
MERJENJE DOLOČENIH VIDIKOV SKLADENJSKE KOMPLEKSNOŠTI V PISNIH BESEDILIH SLOVENSКИH OSNOVNOŠOLCEV

Kompleksnost je v jezikoslovju pogosto, a premalo dosledno obravnavan pojem. V uporabnem jezikoslovju je večinoma opredeljen kot ena izmed dimenzij aktivnega znanja jezika ter jezikovne performance. Redkeje so raziskave, ki bi merile kompleksnost materinščine ter presegle enačenje kompleksnosti z jezikovnim razvojem. Prispevek na podlagi na novo oblikovanih meril kompleksnosti, ki zajamejo kar največ vidikov tega pojma, preveri že znane metode za merjenje skladijske kompleksnosti na pisnih besedilih slovenskih osnovnošolcev.

Ključne besede: pragmatika, kompleksnost v pisnih besedilih, skladijska kompleksnost, dolžina skladijskih struktur, strukturalna kompleksnost, pogostnost skladijskih struktur, raznolikost skladijskih struktur

1 Pojem kompleksnosti v jezikoslovju

1.1 Opredelitev kompleksnosti

Kompleksni sistemi, kamor uvrščamo tudi jezik, so sestavljeni iz preprostih elementov, ki med seboj tvorijo povezave, se razvijajo, razvrščajo in tvorijo vzorce (Mitchell 2009: 4). Zaradi poenotenja znanstvenega pojma pojmem kompleksnost kot eno izmed dimenzij aktivnega znanja jezika ter jezikovne rabe oziroma performance¹ (Housen, Kuiken, Vedder 2012) ter sledim že vzpostavljeni opredelitvi kompleksnosti: »število in lastnosti samostojnih sestavnih delov neke

¹ Zaradi poenotenja pojmov uporabljam po Ferbežar (1997) za angleški izraz *lanugage proficiency* prevod **aktivno znanje jezika**, za *language performance* pa **jezikovna raba** oziroma **performanca**.

danosti« ter »število in narava razmerij med sestavnimi deli« (Bulté 2012–2013: 58). Kompleksnost je že sam po sebi precej večdimenzionalen pojem, ki ga raziskovalci pogosto uporabljajo, redko pa predhodno opredelijo. V tem članku strogo sledim le navedeni opredelitvi, ki je v povezavi s pisnimi besedili najpogosteje prisotna na področju uporabnega jezikoslovja. Zavedam pa se, da je kompleksnost možno obravnavati tudi z drugih vidikov.

1.2 Kompleksnost z uporabnojezikoslovnega vidika

Kompleksnost je v teoriji dinamičnih sistemov poleg pravilnosti in tekočnosti² obravnavana kot ena izmed dimenzij aktivnega znanja jezika ter jezikovne performance. Kot takšno jo je moč opazovati ter meriti na podlagi jezikovne produkcije (pisne in govorjene). Najpogostejše so longitudinalne študije, ki na podlagi obsežnejših korpusov opazujejo povečevanje kompleksnosti, saj jo eksplicitno ali implicitno enačijo z razvitostjo (Vercellotti 2012, Verspoor, Schmid, Xu 2012, Bulté 2012–13). Kljub temu da učenci načeloma sčasoma tvorijo vedno bolj kompleksne strukture, se kompleksnosti ne sme enačiti z boljšim znanjem oziroma rabo jezika. Dobro znanje jezika lahko deluje celo v nasprotju z jezikovno kompleksnostjo, saj imajo učenci višje stopnje več izkušenj z reševanjem določenega tipa nalog; na podlagi teh izkušenj lahko pri opravljanju nalog uporabijo minimalna – le za tisto nalogo nujno potrebna – jezikovna sredstva (Robinson, White 1995). Izkušeni govorci ali pisci kompleksne ideje pogosto izrazijo na enostavnejši način kot novinci, saj jim aktivno znanje jezika omogoča učinkovitejše izražanje (Lambert 2014).

2 Pregled dosedanjih raziskav s področja kompleksnosti pisnih besedil

Oblikovanje celovitega pregleda raziskav kompleksnosti je težavna naloga, saj je pojem v raziskavah pogosto različno opredeljen. Nekatere raziskave merijo kompleksnost, a je ne poimenujejo s tem izrazom, druge govorijo o kompleksnosti, čeprav opazujejo druge vidike. Raziskava Jarvis, Bikowski in Ferris (2003) (v Verspoor idr. 2017: 17) denimo izpostavi povprečno dolžino besed kot enega izmed kazalcev kompleksnosti, vendar omenjeni avtorji nikoli ne uporabijo izraza kompleksnost, temveč govorijo o najboljše ocenjenih besedilih.

Ker se v članku ukvarjam predvsem s kompleksnostjo v smislu števila elementov in odnosov med njimi, se omejujem le na tiste vidike kompleksnosti, ki ustrezajo tej opredelitvi. Kot prvi in najpogostejši vidik kompleksnosti naj navedem dolžino. Meritve tako pisnih (Bulté 2012–13, Ortega 2003, Housen, Kuiken 2009 idr.) kot

² Tekočnost (angl. *fluency*) se v slovenskem jezikoslovju uporablja predvsem v povezavi z bralno tekočnostjo (Jurišić 2016). Izraz se uporabi tudi v SEJO, in sicer ga uvršča pod jezikovne vidike. V tem članku pa izraz označuje eno izmed dimenzij aktivnega znanja jezika in je kot taka predvsem fonološki pojav jezikovne produkcije (hitrost in gostota jezikovnih enot; število, dolžina in mesto premorov; ponovitve, samopopravljanje, napačne formulacije, napačni začetki) (Housen, Kuiken, Folkert, Vedder, Ineke 2012: 5).

govorjenih (Kranjc 1999: 106, Vercellotti 2012: 107) besedil dokazujejo, da se z leti učenja/usvajanja jezika daljšajo stavki (govorne enote), T-enote³ (Hunt 1966, Crowhurst 1983) ter povprečna dolžina povedi, ki naj bi imela celo prednost pred T-enoto, saj je povedi računalniško lažje prepoznati, hkrati pa se upošteva še tvorčevo zavestno odločitev, kaj bo pojmoval kot samostojno enoto, medtem ko naj bi bilo ločevanje na T-enote umeten poseg v besedilo (Vyatkina 2012: 4). Povedi se z leti učenja/usvajanja daljšajo predvsem na račun večjega deleža kompleksnih (podredno zloženih) ter enostavnih (enostavnih) povedi s polstavčnimi strukturami (Verspoor idr. 2017), med daljšanjem povedi in deležem priredno zloženih povedi pa obstaja celo negativna korelacija (Verspoor idr. 2017), kar pomeni, da se povedi ne daljšajo na račun priredij, temveč podredij (podredne strukture se torej v pisanju razvijejo kasneje). Če na besedilo gledamo tudi s pragmatičnega, ne le s strukturnega vidika, naletimo še na pojem stavčnih paketov.⁴ Stavčni paketi veliko povedo o večji ali manjši zgoščenosti informacij skozi besedilo.

Kot eden izmed pokazateljev razvoja jezikovne zmožnosti se poleg kompleksnosti omenja tudi raznolikost, predvsem leksikalna (Malvern et. al. 2004, Bulté 2012–13). Raznolikost pomeni relativno pestrost struktur v besedilu, zato je tesno povezana s kompleksnostjo v smislu večjega števila elementov in različnih razmerij med njimi. Ker se v raziskavi omejujem le na skladnjo, sem v meritev vključila skladenjsko raznolikost (podrobneje razloženo v 4.4). Znano je, da obstaja pomembna korelacija med številom različnih vezniških besed ter raznolikostjo semantičnih razmerij med stavki (Doleschal, Robatsch 2015: 175). Tvorci raje uporabljajo tiste stavčne strukture, ki so bile v besedilu že uporabljene (Bock, 1986; Pickering & Branigan, 1999), kar je posledica skladenjske oziroma strukturne aktivacije⁵ (Gries 2005: 365).

3 Cilj raziskave in metodologija

3.1 Cilj raziskave

Cilj raziskave je na podlagi izoblikovanih meril (predstavljenih v analizi), ki upoštevajo določene vidike skladenjske kompleksnosti, izmeriti skladenjsko kompleksnost slovenskih pisnih besedil. V analizi so opisane metode za merjenje

³ Minimalna končna enota oziroma T-enota (ang. *Minimal terminable unit – T-unit*) je glavni stavek skupaj z vsemi stavki, ki so glavnemu podrejeni ter z vsemi nestavčnimi strukturami, ki so od njega odvisne oziroma so mu priključene (Hunt 1965: 21). Dva priredno povezana stavka torej pomenita dve T-enoti.

⁴ Stavčni paket je enota, sestavljena iz več stavkov, ki so med seboj povezani z abstraktnimi povezavami, ki niso nujno identificirane s skladenjskimi kriteriji; največkrat se razmerja odražajo preko prirednih oziroma podrednih struktur, včasih pa se odkrijejo preko tematske progresije besedila; upošteva se tudi menjanje teme v diskurzu (Berman, Nir 2008: 150). Stavčnega paketa torej ne moremo popolnoma enačiti s tematsko enoto.

⁵ Zaradi poenotenja slovenske terminologije uporabim prevod *aktivacija* (angl. *priming*) iz Verdonik 2015.

skladenjske kompleksnosti. Uporabim metode, najpogosteje omenjene v raziskavah, ki merijo kompleksnost pisnih besedil v različnih jezikih, in se do njih kritično opredelim. Na podlagi rezultatov razložim, kako se skladdenjska kompleksnost kaže v posameznih pisnih besedilih, in ugotavljam, ali so besedila, ki so bolj (manj) kompleksna, bolj (manj) kompleksna po vseh kriterijih kompleksnosti. Za razliko od večine raziskav, ki preučujejo kompleksnost, se moja ne osredotoča na kompleksnost v smislu razvoja jezikovne zmožnosti, temveč opazuje, kako se kompleksnost kaže v besedilu (torej z vidika jezikovne performance). Tokrat sem se osredotočila zgolj na pisna besedila, saj bi morala biti pri merjenju skladdenjske kompleksnosti govorjenih besedil merila precej drugačna.

3.2 Metodologija

3.2.1 Vzorec

V raziskavi je sodelovalo 20 osmošolcev⁶, starih med 13 in 14 let, ki obiskujejo slovensko OŠ s slovenščino kot učnim jezikom. Materni jezik dveh učencev ni bil slovenščina, temveč hrvaščina (B3) in albanščina (B18). Učenca sta imela sicer precej visoko zmožnost vsakdanjega sporazumevanja v slovenščini, v pisnih besedilih pa je bil vpliv maternega jezika bolj očiten.

3.2.2 Instrumentarij

Osnovnošolci so tvorili pisno besedilo na podlagi posnetka – kratkega nemega animiranega filma. Film je prikazoval preprosto zgodbo dekleta, ki je gradila modele letal in rešila težavo mladega pilota, ki se mu je na poti domov pokvarilo letalo; film je trajal 6,15 minut. Vsi udeleženci so si ga na dan raziskave ogledali prvič. Besedila niso bila napisana v skladu s pravili besedilne vrste (npr. obnove). Učenci so dobili navodilo (glej Prilogo), naj pišejo v sedanjiku, vendar se je pri pregledu vzorcev izkazalo, da se niso vsi držali navodil. Po ogledu so imeli dobrih 30 minut časa, da napišejo natančen potek dogajanja. Film so si ogledali dvakrat, enkrat v celoti brez pavz, drugič sem film ustavljala (približno na 1 minuto, da so se natančno spomnili dogajanja in scenografije); naloga je zahtevala natančno obnovo dogajanja in opis okolice, glavnih junakov. Prvo poved sem napisala sama, s čimer sem jim ponudila izhodišče. Druge pol povedi sem prav tako napisala sama, udeleženci so lahko poved nadaljevali ali jo zaključili in začeli z novo. Vzroci so bili anonimni, označila sem jih s številčnim indeksom: B1, B2 ... B20.

⁶ Zaradi zagotavljanja anonimnosti uporabljam pri naslavljanju učencev, vključenih v raziskavo, neznanomani moški spol (osnovnošolec, učenec).

3.2.3 Priprava gradiva za analizo

Besedila učencev sem pretipkala v takšni obliki, kot so nastala. Pri analizi vzorcev sem vključila drugo predhodno napisano poved, ki je bila pri vseh vzrocih enaka: *V roki drži ročno izdelano letalo*. Pri analizi posameznih meritev je navedeno, na kakšen način sem pri prepisu pristopala do pravopisnih/slovničnih napak – navsezadnje sem merila le dimenzijo kompleksnosti (in ne pravilnosti).

4 Analiza merjenja skladenjske kompleksnosti

Če upoštevam opredelitev kompleksnosti iz uvoda tega članka, ki v ospredje postavlja predvsem število elementov in odnose med njimi, potem lahko oblikujem tri, v tujih raziskavah že preizkušena, merila kompleksnosti. Poleg tega uvajam še četrto, tj. merilo raznolikosti, ki se mi je v raziskovalnem postopku pokazalo kot relevantno.

Merilo dolžine: daljše skladenjske strukture so kompleksnejše od krajših (daljše povedi so bolj kompleksne od krajših, daljši stavki so bolj kompleksni od krajših, daljše T-enote so bolj kompleksne od krajših, daljši stavčni paketi so bolj kompleksni od krajših).

Merilo strukturne kompleksnosti: nekatere strukture so kompleksnejše od drugih, ker so kasneje oziroma težje usvojljive/naučene (podredno zložene povedi so bolj kompleksne od priredno zloženih; samostalniške zveze z več določili so bolj kompleksne od tistih z manj določili).

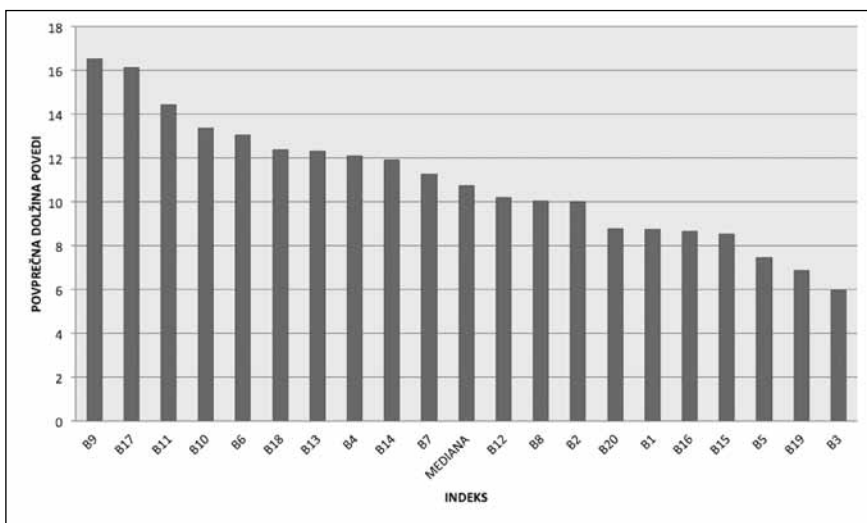
Merilo pogostnosti: manj pogoste skladenjske strukture so kompleksnejše od bolj pogostih. Opazujem večjo/manjšo pogostnost le tistih struktur, ki se pojavijo v uporabljenem korpusu besedil (najpogostejše in najmanj pogoste skladenjske strukture so navedene v 4.3).

Merilo raznolikosti: večja razpršenost različnih skladenjskih struktur po besedilu pomeni večjo raznolikost (pestrejšje menjavanje struktur), večja raznolikost pa pomeni večjo kompleksnost.

4.1 Dolžina kot merilo skladenjske kompleksnosti

Povprečna dolžina povedi je razmerje med številom besed in številom povedi v besedilu. Namesto števila besed je enota lahko tudi število stavkov, saj dobimo pri obeh meritvah podobne rezultate – korelacija med obema je dovolj močna ($\rho = 0.8185$). Težava se pojavi pri pojmovanju pojma povedi – če jo pojmujejo kot besedilo, ki ga zamejujeta velika začetnica in končno ločilo, še ne zagotovimo popolne objektivnosti, saj zanemarimo dejstvo, da predvsem mlajši in manj izkušeni pisci ločila postavljajo nedosledno (Hunt 1965: 24). Po drugi strani pa lahko ravno

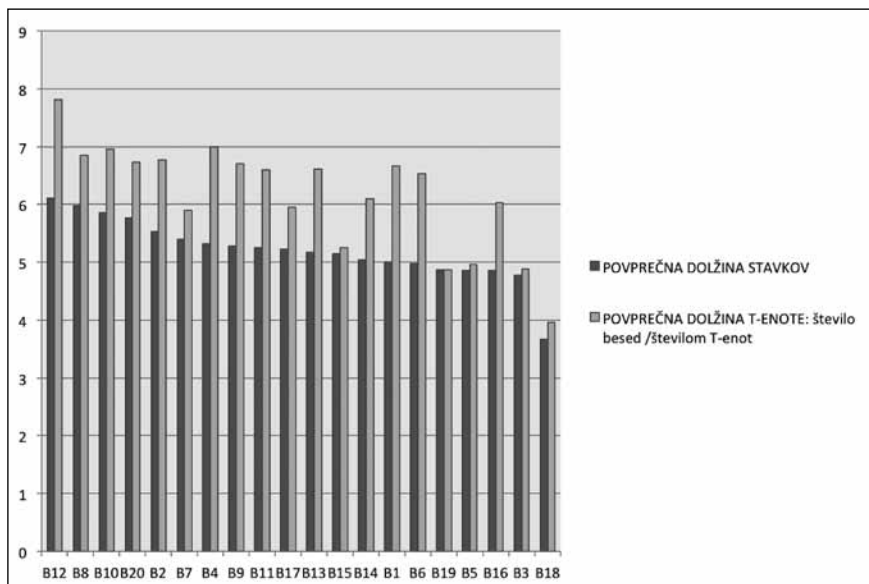
omenjeno dejstvo pokaže na večje/manjše obvladovanje enot besedila. Poleg tega se učenci v višjih razredih osnovne šole srečujejo z zgoraj omenjeno definicijo povedi, zato sem se je držala tudi pri analizi. Dolgih povedi nisem razbijala, kljub temu da je bilo jasno, da gre v nekaterih primerih za več povednih enot, kot jih je pisec označil (*Dekle meče letalo ampak ne vzleti, dekle je zelo užaljena, vznemirjena, zelo ima rada letala saj se vidi po puloverju da ima narisano letalo.*). Kot novo poved sem štela le tisto, ki je bila zaznamovana s končnim ločilom, razen v primeru, da sem sklepala, da je pisec ločilo pozabil zaradi naglice (saj je vseeno uporabil veliko začetnico) ali namesto končnega ločila naredil vejico (*deček ima oblečeno zelene hlače, Deček vzame svinčnik; Ko je prišel na nebo je iz ruzaka vzel zvezdo in nato se mu je prižgalo Deček si je je zahvalu*).



Graf 1: Povprečna dolžina povedi.

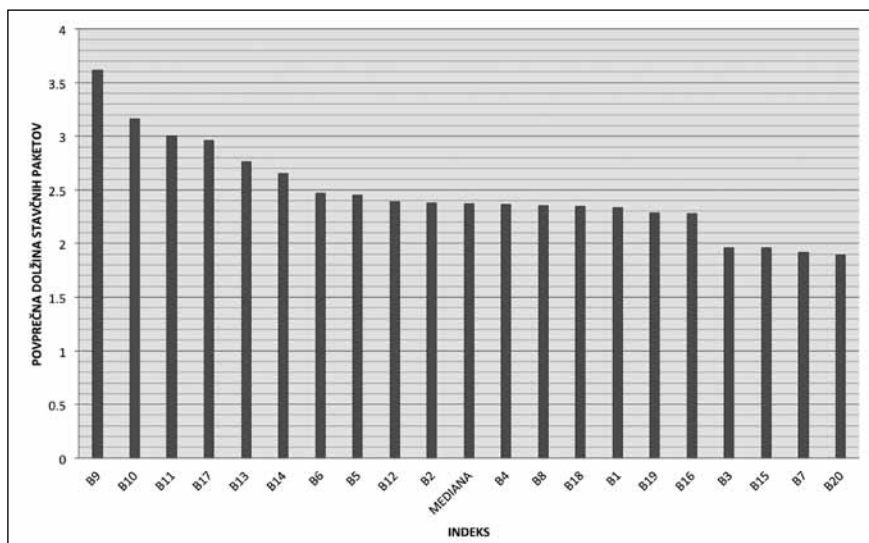
Povprečna dolžina stavkov je razmerje med številom besed in številom stavkov. Povprečje povprečnih dolžin vseh stavkov, ki so jih za pričujočo raziskavo napisali osnovnošolci, je 5.1409.

Poleg povprečne dolžine povedi je pri merjenju skladske kompleksnosti najpogostejša še povprečna dolžina T-enote (razmerje med številom besed v T-enoti ter številom T-enot v besedilu).



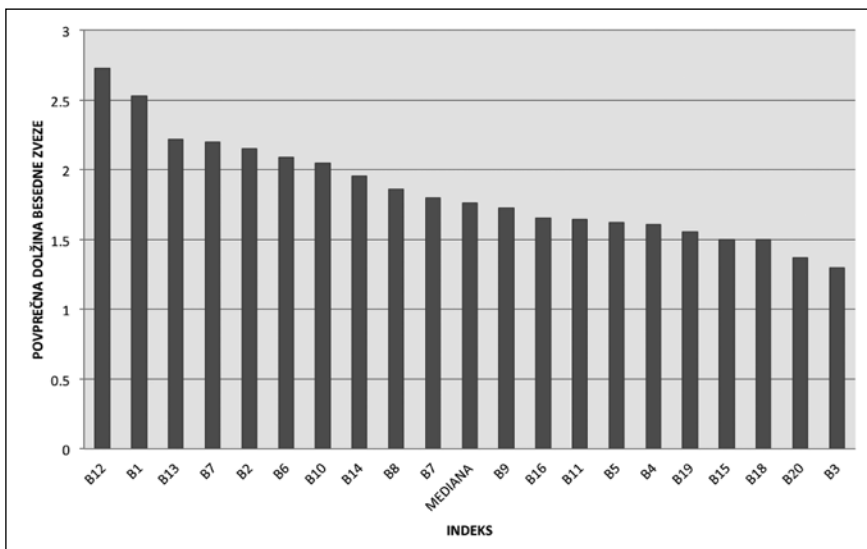
Graf 2: Povprečna dolžina stavka ter T-enote.

Stavčni paketi presegajo T-enoto v tem smislu, da upoštevajo tudi pragmatični vidik, predvsem tematsko razporejenost elementov po besedilu.



Graf 3: Povprečna dolžina stavčnih paketov.

Povprečna dolžina besedne zveze (samostalniške in predložne, npr. *prestrašen palček, ki nosi rjavo kapo z letalskimi očalmi, zelene hlače z naramnicami in belo majico / na klopico pod drevo*) močno korelira s povprečno dolžino stavčnega paketa ($\rho = 0.8139$).

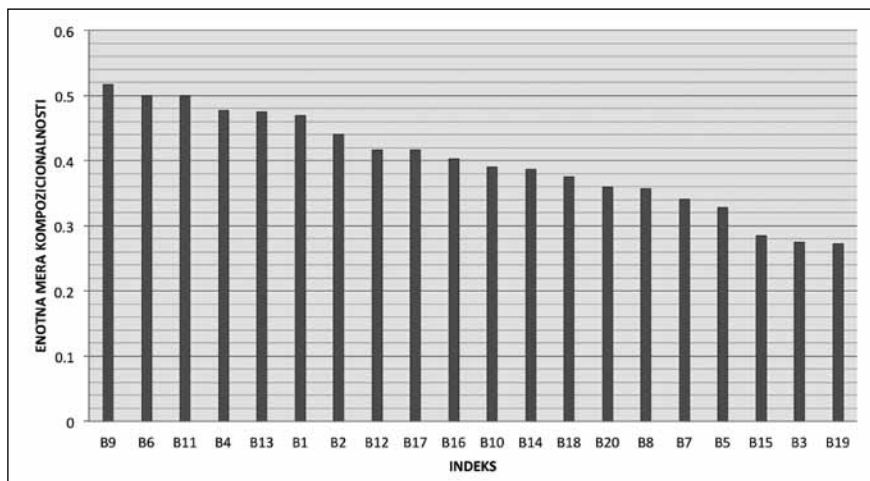


Graf 4: Povprečna dolžina samostalniške in predložne besedne zveze.

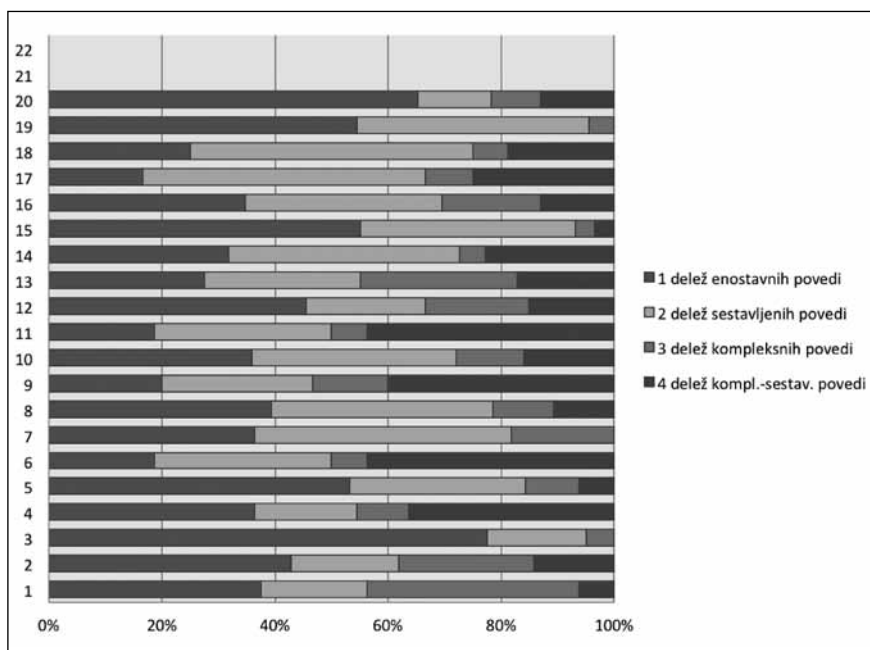
4.2 Delež kompleksnejših struktur v besedilu

Pri uvrščanju različnih povedi med bolj ali manj kompleksne sem sledila Bultéju (2012–2013: 80, 81), ki povedi (besedilo med veliko začetnico in končnim ločilom) razdeli (od najmanj do najbolj kompleksnih) na: enostavne (enostavčne) < sestavljene (priredno zložene) < kompleksne (podredno zložene) < kompleksno-sestavljene (priredno in podredno zložene). Večji delež kompleksnih in kompleksno-sestavljenih povedi pomeni večjo kompleksnost na kompozicionalni ravni. V uporabljenem korpusu ima 40 % besedil največji delež enostavnih povedi, 30 % besedil največji delež sestavljenih povedi, 25 % besedil ima iznačen delež enostavnih in sestavljenih povedi.

Enotno mero, s katero se izoblikuje profil vsakega besedila glede na kompozicionalnost, kjer je 1 največja kompleksnost, 0 pa najmanjša, sem računala po formuli: $(1 - \text{delež enostavnih povedi}) + (1 - \text{delež sestavljenih povedi}) + \text{delež kompleksnih} + \text{delež kompleksno-sestavljenih povedi} / 4$.



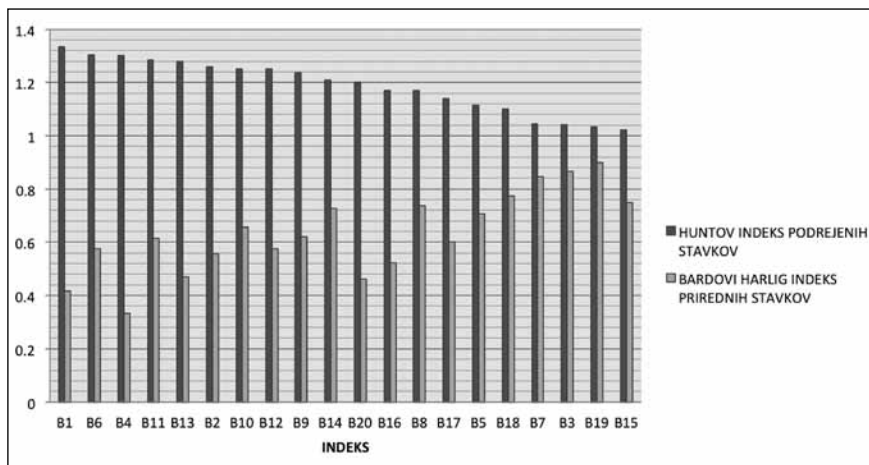
Graf 5: Enotna mera kompozicionalnosti.



Graf 6: Delež posameznih vrst povedi (vodoravna os) v besedilih (navpična os).

Ker veljajo priredne strukture za manj kompleksne od podrednih, so besedila z večjim deležem prirednih struktur manj kompleksna. Za izračun deleža prirednih struktur sem uporabila indeks priredij Bardovi Harligove. Gre za stopnjo, do katere pisec doseže skladijsko kompleksnost s prirednimi strukturami (število vseh povedi se odšteje od števila vseh stavkov – s tem se dobi le stavke, ki so preko

priređja oziroma podređja dodani glavnim stavkom; nato se izračuna še količnik med stavki, ki so priredno dodani in prejšnjo razliko) (Bardovi Harlig 1992: 393); pomanjkljivost indeksa je, da ne upošteva prirednih razmerij med povedmi, npr. (*Omenjena deklica bi rada v zrak poslala doma narejeno letalo. A letalo noče poleteti, deklica je obupana, kar naenkrat na mizo prileti majhno leteče prevozno sredstvo, iz letala pride majhen deček, ki v roki drži kovček.*; označeno je protivno razmerje med prvo povedjo in začetkom druge).



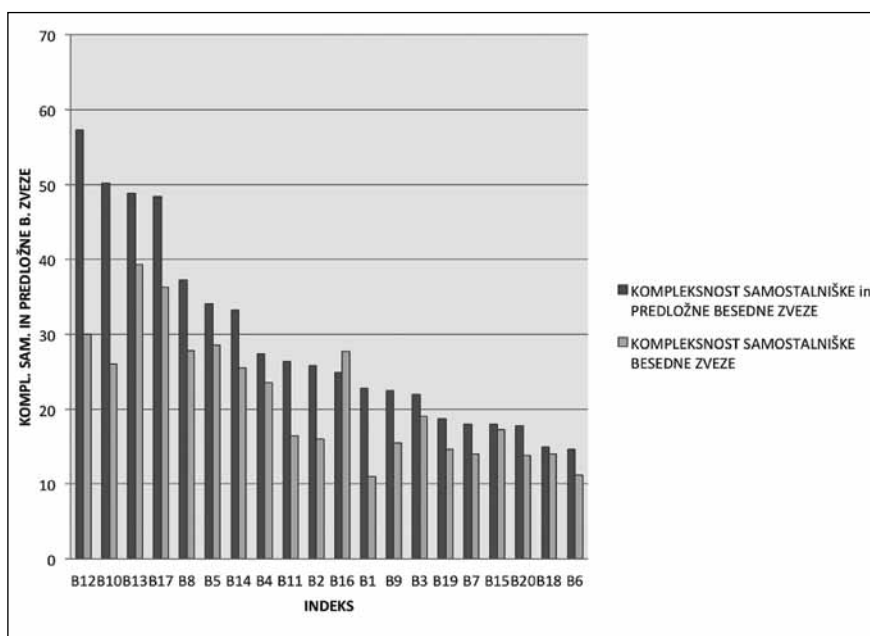
Graf 7: Huntov indeks podređnih struktur in indeks priredij Bardovi Harligove.

Prisotnost kompleksnejših struktur je možno preveriti tudi na besednozvezni ravni. Bulté (2012–13) je posamezna določila besedne zveze obtežil s točkami (določila, ki več prispevajo h kompleksnosti, so višje točkovana), nato pa število točk, pridobljenih z določili, delil s številom jeder. Navajam nekaj primerov iz analiziranih besedil s pripadajočim številom točk (jedra so odebeljena):

- Vsako določilo, ki je neposredno vezano na jedro samostalniške besedne zveze, je vredno **1 točko**: npr. *veliko, belo, prozorno svetlečo zvezdo* – 4 določila, neposredno vezana na jedro = 4×1 točka / 1 jedro = 4 točke.
- Samostalniška besedna zveza, vstavljena v predložno zvezo, je vredna **dodatne 0,5 točke**: npr. *ob punčkinem letalu* – 1 sam. besedna zveza, vstavljena v predložno zvezo = $1 \times 0,5$ + določilo, neposredno vezano na jedro sam. b. z. = 1×1 / 1 jedro = 1,5 točke.
- Priredno zloženi samostalniški besedni zvezi dobita **dodatno 1 točko**: npr. temno rdeče **krilo** in **nogavice** – 2 določila + 1 točka za priredno zloženost / 2 jedri = 1,5 točke.
- Vstavljena sam. besedna zveza v že obstoječo sam. besedno zvezo pomeni **dodatni 2 točki**: npr. moder **kovček**, z belimi **črtami** – 3 določila + 2 dodatni točki za vstavljeno besedno zvezo / 2 jedroma = 2,5 točk

- Vstavljene stavke pomenijo **3 dodatne točke**: npr. ročno izdelano **letalo**, ki noče vzleteti = 2 določili + 3 dodatne točke za vstavljen stavek / 1 jedro = 6 točk.

Obtežitev je teoretično smiselna – največ točk dobi stavčno določilo, najmanj enobesedno določilo – a v praksi je opazno, da ima na videz kompleksnejša besedna zveza *majhen deček v zelenih hlačah, svetlo zeleni majici in klobučkom na glavi* manj točk (4,5) kot na videz preprostejša besedna zveza: *ročno izdelano letalo, ki noče vzleteti* (6 točk). Hkrati dobi jezikovno ustrežnejša različica *punčkino letalo* manj točk (1) od slogovno zaznamovane različice *letalo od punčke* (2). Kljub temu da merimo kompleksnost in ne pravilnosti oziroma slogovne ustreznosti, ponujajo omenjeni primeri izhodišče za razmislek o tem, ali bi se dalo besedne zveze drugače obtežiti, da bi bile točke razporejene ustrežneje. Opozoriti pa je treba še na eno težavo, ki se pojavi zaradi neenotnosti definicij – iz Bultéjevih primerov je razvidno, da v meritev kompleksnosti samostalniške besedne zveze vključi še predložno besedno zvezo. Vendar pa neenotne definicije v tem primeru ne vplivajo na rezultat, saj je korelacija med uteženo kompleksnostjo vseh besednih zvez ter uteženo samostalniško besedno zvezo (brez desnih stavčnih prilastkov) zelo močna ($\rho = 0.8502$).

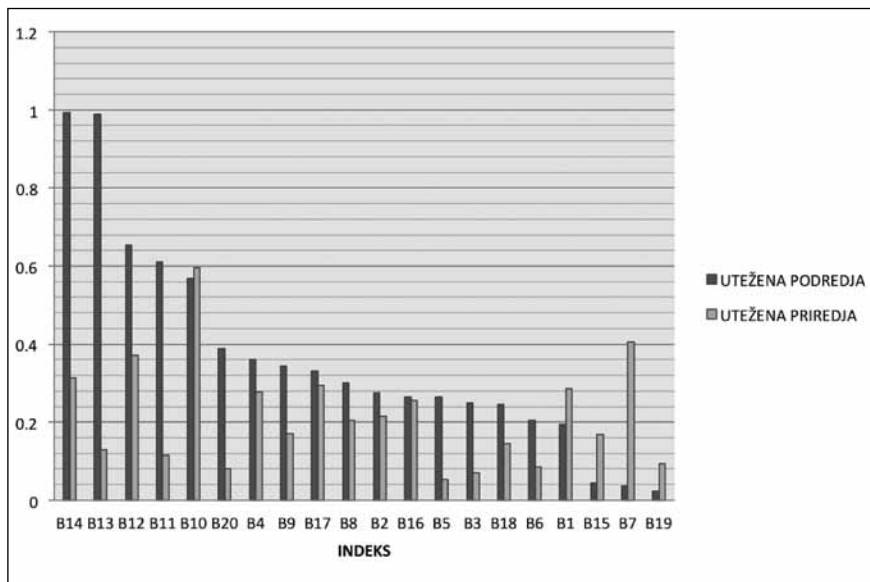


Graf 8: Kompleksnost besedne zveze in samostalniške besedne zveze.

4.3 Kompleksnost besedil z vidika deleža manj oziroma bolj pogostih struktur

Pogostnost struktur sem določila le na podlagi izbranega korpusu besedil osnovnošolcev. Na podlagi statistike (vsote) sem za celotno populacijo za vsak tip razmerja med stavki izračunala utež, ki je obratno sorazmerna številu pojavitev tega tipa razmerja v vseh besedilih skupaj. Z rabo uteži sem za vsako besedilo izračunala ponderirano (uteženo) vsoto pojavitev vseh tipov razmerij. Razmerja, ki se pojavljajo v manjšem deležu, imajo večji vpliv na končno vrednost – več redkejših struktur, kot ima besedilo, večja je vsota. Razmerje je manj kompleksno, če se v besedilu pojavi večkrat.

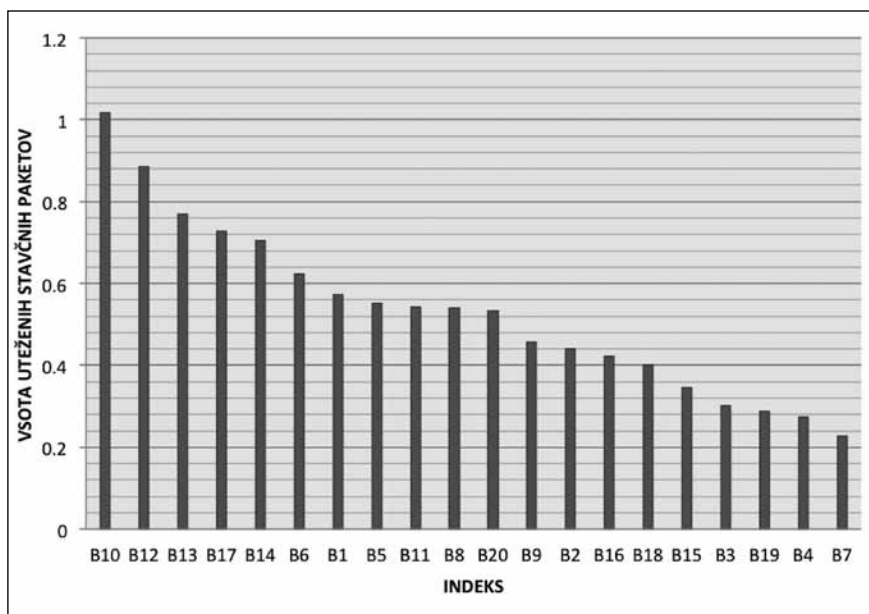
Razmerja (ki so hkrati deli stavčnih paketov) si glede na pogostnost sledijo v navedenem vrstnem redu (od najbolj do najmanj pogostih): stavek brez odkritih povezav > priredno razmerje z istim osebkom > priredno razmerje z različnim osebkom > soredno razmerje > glavni stavek > prislovni odvisni stavek > oziralni odvisni stavek > predmetni odvisnik > priredno razmerje z elipso glagola ali osebka > vloženi oziralni stavek > vloženi prislovni odvisnik. Pri tem naj opozorim, da sem vsak stavek označila zgolj glede na eno izmed povezav, ki jo tvori z okolico; oznake stavkov sem povzela po Berman, Nir 2010: 749–751.



Graf 9: Utežena vsota stavčnih paketov.

Utežena priredja in podredja sem merila na enak način kot stavčne pakete; če besedilo uporablja le najpogostejša priredja oziroma podredja (med priredji je to vezalno, med podredji pa prilastkov odvisnik), potem bo na grafu uteženih priredij/

podredij dosegalo nižje vrednosti in obratno; če uporabi najmanj pogosta, bo dosegalo najvišje vrednosti. Iz Grafa 10 je razvidno neujemanje med posameznimi besedili.



Graf 10: Utežena vsota podredij in priredij.

4.4 Raznolikost kot kazalec kompleksnosti v besedilih

Metod za opazovanje raznolikosti skladenjskih struktur med jezikoslovnimi raziskavami še nisem zasledila, zato sem izhajala iz aktivacije, opredeljene v Gries 2005: 365 in Szmrecsanyi 2005: 113, 114 ter oblikovala svojo metodo.

Raznolikost na povedni ravni sem opazovala tako, da sem v besedilih, ki so v prejšnjih meritvah na tej ravni izstopala, pregledala razporejenost različnih vrst povedi v celotnem besedilu (če so istovrstne povedi nakopičene v enem delu, je učinek aktivacije večji in kompleksnost manjša).

Besedila sem razvrstila v skupine glede na to, koliko zaporednih istovrstnih povedi vsebujejo. V I. skupini so besedila, ki imajo zaporedoma skupni le dve istovrstni povedi (npr. dve enostavni povedi na začetku besedila). Tabela 1 prikazuje omenjene skupine. Številki v razdelku Število zaporednih istovrstnih povedi prikazujeta, kolikokrat in koliko istovrstnih povedi se v celotnem besedilu pojavi ena za drugo, npr. 1 x 2 v I. skupini pomeni, da je v B4 in B9 le 1 primer, ko se 2 istovrstni povedi v besedilu nahajata ena za drugo.

OZNAKA SKUPINE	ŠTEVILO ZAPOREDNIH ISTOVRSTNIH POVEDI	INDEKS BESEDILA
I. skupina	1 x 2	B4 B9
II. skupina	1 x 3	B7 B16 B11 B6 B10 B13
III. skupina	2 x 3 3 x 3	B1 B12
IV. skupina	1 x 4	B15 B18
V. skupina	1 x 5	B2 B19 B5
VI. skupina	2 x 5	B17
VII. skupina	1 x 6	B8
VIII. skupina	1 x 7	B20 B14
IX. skupina	1 x 7, 1 x 8	B3

Tabela 1: Skupine, v katero so razvrščena besedila glede na zaporedno rabo istovrstnih povedi.

K manjši skladenjski kompleksnosti z vidika raznolikosti pripomorejo tudi paralelne strukture oziroma ponavljanja. Že ob površnem pregledu so v B2, B5, B6, B7 in B15 opazne aktivacije časovnih prislovov. Pretežno se v teh besedilih pojavlja en prevladujoči prislov, ki je občasno prekinjen s sopomenkami. Aktivacija je v največji meri prisotna v B15 – vsebuje 15 aktivacij časovnih prislovov: nato (11x), ki jih lahko označim za α -obstojnost ter 5 primerov β -obstojnosti: potem (3x), ter dvakrat in (v vlogi časovnega prislova). B5 uporabi kot leksikalno aktivacijo 6 časovnih prislovov: nato (4x), ki ga zamenja s *takrat* in *kasneje*. B7 namesto *nato* uporablja *potem* (4x), enkrat prekinjen z *nakar*. B6 je najbolj raznoliko izmed naštetih, saj sta le 2 prislova *nato*, ostali so raznoliki: *potem*, *zatem*, *nakar*.

5 Diskusija

Analiza rezultatov pokaže, da nekatera besedila izstopajo v smeri daljših (B10 in B12), druga pa v smeri krajših enot (B3 (materni jezik hrvaščina), B19, B15 in B18 (materni jezik albanščina)). Nekatere korelacije med meritvami je moč razložiti s strukturno podobnostjo enot, predvsem T-enot ter stavkov ($\rho = 0.8185$). Besedila z daljšimi povedmi imajo daljše stavčne pakete ($\rho = 0.6745$), saj stavčni paketi tako kot poved predstavljajo tematsko zaključeno enoto. Glede zaznamovanja meja med povedmi lahko potrdim že znano ugotovitev: meje med povedmi pri procesu tvorjenja

nastajajo sproti, na proces pa med drugim vpliva tudi pozornost (Chafe 1994: 144) – količina informacij, ki so jih posamezniki sposobni zadržati v delovnem spominu, se med učenci razlikuje. Če primerjam B18 (materni jezik makedonščina) in B3 (materni jezik hrvaščina), ugotovim, da ima B18 občutno daljše povedi kot B3, kar gre najverjetneje bolj kot maternemu jeziku pripisati kognitivnim sposobnostim.

Povprečna vrednost dolžine besedne zveze vseh besedil je 20.275, standardna deviacija pa $\sigma = 12.9097$, kar pomeni, da so si besedila glede na kompleksnost besedne zveze med seboj kar različna. B12 in B1 imata najdaljše besedne zveze in hkrati tudi največji delež podredij, kar je zanimivo, saj naj bi bolj izkušeni pisci podredne strukture pogosteje zamenjevali s kompleksnejšimi besednimi zvezami. Glede povprečne dolžine stavkov se besedila med seboj manj razlikujejo, a stavki so precej krajši od povprečja drugih raziskav:⁷ 6.79 (Hunt 1970 – materni jezik angleščina); 7.25 (Véliz 1999 – materni jezik španščina); 5.59 (Torres 1996 – materni jezik španščina); 5.56 (Bartolomé Rodríguez 2007 – tuji jezik angleščina), kar lahko pripišem tipološkim razlikam med naštetimi jeziki. Podobno je s Huntovim indeksom podrednih stavkov, ki v raziskavi istoimenskega avtorja znaša 1.4 (Hunt: 1965), v tej pa 1.1869.

Metoda uteženih stavčnih paketov, priredij in podredij je občutljiva na vzorec, zato lahko rezultate razlagam le znotraj okvirov te raziskave. Med vrstami odvisnikov je najpogostejši prilastkov, saj so se izbrana besedila bolj približala opisu dogajanja in naštevanju dogodkov kot pravi pripovedi. Besedila se v večji meri razvijajo bolj preko dodelave⁸ (Matthiessen 2001: 257) – opisovanje okoliščin, karakterizacija. Takoj za *ki* je naslednja najpogostejša vezniška beseda *da*, ki se pri usvajanju pojavi precej zgodaj (Kranjc 1999: 106). Če primerjam meritve pogostnosti z drugimi merili, opazim, da so besedila, ki so kompleksnejša v smislu dolžine skladenjskih struktur, kompleksnejša tudi v smislu večjega deleža manj pogostih struktur, a nimajo večjega deleža kompleksnejših struktur.

Graf 10 pokaže razliko v kvaliteti dolgih stavčnih paketov – B10 ima dolge stavčne pakete na račun prislovnih in oziralnih stavkov – med slednjimi uporablja tiste, ki so najmanj pogosti (vloženi: *majica, ki jo nosi, je bela* ter predmet znotraj oziralnega stavka: *od predmeta, ki ga je deklica držala*). Na drugi strani ima B9 dolge stavčne pakete na račun priredij.

Meritev raznolikosti se najmanj povezuje z ostalimi; besedila, ki so prej bila povprečna, tu izstopajo v smeri večje (B16, B6, B9) ali manjše (B14, B8) kompleksnosti. Meritev se slabo povezuje z dolžino struktur (B17 ima dolge enote, a kaže zelo malo raznolikosti), pa tudi z deležem bolj kompleksnih struktur (B4 ima večji delež manj kompleksnih enot, ki so hkrati bolj razpršene po celotnem besedilu).

⁷ Povprečne dolžine stavkov se v mojem korpusu bolj približajo Chafovim intonacijskim enotam, ki so v angleščini dolge med 5 in 6 enot (Chafe 1988: 10).

⁸ Uporabljam prevod *dodelava* za angl. *elaboration* po Snoj 2010.

Rahla korelacija med dolžino povedi ter deležem kompleksno-sestavljjenih povedi ($\rho = 0.5958$) kaže na to, da povedi pridobivajo na dolžini večinoma na račun priredij. Korelacija med povprečno dolžino T-enote in deležem podrednih stavkov je zelo visoka ($\rho = 0.8072$) – besedila z večjim deležem podredij imajo hkrati tudi daljše enote, ki podredja vsebujejo. Korelacija med kompleksnostjo in dolžino besedne zveze je razmeroma visoka ($\rho = 0.6180$); izbrana mera za kompleksnost torej v veliki meri upošteva tudi dolžino besedne zveze. Podobno je z razmerjem med dolžino stavčnih paketov in deležem manj pogostih stavčnih paketov ($\rho = 0.6179$).

6 Sklep

V članku na podlagi različnih pojmovanj kompleksnosti oblikujem štiri merila, na podlagi katerih z že znanimi metodami izmerim skladenjsko kompleksnost. Precejšnja neskladja med posameznimi meritvami kažejo na to, da je kompleksnost večdimenzionalen pojav, zato je potrebno pri vsakokratnem merjenju natančno opredeliti metode ter navesti strukture, ki so pojmovane kot bolj ali manj kompleksne. Vsak tvorec namreč kompleksnost izrazi na svojstven način. Večja je povezanost znotraj posameznih meritev (npr. ista besedila uporabljajo daljše povedi, daljše stavke, daljše stavčne pakete ...). Manj kompleksna besedila so v večini krajša in uporabljajo manjši delež kompleksnih struktur. Za merjenje raznolikosti sem predlagala metodo opazovanja razporeditve skladenjskih struktur po besedilu in ugotovila, da je tovrstno merjenje raznolikosti najmanj povezano s preostalimi meritvami.

V pričujočem sestavku sem že opozorila na nekatere težave in pomanjkljivosti pri uporabi oziroma prevzemanju metod za merjenje skladenjske kompleksnosti, ki so jih razvili na drugih jezikih (ne na slovenščini). Pri nadaljnjem raziskovanju skladenjske kompleksnosti na slovenskih pisnih besedilih nameravam kritičnemu ovrednotenju teh metod nameniti posebno pozornost.

Raziskave na temo kompleksnosti, ki kompleksnost pojmujejo kot dimenzijo aktivnega znanja jezika in performance ter jo merijo na besedilih maternih govorcev, so redke. Članek je le poskus prenosa tovrstnih raziskav v slovenski prostor in preverjanja njihove ustreznosti na slovenskem gradivu ter ponuja izhodišče za nadaljnje delo. Raziskava se namreč omejuje na kompleksnost v najožjem smislu ter zanemari ostali dve dimenziji (pravilnost in tekočnost), poleg tega z osredotočanjem na število elementov in odnose med njimi zanemari nekatere pomembne vidike kompleksnosti, predvsem smiselno spetost besedila (Ferbežar 2012). Omejena je le na določeno besedilno zvrst, saj bi pri merjenju kompleksnosti govorjenega jezika zaradi drugačne skladenjske podobe potrebovali drugačna merila.

Literatura

- Bardovi-Harlig, Kathleen, 1992: A Second Look at T-Unit Analysis: Reconsidering the Sentence. *TESOL Quarterly* 26/2. 390–395.
- Bartolomé Rodríguez, Rocío, 2008: La adquisición y el aprendizaje de los modificadores nominales en el discurso escrito de estudiantes de español como L1 y L2. *Revista Electrónica de Lingüística Aplicada* 7. 45–54.
- Berman, Ruth A., in Nir, Bracha, 2010: Complex syntax as a window on contrastive rhetoric. *Journal of Pragmatics* 42. 744–765.
- Bock, Kathryn, 1986: Syntactic Persistence in language production. *Cognitive Psychology* 18. 355–387.
- Bulté, Bram, 2012–2013: *The development of complexity in second language acquisition. A dynamic systems approach*. Doktorska disertacija. Bruselj: Vrije Universiteit Brussel.
- Chafe, Wallace, 1988: Punctuation and the Prosody of Written Language. *Written Communication* 5/4. 396–426.
- Chafe, Wallace, 1994: *Discourse, Consciousness, and Time*. Chicago. University of Chicago Press.
- Crowhurst, Marion, 1983: Syntactic Complexity and Writing Quality: A Review. *Canadian Journal of Education / Revue canadienne de l'éducation* 8/1. 1–16.
- Doleschal, Ursula, in Robatsch, Gerald, 2015: Using Quantitative Linguistics to Assess Pupils' Language Proficiency in a Bilingual Context: The Case of Slovene in Carinthia. Smolej, Mojca (ur.): *Slovnica in slovar – aktualni jezikovni opis 1. del*. Obdobja 34. 173–183.
- Ferbežar, Ina, 1996: Jezikovno testiranje na tečajih slovenščine kot drugega/tujega jezika: (teoretični okvir: prva spoznanja). *Jezik in slovstvo* 42/7. 279–292.
- Ferbežar, Ina, 2012. *Razumevanje in razumljivost besedil*. Znanstvena založba Filozofske fakultete.
- Gries, Stefan, 2005: Syntactic Priming: A Corpus-based Approach. *Journal of Psycholinguistic Research* 34/4. 365–399.
- Housen, Alex, in Kuiken, Folkert, 2009: Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics* 30/4. 461–473.
- Housen, Alex, Kuiken, Folkert, in Vedder, Ineke (ur.), 2012: *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hunt, Kellogg, W., 1965: *Grammatical Structures Written at Three Grade Levels*. NCTE Research Report No. 3. Illinois: National Council of Teachers of English.
- Hunt, Kellogg, W., 1966: Recent measures in syntactic development. *Elementary English* 43/7. 732–739.
- Jarvis, Scott, Grant, Leslie, Bikowski, Dawn, in Ferris, Dana, 2003: Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing* 12/ 4. 377–403.
- Kranjc, Simona, 1999: Skladnja in otroški govor. *Jezik in slovstvo* 44/6. 197–212.
- Lambert, Kormos, 2014: Complexity, Accuracy, and Fluency in Task-based L2 Research: Toward More Developmentally Based Measures of Second Language Acquisition. *Applied Linguistics* 2014: 1–9.

- Malvern, David, Richards, Brian, Chipere, Ngoni, in Durán, Pilar, 2004: *Lexical Diversity and Language Development. Quantification and Assessment*. New York: Palgrave Macmillan.
- Matthiessen, Christian M. I. M., 2001: Combining clauses into clause complexes. A multifaceted view. Bybee, Joan L.; Noonan, Michael (ur.): *Complex Sentences in Grammar and Discourse: Essays in honor of Sandra A. Thompson*. 235–319.
- Mitchell, Melanie, 2009. *Complexity. A guided tour*. New York: Oxford University Press.
- Ortega, Lourdes, 2003: Syntactic Complexity Measures and their Relationship to L2 Proficiency: A Researchy Synthesis of College-level L2 Writing. *Applied Linguistics* 24/2. 492–518.
- Pickering, Martin, in Branigan, Holly, 1999: Syntactic priming in language production. *Trends in Cognitive Science* 3. 136–141.
- Robinson, Peter, in White, Ron, 1995: Current Approaches to Syllabus Design: A Discussion with Ron White. *Guidelines* 17/1. 93–101.
- Szmrecsanyi, Benedikt, 2005: Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1/1. 113–150.
- Snoj, Jerica, 2010: *Metafora v leksikalnem sistemu*. Ljubljana: Založba ZRC.
- Torres, Antonia Nelsi, 1996: *Complejidad sintáctica en estudiantes de Tenerife*. Santa Cruz de Tenerife: Centro de la Cultura Popular Canaria.
- Véliz, Mónica, 1999: Complejidad sintáctica y modo del discurso. *Estudios Filológicos* 34: 181–192.
- Vercellotti, Mary Lou, 2012: *Complexity, Accuracy, and Fluency as Properties of Language Performance: The Development of the Multiple Subsystems over Time and in Relation to Each Other*. Doktorska disertacija. Pittsburgh: University of Pittsburgh.
- Verdonik, Darinka, 2015: Jezikovnoteoretska načela v korpusnem jezikoslovju. *Slovenščina 2.0* 3/1. 1–27.
- Verspoor, Marjolijn, Schmid Monika S., in Xu, Xiaoyan, 2012: A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing* 21. 239–263.
- Verspoor, Marjolijn, Lowie, Wander, Ping Chan, Hui, in Vahtrick, Louisa, 2017: Linguistic complexity in second language development: variability and variation at advanced stages. *Recherches en didactique des langues et des cultures* 14/1. 1–27.
- Vyatkina, Nina, 2012: The Development of Second Language Writing Complexity in Groups and Individuals: A Longitudinal Learner Corpus Study. *The Modern Language Journal* 96/4. 576–598.