

JANA SOKOLOVÁ – AUGUSTÍN SOKOL*

Status lingvistických premenných v metodológii analýz textu¹ (Náhl'ad do pomenovacích databáz slovenčiny)

SOKOLOVÁ, J. – SOKOL, A.: Linguistic variables in the methodology of text analysis. (A probe into Slovak terminological databases). *Slavica Slovaca*, 57, 2022, No 2, pp. 137-148 (Bratislava).

The subject of the study is the definitional understanding and reflection on the *linguistic variables* used in the methodology of text analysis. The impetus to write this paper was the definition of the very linguistic entities that text is made out of, and their conceptual and terminological designation from the point of view of text linguistics. The term *variable* was chosen for the flexibility of its status and value. *Linguistic variables* are presented from the perspective of systemic, textual, quantitative and corpus linguistics. Interpretation, which is used in the sense of explanation, is the methodological framework. The definition of *linguistic variables* is connected with the operational steps of segmentation, identification and classification. A distinction is made between the *linguistic variables* with a primarily systemic value, which traditionally show the linguistic (*langue*) and speech (*parole*) status, and the *linguistic variables* with a primarily textual and pragmatic value.

Variable, Segmentation, Status, Value, Text, Linguistics.

Úvod

V metodológii analýz textu sa často stretávame s rozkolísanosťou obsahovej náplne uplatňovanej terminológie. Pojmová nejednotnosť sa dotýka predovšetkým vymedzenia samotných lingvistických entít konštituujúcich text² a následne ich pojmovo-terminologického označenia z pohľadu lingvistiky textu. *Lingvistické premenné* sú spravidla bežne používané odborné výrazy, ktoré sú vystavované prehodnocovaniu, reinterpretácii a prispôsobovaniu príslušným metodológiám a teóriám. Mnohé patria k výrazom s interdisciplinárnym presahom a výstupy lingvistov, filozofov, logikov neraz ponúkajú čiastočne odlišné riešenia.

Predmetom našej štúdie bude definíčné uchopenie a reflektovanie *lingvistických premenných* v širších súvislostiach, čo značí, že budeme ad hoc aplikovať poznatky systémovej, textovej, kvantitatívnej a korpusovej lingvistiky. Sústredíme sa na vymedzenie nosných pojmov a termínov, uvedenie ich informačného obsahu a na ich systematizáciu. Kľúčové pojmy predstavíme

* Prof. PhDr. Jana Sokolová, CSc., Katedra rusistiky FF UKF v Nitre, Štefánikova 67, 949 74 Nitra, Slovakia; jsokolova@ukf.sk; Mgr. Augustín Sokol, interný doktorand FF UKF v Nitre, Štefánikova 67, 949 74 Nitra, Slovakia; augustin.sokol@ukf.sk.

¹ Štúdia bola vypracovaná v rámci riešenia projektu VEGA 1/0650/22 Masmediálne komunikáty v digitálnej a printovej forme a ich komprehenzia u rôznych cieľových skupín a projektu UGA III/14/2022 K identifikácii úrovne porozumenia cudzojazyčných textov.

² Termín *text* tu uplatňujeme v zmysle verbalizovaného faktu, t. j. výsledátu rečovej činnosti. Zameriavame sa na písaný prozaický text.

v ponímaní etablovaných autorít slovenskej a českej jazykovedy. Definičné vyjadrenia uvedieme v preklade do slovenčiny.

Zastrešujúcim pojmom štúdie je pojem *premennej*, ktorá sa posudzuje ako prvok nadobúdajúci (a) rôzne hodnoty; (b) rôzne stavy (úrovne).

Na *lingvistické premenné* budeme nazerať v zmysle (a) jazykových premenných (entít jazykového systému konštituujuúcich text) a v zmysle (b) textových premenných (entít textu uplatňovaných s oporou aj bez opory premenných v jazykovom systéme). Jazykové premenné nebudeme striktné posudzovať len na pozadí protikladu jazyka ako *langue* a reči ako *parole*, ale aj ako na prvky jazyka v zmysle *langage*, teda nadradeného pojmu zahrnujúceho rečovú činnosť i jazyk ako potenciálny systém. Metodologickým rámcom sledovaných aspektov sa stane interpretácia, ktorú budeme uplatňovať v zmysle vysvetlenia, explanácie. Z priestorových dôvodov k *lingvistickým premenným* nebudeme uvádzať ilustračné príklady.

Vymedzenie *lingvistických premenných* sa spája s operačnými krokmi segmentácie, identifikácie a klasifikácie. Segmentácia je semaziologická operácia, v ktorej sa postupuje od formy k obsahu. Znamená proces (automatického) členenia celku na segmenty a výsledok tohto procesu. Segment je „uzuálne stanovený prvok ako jednotka analyzovanej roviny“.³ Spravidla býva najmenšou jednotkou analyzovanej roviny. Identifikácia *lingvistických premenných* súvisí s dvoma komplementárnymi pohľadmi, ktorým zodpovedajú dve smerovania: „od vety k textu“, t. j. analytický pohľad „zdola“ (tento smer uvažovania zaznamenávame napr. u F. Daneša) a smerovanie „od textu k slovám“, t. j. holistický pohľad „zhora“ (tento smer uvažovania zaznamenávame napr. u P. Karlíka). Klasifikácia *lingvistických premenných* znamená ich triedenie a organizáciu. Výsledkom je rozlišovanie *lingvistických premenných* s primárnou systémovou hodnotou tradične vykazujúcich jazykový (*langue*) a rečový (*parole*) status a *lingvistických premenných* s primárnou textovopragmatickou hodnotou. Systematizácia *lingvistických premenných* s primárnou systémovou hodnotou (ďalej jazykových premenných) odráža princíp usporiadania vysvetľujúci závislosť štruktúry medzi usporiadanými jazykovými prvkami. Prevažne reflektuje binárnu interpretáciu, ktorá vychádza z predpokladu, že každá opozícia je konštruovaná ako premenná, ktorá môže nadobúdať dve hodnoty. Zohľadňuje perцепčný prístup a princíp štruktúrovanosti, napr. *fonéma*, *morféma*, *lexéma*, *syntagméma*, *veta/vetný vzorec* a i. sú prvky so statusom *langue*, *fóna*, *morfa*, *lexa*, *výpoveď* a i. sú prvky so statusom *parole*. Systematizácia *lingvistických premenných* s primárnou textovopragmatickou hodnotou (ďalej textových premenných) má povahu nomenklatúrneho zaradenia do súboru (zoznamu) bez hlbších vzájomných vnútorných súvislostí, napr. *agregát*, *odsek*, *type*, *token*, *lema*, *textové slovo*, *základná textová jednotka*, *textéma*, *konektor*, *textový orientátor*, *odkazovací prostriedok* a i. V kvantitatívnej lingvistike sa vymedzenie *lingvistických premenných* v texte poníma napr. aj v zmysle náhodných veličín majúcich lineárny rozmer. *Lingvistickou premennou* je napr. aj dĺžka slova, dĺžka vety, ale aj dĺžka slova a pozícia slova vo vete vo vzájomnej súvislosti a pod.

1. Jazykové premenné

Jazykové premenné tvoria databázu systémovej lingvistiky. Prezентujú stavebné zložky založené na predpoklade existencie abstraktných entít, ktoré nie sú priamo pozorovateľné a majú povahu teoretických konštruktov. Jazykové premenné majú status *jazykových výrazov*, *jazykových prvkov*, *jazykových prostriedkov* a *jazykových jednotiek*.

³ Šimková, M.: Segment. Výberový slovník termínov z počítačovej a korpusovej lingvistiky. 2006. Dostupné na: <https://docplayer.cz/105183454-Vyberovy-slovník-termínov-z-pocitacovej-a-korpusovej-lingvistiky.html> [cit. 10. 12. 2021].

Jazykový výraz sa v metodológii vedeckého výskumu posudzuje ako plnovýznamová jednotka jazyka, ktorá existuje v systéme daného jazyka a riadi sa určitými syntaktickými a sémantickými pravidlami.⁴ *Jazykové výrazy* prezentujú statusovú funkciu, ktorá v danom jazyko-vo-kultúrnom prostredí reťazcu fyzických zvukov pripisuje intencionalitu. Vzťah medzi *jazykovým výrazom* a tým, čo označuje je vzťahom označovania a základnou sémantickou reláciou.

Jazykový prvok je entita všeobecnejšieho rázu, ktorá podľa spôsobu použitia môže byť prevažne významovej alebo výrazovej povahy.⁵ Inventár jazykových prvkov a sieť vzťahov, zviazanosť týchto prvkov na základe ich podobnosti a rozdielnosti sú faktory systémovej koncepcie jazyka.

Jazykový prostriedok je prvkom výrazovej povahy patriaci do výrazového systému jazyka a majúci nástrojovú povahu.⁶ Jazyková premenná v zmysle *jazykového prostriedku* je výraz majúci určitý funkčný status, porov.: „elementy a jednotky majúce smerom hore platnosť prostriedkov, smerom dole platnosť funkcie“.⁷ F. Daneš rozlišuje dve triedy *jazykových prostriedkov*: (a) predznakové (elementárne) prostriedky a (b) znakové prostriedky.⁸ V korpusovej lingvistiky je to „akýkoľvek prvok jazykového systému znakovkej (slovo, číslica, interpunkcia) aj neznakovkej (fonéma, slovosled) povahy“.⁹ Súpis, súbor jazykových prostriedkov rovnakého typu tvorí inventár.

Jazyková jednotka je vymedzená v zmysle bilaterálnej, t. j. významovo-výrazovej entity. Ako najmenšia jednotka je identifikovaná na príslušnej jazykovej rovine formou, obsahom a funkciou. Každá *jazyková jednotka* je systémovo ukotvená a vyznačuje sa statusom abstraktnej jednotky systému jazyka a jeho realizáciou v reči, napr. *fonéma* a *fóna*, *morféma* a *morfa*, *lexéma* a *lexa*, *veta* a *výpoveď* a pod. *Jazykové jednotky* nejakej roviny v ponímaní F. Daneša¹⁰ sú priesečníkom funkcií (z vyšších rovín) a stavebných prostriedkov (z nižších rovín a z oblasti elementárnych prostriedkov). Inými slovami: pri tvorení jazykových jednotiek danej roviny sa ako stavebné prostriedky používajú jednotky nižších rovín a niektoré elementárne prostriedky. Tieto elementárne prostriedky sa stávajú konštitutívnou zložkou týchto jednotiek a na vyššej rovine (vyšších rovinách) vystupujú už ako integrálna zložka komplexného znakového prostriedku (znakovkej jednotky). Funkčný vzťah medzi jednotkami nižšej a vyššej (vyšších) rovín F. Daneš¹¹ formuloval ako možnosť (a) normálnej funkčnej postupnosti (čes. poslounosti), (b) preskakovania rovín, (c) simultánneho fungovania na niekoľkých rovinách, (d) rekurzívneho fungovania, (e) regresívneho fungovania (prevrátenej funkčnej postupnosti). Hovorí sa napr. o fonologickej (fonematickej), morfolologickej, lexikálnej (lexematickej), syntactickej rovine. Každá rovina sa javí vzhľadom k nižšej rovine ako oblasť jej funkcií a zároveň vzhľadom k nadradeným rovinám sa javí ako oblasť ich prostriedkov. Čím je rovina vyššia, tým sú jej jednotky zložitejšie a komplexnejšie. Jazykové prvky a jednotky majú smerom nahor platnosť prostriedkov a smerom nadol platnosť funkcií.

Jazyková jednotka z pohľadu korpusovej lingvistiky značí „najmenšiu jednotku jazykového systému na konkrétnej rovine“.¹²

⁴ Bielik, L. – Halas, J. – Hanzel, I. – Kostelec, M. – Marko, V. – Zouhar, M.: Slovník metodologických pojmů. Bratislava: Univerzita Komenského v Bratislave, 2016, s. 62.

⁵ Kačala, J.: Syntaktický systém jazyka. Pezínok: Formát, 1998, s. 27.

⁶ Kačala, J.: Syntaktický systém jazyka, c. d., s. 27.

⁷ Daneš, F.: O pojmu „jazykový prostředek“. In: Slovo a slovesnost, 1967, roč. 28, č. 4, s. 344.

⁸ Daneš, F.: O pojmu „jazykový prostředek“, c. d., s. 346.

⁹ Porov.: <https://terminologickyportal.sk/> [cit. 22. 07. 2022].

¹⁰ Daneš, F.: O pojmu „jazykový prostředek“, c. d., s. 348.

¹¹ Daneš, F.: O pojmu „jazykový prostředek“, c. d., s. 347.

¹² Šimková, M.: Jazyková jednotka. Výberový slovník termínů z počítačové a korpusové lingvistiky, c. d., [cit. 22. 07. 2022].

Ako bolo spomínané, v rámci *jazykových jednotiek* so systémovou platnosťou sa rozlišujú (a) predznakové a (b) znakové jazykové jednotky, resp. dištinktémy a signémy.¹³ Predznakovou jazykovou jednotkou (segmentálnou dištinktémou) je **fonéma** ako minimálna funkčná jednotka jazykového systému (na fonematickej rovine) schopná rozlíšiť formálno-sématické jednotky a **graféma** (písmeno) ako základná jednotka grafického systému jazyka. *Fonémy* sú vyabstrahované zo súborov *alofón* s prihliadnutím na ich vzťahy a majú dištinktívnu funkciu. Vzťah medzi *grafémou* a *fonémou* môže byť obojstranne jednoznačný, t. j. jednej *graféme* zodpovedá jedna *fonéma* a naopak, alebo jednostranný, t. j. jednej *graféme* zodpovedá viac *foném* alebo jednej *fonéme* zodpovedá viac *grafém*. K neznakovým jednotkám sa zaraďuje aj **slabika** ako sled foném, z ktorých jedna je nosičom slabičnosti. Je to fonicko-rytmická jednotka s jedným vrcholom zvučnosti fungujúca ako nositeľ všetkých suprasegmentálnych javov príslušného jazyka.¹⁴ *Slabika* v definičnom vymedzení J. Sabola je „médiom prepájajúcim segmentálny a suprasegmentálny podsystém (je zložená zo segmentov, v jadrovej sylabickej štruktúre realizujúcich „nerozčleniteľné“ spojenie konsonantickej, striktúrnej a vokalickej, apertúrnej fázy, a je nositeľom všetkých suprasegmentálnych javov“.¹⁵ *Slabika* je aj segmentálnou rečovou jednotkou aj z fonologického hľadiska jazykovou jednotkou. Poníma sa tiež v zmysle technickej jednotky, ktorá sa len v niektorých jazykoch (napr. v čínštine) prekrýva s *morfom*.¹⁶ *Slabika* je jedným z nosných pojmov versológie. Izosylabizmus predstavuje metrickú konštantu sylabickeho a sylabotonického veršového systému.¹⁷ Dĺžka *slabiky* nie je konštantná, je premenlivá a nie sú vypracované jednotné kritériá, ktoré by presne definovali hranice *slabík*. *Slabika* je štandardná premenná, ktorou sa v kvantitatívnej lingvistike meria dĺžka slova.

K jazykovým znakovým bilaterálnym jednotkám majúciom nominatívnu platnosť štandardne patrí *morféma*, *lexéma*, *frázéma*.

Morféma je najmenšia významovo-výrazová jednotka morfematickej roviny. Podľa J. Horeckého obsahom morfémy je odkaz na isté prvky v oblasti významu, kým formou morfémy sú jednotlivé fonémy a súhrn vzťahov, do ktorých môže daná morféma vstupovať.¹⁸ Z pohľadu funkcií sa rozlišujú (i) lexikálne (koreňová, derivačná), (ii) gramatické (relačná, samostatná gramatická) a (iii) lexikálno-gramatická (modifikačná) morféma.¹⁹ Pre účely deskripcie celej morfematickej zásoby jazyka V. Krupa navrhol súbor štyroch parametrov umožňujúcich multidimenzionálne hodnotenie morfémy. Sú to: parameter formy, výskytu, významu a segmentovateľnosti. Každý parameter možno považovať za premennú, ktorá nadobúda najmenej dve rozličné hodnoty na nominálnej alebo ordinálnej škále.²⁰ **Morfonéma** je základná jazyková jednotka na morfonologickej rovine. Je zložená z morf a reprezentovaná morfou. Rozpoznáva sa vtedy, keď dve a viac morf jednej morfémy nie sú odlišené v dôsledku syntagmatickej alternácie, ale keď táto alternácia nesie gramatický alebo derivačný význam.²¹

¹³ Dolník, J.: Základy lingvistiky. Bratislava: Stimul, 1999, s. 67.

¹⁴ Mistrík, J. a kol.: Encyklopédia jazykovedy. Bratislava: Obzor, 1993, s. 384-385.

¹⁵ Sabol, J.: Semiotické pozadie segmentov a suprasegmentov. In Bohunická, A. (ed.): Jazykoveda v pohybe. Bratislava: Univerzita Komenského Bratislava, 2012, s. 62.

¹⁶ Hřebíček, L.: Vyprávění o lingvistických experimentech s textem. Praha: Academia, 2002, s. 51.

¹⁷ Porov.: <https://beliana.sav.sk/heslo/izosylabizmus> [cit. 22. 08. 2022].

¹⁸ Horecký, J.: Morfematická štruktúra slovenčiny. Bratislava: Vydavateľstvo Slovenskej akadémie vied, 1964, s. 20.

¹⁹ Dolník, J.: Základy lingvistiky, c. d., s. 67.

²⁰ Krupa, V.: Jednota a variabilita jazyka. Bratislava: Veda, 1980, s. 122-123.

²¹ Šefčík, O.: Morfoném. In Karlík, P. – Nekula, M. – Pleskalová, J. (eds.): CzechEncy – Nový encyklopedický slovník češtiny online. Dostupný na: <https://www.czechency.org/slovník/MORFONÉM> [cit. 30. 08. 2022].

S *morfémou* a *lexémou* (najmenšou významovo-výrazovou jednotkou lexikálnej roviny) sa spájajú výrazy *seméma* ako lexikálny význam morfémy a *lexéma* ako gramatický význam morfémy a lexémy. S *lexémou* sú prepojené výrazy *pomenovanie* a *slovo*. Tvoria tri druhy lexikálnych jednotiek vyčlenených podľa rôznych aspektov.

Lexéma je jednotka lexikónu. Ako samostatná lexikálna jednotka je v lineárnej postupnosti vo vete voľne premiestniteľná. Na *lexému* (slovo, pomenovanie) možno nazerať z troch aspektov, a síce ako na: (a) **útvár** zložený z komponentov nižšej roviny; (b) **jednotku** vo vzťahu k iným jednotkám: (i) rovnakej funkčnej alebo sémantickej triedy: t. j. vertikálne vzťahy v systéme jazyka (oblasť lexikálnej paradigmatiky); (ii) inej funkčnej alebo sémantickej triedy in absentia: t. j. horizontálne vzťahy v systéme jazyka (oblasť lexikálnej syntagmatiky – kolokabilita); (iii) inej funkčnej alebo sémantickej triedy in praesentia (v texte), ktorá je buď sémanticky pravidelná (oblasť bežnej syntaxe a oblasť odbornej terminológie), alebo sémanticky nepravidelná (oblasť idiomatiky a frazeológie, oblasť autorských metafor); (c) **komponent** – vyššie jednotky ponímané ako štruktúrny model, v ktorom plní funkciu povahy: (i) *kolokačné štruktúry* predstavujúce horizontálne vzťahy lexikálnej kolokability; (ii) *vetné štruktúry* ako horizontálne syntaktické vzťahy; (iii) *nadvetné (viacvetné) štruktúry* prezentujúce syntaktické textové vzťahy.²² V korpusovej lingvistike sa *lexéma* posudzuje ako „formálna a významová (aj viacslvná) jednotka slovnjej zásoby reprezentovaná základným, slovníkovým tvarom“.²³

Pomenovanie je označenie javu reality alebo myšlienkového obsahu formou, jazykovým názvom (jednoslovným alebo viacslvným), ponímané ako proces alebo jeho výsledok.²⁴ Pomenovania v zmysle lexikálnych jednotiek majú podobu (i) jednoslovných pomenovaní (lexémy, slová) a (ii) viacslvných pomenovaní (združené pomenovania; výraz sa používa tak na neterminologické, ako aj terminologické slovné spojenia). Zoskupenie jednoslovných a viacslvných pomenovaní do synonymických radov (synsetov) a ich prepojenie sémantickými vzťahmi sa v počítačovej lingvistike nazýva **wordnet**. Ide o sémantické vzťahy najmä medzi hyponymami a hyperonymami, meronymami a holonymami a o paradigmatické vzťahy medzi opozitami. V lexikológii a lexikografii sa uplatňuje pojem **tezaurus**, ktorý v slovníkovej forme prezentuje onomaziologicky organizovaný lexikálny systém príslušného jazyka.

Pri viacslvných pomenovaniach sa v procesoch idiomatizácie konštituuju *frazémy*. **Frazéma** (frazeologická jednotka, frazeologický zvrät) je spravidla viacslvné pomenovanie, ktoré charakterizuje obraznosť a petrifikovanosť lexikálnej a gramatickej formy. *Frazéma* je funkčnou a formálnou jednotkou frazeológie, ktorá je podľa F. Míka unikátnym prípadom „kríženia parole a langue“, kontamináciou „výpovede a prostriedku výpovede“.

Slovo patrí k tradičným a zároveň najviac diskutovaným výrazom. Označuje sa ním formálna, celostná a intuitívne vymedzená jednotka. Kritériami delimitácie slova sú tieto určenia: (a) minimálna dištinktívna sémantická jednotka; (b) premiesniteľnosť; (c) určité fonetické stvárnenie a ustálená neprerušiteľná forma; (d) izolovateľnosť v texte (možnosť pauzy na jeho oboch koncoch); (e) nositeľ prízvuku.²⁵ V závislosti od slovných druhov sa delia na slová s funkciou (a) **deznátora** – majú funkciu označovania (substantíva, adjektíva, verbá a adverbia) a slová s funkciou (b) **operátora** – abstraktné výrazy relačnej povahy (v podstate synsématická).²⁶

²² Čermák, F. – Holub, J.: Syntagmatika a paradigmatika českého slova. I. Valence a kolokabilita. 3. vydání. Praha: Karolinum, 2005, s. 10.

²³ Porov.: <https://terminologickyportal.sk/> [cit. 22. 07. 2022].

²⁴ Čermák, F.: Lexikon a sémantika. Praha: Lidové noviny, 2010, s. 272.

²⁵ Čermák, F.: Jazyk a jazykověda. Přehled a slovníky. Praha: Karolinum, 2001, s. 174.

²⁶ Čermák, F. – Holub, J.: Syntagmatika a paradigmatika českého slova, c. d., s. 18.

Za elementárne syntaktické jednotky sa považuje *syntagma, fráza a veta*.²⁷

Syntagma (syntagméma) je syntaktické spojenie minimálne dvoch plnovýznamových lexikálnych, sémanticky kompatibilných jednotiek s rôznou mierou syntaktickej celostnosti. *Syntagma* je štruktúra skladajúca sa z jedného alebo niekoľkých (v prípade valenčnej *syntagmy*) určených členov a z člena bezprostredne určujúceho na identickom princípe.²⁸ *Syntagma* je celkom utvoreným zo spojenia prvkov patriacich do rozličných paradigiem.²⁹ *Syntagma* je syntaktický útvar, ktorý tvorí konfigurácia jazykových jednotiek, spravidla slov, takže sa hovorí o *slovnom spojení* a v užšom ponímaní na úrovni lexém o *kolokáciách*.

Kolokácia je v lexicológii a lexikografii typické, zmysluplné, ustálené, syntagmatické spojenie dvoch (resp. viacerých) slovných tvarov. V korpusovej lingvistike sa takto označuje spoločný výskyt dvoch, troch, výnimočne viacerých syntagmaticky a sémanticky spätých (lexikálnych) jednotiek.³⁰ V základe *kolokácie* je aplikácia princípu predurčeného výberu. Obmedzená kolokabilita výrazov je markantným rysom frazeologických jednotiek.

Fráza (zložka, konštituent)³¹ je v porovnaní so *syntagmou* všeobecnejším pojmom. Rozlišujú sa jednoduché a komplexné (zložené) *frázy*. *Fráza* môže byť (teoreticky neobmedzene) rekurzívne rozvinutá.³² *Fráza* v zmysle zložky vo formálnej gramatike prezentuje hierarchickú syntaktickú štruktúru XP, napr. NP je (nominálna fráza), VP (verbálna fráza), AdjP (adjektívna fráza), AdvP (adverbiálna fráza), PP (predložková fráza) a pod. Zapisuje sa buď v podobe závislostného stromu alebo vo formáte anotovaného zátvorkovania. Oba formáty zaznamenávajú syntaktickú štruktúru vety z hľadiska hierarchie fráz. Zložka VP sa považuje za najdôležitejšiu súčasť klauzy. Rozdiel medzi frázou a klauzou spočíva v tom, že každá klauza je fráza, ale nie každá fráza je klauza.³³

Veta je jazyková jednotka abstraktného jazykového systému; veta ako abstraktný vzorec je lexikálne realizovanou **vetnou schémou**.³⁴ *Veta* ako prirodzený gramaticko-sémantický systém majúci komunikačný rozmer uplatňuje tieto výstavbové princípy: (i) aktualizačný princíp; (ii) gramatický princíp; (iii) (lexikálno)sémantický princíp; (iv) informačný princíp; (v) princíp jadra a periférie vety.³⁵

Klauza je predikačná jednotka, „podveta súvetia“.³⁶ To znamená, že prirad'ovacie aj podrad'ovacie súvetie sú biklauzálnymi útvarmi (klauzami sú aj hlavná veta aj podradená veta). Klauza môže byť jednočlenná aj dvojčlenná veta. Pojem klauzy sa používa vtedy, keď sa nevyjadrujú vzťahy viet medzi sebou.³⁷

Propozícia je základ sémantickej stavby vety. Vo funkčnej gramatike propozícia je konfigurácia obsahových entít predstavujúcich odraz nejakého výseku skutočnosti (nazývame ju denotatívnou situáciou alebo „stavom veci“).³⁸ Vo funkčne generatívnom opise (FGP) sa chápe ako

²⁷ Mluvnice češtiny 3. Praha: Academia, 1987, s. 18.

²⁸ Mluvnice češtiny 3, c. d., s. 18.

²⁹ Krupa, V.: Metafora na rozhraní vedeckých disciplín. Bratislava: Tatran, 1990, s. 161.

³⁰ Porov. <https://terminologickyportal.sk/> [cit. 22. 07. 2022].

³¹ Termíny nie sú plne synonymné. Platí, že každá fráza je zložkou alebo konštituentom, ale nie každá zložka je frázou.

³² Mluvnice češtiny 3, c. d., s. 20.

³³ Panevová, J. – Karlík, P.: Klauze. In Karlík, P. – Nekula, M. – Pleskalová, J. (eds.): CzechEncy – Nový encyklopedický slovník češtiny, c. d., [cit. 05. 07. 2022].

³⁴ Dolník, J.: Všeobecná jazykoveda. Druhé, doplnené a upravené vydanie. Bratislava: Veda, 2013, s. 38.

³⁵ Kačala, J.: Syntaktický systém jazyka, c. d., s. 21.

³⁶ Wimmer, G. – Altmann, G. – Hřebíček, L. – Ondrejovič, S. – Wimmerová, S.: Úvod do analýzy textov. Bratislava: Veda. 2003, s. 43.

³⁷ Mistrík, J. a kol.: Encyklopédia jazykovedy, c. d., s. 237.

³⁸ Mluvnice češtiny 3, c. d., s. 9.

pravdivostná podmienka vetnej výpovedi a využíva sa ako pojmový nástroj k uchopeniu zmyslu vety.³⁹ V modálnej logike je jedným z jej kľúčových pojmov. V intenzionálnom modeli významu sa propozície ponímajú „ako významy či sémantické obsahy viet, prípadne ako ich denotáty či to, o čom dané vety hovoria.“⁴⁰

Status bilaterálnej jednotky patriacej do langue aj parole má **štyléma**. Vymedzenie *štylémy* súvisí s postulovaním roviny textu „in abstracto“ opierajúceho sa o funkčnú komplexnosť textových útvarov a predpokladom uplatňovania určitého súboru textových vzorcov a špecifických pravidiel pri produkcii a recepcii textu.⁴¹ *Štyléma* sa poníma ako základná jednotka štylistického plánu jazyka (štylistickej roviny jazyka) reprezentujúca alebo tvoriaca typ textu, ktorý sa realizuje ako konkrétny text. Z lingvistického hľadiska J. Mistrík rozlišuje jazykové a kompozičné *štylémy*. *Jazykové štylémy* ďalej delí na lexikálne, syntaktické (syntagmatické a vetné), morfológické (morfonologické) a fónické. *Kompozičné štylémy* delí na makrokompozičné a mikrokompozičné.⁴² *Štyléma* v ponímaní D. Slančovej je jednotkou štylistickej a štylovej kvality, ktorá má svoj obsah, formu a funkciu. Obsahovú zložku tvorí kategoriálna kvalita ako individuálne a/alebo socializovane konvencionalizované zovšeobecnené vedomie čiastkovej alebo komplexnej kvality rozličného typu. Forma štylémy je vnútorná a vonkajšia. Vnútornú formu štylémy tvorí vedomie asociatívne (napríklad a hlavne na báze synonymie) usporiadaných konceptuálnych (aj tematických), nominačných (vlastných jazykových) a interakčných (teda parajazykových) prostriedkov. Vonkajšiu formu štylémy tvoria konkrétne tematické, jazykové a parajazykové prostriedky, ktorými sa príslušná kategória realizuje v konkrétnom komunikačnom akte. Funkciou štylémy je vyvolať zamýšľaný štylistický účinok.⁴³

2. Textové premenné

Textové premenné sú textové jednotky (jednotky textu). Sú to prostriedky výstavby textu, ktoré možno adekvátne interpretovať len na pozadí textového celku. Jadro pojmového aparátu tvoria označenia jazykových premenných, ktoré sú základné operačné jednotky pre tvorenie textov (komunikátov).

Z definícií *textu*, v ktorých sa reflektuje jeho štruktúracia spomenieme ponímanie *textu* v zmysle komunikátu.⁴⁴ *Text* sa tu posudzuje ako istý komunikačný útvar, prvky ktorého sa z hľadiska štruktúry nachádzajú na troch rovinách: (a) vnútorná štruktúra sa týka jazykovej/semiotickej podoby textu; zahŕňa obsahové prvky (tému, motívy),⁴⁵ lexikálne a gramatické prostriedky, štylistické prostriedky, grafické a typografické prostriedky a i.; (b) vonkajšia štruktúra sa týka vzťahu medzi komunikačným konaním a sociálnymi štruktúrami;⁴⁶ zasahuje sociálne aspekty textu/komunikátu: spätosť s určitým komunikačným prostredím a s určitými typmi adresátov a inštitucionálnou distribúciou; (c) interakčná rovina sa týka situačnej realizácie (je

³⁹ Karlík, P. – Nekula, M. – Pleskalová, J. (eds.): Encyklopedický slovník češtiny. Praha: Nakladatelství Lidové noviny, 2002, s. 347.

⁴⁰ Glavaničová, D.: Explikácia modálnych pojmov. In: Filozofia, 2017, roč. 72, č. 7, s. 558.

⁴¹ Mluvnice češtiny 3, c. d., s. 632.

⁴² Mistrík, J. a kol.: Encyklopédia jazykovedy, c. d., s. 442.

⁴³ Slančová, D.: K vymedzeniu pojmu a termínu štyléma. In: Slovo o slove 12. Prešov: Katedra komunikačnej a literárnej výchovy Pedagogickej fakulty Prešovskej univerzity v Prešove, 2006, s. 14.

⁴⁴ Berger, P. L. – Luckmann, T.: Sociální konstrukce reality. Pojednání o sociologii vědění. Praha: Centrum pro studium demokracie a kultury, 1999. 214 s.

⁴⁵ Text nie je len usporiadaný sled textových jednotiek, ale aj sled motívov, pričom niet priamej zhody medzi jednotkami textu a motivickými jednotkami.

⁴⁶ Luckmann, T.: Observations on the Structure and Function of Communicative Genres. In: Semiotica, 2009, 173, pp. 267-282. DOI:10.1515/SEMI.2009.011.

odlišná v ústnych textoch a písaných/tlačených textoch). V textovej lingvistike sa vymedzuje sedem konštitutívnych princípov, ktoré vo vzájomnej interakcii tvoria *textovosť* v zmysle základnej vlastnosti *textu*: (i) kohézia; (ii) koherencia; (iii) intencionálnosť; (iv) informatívnosť; (v) akceptovateľnosť; (vi) situatívnosť; (vii) intertextovosť.⁴⁷ Z pohľadu kvantitatívnej lingvistiky *text* predstavuje množinu, ktorá je dynamickým systémom s fraktálovou štruktúrou.⁴⁸ Pochopenie textu tu vychádza z Mandelbrotovej charakteristiky fraktálov ako množín majúcich nejakú podobnosť medzi štruktúrnymi časťami každej množiny a celkom.⁴⁹ Operuje s tromi základnými pojmami: (a) podobnosť, (b) časť a (c) celok. Ústredným bodom teórie fraktálov je pojem dimenzie. *Text* je útvar ľubovoľnej dĺžky. Môže byť jednoslovný, ale text je i útvar, ktorý má dĺžku aj niekoľkokozväzkového románu. Lingvistický vzťah pomeru počtu lexikálnych jednotiek a dĺžky textu je nositeľom informácie týkajúcej sa Herdanovho vzťahu a výskytu koreferencií, ktorý je týmto vzťahom ovplyvnený.⁵⁰

2.1 Segmentácia textu

Vymedzenie *textových* premenných sa spája s problematikou *segmentácie textu* v zmysle funkčnej štruktúracie. Vychádza sa z formy, ktorá leží akoby na povrchu textu a je ľahšie dostupná kvantitatívnej analýze. *Segmentácia textu* znamená: (1) *rozloženie* do funkčne diskretných úsekov (segmentov) a ich (2) *vťahnutie* do významu, resp. funkcie (spravidla vo vete). Rozloženie na funkčné úseky zvyčajne korešponduje s vymedzením pomenovaní (jednoslovných a viacslovných). Vťahnutie do významu, resp. funkcie sprevádza odstránenie dvojznačnosti a homonymie a zistenie skutočnej referencie pomenovania.⁵¹ Odstránenie ambiguity – sémantiky alebo syntakticky podmienenej viacvýznamovosti jazykových výrazov na rozličných jazykových rovinách v texte – sa nazýva *dezambiguácia*.

Text je mnohovrstevný útvar. Z lineárneho charakteru textu vyplýva jeho horizontálne (lineárne) členenie na *začiatok* – *stred* – *koniec*. Výpovede realizované na začiatku a na konci textu tvoria rámec textu.⁵² Nelineárne, vertikálne členenie textu je založené na obsahovej hierarchii jednotlivých segmentov textu. Horizontálne (plošné) a vertikálne (plastické) členenie textu sa poníma ako **tektonika textu**.⁵³

Segmentácia textu zahŕňa (i) paratextovú segmentáciu: titul, motto, poznámky pod čiarou a i.; (ii) grafickú segmentáciu: typy grafém (písma) a ich veľkosť; (iii) tematickú segmentáciu: vyčlenenie tematicky homogénnych úsekov – tematických blokov; (iv) obsahovo-pragmatickú segmentáciu: vyčlenenie obsahovo-pragmatických jednotiek, t. j. úsekov komunikačnej udalosti s istou komunikačnou funkciou – elementárnou komunikačnou intenciou; (v) sémantickú segmentáciu: členenie obsahovo-pragmatických jednotiek na základné významové vzťahy (predovšetkým slovesno-menné vzťahy a ich modifikácie); (vi) syntagmatickú segmentáciu: vyčlenenie syntaktických konštrukcií.⁵⁴ V kvantitatívnej lingvistike sa *segmentácia textu* spája s vymedzením rovin textu, pri ktorom sa uplatňuje (a) hľadisko jazykových prostriedkov a (b) hľadisko vzťahov.

⁴⁷ Nekula, M.: *Textovosť*. In Karlík, P. – Nekula, M. – Pleskalová, J. (eds.): *CzechEncy – Nový encyklopedický slovník češtiny*, c. d., [cit. 24. 08. 2022].

⁴⁸ Hřebíček, L.: *Vyprávění o lingvistických experimentech s textem*, c. d., s. 65.

⁴⁹ Porov.: „Fraktál je štruktúra zložená z častí, ktoré sú v nejakom zmysle podobné celku“ (Mandelbrot, 1988; cit. podľa Hřebíček, 2002, s. 109).

⁵⁰ Hřebíček, L.: *Vyprávění o lingvistických experimentech s textem*, c. d., s. 134-135.

⁵¹ Čermák, F.: *Jazyk a jazykoveda. Přehled a slovníky*, c. d., s. 91.

⁵² *Mluvnice češtiny 3*, c. d., s. 591-592.

⁵³ Mistrík, J. a kol.: *Encyklopédia jazykovedy*, c. d., s. 446-447.

⁵⁴ Dolník, J. – Bajžiková, E.: *Textová lingvistika*. Bratislava: Stimul, 1998. 134 s.

Z pohľadu jazykových prostriedkov sa vyčleňujú entity, ktoré sú súčasťou aj jazykových rovín textu a ktoré nemusia nevyhnutne korešpondovať s jazykovými rovinami v lingvistike. K *textovým premenným* patria: agregát (hreb) – veta – klauza – slovo – (morféma – slabika) – hláska.⁵⁵ **Agregát (hreb)** predstavuje nadvetnú úroveň. Hreby sú jazykový konštrukt, ktorého konštituentom sú vety.⁵⁶ Tvoria samostatnú rovinu medzi textom a vetou, a vytvárajú špeciálny typ linearizácie medzi mentálnou a gramatickou linearizáciou. Hreb je v zásade nespojitá jednotka textu. Jedna veta môže patriť súčasne k viacerým hrebom. Premenné *veta, klauza, slovo, morféma, slabika, hláska* sú jazykové jednotky, ktoré zmenou statusu nadobudli hodnotu *textových premenných*.

Z hľadiska vzťahov boli zásluhou F. Daneša⁵⁷ vymedzené tieto roviny textu: (i) rovina sémantických medzi-propozičných vzťahov, t. j. kognitívne jazyková reprezentácia vzťahov medzi predmetmi a situáciami (stavy vecí) v reálnom svete; (ii) izotopické vzťahy (totožnosť alebo afinita predmetu reči);⁵⁸ (iii) tematické vzťahy (v zmysle aktuálneho členenia a tematických postupností, tematických trás operujúcich sčasti na sieti izotopických vzťahov); (iv) kompozičné vzťahy, t. j. vzťahy medzi jednotlivými výpoveďami (alebo ich blokmi) v texte (výstavbové funkcie jednotlivých informačných častí v celkovom pláne textu); (v) rovina štruktúry performatívnej interakcie (funkcie výpovedí či zámery).⁵⁹ Vzťahy medzi jazykovými rovinami textu zachytáva Menzerath-Altmanov zákon: vyjadruje sa ku vzťahu dĺžky jazykových jednotiek a ich konštituentov.⁶⁰ Každá jazyková entita je vo vzťahu k vyšším jazykovým úrovňam konštituentom a vo vzťahu k nižším je konštruktom.

S procesom segmentácie súvisí aj uplatňovanie **parovania** (parsing). Pojem je z oblasti počítačnej lingvistiky a znamená syntaktickú analýzu textu, automaticky (na počítači uskutočňovaný) proces a výsledok priradovania syntaktických štruktúr vetám prirodzeného jazyka zadaným na vstupe – výsledné syntaktické štruktúry. Podobne v korpusovej lingvistike značí proces tvorby a aj výsledok procesu tvorby syntaktických štruktúr (vetných celkov a vzťahov medzi vetnými členmi) napr. v podobe závislostného stromu.⁶¹ *Parovanie* v psycholingvistike označuje psychický proces spracovania výpovede.

2.2 Segmenty textu (textové segmenty)

K segmentom textu, ktoré nemajú status textových premenných možno zaradiť napr. textové pasáže, textové úseky, textové fragmenty, textové celky, úryvky textu a pod. Sú to časti textov, ktoré sú delimitované na základe osobného rozhodnutia, iniciatívy, preferencií niekoho, kto narába s textom. Textové jednotky so statusom textových premenných majú isté vlastnosti, majú vzťahy, tvoria štruktúry a podliehajú procesom. Pre textové jednotky platia tieto princípy: (a) hranice – textová jednotka musí byť relatívne dobre odlišiteľná od svojho okolia; hranice textových jednotiek nemusia byť však fixné – v jednom texte je tá istá jednotka pri analýze považo-

⁵⁵ Wimmer, G. – Altmann, G. – Hřebíček, L. – Ondrejovič, S. – Wimmerová, S.: Úvod do analýzy textov, c. d., s. 212.

⁵⁶ Wimmer, G. – Altmann, G. – Hřebíček, L. – Ondrejovič, S. – Wimmerová, S.: Úvod do analýzy textov, c. d., s. 209.

⁵⁷ Daneš, F.: Odstavec jako centrální jednotka tematicko-kompoziční výstavby textu (na materiále textů výkladových). In: Slovo a slovesnost. Časopis pro otázky teorie a kultury jazyka, 1994, roč. LV, č. 1, s. 1-2.

⁵⁸ *Predmet reči* je podľa F. Daneša (Daneš, F.: Veta a text. Praha: Academia, 1985) nešpecifikovaná entita, čokoľvek, čoho sa jazykový prejav týka, čo má hovoriaci na mysli, čo v prehovore uvádza a pomenúva. Je to vedomá alternatíva k pojmu referent.

⁵⁹ *Poznámka*: elementy rovín (i) a (v) zobrazujú mimotextové vzťahy, elementy rovín (ii), (iii) a (iv) vyjadrujú vnútrotextové vzťahy.

⁶⁰ Porov.: <https://wiki.korpus.cz/doku.php/pojmy:menzerath> [cit. 22. 07. 2022].

⁶¹ Porov.: <https://terminologickyportal.sk/> [cit. 30. 07. 2022].

vaná za jednu textovú jednotku, v inom texte pri inej analýze je považovaná za dve; (b) identita – textová jednotka musí mať aspoň vágnu identitu, t. j. istú existenčnú stabilitu; (c) integrácia – textová jednotka musí byť súčasťou aspoň jedného systému.⁶²

Vymedzenie textových jednotiek v zmysle *textových segmentov* sa v korpusovej lingvistike spája so segmentáciou textu ako procesom a najmä výsledkom „(automatizovaného) členenia celku na segment“.⁶³ V korpusovej lingvistike sa spája prevažne s (a) tokenizáciou a (b) vetnou segmentáciou.

Tokenizácia je automatický proces, ktorý člení text na jednotlivé *tokeny*, t. j. na slovné tvary a interpunkčné znamienka pre účely ďalšieho (spravidla počítačového) spracovania.⁶⁴

Token (token) je najmenšia jednotka textu, spravidla slovo, resp. jeho konkrétna realizácia (slovný tvar), ktorá je vymedzená v opozícii k **type** (ako dekontextualizovanej jednotke, jednotke abstrakcie). V počítačovej a korpusovej lingvistike sa *token* poníma ako „základná textová jednotka použitá v analýze alebo spracovaní textu, znak alebo súbor znakov formálne oddelených od iného znaku alebo súboru znakov napr. medzerou (tvar slova v texte, interpunkčné znamienko, číslica a pod.)“.⁶⁵ *Token* sa najčastejšie asocjuje so slovným tvarom, t. j. „slovným útvarom medzi dvoma medzerami“.⁶⁶ *Typom* môže byť ľubovoľná jednotka, najčastejšie sa asocjuje s *lemou*.

Identifikovanie pomeru *typov* a *tokenov* vypovedá o povahe príslušného textu. Vysoký pomer naznačuje lexikálnu bohatosť a rôznorodosť textu, nízky pomer naznačuje vysokú mieru opakovania *lexém*. Pomer *typov* a *tokenov* sa využíva pri meraní **lexikálnej hustoty textu** (počet plnovýznamových lexikálnych jednotiek sa vydělí počtom všetkých *tokenov* v texte a vynásobí sa číslom 100) aj pri meraní **slovného bohatstva textu** – vzťah medzi počtom *typov* a počtom *tokenov* približuje index označovaný ako type-token pomer (TTR).

Lema je predovšetkým abstraktná entita korešpondujúca s jednoslovnou lexikálnou jednotkou vo svojom východiskovom tvare; slovníkové slovo; „kánonický tvar lexémy“.⁶⁷ V korpusovej lingvistike sa definuje ako „základný, slovníkový tvar, pri ohybných slovných druhoch reprezentujúci všetky tvary príslušného slova z ohybných slovných druhov“.⁶⁸ Súbor všetkých slovných tvarov príslušnej *lemy* tvorí *paradigmum*. V prípade slovenčiny ako flektívneho jazyka *lemy* ohybných slovných druhov reprezentujú celú paradigmu, napr. *lema* slovíes má tvar neurčitku, *lemy* menných slovných druhov majú tvar nominatívu singuláru a v prípade gramatického rodu – tvaru mužského rodu. Určenie základného tvaru slova (*lemy*), resp. formálne priradenie istého unikátneho reťazca znakov identifikujúceho lexému sa posudzuje ako **lematizácia**.

Slovný tvar je jeden konkrétny tvar *lemy*.⁶⁹

Textové slovo – konkrétna manifestácia slova v texte; jeho delimitácia sa riadi medzerami v grafickom zázname textu. *Textové slovo* je konštrukt, ktorý má dva rôzne typy konštituentov – morfy a slabiky.⁷⁰ V kvantitatívnej lingvistike sa dĺžka textového slova v zmysle lineárneho rozmeru textovej premennej spravidla vymedzuje počtom slabík alebo počtom grafém v slove.

⁶² Wimmer, G. – Altmann, G. – Hřebíček, L. – Ondrejovič, S. – Wimmerová, S.: Úvod do analýzy textov, c. d., s. 18-19.

⁶³ Porov.: <https://terminologickyportal.sk/> [cit. 22. 07. 2022].

⁶⁴ Petkevič, V.: Tokenizace. In Karlík, P. – Nekula, M. – Pleskalová, J. (eds.): CzechEncy – Nový encyklopedický slovník češtiny, c. d., [cit. 24. 08. 2022].

⁶⁵ Šimková, M.: Token. In: Výberový slovník termínov z počítačovej a korpusovej lingvistiky, c. d., [cit. 22. 07. 2022].

⁶⁶ Uhlířová, L.: O vztahu mezi délkou slova a jeho polohou ve větě. In: Slovo a slovesnost, 1997, roč. 58, č. 3, s. 174-184.

⁶⁷ Čermák, F.: Lexikon a sémantika, c. d., s. 233.

⁶⁸ Porov.: <https://terminologickyportal.sk/> [cit. 22. 07. 2022].

⁶⁹ Porov.: <https://terminologickyportal.sk/> [cit. 30. 07. 2022].

⁷⁰ Wimmer, G. – Altmann, G. – Hřebíček, L. – Ondrejovič, S. – Wimmerová, S.: Úvod do analýzy textov, c. d., s. 212.

Textová jednotka v korpusovej lingvistike korešponduje s *tokenom*.⁷¹

V korpusovej a počítačovej lingvistike sa používa spojenie **regulárny výraz** ako „textový reťazec pozostávajúci z alfanumerických znakov a špeciálnych symbolov utvorený na základe určitých konkrétnych syntaktických pravidiel, ktorý opisuje istú množinu reťazcov vyhovujúcich danému regulárnemu výrazu, napr. umožňuje vyhľadávať slovo vo všetkých jeho tvaroch v prípade nelematizovaného korpusu a/lebo rôzne kombinácie slov, tvarov a iných reťazcov znakov“.⁷² *Reťazec znakov* je „v programovaní a formálnych jazykoch usporiadaná postupnosť symbolov z danej abecedy“.⁷³

Vetná segmentácia v korpusovej lingvistike je automatický proces, ktorý člení text na *vety/výpovede*. Spravidla je to druhá fáza automatického spracovania textu.

Veta v zmysle *výpovede* v ponímaní M. Foucaulta predstavuje „atóm“ diskurzu.⁷⁴ V písaných textoch sú hranice medzi výpoveďami signalizované grafickými prostriedkami. Komunikačná platnosť *výpovede* je daná jej situačným ukotvením, t. j. jej určením z hľadiska determinantov komunikačnej situácie: miesta, času a aktérov. Vzťah *slova* a *výpovede* v prehovore aj v texte reflektuje *princíp kontextu*, ktorý sformuloval H. Putnam na základe sémantického holizmu G. Fregeho a podľa ktorého *slová* nemajú význam a mimo kontextu *výpovede*, ktorej sú súčasťou, na nič neodkazujú; význam určitej *výpovede* však nie je súhrnom významov slov, z ktorých je táto výpoveď zložená.⁷⁵ Počet *slov* vo *vete* sa v kvantitatívnej lingvistike spravidla vzťahuje na dĺžku *vety* v zmysle lineárneho rozmeru textovej premennej. Vzťah *vety/výpovede* a textu je vzťahom *vety/výpovede* a jej textového okolia.

Základnými pojmami komunikačného a textologického členenia *vety/výpovede* sú *téma* a *réma*. *Téma* (východisko, topik) predstavuje informáciu známu účastníkom výpovedného aktu, teda časť nadväzujúcu na predchádzajúci známy text. *Réma* (jadro, ohnisko, fokus) je aktuálna zložka *vety/výpovede* prinášajúca novú informáciu. Súvislosti medzi tematicko-rematickou štruktúrou výpovede v texte a izotopickými (anaforickými) vzťahmi v texte F. Daneš⁷⁶ označil ako *tematické postupnosti* a opísal ich typy, ku ktorým patrí (i) jednoduchá tematická postupnosť; (ii) tematická postupnosť s konštantnou témou; (iii) tematická postupnosť s odvodenou témou; (iv) tematická postupnosť s rozštiepenou rémou.⁷⁷ Pre rovinu významovej (hlbkovej) roviny vetnej stavby sa vo funkčne generatívnej gramatike používa termín *tektogramatická reprezentácia*, ktorá je dezambiguovaným zápisom významu *vety*.⁷⁸

Záver

V štúdií sme prezentovali výber a kvalitatívnu interpretáciu lingvistických premenných uplatňovaných v metodológii lingvistických analýz písomných textov. Impulzom bol nárast

⁷¹ Šimková, M.: Textová jednotka. In: Výberový slovník termínov z počítačovej a korpusovej lingvistiky, c. d., [cit. 22. 07. 2022].

⁷² Šimková, M.: Regulárny výraz. In: Výberový slovník termínov z počítačovej a korpusovej lingvistiky, c. d., [cit. 10. 12. 2021].

⁷³ Šimková, M.: Reťazec znakov. In: Výberový slovník termínov z počítačovej a korpusovej lingvistiky, c. d., [cit. 22. 07. 2022].

⁷⁴ Cingerová, N. – Motyková, K.: Úvod do diskurznej analýzy. Bratislava: Univerzita Komenského v Bratislave 2017, s. 38.

⁷⁵ Porov.: <https://beliana.sav.sk/heslo/kontext> [cit. 24. 08. 2022].

⁷⁶ Daneš, F.: Veta a text. Praha: Academia, 1985, s. 198-207.

⁷⁷ Daneš, F.: Typy tematických poslopností v textu. In: Kultura a struktura českého jazyka. Praha: Karolinum, 2009, s. 420-440.

⁷⁸ Panevová, J.: Tektogramatika. In: Karlík, P. – Nekula, M. – Pleskalová, J. (eds.): CzechEncy – Nový encyklopedický slovník češtiny, c. d., [cit. 30. 07. 2022].

a prehodnocovanie pojmov v dôsledku postupne akcelerujúceho procesu konštituovania databáz slovenských terminologických sústav počítačovej, kvantitatívnej a korpusovej lingvistiky a definičného vymedzovania ich aparátu, ktoré vychádza a prirodzene využíva pojmoslovie systémovej lingvistiky. Výraz *premenná* sme si zvolili pre flexibilitu jeho hodnotového a statusového nastavenia. *Lingvistické premenné* sme posudzovali z pohľadu (a) jazykových premenných (entít jazykového systému potenciálne konštituujujúcich text) a (b) textových premenných (entít textu realizovaných s oporou aj bez opory jazykových premenných). Vychádzali sme zo skutočnosti, že vymedzenie textových premenných sa spája s problematikou segmentácie textu, ktorá je určená a limitovaná počítačovým spracovaním textu. Uchopovanie segmentov textu v zmysle jazykových prvkov, prostriedkov a jednotiek sa tu realizuje prioritne z aspektu povrchovej štruktúracie textu. Výrazová sústava textovej lingvistiky vo vzťahu s výrazovou sústavou systémovej lingvistiky vykazuje: (i) *aktualizáciu* výrazov systémovej lingvistiky, ktoré nepredstavujú jej nosné pojmy a nie sú v nej binárne ukotvené: *graféma* (v zmysle grafického znaku), *slabika*, *slovo*; (ii) *neutralizáciu* binárnej kategorizácie výrazov systémovej lingvistiky a používanie pojmov z oblasti langue vo funkcii parolových výrazov: *veta*, *morféma*; (iii) *nepotvrdenie* výrazov systémovej lingvistiky: *seméma*, *pomenovanie*, *frazéma*, *propozícia*, *štyléma*, *textéma*; (iv) *implementovanie* nových výrazov, ktoré absentujú vo výrazovej sústave systémovej lingvistiky: *token*, *lema*, *agregát (hreb)* a i. Je zrejmé, že pojmoslovia systémovej aj textovej lingvistiky prejavujú aj konzistentnosť aj akomodačné tendencie, t. j. asimilačné tendencie a prispôsobovanie meniacim sa metodologickým potrebám.

Linguistic variables in the methodology of text analysis.

(A probe into Slovak terminological databases)

Jana Sokolová – Augustín Sokol

In the study, we presented the selection and qualitative interpretation of linguistic variables, which are used in the methodology of linguistic analyses of written texts. The impetus to write this paper was the increasing number and reevaluation of terms as a result of the gradually accelerating establishment of Slovak terminological databases of computer, quantitative and corpus linguistics and the definitions used in these databases, which are based on and naturally use the conceptual system (terminology) of systemic linguistics. The term *variable* was chosen for the flexibility of its status and value. We assessed *linguistic variables* from the point of view of (a) linguistic variables (entities of the language system potentially constituting the text) and (b) textual variables (entities of the text realized with and without the help of linguistic variables). We followed from the fact that the definition of text variables is connected with the issue of text segmentation, which is determined and limited by computational processing of text. In relation to the terminological system of systemic linguistics, the terminological system of textual linguistics shows (i) *updated meanings* of the expressions from systemic linguistics that do not represent its key terms and are not anchored in it; (ii) *neutralization* of the binary categorization of terms from systemic linguistics and the use of terms from the field of langue in the function of parole terms; (iii) *failure to confirm* the following expressions from systemic linguistics; (iv) *implementation* of new expressions that are absent in the terminological system of systemic linguistics. It is obvious that the terminology of both systemic and textual linguistics shows both consistency and accommodation tendencies, i.e. assimilative tendencies and adaptation to the changing methodological needs.