

# RJEČNIČKA RAZNOLIKOST PISANIH TEKSTOVA OSOBA S RAZVOJNIM JEZIČNIM POREMEĆAJEM

GORDANA HRŽICA, SARA KOŠUTAR, MATEJA KRAMARIĆ

Odsjek za logopediju, Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu, kontakt: gordana.hrzica@erf.unizg.hr

Primljeno: 05.11.2019.  
Prihvaćeno: 19.12.2019.

Izvorni znanstveni rad  
UDK 81`232+81`24  
<https://doi.org/10.31299/hrri.55.2.2>

**Sažetak:** Razvojni jezični poremećaj (RJP) jedna je od najučestalijih teškoća koja se javlja u jezičnom usvajanju tijekom predškolske i školske dobi, a posljedično se zadržava i u kasnijoj životnoj dobi. Odstupanja u jezičnim sposobnostima osoba s RJP-om mogu se pratiti na svim jezičnim razinama, uključujući i rječnik. Cilj ovog istraživanja bio je ispitati razlike između osoba s RJP-om i osoba urednoga jezičnog razvoja (UJR) u rječničkoj raznolikosti. Za razliku od ranijih istraživanja koja su ispitivala rječničku raznolikost u govorenome diskursu mladih osoba, ovim se istraživanjem željelo na uzorku osoba širega raspona kronološke dobi provjeriti je li ova razlika prisutna i u tekstovima koji nastaju kao rezultat pisanoga diskursa. Za istraživanje su iz Hrvatskoga korpusa neprofesionalnoga pisanog jezika (Kuvač Kraljević, Hržica i Kologranić Belić, u tisku) izdvojeni jezični uzorci govornika s RJP-om ( $n=20$ ) i govornika UJR-a ( $n=19$ ) uparenih prema spolu i kronološkoj dobi. Jezični su uzorci elicitirani pomoću pripovjednoga slikovnog predloška Priča o plaži koji čini dio testnoga materijala Expression, Reception and Recall of Narrative Instrument (ERRNI; Bishop, 2004). Za svakoga su sudionika izračunate četiri mjere rječničke raznolikosti: broj različitih riječi (NDW) i omjer različenica i pojavnica (TTR), pomični prosječni omjer različenica i pojavnica (MATTR) i indeks leksičke raznolikosti D. Podatci su obrađeni t-testom za nezavisne uzorke. Dobiveni rezultati ukazuju na statistički značajno nižu razinu rječničke raznolikosti osoba s RJP-om od osoba UJR-a na svim mjerama. Na temelju dobivenih rezultata može se zaključiti kako testirane mjere rječničke raznolikosti mogu razlikovati osobe različitoga jezičnog statusa u zahtjevnim zadacima pisanoga pripovjednoga diskursa. Dobiveni se rezultati razlikuju od ranijih istraživanja u kojima se tvrdi da su razlike u rječničkoj raznolikosti posredni rezultat sintaktičke složenosti (pregled: Charest i Skoczylaz, 2019), a ova se razlika objašnjava utjecajem žanra, primjenom preciznijih mjera te dobi sudionika.

**Ključne riječi:** razvojni jezični poremećaj, pisani jezični uzorci, rječnička raznolikost

## UVOD

Razvojni jezični poremećaj (u nastavku teksta RJP) jedna je od najučestalijih teškoća koja se javlja u jezičnom usvajanju tijekom predškolske i školske dobi, a posljedično se zadržava i u kasnijoj životnoj dobi. Za RJP je karakteristično da su jezične sposobnosti djeteta slabije u odnosu na njegove verbalne sposobnosti, dok istovremeno ne postoje druge razvojne teškoće, intelektualni ili senzorni nedostaci oštećenja mozga ili drugi medicinski faktori za koje je poznato da dovode do jezičnih teškoća (Reilly i sur., 2014; Leonard, 2014; Bishop, 2006). Prema najnovijim podacima procjenjuje se da 7.5% djece predškolske dobi ima RJP, u školskoj dobi RJP pogađa prosječno dvoje od tridesetero djece te u prosjeku 7% opće populacije ima RJP (Norbury i sur., 2016). Unatoč tomu

što RJP zahvaća velik dio populacije različite dobi, njegovo je sustavno proučavanje relativno novo, zbog čega i danas izostaju odgovori na mnoga pitanja o obilježjima RJP-a. Odstupanja se u jezičnim sposobnostima osoba s RJP-om mogu pratiti na svim jezičnim razinama, od kojih se rječnik (leksik) može smatrati temeljem koji omogućuje uporabu ostalih jezičnih sastavnica. Na to upućuju i uočena odstupanja ove populacije u proizvodnji, razumijevanju i učenju riječi, od samih početaka jezičnog usvajanja te kroz kasniji razvoj (pregled: Nation, 2014). Iz toga je razloga važno pronaći prikladne načine kako vrednovati jezična postignuća na rječničkoj razini u osoba urednoga jezičnoga razvoja (UJR u nastavku teksta) te ih usporediti s postignućima osoba s RJP-om. Osim primjenom standardiziranih psihometrijskih testova rječničke se sposobnosti mogu procijeniti uporabom razli-

čitih jezičnih mjera za analizu jezičnih uzoraka govorenoga ili pisanoga diskursa<sup>1</sup>. Jedna od skupina mjera uključuje procjenu rječničke raznolikosti, koja se pokazala snažnim prediktorom kvalitete pisanoga teksta (pregled: Yu, 2009). No mnoga pitanja glede rječničke raznolikosti osoba s RJP-om ostaju i dalje otvorena.

### Mjere rječničke raznolikosti

Rječnik je temeljna jezična sastavnica neophodna za jezičnu proizvodnju, kako govorenu tako i pisanu. Tijekom formalnog obrazovanja, s razvojem pismenosti, opseg djetetova rječnika se proširuje, a postojeći se usložnjava. Sofisticiranije promjene u rječniku, iako kvantitativno sporije, nastavljaju se i nakon školske dobi, u skladu s daljnjim obrazovanjem i akademskim vještinama (Nippold, 2016). Stoga se u jezičnoj proizvodnji razvoj rječnika može pratiti čitavoga života, odnosno dok god čovjek govori i piše. Rječničko bogatstvo (engl. *lexical richness*) kao pokazatelj rječničkoga znanja uključuje nekoliko aspekata koji se mogu mjeriti, a to su rječnička raznolikost (engl. *lexical diversity*), rječnička gustoća (engl. *lexical density*), rječnička sofisticiranost (engl. *lexical sophistication*) i broj pogrešaka (engl. *number of errors*) (Read, 2000). Rječnička raznolikost kao parametar procjene kvalitete teksta odnosi se na broj različitih riječi upotrijebljenih u tekstu, pisanomu ili govorenomu (Johansson, 2008). U dosadašnjim se istraživanjima rječnička raznolikost pokazala snažnim prediktorom općih jezičnih sposobnosti te se smatra važnim pokazateljem kvalitete pisanja i govorenja (pregled: Yu, 2009).

Jedna od najranijih i najjednostavnijih mjera rječničke raznolikosti uključivala je *broj različitih riječi* (engl. *number of different words* – NDW; Miller, 1991; Klee, 1992). Prijašnja su istraživanja pokazala da NDW snažno korelira s dobi te razlikuje djecu UJR-a od djece s RJP-om (Klee, 1992). Osim toga, rječnička raznolikost često se mjerila omjerom različenica i pojava (engl. *type-token ratio* – TTR; Templin, 1957 – u hrvatskome

je jeziku ustaljena i kratica ORP (vidi Palmović i Kuvač Kraljević, 2007), ali će se radi lakšeg praćenja u nastavku teksta koristiti engleskom kraticom). Pojmom *različnica* označuje se svaki glasovno različiti oblik riječi, a *pojavnica* svako ponavljanje jedne riječi u tekstu (Jelaska, 2005). Veći broj riječi te viši omjer različenica i pojava pokazatelji su veće rječničke raznolikosti, dok manji broj riječi te niži omjer različenica i pojava upućuju na manju rječničku raznolikost. Međutim, obje mjere, i NDW i TTR, pokazale su metodološka ograničenja zato što zavise o veličini uzorka. Utvrđeno je da jezični uzorci s većim brojem riječi imaju manji TTR. Razlog tomu je što se broj pojava u tekstu može povećavati, a kako se nove leksičke riječi proizvode, automatski se povećava broj gramatičkih riječi (Johansson, 2008). Stoga će manji jezični uzorci dati višu vrijednost TTR-a, dok će niže vrijednosti biti rezultatom većih uzoraka (Richards, 1987). Ograničenja vezana uz veličinu uzorka nastojala su se premostiti standardiziranjem jezičnih uzoraka na način da bi se, primjerice NDW, izračunavao na prvih 50 ili 100 iskaza. No NDW ne samo da je pod izravnim utjecajem veličine uzorka već je u snažnoj korelaciji i s mjerom *prosječne duljine iskaza* (engl. *mean length of utterance* – MLU; Brown, 1983; Miller, 1991). Poznato je da se MLU s porastom dobi povećava (vidi Wells, 1985), pa se zbog korelacije MLU-a i NDW-a ne mogu usporediti djeca različite kronološke dobi ili djeca različitoga jezičnog statusa.

Budući da su jezične mjere NDW i TTR ovisne o veličini uzorka, osmišljene su druge mjere koje su se pokazale mnogo robusnijima. Tako se rječnička raznolikost može mjeriti i *indeksom rječničke raznolikosti D* (engl. *vocabulary diversity*; Richards i Malvern, 1997; McKee, Malvern i Richards, 2000; Malvern, Richards, Chipere i Durán, 2004), za koji je utvrđeno da nije ovisan o veličini jezičnog uzorka. Mjera D izračunava se kao matematička krivulja koja se temelji na predviđenome smanjivanju TTR-a pri porastu veličine jezičnog uzorka. Viša vrijednost parametra D pokazatelj je veće rječničke raznolikosti. Za izračunava-

1 Za potrebe ovoga rada *diskurs* se shvaća kao tekst u kontekstu, odnosno diskurs uključuje komunikacijski okvir unutar kojega pošiljalac (govornik ili pisac) ostvaruje komunikacijsku namjeru odašiljući poruku (tekst) primatelju (slušatelju ili čitatelju). Diskurs je stoga proces (*discourse-as-process*), dok je tekst konačni proizvod diskursa (*text-as-product*) (Brown i Yule, 1983; usp. Stillar, 1998). Različite tekstne vrste ostvaruju se na planu jezične realizacije kao pisane ili govorene i, tako koncipirane, situiraju se u određeni diskursni model, primjerice narativni diskurs (vidi više Badurina i Kovačević, 2001; usp. Ivanetić, 2003).

nje parametra D upotrebljavaju se podatci iz čitavoga jezičnog uzorka, a prijašnjim istraživanjima (primjerice, McCharty i Javis, 2007) pokazalo se da mjera D najbolje funkcionira na tekstovima u kojima se broj pojavnica kreće između 100 i 400. Iako je mjera D naišla na kritike zbog tvrdnje o neovisnosti o veličini uzorka (primjerice, Daller, van Hout i Treffers-Daller, 2003; McCharty i Javis, 2007), pružila je pouzdane rezultate o rječničkome razvoju populacije UJR-a (primjerice, Malvern i sur., 2004; Yu, 2009; ), kao i populacije s RJP-om (primjerice, Owen i Leonard, 2002; Klee, Stokes, Wong, Fletcher i Gavin, 2004). Još snažnijim indikatorom rječničke raznolikosti pokazala se mjera *pomični prosječni omjer različnica i pojavnica* (MATTR; engl. *moving average type-token ratio*; Köhler i Galle, 1993; Covington i McFall, 2010; Fergadiotis, Wright i Green, 2015). MATTR izračunava rječničku raznolikost na način da se tekst razlomi na okvire željene veličine (engl. *window length*), primjerice odredi se okvir od 500 riječi. MATTR pruža vrijednosti za svaki okvir teksta na način da izračunava TTR za raspon 1-500 riječi, zatim raspon 2-501 riječi, 3-502 riječi i tako sve do kraja teksta. Srednja vrijednost svih TTR-a u tekstu mjera je rječničke raznolikosti cjelokupnoga teksta. Prednost je ove mjere što nije ovisna o veličini uzorka te omogućava usporedbu TTR-a pojedinačnih dijelova teksta.

### **Rječnička raznolikost s obzirom na razlike u modalitetu**

Iako se govoreni i pisani modalitet najčešće opisuju na temelju razlika koje proizlaze iz njihovih formalnojezičnih obilježja (vidi Brown i Yule, 1983; Miller i Weinert, 1998), mjere sintaktičke složenosti i rječničke raznolikosti pokazale su da granice između modaliteta nisu uvijek tako fluidne. Načelno se pisani jezik pokazao kao izrazito strukturiran i sintaktički složeniji, za razliku od govorenoga jezika koji je manje strukturiran i ima fragmentiranu i jednostavniju sintaksu (Baron, 2000; Paltridge 2006; Zhang 2013). Ipak, pojedina su istraživanja ukazala na to da upravo govoreni jezik sadrži više subordiniranih rečenica i složenijih sintaktičkih struktura (Poole i Field 1976; Halliday 1989). Usporedbom pak modaliteta na rječničkoj razini uočeno je da pisani jezik sadrži

više vrsta riječi (Ure, 1979; Allwood, 1998), od kojih su češće punoznačne riječi (Henrichsen i Allwood, 2005; Olujić i Matić, 2017), dok u govorenome jeziku pretežu nepunoznačne riječi (Biber, 1988; Paltridge, 2006; usp. Olujić i Matić, 2017). Zatim uporabom mjera TTR i D utvrđeno je da se modaliteti razlikuju u rječničkoj raznolikosti, i to na način da je pisanomu jeziku svojstvena veća rječnička raznolikost za razliku od govorenoga jezika (Biber, 1995; Strömquist i sur., 2002; Johansson, 2008; Olujić i Matić, 2017). Štoviše, razlike, u smislu veće rječničke raznolikosti u pisanome jeziku, ostaju perzistentne i tijekom jezičnog razvoja (Strömquist i sur., 2002; Johansson, 2008). S druge strane, provedena su istraživanja u kojima je na mjeri D dobivena vrlo slična rječnička raznolikost s obzirom na modalitet (Yu, 2009), dok u drugim istraživanjima razlike u rječničkoj raznolikosti na mjeri TTR nisu pronađene (Berman i Verhoeven, 2002). Ipak, kontradiktorni se rezultati prethodnih istraživanja ne mogu pripisati isključivo razlikama u modalitetu, već u obzir treba uzeti druge čimbenike koji utječu na izbor rječnika, primjerice razinu formalnosti i žanr u kojem se tekst ostvaruje (vidi Biber, 1995; Baron, 2000; Johansson, 2008). Biber (1988) ističe da konačne razlike između pisanoga i govorenoga jezika ne postoje, već se eventualne varijacije mogu pripisati različitim tekstovima i žanrovima kao višedimenzionalnim konstruktima. Osim toga, morfološke specifičnosti jezika mogu značajno utjecati na mjerenje rječničke raznolikosti (Stromqvist i sur., 2002). Pretpostavlja se da će jezici s manje složenom morfologijom imati veći udio punoznačnih riječi (Ure, 1971). Rječnička raznolikost također zavisi o poznatosti teme o kojoj sudionik piše ili govori, primjerice viša je rječnička raznolikost dobivena onda kada je sudionik upoznat s temom (vidi Yu, 2009). Također, na jezičnu proizvodnju mogu utjecati vrijeme koje sudionik ima za planiranje iskaza, vrsta publike i sl. (Redeker, 1984). Dulje vrijeme planiranja općenito rezultira većom integriranošću ideja u jedinstvenu jezičnu cjelinu (Chafe, 1984). Svi su spomenuti parametri važni žele li se jezični uzorci različitih modaliteta usporediti precizno, a u obzir treba uzeti i činjenicu da pri usporedbi modaliteta podatci budu prikupljeni od istih sudionika i elicirani jednakim zadacima.

## RJEČNIK OSOBA S RAZVOJNIM JEZIČNIM POREMEĆAJEM

Najčešći je prvi pokazatelj RJP-a kasna pojava prve riječi, a nakon toga i kašnjenje u povezivanju riječi u rečenice. Ipak, za razliku od zakašnjeloga jezičnog razvoja, za koji je poznato da se odvija sporijim tempom, ali većina djece s vremenom dostiže svoje vršnjake, RJP je perzistentan poremećaj, pa se odstupanja u jezičnim sposobnostima zadržavaju i u odrasloj dobi (Leonard, 2014). Radi se o vrlo heterogenome poremećaju koji se klinički manifestira na različite načine. Simptomi koji obično prate RJP očituju se u otežanom usvajanju novih riječi, kašnjenju s oblikovanjem dvočlanih i višočlanih iskaza, usporenome razvoju fonološkoga sustava, kašnjenju u usvajanju gramatike, pogreškama u glagolskoj i imenskoj morfologiji te otežanom usvajanju prostornih odnosa (Kologranić Belić, Matić, Olujić i Srebačić, 2015). Djeca s RJP-om u prosjeku znaju manje riječi nego njihovi vršnjaci UJR-a, pri čemu se taj nedostatak zadržava i u kasnijoj dobi (Rice i Hoffman, 2015).

U opisima jezičnih sposobnosti osoba s RJP-om češće se spominju teškoće s morfologijom ili sintaksom (pregled dosadašnjih definicija: Leonard, 2014), dok se nešto rjeđe piše o njihovu rječniku. No upravo je zakašnjeli i spori rječnički razvoj jedan od znakova RJP-a (McGregor, Newman, Reilly i Capone, 2002; Nash i Donaldson, 2005; Thal, Reilly, Seiberd, Jeffries i Fenson, 2004). Brojna su istraživanja različitih metodologija ukazala na njihove teškoće u rječničkome razvoju (pregled: McGregor, 2017; Schwartz, Botwinik-Rotem i Friedman, 2017; Leonard, 2014). Djeca s RJP-om progovaraju kasnije od djece UJR-a (Clarke i Leonard, 1996) i razlikuju se od njih u razumijevanju i učenju riječi (pregled: Nation, 2014), kako kroz predškolsko razdoblje (Thal, O'Hanlon, Clemmons i Fralin, 1999; Watkins, Kelly, Harbers i Hollis, 1995), tako i kroz školsko razdoblje kada se razlike u jezičnim sposobnostima samo produbljuju (Haynes, 1992; Stothard, Snowling, Bishop, Chipchase i Kaplan, 1998).

S obzirom na to da rezultati eksperimentalnih istraživanja pokazuju da djeca s RJP-om imaju teškoće u usvajanju i obradi riječi, moguće je pretpostaviti da će se one odraziti i na njihovu jezičnu

proizvodnju, posebice na samostalnu ili poticanu proizvodnju većih jezičnih cjelina, to jest diskursa. U analizi jezičnih uzoraka djece s RJP-om najčešće su proučavane mjere rječničke raznolikosti kao odraz rječničkoga znanja u uporabi. Unatoč brojnim dokazima da djeca s RJP-om imaju teškoća u razvoju i uporabi rječnika, nije se uvijek pokazalo da ih mjere rječničke raznolikosti razlikuju od vršnjaka UJR-a. Već su rana istraživanja pokazala da različite metode procjene mogu dovesti do različitih rezultata. Tako Watkins i sur. (1995) rabe mjere TTR i NDW kako bi usporedili djecu s RJP-om te djecu UJR-a. Dok se NDW pokazao kao dobra i robusna mjera, autori upozoravaju na neopravdanost uporabe mjere TTR. Mnoga su istraživanja potvrdila diskriminativnu vrijednost mjere NDW (primjerice, Klee, 1992; Hewit, Hammer, Yont i Tomblin, 2005), ali većina ih je provedena na jezičnim uzorcima ujednačenima prema broju iskaza. S obzirom na to da se djeca s RJP-om redovito razlikuju od djece UJR-a prema prosječnoj duljini iskaza (vidi Klee i sur., 2004), ujednačavanje prema broju iskaza ne mora nužno značiti da je problem razlike u duljini jezičnog uzorka razriješen. Tako usporedbe uzoraka ujednačenih prema broju riječi često ne pokazuju razlike između djece s RJP-om i djece UJR-a (Owen i Leonard, 2002; Thordardottir i Ellis Weismer, 2001; za suprotne rezultate vidi Watkins i sur., 1995).

Prije no što se utvrdi kako istraživanja idu u smjeru potvrđivanja nepostojanja razlika između djece UJR-a i djece s RJP-om u rječničkoj raznolikosti, potrebno je u obzir uzeti dvije činjenice. Prva je da su i sami istraživači neskloni svoje rezultate interpretirati na takav način. Primjerice, Charest i Skoczylas (2019) upozoravaju da je, s obzirom na sve što nam je poznato, teško ustvrditi da se ove dvije skupine ne razlikuju u uporabi riječi u jezičnim uzorcima. No autori upozoravaju da bi možda trebalo prilagoditi način mjerenja ili općenito pristup mjerenju. Druga se činjenica tiče problema ujednačavanja jezičnih uzoraka prema duljini. U ujednačavanju duljine potrebno je skraćivanje svih jezičnih uzoraka prema unaprijed određenom broju riječi ili prema najkraćem uzorku (Thordardottir i Ellis Weismer, 2001). Uporaba određenoga broja riječi znači da će se neki od jezičnih uzoraka skratiti malo, a neki drugi znatno više, i to obično tako



da se odabere dio teksta s početka ili sredine uzorka. Dakle, nekim će se sudionicima, primjerice, procijeniti dio nekog opisa ili priče, a nekim drugima čitava priča. S obzirom na to da se rječnička raznolikost može mijenjati u različitim dijelovima teksta, ovisno o njihovoj svrsi i retoričkim sredstvima potrebnima da se oblikuje (Jarvis, 2002), odabirom samo jednoga dijela teksta koji po svojoj ulozi u tekstu ima nižu rječničku raznolikost ponovno se može narušiti objektivnost usporedbe. Također, u dužim jezičnim uzorcima neminovno je da se sudionik dulje vrijeme zadrži na jednoj temi. Tako opis lika ili događaja traži ponavljanje određenih riječi, primjerice naziva ili imena lika. U kraćim se jezičnim uzorcima teme ili događaji brzo izmjenjuju, a time i riječi kojima se imenuju likovi ili opisuju radnje.

Prethodno spomenuti čimbenici dovode do potrebe primjena novijih mjera koje nude drugačiji način izračuna rječničke raznolikosti te se mogu pokazati prikladnijima za usporedbu različitih skupina sudionika. Takve mjere ne traže kao preduvjet ujednačavanje jezičnih uzoraka, to jest odbacivanje dijela podataka. Upravo su mjera MATTR i mjera D dobri primjeri toga. U validacijskim istraživanjima ove mjere pokazuju slabu ovisnost o duljini uzorka (primjerice, McCarthy i Jarvis, 2010; Fergadiotis, Wright i Greenc, 2015). Uz to se pri izračunu mjere D, riječi iz uzorka odabiru nasumično te se tako izbjegava utjecaj zadržavanja na istoj temi na mjerenje. Za razliku od mjere D, u izračunu mjere MATTR uzima se određen niz riječi. No malobrojna dosadašnja istraživanja djece s RJP-om nisu pokazala jednoznačne rezultate u uporabi ovih mjera. Klee i sur. (2004) izvještavaju o značajnim razlikama u rezultatima na mjeri D između djece s RJP-om i djece UJR-a, što je potvrđeno i u narednom istraživanju iste skupine autora (vidi Wong, Klee, Stokes, Fletcher i Leonard, 2010). Slične su rezultate dobili Owen i Leonard (2002) kada su uspoređivali 100 iskaza djece s RJP-om i djece UJR-a. Međutim, ujednačavanje prema broju pojavnica pokazuje sličan obrazac rezultata, ali bez statistički značajnih razlika. Autori zaključuju da na mjeru D duljina transkripta ipak ima utjecaja, iako je on znatno manji od utjecaja na tradicionalne mjere. Na temelju istraživanja druge kliničke skupine, to jest odraslih osoba s afazijom, slično

zaključuju i Wright, Silverman i Newhoff (2003). Ova se mjera pokazala značajnom i u razlikovanju dvojezične djece UJR-a od dvojezične djece s RJP-om, i to u oba njihova jezika (vidi Jacobson i Walden, 2013). S druge strane, mjera MATTR rijetko se upotrebljavala u istraživanjima djece s RJP-om. Charest i Skoczylaz (2019) ne nalaze razlike između djece UJR-a i djece s RJP-om na toj mjeri.

### **Pisani jezik osoba s razvojnim jezičnim poremećajem**

Prijašnja istraživanja jezičnih sposobnosti govornika s RJP-om većinski su, ako govorimo o jezičnoj proizvodnji, usmjerena na govoreni jezik (Williams, Larkin i Blaggan, 2013). U manjem se broju istraživanja proučavaju karakteristike pisanoga jezika osoba s RJP-om (primjerice, Dockrell, Lindsay i Connelly, 2009; Suddarth, Plante i Vance, 2012; Williams, Larkin i Blaggan, 2013) ili se pisani jezik uspoređuje s govorenim (primjerice, Gillam i Johnston, 1992; Scott i Windsor, 2000). Rezultati tih istraživanja pokazuju da i u pisanome jeziku osobe s RJP-om zaostaju za vršnjacima UJR-a na nizu mjera i pokazatelja jezičnih sposobnosti (pregled: Williams i sur., 2013). Posebno su zanimljivi rezultati istraživanja u kojima se uspoređuju govoreni i pisani jezični uzorci. Scott i Windsor (2000) pokazuju da je modalitet ispitivanja imao značajan utjecaj na produktivnost, broj pogrešaka te sintaktičku složenost djece s RJP-om. U usporedbi s urednim govornicima izjednačenima po kronološkoj dobi, djeca s RJP-om u pisanju proizvode kraće jezične uzorke, rabeći pri tome jednostavnije rečenice te čineći više gramatičkih pogrešaka. Dockrell i sur. (2009) u longitudinalnom istraživanju pokazuju da, u usporedbi s nacionalnim normama, djeca s RJP-om pokazuju kontinuirano niske rezultate. Štoviše, s porastom dobi, to jest u tinejdžersko i adolescentsko doba, rezultati su sve niži.

Istraživanja u kojima se proučavalo pisanje osoba s RJP-om nisu bila usredotočena isključivo na rječničku raznolikost, ali su neka od njih uključivala procjenu rječničke raznolikosti ili neke druge mjere rječničkoga razvoja. Tako Williams i sur. (2013) potvrđuju da u pisanome jeziku djeca

s RJP-om upotrebljavaju manji raspon riječi od djece UJR-a. Rezultati istraživanja Scott i Windsor (2000) pokazuju da se skupina s RJP-om od skupine UJR-a razlikuje u rječničkoj raznolikosti (mjerenoj mjerom NDW) samo u usporedbi pisanih pripovjednih uzoraka, ali ne i govorenih.

## CILJ I PRETPOSTAVKE ISTRAŽIVANJA

Dosadašnja istraživanja ostavljaju otvorenima pitanja o razlikovanju osoba s RJP-om od osoba UJR-a u rječničkoj raznolikosti te o prikladnosti uporabe različitih mjera. Nije sasvim jasno razlikuju li se dvije skupine u rječničkoj raznolikosti te koje mjere tu razliku mogu pokazati. Uz to je rječnička raznolikost rijetko istraživana na pisanim tekstovima osoba s RJP-om (ali vidi, primjerice, Williams i sur., 2013; Scott i Windsor, 2000). Cilj je ovog istraživanja ispitati rječničku raznolikost osoba s RJP-om u odnosu na rječničku raznolikost osoba UJR-a. Za razliku od ranijih istraživanja koja su ispitivala rječničku raznolikost u govorenome diskursu mlađih osoba, ovim se istraživanjem željelo na uzorku osoba širega raspona kronološke dobi provjeriti je li ova razlika prisutna u tekstovima koji nastaju kao rezultat pisanoga diskursa. U skladu s time postavljena su sljedeća istraživačka pitanja i pretpostavke:

P1: Razlikuju li se jezični uzorci osoba s RJP-om od jezičnih uzoraka osoba UJR-a u rječničkoj raznolikosti na mjerama koje pretpostavljaju ujednačavanje duljine uzoraka?

H1: Na mjeri broj različitih riječi (NDW) neće biti razlike u razini rječničke raznolikosti između osoba s RJP-om i osoba UJR-a ako se pri izračunu ujednači broj riječi.

H2: Na mjeri omjer različnica i pojava (TTR) neće biti razlika u razini rječničke raznolikosti između osoba s RJP-om i osoba UJR-a ako se pri izračunu ujednači broj riječi.

P2: Razlikuju li se jezični uzorci osoba s RJP-om od jezičnih uzoraka osoba UJR-a u rječničkoj raznolikosti na mjerama koje ne pretpostavljaju ujednačavanje duljine uzoraka?

H1: Na mjeri *indeks rječničke raznolikosti D* osobe s RJP-om postizat će značajno niži rezultat od osoba UJR-a.

H2: Na mjeri *pomični prosječni omjer različnica i pojava (MATTR)* osobe s RJP-om postizat će značajno niži rezultat od osoba UJR-a.

## METODOLOGIJA

### Sudionici istraživanja

Jezični uzorci za ovo istraživanje izdvojeni su iz *Hrvatskoga korpusa neprofesionalnoga pisanog jezika* (Kuvač Kraljević, Hržica i Kologranić Belić, u tisku)<sup>2</sup>. Korpus sadrži više od pola milijuna pojava pisanoga teksta jednojezičnih hrvatskih govornika hrvatskoga jezika koji se ne bave profesionalnim pisanjem. Za ovo su istraživanje izdvojene dvije skupine govornika: skupina govornika UJR-a te skupina govornika s RJP-om. Sudionici s RJP-om ( $n=20$ ) upareni su sa sudionicima UJR-a ( $n=20$ ) prema spolu i kronološkoj dobi (u rasponu od +/- 12 mjeseci). Sudionici s RJP-om ispitivani su u klinikama opremljenima za logopedsku dijagnostiku i terapiju. Riječ je o osobama koje su prema *Međunarodnoj klasifikaciji bolesti i srodnih zdravstvenih stanja* (MKB-10; Svjetska zdravstvena organizacija, 2012) klasificirani pod kategorijama F80.1 (*Poremećena sposobnost jezičnog izražavanja*) ili F80.2 (*Poremećaji razumijevanja*). Podatci o broju, kronološkoj dobi i spolu sudionika priloženi su u Tablici 1.

Tablica 1. Demografske karakteristike sudionika

Skupine	n	Kronološka dob			Spol	
		Raspon	M	SD	m	f
Uredni jezični status	20	10;1 – 20;6	13;6	2.62	12	8
Razvojni jezični poremećaj	20	10;2 – 21;6	13;6	2.77	12	8

### Materijali i procedura

Za usporedbu jezičnih uzoraka u ovom istraživanju izabrana je priča *Priča o plaži*, elicitirana pomoću slikovnog predloška koji čini dio

2 Korpus je nastao u okviru projekta *Računalni asistent za podršku pri unosu teksta osobama s jezičnim poremećajima* (RAPUT; RC.2.2.08 - 0050) pod voditeljstvom izv. prof. dr. sc. Jelene Kuvač Kraljević.

testnoga materijala *Expression, Reception and Recall of Narrative Instrument* (ERRNI; Bishop, 2004). Sudionici su tekstove pisali na posebnom obrascu rukom. Ispitivači su potom tekstove unosili u program za prikupljanje jezičnih uzoraka. Iako su tekstovi naknadno lematizirani te morfološki označeni (Štefanec, Ljubešić i Kuvač Kraljević, 2016), za ovu su prigodu izdvojeni neoznačeni tekstovi u .txt formatu. Tekstovi su potom prebačeni u računalni program *Computerised Language Analysis* (CLAN) koji je dio baze TalkBank (MacWhinney, 2000) i kodirani su prema sustavu *Codes for the Human Analyses of Transcripts* (CHAT; MacWhinney, 2000). Budući da na izračun mjera rječničke raznolikosti mogu utjecati ortografske pogreške (isti se sadržaj izriče različitim oblicima), one su kodirane tako da ih program tretira kao ispravan oblik riječi u svim jezičnim uzorcima. Takve prilagodbe čine 0.28% od ukupnoga broja riječi u skupini sudionika UJR-a te 1.04% od ukupnoga broja riječi u skupini sudionika s RJP-om. Pomoću naredbi u programu CLAN izračunate su četiri mjere rječničke raznolikosti: *broj različitih riječi* (NDW), *omjer različenica i pojavnica* (TTR), *pomični prosječni omjer različenica i pojavnica* (MATTR) te *indeks rječničke raznolikosti D*. Prve su tri mjere, to jest NDW, TTR i MATTR, izračunate uporabom programa *FREQ*. Mjeru NDW program računa tako da za svaki transkript izračuna ukupan broj različenica, to jest zapisanih različitih oblika riječi u tekstu, dok se vrijednost TTR-a temelji na automatskom izračunu omjera različenica i pojavnica, to jest svake pojave riječi, uključujući i ponavljanja (Jelaska, 2005). Kako su mjere NDW i TTR ovisne o veličini uzorka, jezični su uzorci ujednačeni prema broju riječi transkripta najmanje veličine i mjere su izračunate na tome broju riječi. Mjeru MATTR program izračunava kao omjer različenica i pojavnica (TTR) za svaki pomični okvir teksta, pri čemu je bilo potrebno u dijelu formule  $+bN$  odrediti  $N$ , to jest broj riječi na temelju kojega će program izračunati TTR unutar svakoga pomičnog okvira. S obzirom na to da u prijašnjim istraživanjima ne postoje zadane preporuke koliki bi  $N$  bio prikladan, s ciljem da se pomicanjem okvira i u najmanjem jezičnom uzorku u obzir uzima rječnička raznolikost različitih dijelova teksta, u ovome je radu u pomični okvir uzeto 30 riječi. Nadalje, mjera  $D$  izračunata je pomoću *VOCD* programa

koji nasumično iz jezičnog uzorka izdvaja riječi na kojima tijekom prosječno 100 ponavljanja izračunava TTR. Usporedbom dobivenih vrijednosti te teorijskoga modela određuje se vrijednost  $D$ . Naposljetku su u zasebnim tabličnim datotekama dobiveni podatci o vrijednostima navedenih mjera za svaki od analiziranih transkripata, ovisno o programu kojim su mjere izračunate. S obzirom da je za jednoga sudionika nedostajao jezični uzorak na priči odabranoj za analizu, podatci su obrađeni na uzorku od 20 osoba s RJP-om te 19 osoba UJR-a.

## REZULTATI

Prije no što se pristupilo obradi podataka, za svaku su skupinu sudionika provjereni normaliteti distribucija dobivenih rezultata. S obzirom na malu veličinu uzoraka na kojima su podatci prikupljeni, normaliteti distribucija testirani su Shapiro-Wilkovim testom, koji u obje skupine sudionika nije pokazao značajno odstupanje distribucija od normalne ni na jednoj ispitivanoj varijabli. Također, na temelju pokazatelja simetričnosti i spljoštenosti distribucija zaključeno je da ne postoje velika odstupanja. Kako su uočena odstupanja mala, a parametrijski testovi robusni te u velikoj mjeri neosjetljivi na odstupanja opaženih rezultata od normalne distribucije, odlučeno je da će podatci biti analizirani  $t$ -testom za nezavisne uzorke. Ipak, zbog oblika pojedinih distribucija rezultata koje ni približno ne prate oblik normalne krivulje te nezadovoljenosti uvjeta homogenosti varijanci na varijabli  $D$ , odlučeno je da se procjene dobivene  $t$ -testom dodatno provjere *bootstrap* metodom. Radi se o statističkom postupku koji se temelji na ponovljenom uzorkovanju elemenata iz dobivenoga skupa podataka s njihovim vraćanjem u osnovni skup. Dobiveni skup podataka u ovome se slučaju tretira kao populacija iz koje se po slučaju u novi uzorak odabire jedan po jedan element, pri čemu se svaki izabrani element prije odabira sljedećega vraća natrag u osnovni skup. Ovaj se postupak uzorkovanja ponavlja dok se ne formira veći broj uzoraka, najčešće između 1000 i 10 000. Za svaki ponovljeni uzorak izračunaju se statističke vrijednosti koje istraživač želi procijeniti (primjerice, aritmetičke sredine), što omogućuje dobivanje njihove distribucije te točnije određivanje intervala pouzdanosti.

*Bootstrap* postupak omogućuje ponavljanje izvorne analize na velikome broju skupova podataka generiranih iz početnoga skupa, što nam omogućuje provjeru stabilnosti dobivenih rezultata. Pritom je njegova glavna prednost što ne počiva na pretpostavkama o obliku populacijske distribucije. Budući da daje točnije statističke procjene u situacijama kada distribucije dobivenih rezultata značajnije odstupaju od normalne, *bootstrap* postupak sve se češće koristi kao alternativa parametrijskim postupcima kada nisu zadovoljeni svi preduvjeti. Kako su *bootstrap* postupkom na 10 000 generiranih uzoraka i uz intervale pouzdanosti od 95% potvrđene procjene dobivene *t*-testom (vrijednosti dobivenih parametara zanemarivo se razlikuju u drugoj ili trećoj decimali), zaključeno je kako ranije navedena odstupanja nisu značajno utjecala na rezultate analiza, zbog čega će u nastavku teksta biti prikazani samo rezultati *t*-testa. U Tablici 2 su za dvije skupine sudionika, osobe UJR-a i osobe s RJP-om, prikazani prosječni rezultati dobiveni na četiri mjere rječničke raznolikosti. Kao što je vidljivo iz Tablice 2 na svim su mjerama sudionici s RJP-om postigli nešto niži rezultat u odnosu na kontrolnu skupinu UJR-a.

**Tablica 2.** Prosječni rezultati na mjerama rječničke raznolikosti (te pripadajuće standardne devijacije) s obzirom na jezični status sudionika

Mjere rječničke raznolikosti	UJR			RJP		
	n	M	SD	N	M	SD
NDW	19	64.0	3.23	20	61.3	4.13
TTR	19	0.720	0.040	20	0.687	0.049
MATTR	19	0.844	0.026	20	0.813	0.028
D	19	39.335	10.008	20	32.205	6.401

*Legenda.* NDW (engl. *number of different words*) = broj različitih riječi; TTR (engl. *type-token ratio*) = omjer različenica i pojavnica; MATTR (engl. *moving average type-token ratio*) = pomični prosječni omjer različenica i pojavnica; D (engl. *vocabulary diversity*) = indeks rječničke raznolikosti; UJR = sudionici urednog jezičnog razvoja; RJP = sudionici s razvojnim jezičnim poremećajem.

Prvi problem ovog istraživanja bio je provjeriti razlikuju li se jezični uzorci osoba s RJP-om od jezičnih uzoraka osoba UJR-a u rječničkoj raznolikosti na mjerama koje pretpostavljaju ujednačavanje duljine uzoraka. Kako bi se odgovorilo na ovo pitanje, izračunata su dva *t*-testa kojima su

testirane razlike između dvije skupine sudionika u rezultatima na mjerama NDW (broj različitih riječi) te TTR (omjer različenica i pojavnica). Dobiveni su rezultati pokazali statistički značajnu razliku na obje mjere ( $t(37) = -2.27$ ;  $p = .03$ ;  $d = 0.73$  za NDW, odnosno  $t(37) = -2.32$ ;  $p = .03$ ;  $d = 0.74$  za TTR), što suprotno pretpostavkama ukazuje da se jezični uzorci osoba s RJP-om i osoba UJR-a značajno razlikuju na mjerama rječničke raznolikosti koje pretpostavljaju ujednačavanje duljine uzoraka. Pritom su sudionici s RJP-om postigli značajno niži rezultat na rječničkoj raznolikosti mjerenoj mjerama NDW i TTR ( $M_{NDW} = 61.3$ ;  $M_{TTR} = 0.687$ ) u odnosu na sudionike UJR-a ( $M_{NDW} = 64.0$ ;  $M_{TTR} = 0.720$ ). Kao pokazatelj veličina učinaka izračunat je Cohenov  $d$  koji ukazuje na srednju veličinu učinka na obje ispitivane varijable. Međutim, procijenjeni intervali pouzdanosti uz razinu rizika od 5% ukazuju da se prava veličina učinka kreće u rasponu od vrlo male do velike veličine učinka (za NDW  $d = 0.73$ , 95% CI: [0.08, 1.37]; za TTR  $d = 0.74$ , 95% CI: [0.09, 1.39]). Ipak, treba uočiti da navedeni rasponi ne uključuju vrijednost nula, na temelju čega možemo zaključiti da postoji značajan efekt jezičnoga statusa na rječničku raznolikost mjerenu ovim mjerama.

Drugi problem na koji se ovim istraživanjem željelo odgovoriti jest razlikuju li se jezični uzorci osoba s RJP-om od jezičnih uzoraka osoba UJR-a u rječničkoj raznolikosti na mjerama koje ne pretpostavljaju ujednačavanje duljine uzoraka. Ponovno su izračunata dva *t*-testa kojima je testirano postoje li statistički značajne razlike između dvije skupine sudionika u rezultatima na mjerama MATTR (pomični prosječni omjer različenica i pojavnica) te D (indeks rječničke raznolikosti). Dobiveni su rezultati pokazali statistički značajnu razliku na obje mjere ( $t(37) = -3.6$ ;  $p = .001$ ;  $d = 1.15$  za MATTR, odnosno  $t(30.37) = -2.64$ ;  $p = .013$ ;  $d = 0.85$  za indeks D). Ovi rezultati potvrđuju pretpostavku da se jezični uzorci osoba s RJP-om i osoba UJR-a značajno razlikuju na mjerama rječničke raznolikosti koje ne pretpostavljaju ujednačavanje duljine uzoraka, pri čemu je niža rječnička raznolikost očekivano dobivena u skupini osoba s RJP-om ( $M_{MATTR} = 0.813$ ;  $M_D = 32.205$ ) u odnosu na kontrolnu skupinu UJR-a ( $M_{MATTR} = 0.844$ ;  $M_D = 39.335$ ). Kao pokazatelj veličina učinaka izraču-

nat je Cohenov  $d$  koji ukazuje na veliku veličinu učinka na obje varijable. Međutim, procijenjeni intervali pouzdanosti uz razinu rizika od 5% ukazuju da se prava veličina učinka kreće u rasponu od srednje do velike za mjeru MATTR ( $d = 1.15$ , 95% CI: [0.47, 1.82]), odnosno od male do velike veličine učinka za indeks D ( $d = 0.85$ , 95% CI: [0.20, 1.51]). Navedeni rasponi ne uključuju vrijednost nula, zbog čega možemo zaključiti da postoji značajan efekt jezičnoga statusa na rječničku raznolikost mjerenu ovim mjerama.

## RASPRAVA

Ovim se istraživanjem željelo ispitati razlikuju li se osobe s RJP-om od osoba UJR-a u rječničkoj raznolikosti. Pri tome se polazilo od činjenice da je rječničku raznolikost moguće mjeriti na različite načine, to jest, različitim mjerama, a rezultati prethodnih istraživanja ističu kao pouzdanije skupinu mjera neovisnih (ili barem manje ovisnih) o duljini jezičnog uzorka. Rječničku je raznolikost moguće mjeriti i na pisanim tekstovima i na transkriptima govorenoga jezika. U istraživanjima osoba s RJP-om rječnička raznolikost znatno se češće promatrala na govorenome jeziku, kao što je i općenito govoreni jezik tih osoba proučavaniji. No istraživanja govorenoga i pisanoga jezika osoba s RJP-om pokazuju da oni, posebice u adolescentskoj dobi, mogu sve više zaostajati za govornicima UJR-a baš u pisanome jeziku. Ovo istraživanje uključuje različite mjere rječničke raznolikosti, slabije istražen pisani modalitet te nešto starije sudionike no što je u većini istraživanja RJP-a, zbog čega je rezultate potrebno promatrati u svjetlu tih činjenica.

Ranija istraživanja rječničke raznolikosti većinom su pokazala da se osobe s RJP-om ne razlikuju od osoba UJR-a na mjerama NDW (broj različitih riječi) te TTR (omjer različenica i pojavnica), ako su one izračunavane na jednakome broju riječi za svakog ispitanika (Owen i Leonard, 2002; Thordardottir i Weismer, 2001). Ovim se pak istraživanjem došlo do drugačijih rezultata. Na obje mjere, mjerene na jednakom broju riječi, dobivena je značajna razlika u rječničkoj raznolikosti između dviju skupina govornika. Ipak, ovo istraživanje nije jedino koje je otkrilo razlike između dviju skupina na mjerama NDW ili TTR. Primjerice, Fey i sur. (2004) pronalaze razlike u broju različitih riječi

kod djece mlađe školske dobi u dvjema dobnim skupinama (drugi razred i četvrti razred). Uz usporedbom napredovanja djece od drugoga do četvrtoga razreda pokazali su da iako djeca s RJP-om općenito pokazuju manji napredak od djece UJR-a, statistički značajno manje napreduju samo u rječničkoj raznolikosti mjerenoj mjerom NDW. Slično potvrđuju i Norbury, Gemmel i Paul (2013) te Mäkinen (2014). Spomenuta se istraživanja razlikuju od istraživanja u kojima nisu dobivene značajne razlike između dvije skupine djece na mjerama NDW i TTR prema načinu elicitanja jezičnog uzorka te dobi sudionika. Razlike između dviju skupina većinom se nisu pokazale značajnima u istraživanjima jezičnih uzoraka spontane konverzacije, a pokazuju se značajnima u istraživanjima pripovjednih uzoraka (za odstupanje od ovih trendova vidi Charest i Skoczylaz, 2019). Osim što se ovi jezični uzorci eliciraju na različite načine, to jest u različitim kontekstima (slobodna ili vođena igra naspram slikovnoga predloška pomoću kojega je potrebno oblikovati priču), razina složenosti teksta koji je potrebno proizvesti znatno je različita. Dok spontana konverzacija elicirana slobodnom igrom uključuje nižu razinu strukture, uporabu manje naprednoga i ne nužno ciljanoga rječnika, pripovijedanje je složena jezična vještina s visokim zahtjevima na različitim jezičnim razinama. U usporedbi ovih dvaju žanrova Westerveld, Gillion i Muller (2004) izvještavaju da su djeca redovito proizvodila više riječi u pričama nego u konverzaciji. Wagner, Nettelbladt, Sahlen i Nilholm (2000) također otkrivaju brojne razlike između ovih dviju vrsta diskursa djece s RJP-om. Oni ih objašnjavaju većom zahtjevnosću planiranja te višom razinom strukture priča. Istraživanja koja rječničku raznolikost mjere u spontanoj konverzaciji usmjerena su na mlađu djecu, za razliku od istraživanja rječničke raznolikosti u pripovijedanju koja su pak usmjerena na stariju djecu. Stoga je moguće da je na dobivene rezultate utjecala upravo dob sudionika. Primjerice, Westerveld i sur. (2004) su pokazali da se četverogodišnjaci i petogodišnjaci ne razlikuju u broju različitih riječi na zadatku pripovijedanja, dok se starija djeca, tinejdžeri i adolescenti razlikuju (Westerveld i Moran, 2013). Dakle, na razlike između osoba s RJP-om i osoba UJR-a u ovom istraživanju mogla je utjecati činjenica da su

sudionici ispitivanja bila starija djeca, tinejdžeri i adolescenti.

Znatno je manje istraživanja rječničke raznolikosti osoba s RJP-om i osoba UJR-a u obzir uzimalo mjere koje nisu NDW ili TTR. Rezultati ovog istraživanja pokazali su da se skupine osoba UJR-a i osoba s RJP-om razlikuju u rječničkoj raznolikosti mjerenoj i mjerama D te MATTR. I u dosadašnjim istraživanjima pokazalo se da postoje razlike između ovih dviju skupina u rječničkoj raznolikosti kada je ona mjerena mjerom D (Klee i sur., 2004; Wong i sur., 2010; Owen i Leonard, 2002). No mjera MATTR, unatoč preporukama uporabe (McCarthy i Jarvis, 2010; Fergadiotis, Wright i Greenc, 2015) te uključenosti u standardne alate za obradu uzoraka govorenoga jezika (*Computerized Language Analysis*, CLAN; *Systematic Analysis of Language Transcripts*, SALT), nije se često upotrebljavala u ovakvim usporedbama. Charest i Skoczylas (2019), u istraživanju u kojem se uspoređuju pripovjedni uzorci priče jednake onoj koja je upotrijebljena u ovom istraživanju, ne nalaze razlike između djece UJR-a i djece s RJP-om na toj mjeri. Iako takvi rezultati mogu djelovati proturječno rezultatima dobivenima u ovom istraživanju, potrebno je u obzir uzeti dvije stvari. Prvo, dob sudionika općenito je bila niža od dobi sudionika ovog istraživanja, to jest mjere su izračunate na jezičnim uzorcima šestogodišnjaka. Drugo, način izračuna mjere MATTR može znatno varirati, što utječe na rezultate. Mjera se izračunava tako da istraživač sam odredi koliki će biti pomični okvir za izračun omjera TTR-a. Ako je za pomični okvir određen broj 10, program će računati TTR na 10 riječi, zatim se pomaknuti za jedno mjesto i izračunati TTR za narednih 10 riječi i tako dok ne dođe do kraja transkripta. No pomični okvir može biti i znatno veći. Naime, Charest i Skoczylas (2019) izabrale su pomični okvir od 100 riječi. S obzirom na to da je najkraći uzorak iznosio 146 riječi, u ovakvom je uzorku svako ponovljeno mjerenje obuhvatilo najveći dio međusobno preklapajućega teksta, računajući TTR na sličnim nizovima riječi. Za razliku od toga odabir manjega broja riječi (u ovom istraživanju 30 riječi) za analizu kraćih tekstova omogućava da se pomicanjem užega pomičnog okvira kroz tekst dobivaju te usrednjuju TTR-i različitih dijelova teksta. Uz ove je metodološke

razlike svakako potrebno u obzir uzeti i činjenicu da je u ovom istraživanju riječ o pisanim jezičnim uzorcima, a u dosadašnjim se istraživanjima pokazalo da osobe s RJP-om ne samo da zaostaju za vršnjacima kada je riječ o pisanju tekstova već je taj zaostatak i veći od onoga u govorenome jeziku (Williams i sur., 2013; Scott i Windsdor, 2000).

Na sve četiri mjere rječničke raznolikosti koje su se rabile u ovom istraživanju dobivene su statistički značajne razlike u rječničkoj raznolikosti između sudionika s RJP-om i sudionika UJR-a. Iako se polazilo od pretpostavke da na mjerama NDW i TTR neće biti razlika u razini rječničke raznolikosti između osoba s RJP-om i osoba UJR-a, rezultati su pokazali suprotno, što se može objasniti zahtjevnim jezičnim zadatkom koji uključuje i pripovijedanje i pisanje, ali i razlikama u dobi sudionika u odnosu na ranija istraživanja. No ranija su istraživanja mjera rječničke raznolikosti pokazala da su mjere D te MATTR pouzdanije od mjera NDW i TTR. Rezultati ovog istraživanja također mogu govoriti u prilog primjene mjera D te MATTR. Iako je razlika među skupinama dobivena na svim mjerama, veličine učinka su se pokazale većima na mjerama D te MATTR, u odnosu na mjere NDW te TTR. Naime, dobivena veličina učinka za mjeru MATTR iznosi  $d = 1.15$ , što znači da su na ovoj mjeri osobe UJR-a, koje su postigle prosječni rezultat, imale bolji rezultat od 87.5% osoba s RJP-om. Za mjeru D dobivena je veličina učinka  $d = 0.85$ , što znači da su na ovoj mjeri osobe UJR-a, koje su postigle prosječni rezultat, imale bolji rezultat od oko 80% osoba s RJP-om. S druge strane, razlikovanje skupina na mjerama NDW i TTR nešto je manje i dobivene veličine učinka ukazuju na to da su osobe UJR-a s prosječnim rezultatom na ovim mjerama bolje od oko 77% osoba s RJP-om. Dobiveni rezultati ukazuju na veću mogućnost razlikovanja ovih skupina u rječničkoj raznolikosti pomoću mjera MATTR i D. Osim toga, mjere NDW i TTR izračunate su na jednakome broju riječi za sve sudionike. Dakle, pri primjeni ovih mjera potrebno je jezične uzorke dodatno kratiti prema uzorku najkraće duljine, uz sve opasnosti koje iz toga proističu.

Za kraj je potrebno osvrnuti se na nekoliko metodoloških nedostataka. Jezični uzorci koji su analizirani prikupljeni su na vrlo malenom uzorku,

što ograničava mogućnost zaključivanja o veličini dobivenih učinaka. Naime, iako su dobivene statistički značajne razlike i značajne veličine učinaka, intervali pouzdanosti su vrlo široki i ukazuju da se prave veličine učinaka nalaze u rasponu od malih do velikih. Kako bi se dobile preciznije veličine učinaka, potrebno je ispitati razlike u rječničkoj raznolikosti između osoba s RJP-om i osoba UJR-a na većem uzorku sudionika. Također, treba uzeti u obzir da su podatci prikupljeni na prigodnom uzorku, zbog čega je ograničena mogućnost generalizacije rezultata na ostatak populacije. Kako bi se kontrolirali efekti dobi i spola na rječničku raznolikost, dvije skupine sudionika izabrane su na način da je svaki sudionik UJR-a uparen sa sudionikom s RJP-om prema ovim karakteristikama. Međutim, budući da nam podaci o socio-ekonomskom statusu sudionika nisu poznati, ne možemo zaključivati o ujednačenosti skupina prema ovom obilježju. S obzirom na to da se socio-ekonomski status također pokazao značajnim u razvoju rječnika, u budućim bi istraživanjima svakako trebalo kontrolirati i ovaj čimbenik.

Bitno je napomenuti da su prilikom uzorkovanja jezičnih uzoraka osoba s RJP-om birani uzorci s većim brojem riječi zato što se pokazalo da pojedine mjere rječničke raznolikosti daju preciznije rezultate kada se računaju na većim jezičnim uzorcima. S druge strane, u skupini osoba UJR-a bilo je dostupno svega 19 jezičnih uzoraka te su oni svi uzeti u analizu bez obzira na broj riječi. Ova pristranost u uzorkovanju vidljiva je u prosječnome broju riječi koji je podjednak u obje skupine sudionika, ali ne bi trebala imati utjecaja na rezultate na mjerama rječničke raznolikosti. Naime, mjere NDW i TTR su izračunate na jezičnim uzorcima ujednačenima prema duljini najkraćeg uzorka, a mjere MATTR i D bi same po sebi trebale biti neosjetljive na duljinu jezičnog uzorka. Dakle, da u uzorak osoba s RJP-om nisu namjerno birani sudionici s duljim jezičnim uzorcima, za očekivati je da bi ova skupina u prosjeku imala manju rječničku produktivnost od osoba UJR-a, ali razlika u rječničkoj raznolikosti bi i dalje trebala postojati u istome smjeru. Ipak, dodatno treba uzeti u obzir da je ovakav odabir sudionika mogao rezultirati time da su u uzorku osoba s RJP-om zastupljeni pojedinci s nešto blažim oblikom poremećaja,

odnosno boljim jezičnim sposobnostima, zbog čega treba biti oprezan prilikom tumačenja dobivenih rezultata na mjerama rječničke raznolikosti u ovoj skupini. Sudionici uključeni u istraživanje već su imali postavljenu dijagnozu razvojnoga jezičnog poremećaja te se nisu provodila dodatna ispitivanja (primjerice, standardiziranim jezičnim testovima) kojima se mogla utvrditi jakost poremećaja. Ukoliko je razina rječničke raznolikosti procijenjena na osobama s blažim poremećajem, možemo pretpostaviti da su dobiveni prosječni rezultati nešto viši od očekivanih za osobe s RJP-om. Unatoč tomu, razlike između skupina dobivene su na svim mjerama rječničke raznolikosti, što dodatno potvrđuje naše zaključke.

Naposlijetku, u obzir treba uzeti dob sudionika. Izuzev toga što su sudionici ovog istraživanja bili stariji u odnosu na sudionike ranijih istraživanja, raspon dobi je vrlo širok. Iako još uvijek ne znamo dovoljno o razvojnim promjenama jezičnih sposobnosti osoba s RJP-om, neka istraživanja (primjerice, Haynes, 1992; Stothard i sur., 1998) ukazuju na to da se razlike u jezičnim sposobnostima između osoba s RJP-om i osoba UJR-a mogu razlikovati s obzirom na dob. Za razliku od ovog istraživanja, u kojem su sudionici različite dobi promatrani kao homogena skupina, budućim bi se istraživanjima mogle ispitati razlike u rječničkoj raznolikosti između osoba s RJP-om i osoba UJR-a s obzirom na dob, čime bi se provjerilo razlikuju li ispitivane mjere u jednakoj mjeri mlađe i starije sudionike.

## ZAKLJUČAK

Rezultati ovog istraživanja nedvojbeno pokazuju da se osobe s RJP-om razlikuju od vršnjaka UJR-a u rječničkoj raznolikosti. Dobiveni se rezultati razlikuju od nekih ranijih istraživanja, u kojima se tvrdi da su razlike u rječničkoj raznolikosti samo posredni rezultat sintaktičke složenosti koja dovodi do razlika u duljini uzoraka (pregled: Charest i Skoczylaz, 2019). U ovom su istraživanju primijenjeni postupci kojima se utjecaj razlika u veličini uzorka neutralizira (ili izračunom na jednakome broju riječi ili uporabom mjera koje ne ovise o veličini uzoraka) te dobiveni rezultati odražavaju samo razlike u rječničkoj raznolikosti između dviju skupina. Ovime se pokazuje da je rječnik jedna od jezičnih sastavnica s kojom osobe s RJP-om zasi-

gurno imaju teškoća, što utječe na njihovu jezičnu proizvodnju, a posebice na zahtjevne zadatke oblikovanja pisanoga pripovjednog diskursa.

Ovakvi rezultati istraživanja imaju implikacije i na jezičnu procjenu, to jest logopedsku praksu. Naime, procjena rječničkoga znanja, kao dio jezične procjene, obično se provodi pomoću standardiziranih testova, a posebice se u Hrvatskoj rijetko uz takve testove upotrebljava i analiza jezičnih uzoraka. Međutim, takva analiza, naročito na razini rječnika, može biti izvrsna dopuna testovima jer ispituje uporabu jezika u uobičajenoj situaciji jezične proizvodnje. Osim toga, testovi u obzir uzimaju poznavanje standardnoga jezika, dok se mjere rječničke raznolikosti temelje isključivo na odnosu različitih riječi i ukupnoga broja riječi, pa će prikladnije procijeniti govornike različitih dijalekata.

## **ZAHVALA**

Ovaj rad nastao je u sklopu projekta “Računalni asistent za pomoć pri unosu teksta osobama s jezičnim teškoćama” (Computer assistant for text input for persons with language impairment; RAPUT, EU, RC.2.2.08-0050) financiran iz Europskih strukturnih i investicijskih fondova.



## LITERATURA

- Allwood, J. (1998). Some frequency based differences between spoken and written Swedish. U: T. Haukioja (ur.), *Proceedings of the 16th Scandinavian Conference of Linguistics* (str. 18-29). Department of Linguistics, University of Turku.
- Badurina, L. i Kovačević, M. (2001). *Raslojavanje jezične stvarnosti*. Rijeka: Izdavački centar Rijeka.
- Baron, N. S. (2000). *Alphabet to Email: How Written English Evolved and Where It's Heading*. London: Routledge.
- Biber, D. (1995). *Variation across Speech and Writing* (3. izdanje). Cambridge: Cambridge University Press.
- Bishop, D. V. (2006). What causes specific language impairment in children? *Current Directions in Psychological Science*, 15(5), 217-221. DOI: 10.1111/j.1467-8721.2006.00439.x.
- Bishop, D. V. M. (2004). *Expression, reception, and recall of narrative instrument [Measurement instrument]*. London, UK: Harcourt Assessment.
- Brown, G. i Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Chafe, W. (1984). Integration and involvement in speaking, writing, and oral literature. U: D. Tannen (ur.), *Spoken and Written Language: Exploring Orality and Literacy* (2. izdanje, str. 35-53). Norwood, NJ: Ablex.
- Charest, M. i Skoczylas, M. J. (2019). Lexical diversity versus lexical error in the language transcripts of children with developmental language disorder: Different conclusions about lexical ability. *American Journal of Speech-Language Pathology*, 28(3), 1275-1282. DOI: [https://doi.org/10.1044/2019\\_AJSLP-18-0143](https://doi.org/10.1044/2019_AJSLP-18-0143).
- Clarke, M. i Leonard, L. (1996). Lexical comprehension and grammatical deficits in children with specific language impairment. *Journal of Communication Disorders*, 29(2), 95-105. DOI: 10.1016/0021-9924(94)00036-0.
- Covington, M. A. i McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type-token ratio (MATTR). *Journal of Quantitative Linguistics*, 17(2), 94-100. DOI: <https://doi.org/10.1080/09296171003643098>.
- Daller, H., van Hout, R. i Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24(2), 197-222. DOI: 10.1093/applin/24.2.197.
- Dockrell, J. E., Lindsay, G. i Connelly, V. (2009). The impact of specific language impairment on adolescent's written text. *Exceptional Children*, 75(4), 427-446.
- Durán, P., Malvern, D., Richards, B. i Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics*, 25(2), 220-242. DOI: 10.1093/applin/25.2.220.
- Fergadiotis, G., Wright, H. H. i Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: Assessing length effects. *Journal of Speech, Language, and Hearing Research*, 58(3), 840-852. DOI: 10.1044/2015\_JSLHR-L-14-0280.
- Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. i Zhang, X. Y. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research*, 47(6), 1301-1318.
- Gillam, R. B. i Johnston, J. R. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research*, 35(6), 1303-1315.
- Guoxing, Y. (2009). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31(2), 236-259. DOI: <https://doi.org/10.1093/applin/amp024V>.
- Halliday, M. A. K. (1989). *Spoken and written language* (2. izdanje). Oxford: Oxford University Press.
- Haynes, C. (1992). Vocabulary deficit: One problem or many? *Child Language Teaching and Therapy*, 8(1), 1-17. DOI: <https://doi.org/10.1177/026565909200800101>.
- Henrichsen, P. J. i Allwood, J. (2005). Swedish and Danish, spoken and written language. A statistical comparison. *International Journal of Corpus Linguistics*, 10(3), 367-399. DOI: <https://doi.org/10.1075/ijcl.10.3.05hen>.
- Hewitt, L. E., Hammer, C. S., Yont, K. M. i Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders*, 38(3), 197-213. DOI: 10.1016/j.jcomdis.2004.10.002.

- Ivanetić, N. (2003). *Uporabni tekstovi*. Zagreb: Zavod za lingvistiku Filozofskog fakulteta Sveučilišta u Zagrebu.
- Jacobson, P. F. i Walden, P. R. (2013). Lexical diversity and omission errors as predictors of language ability in the narratives of sequential spanish–english bilinguals: A cross-language comparison. *American Journal of Speech-Language Pathology*, 22(3), 554-565. DOI: 10.1044/1058-0360(2013/11-0055).
- Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19(1), 57-84. DOI: <https://doi.org/10.1191/0265532202lt220oa>.
- Jelaska, Z. (2005). Oblici hrvatskih riječi. U: Z. Jelaska (ur.), *Hrvatski kao drugi i strani jezik* (str. 136-143). Zagreb: Hrvatska sveučilišna naklada.
- Klee, T. (1992). Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language Disorders*, 12(2), 28-41. DOI: <http://dx.doi.org/10.1097/00011363-199202000-00005>.
- Klee, T., Stokes, S. F., Wong, A. M., Fletcher, P. i Gavin, W. J. (2004). Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, 47(6), 1396-1410. DOI: 10.1044/1092-4388(2004/104).
- Köhler, R. i Galle, M. (1993). Dynamic aspects of text characteristics. U: L. Hřebíček i G. Altmann (ur.), *Quantitative text analysis* (str. 46-53). Trier: Wissenschaftlicher Verlag Trier.
- Kologranić Belić, L., Matić, A., Olujić, M. i Srebačić, I. (2015). Jezični, govorni i komunikacijski poremećaji djece predškolske i školske dobi. U: J. Kuvač Kraljević (ur.), *Priručnik za prepoznavanje i obrazovanje djece s jezičnim teškoćama* (str. 64-76). Zagreb: Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu.
- Kuvač, J. i Palmović, M. (2007). *Metodologija istraživanja dječjega jezika*. Jastrebarsko: Naklada Slap.
- Leonard, L. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.
- Mäkinen, L. (2014). *Narrative language in typically developing children, children with specific language impairment and children with autism spectrum disorder*. Oulu: University of Oulu.
- Malvern, D. D., Richards, B., Chipere, N. i Durán, P. (2004). *Lexical diversity and language development quantification and assessment*. New York: Palgrave Macmillan.
- McCarthy, P. M. i Jarvis, S. (2007). Vocd: A theoretical and empirical evaluation. *Language Testing*, 24(4), 459-488. DOI: 10.1044/1092-4388(2004/104).
- McCarthy, P. M. i Jarvis, S. (2010). MTLT, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392. DOI: [doi.org/10.3758/BRM.42.2.381](http://doi.org/10.3758/BRM.42.2.381).
- McGregor, K. K. (2017). Semantics in child language disorder. U: R. G. Schwartz (ur.), *Handbook of child language disorders* (2. izdanje, str. 365-387). New York: Routledge.
- McGregor, K. K., Newman, R. M., Reilly, R. M. i Capone, N. C. (2002). Semantic representation naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 45(5), 998-1014. DOI: [http://dx.doi.org/10.1044/1092-4388\(2002/081\)](http://dx.doi.org/10.1044/1092-4388(2002/081)).
- McKee, G., Malvern, D. D. i Richards, B. J. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15(3), 323-338.
- Miller, J. F. (1991). Quantifying productive language disorders. U: J. F. Miller (ur.), *Research on child language disorders: A decade of progress* (str. 211-220). Austin, TX: Pro-Ed.
- Miller, J. E. i Weinert, R. (1998). *Spontaneous spoken language: syntax and discourse*. Oxford: Clarendon Press.
- Nash, M. i Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research*, 48, 439-458. DOI: 10.1044/1092-4388(2005/030).
- Nation K. (2014). Lexical learning and lexical processing in children with developmental language impairments. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369(1634), 20120387. DOI: 10.1098/rstb.2012.0387

- Nippold, M. (2016). *Later language development: School-age children, adolescents, and young adults* (4. izdanje). Austin, TX: PRO-ED, Inc.
- Norbury, C. F., Gemmel, T. i Paul, R. (2013). Pragmatic abilities in narrative production: A cross-disorder comparison. *Journal of Child Language*, 41(3), 485-510.
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., ... Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 57(11), 1247-1257. DOI: <http://doi.org/10.1111/jcpp.12573>.
- Olujčić, M. i Matic, A. (2017). Govorni i pisani jezik odraslih: koliko se razlikuju? *Govor*, 34(1), 33-52. DOI: <https://doi.org/10.22210/govor.2017.34.02>.
- Owen, A. J. i Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children with specific language impairment: Application of D. *Journal of Speech, Language, and Hearing Research*, 45(5), 927-937. DOI: [http://dx.doi.org/10.1044/1092-4388\(2002/075\)](http://dx.doi.org/10.1044/1092-4388(2002/075)).
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Redeker, G. (1984). On differences between spoken and written language. *Discourse Processes*, 7(1), 43-55. DOI: <https://doi.org/10.1080/01638538409544580>.
- Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., ... Eleanor, W. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders*, 49(4), 416-51. DOI: 10.1111/1460-6984.12102.
- Rice, M. L. i Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2;6 to 21 years of age. *Journal of Speech, Language, and Hearing Research*, 58(2), 345-359. DOI: 10.1044/2015\_JSLHR-L-14-0150.
- Richards, B. (1987). Type/token ratios: What do they really tell us? *Journal of Child Language*, 14(2), 201-209. DOI: <http://dx.doi.org/10.1017/S0305000900012885>.
- Richards, B. J. i Malvern, D. D. (1997). *Quantifying lexical diversity in the study of language development*. Reading: The New Bulmershe Papers.
- Roger, B. (1973). *A first language: The early stages*. London: Allen i Unwin.
- Rosamond, M., Tracy-Ventura, N. i McManus, K. (2017). *Anglophone students abroad: Identity, social relationships and language learning*. New York: Routledge.
- Schwartz, R. G., Botwinik-Rotem, I. i Friedman, N. (2017). Linguistics in child language disorder. U: R. Schwartz (ur.), *Handbook of child language disorders* (2. izdanje). New York: Routledge.
- Scott, C. M., i Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research*, 43(2), 324-339.
- Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B. i Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, 41(2), 407-418. DOI: <https://doi.org/10.1044/jslhr.4102.407>.
- Suddarth, R., Plante, E. i Vance, R. (2012). Written narrative characteristics in adults with language impairment. *Journal of Speech, Language, and Hearing Research*, 55(2), 409-420. DOI: 10.1044/1092-4388(2011/10-0295).
- Stillar, G. F. (1998). *Analyzing everyday texts: discourse, rhetoric, and social perspectives*. Thousand Oaks: SAGE Publications, Inc.
- Stromqvist, S., Johansson, V., Kriz, S., Ragnarsdottir, H., Aisenman, R. i Ravid, D. (2002). Toward a crosslinguistic comparison of lexical quanta in speech and writing. *Written language and Literacy*, 5(1), 45-67. DOI: <https://doi.org/10.1075/wll.5.1.03str>.

- Svjetska zdravstvena organizacija (2012). *Međunarodna klasifikacija bolesti i srodnih zdravstvenih stanja - MKB-10* (2. izdanje). Zagreb: Medicinska naklada Zagreb.
- Štefanec, V., Ljubešić, N. i Kuvač Kraljević, J. (2016). Croatian error-annotated corpus of non-professional written language. U: N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, ... S. Piperidis (ur.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (str. 3220-3226). Portorož: The European Language Resources Association.
- Templin, M. (1957). *Certain language skills in children*. Minneapolis: University of Minneapolis Press.
- Thal, D., J. O'Hanlon, L., Clemmons, M. i Fralin, L. (1999). Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language, and Hearing Research*, 42(2), 482-496. DOI: <https://doi.org/10.1044/jslhr.4202.482>.
- Thal, D. J., Reilly, J., Seibert, L., Jeffries, R. i Fenson, J. (2004). Language development in children at risk for language impairment: Cross-population comparisons. *Brain and Language*, 88(2), 167-179. DOI: [https://doi.org/10.1016/S0093-934X\(03\)00096-8](https://doi.org/10.1016/S0093-934X(03)00096-8).
- Thordardottir, E. T. i Weismer, E. S. (2001). High-frequency verbs and verb diversity in the spontaneous speech of school age children with specific language impairment. *International Journal of Language & Communication Disorders*, 36(2), 221-244. DOI: 10.1080/13682820152410872.
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E. i O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40(6), 1245-1260. DOI: 10.1044/jslhr.4006.1245.
- Ure, J. (1971). Lexical density and register differentiation. U: G. Perren i J. L. M. Trim (ur.), *Applications of Linguistics* (str. 443-452). Cambridge: Cambridge University Press.
- Wagner, C. R., Nettelbladt, U., Sahlen, B. i Nilholm, C. (2000). Conversation versus narration in pre-school children with language impairment. *International Journal of Language & Communication Disorders*, 35(1), 83-93.
- Watkins, R. V. Kelly, D. J., Harbers, H. M. i Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Disorders*, 38(6), 1349-1355. DOI: 10.1044/jslhr.3806.1349.
- Wells, C. G. (1985). *Language development in the pre-school years*. Cambridge: Cambridge University Press.
- Westerveld, M. F. i Moran, C. (2013). Spoken expository discourse of children and adolescents: Retelling versus generation. *Clinical Linguistics & Phonetics*, 27(9), 720-734.
- Westerveld, M. F., Gillon, G. T. i Miller, J. F. (2004). Spoken language samples of New Zealand children in conversation and narration. *Advances in Speech-Language Pathology*, 6(4), 195-208.
- Williams, G. J., Larkin, R. F. i Blaggan, S. (2013). Written language skills in children with specific language impairment. *International Journal of Language & Communication Disorders*, 48(2), 160-171.
- Wong, A. M.-Y., Klee, T., Stokes, S. F., Fletcher, P. i Leonard, L. B. (2010). Differentiating Cantonese-speaking preschool children with and without SLI using MLU and lexical diversity (D). *Journal of Speech, Language and Hearing Research*, 53(3), 794-799. DOI: 10.1044/1092-4388(2009/08-0195).
- Wright, H. H., Silverman, S. i Newhoff, M. (2003). Measures of lexical diversity in aphasia. *Aphasiology*, 17(5), 443-452. DOI: 10.1080/02687030344000166.
- Yu, G. (2010). Lexical diversity in writing and speaking task performances. *Applied Linguistics* 31(2), 236-259. DOI: <https://doi.org/10.1093/applin/amp024>.

## LEXICAL DIVERSITY IN WRITTEN TEXTS OF PERSONS WITH DEVELOPMENTAL LANGUAGE DISORDER

**Abstract:** *Developmental language disorder (DLD) is one of the most common language disorders in preschool and school age, and it also persists later in life. Children with DLD show a range of expressive and/or receptive difficulties in language, including vocabulary acquisition. The goal of this research was to explore the differences between persons with developmental language disorder and persons with typical language development (TLD) in lexical diversity. Earlier research focused on spoken discourse of younger speakers. In the present research, written discourse of speakers covering a broad age range was explored. Twenty participants with DLD and 19 with TLD were selected from the Croatian Corpus of Non-professional Written Language (Kuvač Kraljević, Hržica and Kologranić Belić, in press). They produced narrative language samples based on the Expression, Reception and Recall of Narrative Instrument (ERRNI; Bishop, 2004). To measure lexical diversity, four measures were calculated: the number of different words (NDW) and type-token ratio (TTR) on a restricted number of words, the moving average type-token ration (MATTR) and lexical diversity D on full-length samples. The independent samples t-test was used to compare the two groups. Participants with DLD had significantly lower results on all four measures. This leads to the conclusion that all four measures can differentiate groups of participants with different language status. Persons with DLD showed difficulties in using vocabulary when producing written narratives, which is a demanding language production task.*

**Keywords:** *developmental language disorder, written language samples, vocabulary diversity*