrávajú nielen v procesoch a možnostiach výučby, ale aj pri tvorivej práci s jazykom v praxi, čo sa prejavuje rozmachom používania strojového prekladu či stále častejším využívaním veľkých jazykových modelov. V informatizovanom svete sa tak bežné vnímanie jazyka začalo čoraz väčšmi zužovať na pragmatické využitie.

Na pozadí týchto skutočností sa objavuje monografia Zuzany Puchovskej s názvom *Slovesný prítomný a minulý čas z francúzsko-slovenskej kontrastívnej perspektívy* (Bratislava: Univerzita Komenského 2024). Autorka si v nadväznosti na nástup nových technológií vši- ma, že ich zavádzanie posúva „do údia potrebu skutočne jazyk poznať a rozumieť mu“ (s. 11), a preto sa rozhodla zvoliť iný prístup. Jazyk nereflektuje len ako systém, ale zdôrazňuje potrebu neutracať zo zreteľa ani jeho filozofický rozmer. V monografii sa tak neobmedzuje na tradičný gramatický opis francúzskeho slovesného systému, ale potenciálnym čitateľom ponúka kontextualizovaný opis francúzskeho prítomného času, minulých časov a čias- točne aj kondicionálu, a to cez prizmu slovenského jazyka.

Takéto uchopenie slovesného času umožňuje do jazyka preniknúť hlbšie, než dovoľí formálna deskripcia morfosyntaktických štruktúr. Monografia má preto potenciál svojím obsahom obohatiť nielen študentov francúzštiny, ktorí si zákonitosti cudzieho jazyka ešte len osvojujú, ale aj skúsených používateľov jazyka, akými sú pedagógovia či prekladatelia, a to práve vďaka kontextualizovanej perspektíve.

Autorka pri opise slovesných časov pracuje primárne s korpusom pôvodných prozaic- kých textov a ich prekladov do slovenčiny, keďže ako sama v predhovore argumentuje, „prá- ve na tomto type textov sa naplno prejaví štylisticko-pragmatický potenciál a komplexná ja- zyková povaha francúzskeho prítomného času a jednotlivých minulých časov“ (s. 9). Čitate- ľovi sa tak do rúk dostáva podnetný materiál, ktorý mu umožní zamyslieť sa nad rôznorodý- mi spôsobmi zachytenia reality v dvoch jazykových systémoch prostredníctvom jedinečných štylistických a pragmatických vyjadrovacích možností gramatickej kategórie času.

Tie sa síce v niečom prelínajú, no v mnohom sa na nich odzrkadľuje odlišný spôsob uvažovania o skutočnosti. Ako vysokoškolská pedagogička sa Z. Puchovská dlhodobo venuje didaktizácii francúzskej gramatiky, lingvistiky a francúzsko-slovenskej kontrastívnej grama- tiky, vďaka čomu problémové javy v monografii nielen identifikuje, ale uvažuje aj o príči-

nách ich ťažkej osvojiteľnosti pre slovenských študentov, zdôrazňuje javy, ktoré pri výučbe zostávajú zanedbávané, pričom zo zreteľa nespúšťa ani prekladové hľadisko, a preto sa systematicky snaží ponúknuť aj možné spôsoby vysporiadania sa s javmi, ktoré predstavujú prekladateľský problém.

Monografia sa skladá z troch častí: prvá nesie názov *Úvod do problematiky*, kde autorka ponúka ukotvenie kategórie slovesného času a spôsobu v systéme jazyka v širších súvislostiach, približuje použité metodologické prístupy a vysvetľuje východiskovú terminológiu. Pri opise uplatňuje prizmu slovenčiny ako materinského jazyka vymedzujúceho východiskovú koncepciu vnímania reality. Autorka v jazykových systémoch hľadá paralely, s väčším dôrazom sa však venuje vysvetleniu jazykovej podstaty javov, ktoré sú odlišné a ťažko osvojiteľné. Hoci úvodná časť monografie môže v súvislosti s prehľadnými kategorizáciami, príkladmi a vysvetleniami pôsobiť učebnicovo, výklad je popretkávaný úvahami o náročnosti jazykových jednotiek, o materinskom jazyku, o jazyku ako duchovnej hodnote. Toto všetko opis gramatických štruktúr posúva do podnetnej reflexívnej roviny.

Druhá časť monografie s názvom *Prítomný čas a minulé časy vo francúzštine* sa venuje rozsiahlemu opisu základnej časovo-vidovej a sekundárnej štylisticko-pragmatickej povahy jednotlivých časov. Autorkiným zámerom nie je objasniť pravidlá formálnej tvorby morfológických tvarov slovies v jednotlivých časoch, túto stránku preto úplne vynecháva, a pozornosť upriamuje na ozrejenie charakteru prítomného času a minulých časov z hľadiska ich podstaty, funkcie a rôznych spôsobov uplatnenia v jazyku, a to vo svetle slovenčiny. Práve túto časť monografie ocenia najmä (no nielen) prekladatelia, keďže ich úloha spočíva v citlivom vnímaní štylisticko-vidového potenciálu použitia časov v texte, aby sa v cieľovom jazyku zachoval porovnateľný efekt na čitateľa. Jednotlivé štylistické využitia časov autorka ilustruje na úryvkoch z prozaických diel francúzskych autorov, ktoré sprevádza ich slovenský preklad, analýza, vysvetlenie, identifikácia potenciálnych prekladateľských problémov a ponuka možností riešení v medziach slovenčiny.

V poslednej časti monografie – *Časová súslednosť v nepriamej a polopriamej reči* – autorka približuje komplexný charakter francúzskeho fenoménu, ktorého základnými výstavbovými prvkami sú práve slovesné časy a kondicionál. Ide o súslednosť časov – *concordance des temps*. Autorka vníma jeho existujúce vysvetlenia v nefrankofónnych gramatikách ako nedostatočné či „nejasné“. Keďže v slovenčine tento fenomén úplne abscentuje, považuje za dôležité časovú súslednosť objasniť „cez prizmu jej úlohy či funkcie v jazyku, a nie cez pravidlá, na ktoré sa v jazykových príručkách veľmi často redukuje“ (s. 201). Analýzou jazykovej skutočnosti, ktorá sa týmto termínom označuje, Z. Puchovská prichádza k záveru, že ide o *syntakticko-pragmatickú kongruenciu slovesného času a spôsobu*. Následne tento jav približuje z hľadiska jeho využitia v prozaických textoch, kde môže predstavovať prekladateľský problém, keďže vystupuje ako jeden z indikátorov nepriamej a polopriamej reči. Od jeho citlivej detekcie a správnej interpretácie závisí zachovanie významu v preklade tak, aby nedošlo k významovým posunom. Aj v tejto časti autorka opis demonštruje na úryvkoch prozaických diel a ich znení v slovenčine, pričom neprestáva reflektovať možnosti, ktorými prekladateľ disponuje, a ponúka analýzou podložené riešenia problematických pasáží.

Celá monografia obsahuje bohatý poznámkový aparát pod čiarou, v ktorom sa okrem iného uvádzajú doplnujúce informácie, odkazy na ďalšiu literatúru, vysvetlenia, ale aj autorčine postrehy či úvahy. Na konci publikácie sa okrem záveru, bibliografie a menného registra nachádzajú aj prílohy pozostávajúce zo sumarizačných tabuliek, ktoré prispievajú k celkovej prehľadnosti. Hoci by sa dalo namietat, že vynechanie opisu budúceho času narúša celistvosť

výkladu, autorka monografie konštatuje, že na rozdiel od prítomného a minulého času, ktoré sa spájajú s vyjadrením reálnej skutočnosti, je budúci čas prostriedkom na vyjadrenie nádejí, obáv, neistôt, ktoré nás posúvajú do „fiktívneho sveta“, pričom argumentuje, že budúci čas preto treba „analyzovať nielen v súvislosti s vyjadrením času, ale aj s vyjadrením modálneho postoja hovoriaceho, a prepojiť ho tak s podmieňovacím spôsobom a subjunktívom“ (s. 240). To je aj dôvod, prečo si opis tohto času vyžaduje samostatný priestor na kontrastívnu analýzu.

Na záver možno konštatovať, že Z. Puchovská vo svojej monografii ponúka čitateľovi detailný opis fungovania prítomného času a minulých časov vo francúzštine na kontrastívnom základe z perspektívy slovenčiny ako materinského jazyka. Autorka sa opiera o svoje skúsenosti s výučbou francúzskeho jazyka a neustále mala na pamäti praktické implikácie, ktoré z konfrontácie týchto dvoch jazykových systémov vyplývajú pre prácu prekladateľa, čím sa zaručuje, že informácie obsiahnuté v monografii budú relevantné rovnako pre študentov, ako aj jazykových profesionálov.

Martina Vladovičová
Ústav svetovej literatúry SAV
Bratislava

VÝSKUM POROZUMENIA TEXTU V MLADŠOM ŠKOLSKOM VEKU
ALEBO
PRIAMO KU KOREŇOM POROZUMENIA TEXTU

LIPTÁKOVÁ, Ľudmila (ed.) – DZIAK, Dávid – GOGOŤÁ, Eva – BRESTOVIČOVÁ, Alexandra – VUŽŇÁKOVÁ, Katarína – LUKÁČ, Juraj – KOPČÍKOVÁ, Marta: *VÝSKUM POROZUMENIA TEXTU V MLADŠOM ŠKOLSKOM VEKU*. Prešov: Prešovská univerzita v Prešove 2023. 304 s.

Kolektívna monografia *Výskum porozumenia textu v mladšom školskom veku* prináša pozoruhodný opis procesov, ktorými sa deti usilujú uchopiť vnímaný text. Autorský kolektív vedený Ľ. Liptákovou sa zameriava na fenomén tvorby inferencií, ktorý je kľúčový pri „identifikovaní“ významov z čítaného či počúvaného textu. Výsledky výskumu ponúkajú nielen cenné teoretické poznatky, ale zároveň predstavujú praktické impulzy pre didaktiku slovenského jazyka a literatúry. Publikácia sa tým snaží zaplniť výraznú medzeru v slovenskom výskumnom prostredí, ktoré doposiaľ trpí deficitom empirických dát v oblasti detského čítania a počúvania s porozumením. Jednotlivé kapitoly knihy systematicky vedú čitateľstvo od teoretického rámca a metodologických východísk až po opis a čiastkovú analýzu bohatého empirického materiálu. Nechýbajú ani konkrétne odporúčania pre vzdelávaciu prax. Pokúsime sa predstaviť obsah jednotlivých kapitol, ale predovšetkým vyzdvihnúť prínosy i limity empirického výskumu, ktorý reaguje na aktuálnu potrebu adekvátneho rozvíjania čitateľskej gramotnosti slovenských detí.

Prvá kapitola s príznačným názvom *Východiská a metodológia výskumu porozumenia textu v mladšom školskom veku* ponúka prvotné výskumné stimuly v sledovanej problematike. Autorky Ľ. Liptáková a M. Kopčíková v nej zrozumiteľným jazykom zasadzujú výskum celého kolektívu do medzinárodného i domáceho vedeckého kontextu. Predstavujú konceptuálny rámec, ktorý pozostáva z troch modelov porozumenia textu: **Simple View of Reading**¹ (autorky v texte neponúkajú slovenský ekvivalent modelu), ktorý však neinterpretujú lineárne („teda najprv dekodovanie a potom porozumenie, pretože deti ešte pred osvojením si čítania disponujú schopnosťami porozumenia počúvaného textu“; s. 12), **konštrukčno-integračný model**² a **interakčný model komponentov porozumenia**³. Inferencie zohrávajú kľúčovú úlohu vo všetkých troch modeloch, pričom práve zmapovanie pro-

¹ GOUGH, Philip B. – TUNMER, William E. (1986): Decoding, reading, and reading disability. In: *Remedial and special education*, roč. 7, č. 1, s. 6 – 10.

² KINTSCH, Walter (1988): The role of knowledge in discourse comprehension: a construction-integration model. In: *Psychological review*, roč. 95, č. 2, s. 163 – 182. KINTSCH, Walter (2013): Revisiting the construction–integration model of text comprehension and its implications for instruction. In: D. E. Alvermann, N. J. Unrau, R. B. Ruddell (eds.): *Theoretical Models and Processes of Reading*. International Reading Association, s. 807–839.

³ TENNENT, Wayne (2015): Understanding reading comprehension. Processes and practices. SAGE, 224 s.

cesu ich utvárania deťmi tretieho ročníka základnej školy tvorí primárny predmet celého výskumu.

Odkazmi na rôznorodosť výskumu v oblasti porozumenia textu jednak priznávajú jeho širokú kontinuitu v medzinárodnom kontexte, jednak poukazujú na jeho hlavný nedostatok u nás – deficit empirických dát. Neprekvapí preto, že jedným z cieľov výskumu bolo získať výskumný materiál orientovaný na proces utvárania inferencií u detí mladšieho školského veku hovoriacich po slovensky. Autorky priznávajú, že „horšie výsledky slovenských žiakov 4. ročníka základnej školy v medzinárodnej štúdii PIRLS⁴ pri porozumení vecného textu v porovnaní s umeleckým textom, ale najmä preukázaná väčšia náročnosť vytvárania inferencií z vecného textu“ (s. 20) sú príčinou, prečo sa vo svojich výskumných textoch, ktorými testovali deti, zamerali na „**vecné texty s dominujúcou informačnou funkciou**“ (ibid.). Za zmienku určite stojí, že pri ich administrácii uplatnili koncept dynamického hodnotenia⁵, resp. tzv. prístup odstupňovaných nápodiev, vďaka ktorému sa im viac-menej úspešne podarilo získať podrobnejší výskumný materiál. Ten umožnil zachytiť individualizovaný spôsob tvorenia koherentných (pomáhajú recipientovi dosiahnuť koherentné porozumenie textu – utvárajú sa automaticky) a interogatívnych (utvárajú sa až po recepcii textu, pričom prehľbujú a obohacujú jeho porozumenie) inferencií u jednotlivých detí.

Detské predstavy o počúvaní a o čítaní opisali v rovnomennej druhej kapitole Ľ. Liptáková a E. Gogová. Reagujú ňou na priznané ontologicko-epistemologické východiská – konštrukcionizmus a fenomenológiu. Kapitola predkladá zaujímavý pohľad do vnútorného sveta detí, pričom sa zameriava na ich spôsob chápania oboch fenoménov (t. j. čítanie i počúvanie). Držia sa tak stanoveného cieľa zabezpečiť empirické dáta, keďže v slovenskom kontexte detské predstavy o počúvaní doposiaľ zmapované neboli. Autorky sa nazdávajú, že zozbieraný a zinterpretovaný materiál nájde svoje uplatnenie nielen v didaktike (ako impulz pri rozvíjaní recepcných schopností detí), ale aj v kognitívnej lingvistiky ako zdroj autentických konceptualizácií čítania a počúvania. Aj napriek svojmu minimalistickému dizajnu (základ pološtruktúrovaného rozhovoru tvorila otázka *Čo podľa teba znamená počúvať/čítať?* s podpornou podotázkou *Čo robíš, keď počúvaš/čítaš?*) a nehomogénnym vzorkám respondentov (oba limity autorky priznávajú) ponúka kapitola pozoruhodné postrehy z individuálnych spôsobov myslenia detí, ktoré majú svoje opodstatnenie v rámci subjektivisticky orientovaného výskum slovenčiny.

Otázkou *Čo robíš, keď počúvaš/čítaš?* sa organicky otvára priestor pre nadväzujúcu otázku *Čo počúvaš/čítaš?*. Po vykreslení detských predstáv o počúvaní a čítaní preto, prirodzene, nasleduje kapitola *Čo deti počúvajú a čítajú*, ktorá opisuje **detské preferencie** v týchto spôsoboch vnímania textu. Kým prisun dát o čitateľskej preferencii slovenských detí je pomerne dobre saturovaný napr. medzinárodnými testovaniami (PIRLS, PISA⁶), detská sféra počúvania tak dobre preskúmaná nie je. Túto medzeru sa pokúšajú zaplniť A. Brestovičová a D. Dziak. Podobne ako v predchádzajúcej kapitole, aj oni si od zistení sľubujú implikácie pre vzdelávanie, ktoré v závere i predkladajú. S ohľadom na veľkosť a špecifickosť vý-

⁴ *Progress in International Reading Literacy Study*. Dostupné na: <https://www.iea.nl/studies/iea/pirls>.

⁵ FEUERSTEIN, Reuven – RAND, Ya'akov A. – HOFFMAN, Mildred B. – EGOZI, Moshe – SHACHAR-SEGEV, Nilly Ben – IDOL, Lorna – JONES, Beau Fly (1980): *Instrumental enrichment program. An intervention program for cognitive modifiability*. Baltimore: University Park Press.

⁶ *Programme for International Student Assessment*, Dostupné na: <https://www.oecd.org/en/about/programmes/pisa.html>.

skumnej vzorky (89 participantov z dvoch prešovských základných škôl) môžu však odporúčania autorov vzťahujúce sa na čitateľskú preferenciu detí pôsobiť miestami až príliš paušalizujúco.

Ľ. Liptáková a A. Brestovičová sa v štvrtej kapitole púšťajú priamo do opisu utvárania **koherentných a interrogatívnych inferencií**. V kapitole s názvom *Porozumenie informačného textu* spracúvajú žánr, ktorý z hľadiska detskej recepcie považujú za náročnejší na porozumenie, keďže si spravidla vyžaduje „špecifické znalosti z istej oblasti poznania, zvyčajne obsahuje nové pojmy alebo neznámu lexiku“ (s. 81). V súlade s výskumným dizajnom, deklarovaným už v prvej kapitole, skúmali vytváranie inferencií osobitne pri počúvaní a čítaní informačného textu. Pri zbere dát používali dva tematicky rozdielne texty, pričom výber textu závisel od aktuálne skúmaného typu vnímania (najprv participanti absolvovali počúvanie textu a minimálne o týždeň neskôr čítanie textu). V oboch prípadoch však výskumný text prešiel ešte pred jeho administráciou niekoľkými úpravami a testom obťažnosti. Celý proces úprav je zrozumiteľne objasnený a zároveň svedčí o dôkladne premyslenom výskumnom dizajne. Autorky v kapitole poukazujú o. i. na odlišnú mieru neporozumenia, ktorá sa preukázala pri verbalizovaní percipovaného textu (tzv. retelling). Pri počúvaní textu sa u participantov prejavila vyššia miera neporozumenia v porovnaní s čítaním textu.

Piata kapitola D. Dziaka a J. Lukáča sprevádza čitateľa zberom dát zameraných na recepciu počúvaného a čítaného naratívneho textu. V úvode kapitoly *Porozumenie naratívneho textu* nechýba krátka charakteristika **naratívneho žánru a naratívnej kompetencie** spolu s jej kľúčovými zložkami **kauzalitou udalostí a motiváciou postáv**. Pri zbere dát boli rovnako použité dva rozdielne výskumné texty, pričom každým sa testoval odlišný spôsob jeho recepcie. Kým naratívny text určený na počúvanie bol vyrozprávaný v prvej osobe, naratívny text určený na čítanie bol vyrozprávaný v tretej osobe. Hoci bola táto rozdielnosť autormi priznaná, nebola však bližšie ozrejmenej (čo by sa vzhľadom na následné porovnávanie recepcie počúvaného a čítaného textu žiadalo zdôvodniť). Vo svojich záveroch odporúčajú, aby sa do výučby zaradili „úlohy na indukčné rozvíjanie hlavnej myšlienky príbehu a jeho prepájanie s predošlými znalosťami a skúsenosťami žiakov, na identifikovanie distraktorov, na usporiadanie naratívnych udalostí podľa časovej postupnosti, na vytváranie anaforických inferencií pri čítanom texte a komparáciu spoločných a odlišných znakov dvoch entít, napr. postáv“ (s. 196), keďže s týmito úlohami mali deti podľa autorov najväčšie problémy. Kapitola ponúka o. i. dôkazy o tom, že vyšší počet slov v retellingu ešte nie je zárukou adekvátneho porozumenia textu.

Prezentáciu nazbieraných empirických dát uzatvára šiesta kapitola *Vytváranie inferencií pri čítaní multimodálneho textu*, v ktorej autorky K. Vužňáková a E. Gogová približujú zber dát a ich interpretáciu. Svojou kapitolou reagujú na zvyšujúci sa podiel multimodálnych a multimediálnych komunikátov, s ktorými vďaka súčasnej digitálnej dobe prichádzajú deti do kontaktu už od útleho veku. Špecifickou kombináciou viacerých semiotických subsystémov, na ktoré výskumné texty u detí cielia, sa tento zber dát od ostatných líši svojou administratívnou náročnosťou. Autorky sa zamerali na kombináciu **dvojitého kódovania: auditívny podnet** (počúvanie vecného textu + autentické zvukové prejavy rôznych entít), **vizuálny podnet** (čítanie vecného textu + rôzne obrázky). Z množstva podnetov, ktoré kapitola predkladá, stojí určite za pozornosť opis utvárania offline interrogatívnych inferencií pri recepcii jazykovej a vizuálnej metafory.

Knihu uzatvára siedma kapitola *Čo vieme a čo ešte potrebujeme vedieť o detskom porozumení textu*. Ľ. Liptáková v nej sumarizuje základné výskumné zistenia, pričom sa zameria-

va na zhodnotenie ôsmich vytýčených cieľov: Získať empirické dáta o porozumení vecného textu (1), analyzovať ich jednak z hľadiska detských predstáv o počúvaní a o čítaní (2), jednak z hľadiska vytvárania koherentných (3) a interrogatívnych (4) inferencií. Ďalej analyzovať úlohu odstupňovanej pomoci pri vytváraní inferencií (5) a pozorovať individuálne rozdiely v ich utváraní so zámerom načrtnúť inferenčné profily detí (6), a napokon preskúmať súvislosť vytvárania inferencií s ďalšími komponentmi porozumenia textu (7) a porovnať vytváranie inferencií pri recepcii počúvaného a čítaného textu (8). Zhodnotením jednotlivých cieľov však autorka nevyčerpáva rozsah tejto kapitoly. S rešpektom k hraniciam súčasného deskriptívneho výskumu otvára perspektívy ďalšieho skúmania zameraného na hĺbkové analýzy a interpretácie získaných empirických dát. V závere pripája výpočet podnetov pre didaktiku slovenského jazyka a literatúry, ako aj pre výučbu slovenčiny v primárnom vzdelávaní.

Opísaná kolektívna monografia predstavuje funkčný výskumný dizajn, ktorý vďaka pomerne detailnému opisu administrácie možno v budúcnosti ľahko replikovať. Transparentným prístupom otvárajú autorky a autori svoje čiastkové závery konštruktívnej kritike a motivujú k ich prípadnému doplneniu či reinterpretácii. Predstavená typológia inferencií spolu s empiricky zameraným výskumom naznačujú sľubné rozvíjanie nášho poznania (nielen predstavy) o procesoch, ktoré u detí prebiehajú pri apercpcii čítaného a počúvaného textu.

Roman Soóky
Jazykovedný ústav Ľ. Štúra SAV
Bratislava

POKYNY PRE AUTOROV

Redakcia JAZYKOVEDNÉHO ČASOPISU uverejňuje príspevky **bez poplatku** za publikovanie.

Akceptované jazyky: všetky slovanské jazyky, angličtina, nemčina. Súčasťou vedeckej štúdie a odborného príspevku je abstrakt v angličtine (100 – 200 slov) a zoznam kľúčových slov v angličtine (3 – 8 slov).

Súčasťou vedeckej štúdie a odborného príspevku v inom ako slovenskom alebo českom jazyku je zhnutie v slovenčine (400 – 600 slov) – preklad do slovenčiny zabezpečí redakcia.

Posudzovanie príspevkov: vedecké príspevky sú posudzované anonymne dvoma posudzovateľmi, ostatné príspevky jedným posudzovateľom. Autori dostávajú znenie posudkov bez mena posudzovateľa.

Technické a formálne zásady:

- ☐ Príspevky musia byť v elektronickej podobe (textový editor Microsoft Word, font Times New Roman, veľkosť písma 12 a riadkovanie 1,5). V prípade, že sa v texte vyskytujú zvláštne znaky, tabuľky, grafy a pod., je potrebné odovzdať príspevok aj vo verzii pdf alebo vytlačený.
- ☐ Pri mene a priezvisku autora je potrebné uviesť pracovisko.
- ☐ Text príspevku má byť zarovnaný len z ľavej strany, slová na konci riadku sa nerozdeľujú, tvrdý koniec riadku sa používa len na konci odseku.
- ☐ Odseky sa začínajú zarážkou.
- ☐ Kurzíva sa spravidla používa pri názvoch prác a pri uvádzaní príkladov.
- ☐ Polotučné písmo sa spravidla používa pri podnadpisoch a kľúčových pojmoch.
- ☐ Na literatúru sa v texte odkazuje priezviskom autora, rokom vydania a číslom strany (Horecký 1956, s. 95).
- ☐ Zoznam použitej literatúry sa uvádza na konci príspevku (nie v poznámkovom aparáte) v abecednom poradí. Ak obsahuje viac položiek jedného autora, tie sa radia chronologicky.

Bibliografické odkazy:

- ☐ knižná publikácia: ONDREJOVIČ, Slavomír (2008): *Jazyk, veda o jazyku, societa*. Bratislava: Veda, vydavateľstvo SAV, 204 s.
- ☐ slovník: JAROŠOVÁ, Alexandra – BUZÁSSYOVÁ, Klára (eds.) (2011): *Slovník súčasného slovenského jazyka. H – L*. [2. zv.]. Bratislava: Veda, vydavateľstvo SAV.
- ☐ štúdiá v zborníku: DUROVIČ, Ľubomír (2000): Jazyk mesta a spisovné jazyky Slovákov. In: S. Ondrejovič (ed.): *Sociolinguistika Slovača 5. Mesto a jeho jazyk*. Bratislava: Veda, vydavateľstvo SAV, s. 111 – 117.
- ☐ štúdiá v časopise: DOLNÍK, Juraj (2009): Reálne vz. ideálne a spisovný jazyk. In: *Jazykovedný časopis*, roč. 60, č. 1, s. 3 – 12. DOI 10.2478/v10113-009-0001-3. [cit. DD-MM-RRRR].
- ☐ internetový zdroj: Slovenský národný korpus. Verzia prim-5.0-public.all. Bratislava: Jazykovedný ústav Ľudovíta Štúra SAV. Dostupné na: <https://korpus.juls.savba.sk> [cit. DD-MM-RRRR].

INSTRUCTION FOR AUTHORS

JOURNAL OF LINGUISTICS publishes articles **free of publication charges**.

Accepted languages: all Slavic languages, English, German. Scientific submissions should include a 100-200 word abstract in English and a list of key words in English (3-8 words).

Scientific articles in a language other than Slovak or Czech should contain a summary in Slovak (400-600 words) – translation into Slovak will be provided by the editor.

Reviewing process: scientific articles undergo a double-blind peer-review process and are reviewed by two reviewers, other articles by one reviewer. The authors are provided with the reviews without the name of the reviewer.

Technical and formal directions:

- ☐ Articles must be submitted in an electronic form (text editor Microsoft Word, 12-point Times New Roman font, and 1.5 line spacing). If the text contains special symbols, tables, diagrams, pictures etc. it is also necessary to submit a pdf or printed version.
- ☐ Contributions should contain the full name of the author(s), as well as his/her institutional affiliation(s).
- ☐ The text of the contribution should be flush left; words at the end of a line are not hyphenated; a hard return is used only at the end of a paragraph.
- ☐ Paragraphs should be indented.
- ☐ Italics is usually used for titles of works and for linguistic examples.
- ☐ Boldface is usually used for subtitles and key terms.
- ☐ References in the text (in parentheses) contain the surname of the author, the year of publication and the number(s) of the page(s). (Horecký 1956, p. 95).
- ☐ The list of references is placed at the end of the text (not in the notes) in alphabetical order. If there are several works by the same author, they are listed chronologically.

References:

- ☐ Monograph: ONDREJOVIČ, Slavomír (2008): *Jazyk, veda o jazyku, societa*. Bratislava: Veda, vydavateľstvo SAV, 204 p.
- ☐ Dictionary: JAROŠOVÁ, Alexandra – BUZÁSSYOVÁ, Klára (eds.) (2011): *Slovník súčasného slovenského jazyka. H – L*. [2. zv.]. Bratislava: Veda, vydavateľstvo SAV.
- ☐ Article in a collection: DUROVIČ, Ľubomír (2000): Jazyk mesta a spisovné jazyky Slovákov. In: S. Ondrejovič (ed.): *Sociolinguistika Slovača 5. Mesto a jeho jazyk*. Bratislava: Veda, vydavateľstvo SAV, pp. 111–117.
- ☐ Article in a journal: DOLNÍK, Juraj (2009): Reálne vz. ideálne a spisovný jazyk. In: *Jazykovedný časopis*, Vol. 60, No. 1, pp. 3–12. DOI 10.2478/v10113-009-0001-3. [cit. DD-MM-RRRR].
- ☐ Internet source: Slovenský národný korpus. Verzia prim-5.0-public.all. Bratislava: Jazykovedný ústav Ľudovíta Štúra SAV. Available at: <https://korpus.juls.savba.sk> [cit. DD-MM-RRRR].

ISSN 0021-5597 (tlačená verzia/print)

ISSN 1338-4287 (verzia online)

MIČ 49263

JAZYKOVEDNÝ ČASOPIS

VEDECKÝ ČASOPIS PRE OTÁZKY TEÓRIE JAZYKA

JOURNAL OF LINGUISTICS

SCIENTIFIC JOURNAL FOR THE THEORY OF LANGUAGE

Objednávky a predplatné prijíma/Orders and subscriptions are processed by:
SAP – Slovak Academic Press, s. r. o., Bazová 2, 821 08 Bratislava
e-mail: sap@sappress.sk

Registračné číslo 7044

Evidenčné číslo 3697/09

IČO vydavateľa 00 167 088

Ročné predplatné pre Slovensko/Annual subscription for Slovakia: 18 €, jednotlivé číslo 6 €
Časopis je v predaji v kníhkupectve Veda, Štefánikova 3, 811 06 Bratislava 1

© Jazykovedný ústav Ľudovíta Štúra SAV, v. v. i., Bratislava