

# Digitalna era logopedije: automatizacija mjera sintaktičke složenosti u aplikaciji MultiDis

---

Kožul, Anđela Elinor

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Education and Rehabilitation Sciences / Sveučilište u Zagrebu, Edukacijsko-rehabilitacijski fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:158:857952>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-17**



Repository / Repozitorij:

[Faculty of Education and Rehabilitation Sciences - Digital Repository](#)



Sveučilište u Zagrebu  
Edukacijsko-rehabilitaciji fakultet

Diplomski rad  
**Digitalna era logopedije: automatizacija mjera sintaktičke  
složenosti u aplikaciji MultiDis**

Anđela Elinor Kožul

Zagreb, lipanj 2024.

Sveučilište u Zagrebu  
Edukacijsko-rehabilitaciji fakultet

Diplomski rad  
**Digitalna era logopedije: automatizacija mjera sintaktičke  
složenosti u aplikaciji MultiDis**

Andela Elinor Kožul

izv. prof. dr. sc. Gordana Hržica

Zagreb, lipanj 2024.

## **Izjava o autorstvu rada**

Potvrđujem da sam osobno napisao/napisala rad *Digitalna era logopedije: automatizacija mjera sintaktičke složenosti u aplikaciji MultiDis* i da sam njegov autor/autorica.

Svi dijelovi rada, nalazi ili ideje koje su u radu citirane ili se temelje na drugim izvorima jasno su označeni kao takvi te su adekvatno navedeni u popisu literature.

Ime i prezime: Anđela Elinor Kožul

Mjesto i datum: Zagreb, lipanj 2024.

## ZAHVALE

Ovaj rad je razvijen u sklopu projekta *Višerazinski pristup govorenom diskursu u jezičnom razvoju (HRZZ-UIP-2017-05-6603)* te se ovim putem želim zahvaliti cijelom timu ovog projekta na pomoći u realizaciji ovog istraživanja. Posebne zahvale bih uputila kolegi Dariju Karlu koji je zaslužan za tehničku pozadinu i ostvaraj ovog rada. Hvala na dostupnosti i strpljenju te prije svega na inicijativi za povezivanje naših na prvi pogled vrlo različitih područja.

Zahvalu dugujem i svojoj mentorici izv. prof. dr. sc. Gordani Hržici, na silnom strpljenju, podršci i brojnim riječima ohrabrenja tijekom ovih mjeseci. Također Vam želim zahvaliti na ukazanom povjerenju s odabranom temom te što ste me konzistentno poticali na nešto više što nisam ni znala da mogu.

Veliko hvala i mojoj obitelji, posebno mom tati i baki što su zaustavili i ublažili sve doline studiranja te sa mnom slavili svaki uspjeh kao da je njihov. Hvala mojoj Luci i Klari na svemu. I veliko hvala mojim kolegicama na podršci tijekom ovih dugih pet godina.

## **Digitalna era logopedije: automatizacija mjera sintaktičke složenosti u aplikaciji MultiDis**

Andela Elinor Kožul

izv. prof. dr. sc. Gordana Hržica

Diplomski studij Logopedija, Edukacijsko – rehabilitacijski fakultet, Sveučilište u Zagrebu

### **SAŽETAK**

Digitalizacija u logopediji predstavlja iduću stepenicu u modernizaciji i unapređenju pružanja logopedskih usluga. Područje jezične procjene zasad ne ispunjava svoj potencijal u tom vidu. No, metoda analize jezičnog uzorka otvara tu priliku. Ova metoda, uz standardizirane jezične testove, sve se više ističe kao važan dio sveobuhvatne jezične procjene. Upravo za ovu metodu razvijeni su računalni programi koji omogućuju bržu i pouzdanu provedbu. Međutim, oni su složeni i nisu prilagođeni svim jezicima (većinom su razvijeni za engleski jezik). Posljednjih godina nastaje nova generacija programa temeljena na sofisticiranim jezičnim tehnologijama. Mrežna aplikacija MultiDis je primjer takvog programa koja za cilj ima omogućiti pouzdanu i preciznu analizu jezičnog uzorka na razini mikrostrukture za hrvatski jezik. Ona je još u razvoju te zahtijeva treniranje i provjeru.

Cilj je ovog rada stvoriti temelje za podučavanje modela i istražiti razinu podudarnosti dviju metoda izračuna mjere prosječne duljine komunikacijske jedinice, računalnog računanja unutar aplikacije MultiDis i ručnog izračuna. Drugi je cilj rada opisati dobivene nepodudarnosti te donijeti zaključke o mogućim razlozima istih kako bi se pospješilo daljnje treniranje modela te tako omogućilo precizniji i pouzdaniji izračun mjera.

Istraživanje je provedeno na 30 transkripata koji su preuzeti iz baze aplikacije MultiDis. Odabrani transkripti prikupljeni su na uzorku od 20 djece, izvornih govornika hrvatskog jezika (devet i deset godina). Transkripti su označeni u alatu Doccano. Označeni transkripti podignuti su u aplikaciju. Dobiveni rezultati uneseni su i uspoređeni u programu za statističku obradu podataka IBM SPSS Statistics, inačica 29. Usporedba dviju metoda testirana je Bland-Altmanovom analizom. Rezultati su pokazali da postoji visoki stupanj podudarnosti između računalnog izračuna mjere PDKJ i zlatnog standarda, to jest ručnog izračuna PDKJ. Navedeni rezultati upućuju na opravdanost uporabe nove i brže metode izračuna jezičnih mjera uz pomoć aplikacije MultiDis jer je pokazala visok stupanj podudarnosti u procesu segmentacije.

Rezultati detaljnije analize primjera neslaganja pokazali su da se kao najčešći uzroci ističu pogreške u procesu transkripcije. Validacija ove mjere važan je korak u procesu automatizacije pa tako i jedan korak manje na dugom putu ka što bržem i bezbolnijem procesu analize jezičnog uzorka za hrvatske logopede.

**KLJUČNE RIJEČI:** aplikacija MultiDis, automatizacija, sintaktička složenost, treniranje modela

## **The digital era of speech-language pathology: automatization of syntactic complexity measures in the application MultiDis**

Anđela Elinor Kožul

Gordana Hržica, PhD, Associate Professor

Graduate Study of Speech and Language Pathology, Faculty of Education and Rehabilitation Sciences, University of Zagreb

### **ABSTRACT**

Digitalization of speech-language pathology represents the next step in the modernization and further development of service provision. The field of language assessment is not yet fully exploiting its potential in this respect. However, language sample analysis is increasingly being emphasized as an important component of comprehensive language assessment. Computer programs have been developed for this method, making it faster and more reliable. However, they are complex and not suitable for all languages (they have mainly been developed for English). Recently, a new generation of programs based on advanced language technologies has emerged. The web-based application MultiDis is an example of such a program, which aims to provide reliable and accurate analysis of language samples at the microstructure level for the Croatian language. It is still under development and needs to be trained and verified. The aim of this paper was to provide a basis for training the model and to investigate the degree of agreement between two methods of calculating the mean length of a communication unit, computer calculation within the MultiDis application and manual calculation. The second aim of the paper was to describe the discrepancies found and to draw conclusions about the possible reasons for them. The study was conducted on 30 transcripts from the MultiDis application database. The selected transcripts were collected from a sample of 20 native-speaking Croatian children (nine and ten years old). The transcripts were analysed using the Doccano tool. The marked transcripts were uploaded to the application. The results obtained were compared in IBM SPSS Statistics, version 29 (Bland-Altman analysis). The results show that there is a high degree of agreement between the computerised calculation of the MLCU measure and the gold standard, i.e. the manual calculation of the MLCU. The above results show that the use of a newer and faster method for calculating language measures with the MultiDis application is justified. The results of a more detailed analysis of examples of discrepancies showed that the most common causes were errors in the transcription process. The validation of this measure is a crucial step in the process of automatization and thus a step towards a faster and painless process of analysing language samples for Croatian speech-language pathologists.

**KEY WORDS:** MultiDis application, automatization, syntactic complexity, model training

# Sadržaj

1. <b>Uvod</b> .....	1
1.1. Analiza jezičnog uzorka.....	1
1.2. Izazovi analize jezičnog uzorka .....	3
2. <b>Programi za računalnu obradu jezika</b> .....	5
2.1. CHAT i CLAN.....	7
2.2. SALT .....	8
2.3. SUGAR.....	9
3. <b>Mrežne aplikacije utemeljene na jezičnim tehnologijama</b> .....	10
4. <b>Aplikacija za obradu dječjeg govornog jezika (MultiDis)</b> .....	11
4.1. Izazovi odjeljivanja govorenog niza i morfološkog označavanja .....	12
4.2. Mjere sintaktičke složenosti.....	13
5. <b>Problemska pitanja i ciljevi</b> .....	14
6. <b>Metodologija istraživanja</b> .....	15
6.1. Materijal.....	15
6.2. Sudionici .....	16
6.3. Postupak.....	17
6.3.1. Priprema transkripata .....	17
6.3.2. Označavanje komunikacijskih jedinica.....	17
6.3.3. Petlja povratne veze .....	19
6.3.4. Obrada podataka .....	19
7. <b>Rezultati</b> .....	20
8. <b>Rasprava</b> .....	24
9. <b>Zaključak</b> .....	27
10. <b>Literatura</b> .....	28



## **1. Uvod**

Digitalizacija u logopediji predstavlja iduću stepenicu u modernizaciji i unapređenju pružanja logopedskih usluga. Logopedija pokriva veliki broj različitih područja te su se neka od njih više priklonila ovom procesu od drugih (na primjer: Kompetencijska mreža ICT-AAC, uporaba računalnih programa u analizi glasa).

Područje jezične procjene zasad ne ispunjava svoj potencijal u tom vidu, no to ne mora biti trajno stanje. Uz standardizirane jezične testove, metoda analize jezičnog uzorka sve se više ističe kao važan dio sveobuhvatne jezične procjene. Ona omogućuje procjenu jezika u prirodnim uvjetima te nudi različite informacije o funkcionalnom jeziku. Upravo za ovu metodu su razvijeni računalni programi koji omogućuju bržu i pouzdanu provedbu. Međutim, oni su složeni i nisu prilagođeni svim jezicima te se brojni stručnjaci ne osjećaju kompetentno upotrebljavati ih u svom radu. Posljednjih godina nastaje nova generacija programa temeljena na jezičnim tehnologijama. Oni su primjer povezivanja brzo rastućeg područja umjetne inteligencije i logopedije. No, ti programi nisu namijenjeni konkretno analizi jezičnih uzoraka s usmjerenjem ka kliničkom radu. Također su trenirani na pisanom tekstu, a ne na govorenim uzorcima. Ali, oni ukazuju na mogućnost primjene novih sofisticiranijih jezičnih tehnologija u analizi jezičnog uzorka. Mrežna aplikacija MultiDis primjer je takvog programa. Ona je još u razvoju te zahtijeva treniranje i provjeru.

Ovaj rad ima za cilj postaviti temelje za izračun te produbiti spoznaje o preciznosti i pouzdanosti izračuna mjera sintaktičke složenosti prilikom računalne obrade unutar aplikacije MultiDis kao korak u smjeru digitalizacije ovog područja hrvatske logopedije.

### **1.1. Analiza jezičnog uzorka**

U logopedskoj procjeni rabe se različite metode, načini i izvori za prikupljanje podataka o jezičnom statusu. Logoped mora odabrati adekvatnu metodu i mjere u skladu s individualnim potrebama osobe koju procjenjuje, a ta će odluka ovisiti o različitim čimbenicima (na primjer: dob, jezični status) (ASHA, n.d.). Jedna od tih metoda je i analiza jezičnog uzorka.

Analiza jezičnog uzorka rabi se za obradu spontanog jezika elicitiranog u različitim komunikacijskim kontekstima (na primjer: u slobodnoj igri, razgovoru, pripovijedanju) te za

izračun različitih mjera s ciljem dobivanja potpunije slike jezičnog statusa uz rezultate na standardiziranim testovima (*ASHA*, n.d.).

Ova metoda pokazala se ekološki valjanom metodom za razumijevanje i procjenu poremećaja jezika (MacWhinney i Fromm, 2022). Dugi niz godina je istraživana i prepoznata kao važan dio sveobuhvatne logopedске procjene jezičnih sposobnosti. Od sedamdesetih godina prošlog stoljeća klinička i znanstvena primjena ove metode značajno su se izmijenile (Scott, 2020). Iako su devedesete godine bile značajne za uspostavu ove metode (Erin, 1990; Leadholm i Miller, 1994; Gavin i Giles, 1996; Stockman, 1996; Sanger i sur., 1997), do danas su se nastavile razvijati nove mjere i načini analize. Ova metoda također omogućuje dublju analizu morfosintaktičkih, semantičkih i pragmatičkih obilježja te se može upotrebljavati s djecom iz različitih jezičnih i kulturoloških sredina (Paul i Norbury, 2012; Košutar i sur., 2022a). Jezični uzorci omogućuju logopedu pristup važnim informacijama poput duljine i složenosti iskaza, obilježja pripovjednih sposobnosti, sposobnosti promjene perspektive te razumijevanja jezika (Prath, 2018). Također omogućuju planiranje terapijskih ciljeva te praćenje napretka u terapiji (Guo i sur., 2018; Pezold i sur., 2020).

Za analizu jezičnih uzoraka mogu se prikupljati govoreni i pisani uzorci. Postupak prikupljanja uzorka govorenog jezika dugotrajan je i složen. Podrazumijeva veći broj koraka: auditivno ili vizualno snimanje, transkribiranje, kodiranje i analiziranje (Kuvač i Palmović, 2007). Prije samog snimanja također treba donijeti odluke o cilju, načinu i vremenskom intervalu snimanja. U procesu transkripcije zapisi se dijele na manje jedinice za analizu (na primjer: na komunikacijske jedinice) (Košutar i sur., 2022a). Nakon toga se transkripti kodiraju i analiziraju ili ručno ili uz računalni program.

Ova metoda omogućuje izračun mjera produktivnosti, sintaktičke složenosti i rječničkog bogatstva. Uobičajene mjere produktivnosti koje se rabe u analizi su: broj komunikacijskih jedinica, ukupan broj riječi i broj različitih riječi (Mäkinen i sur., 2013). Mjera različitih riječi se može promatrati i kao mjera rječničkog bogatstva (Heilman i sur., 2010). Mjere rječničkog bogatstva koje se rabe u analizi su: broj pojavnica, omjer različenica i pojavnica, indeks rječničke raznolikosti D i mjera pomični prosječni omjer različenica i pojavnica (Kelić i sur., 2012; Hržica i sur., 2019).

Od mjera sintaktičke složenosti često se rabe: mjere prosječne duljine (iskaza, komunikacijske jedinice), gustoća klauze, udio zavisno/nezavisno složenih veza te različiti indeksi (na primjer: indeks produktivne sintakse).

Važno je istaknuti da ova metoda omogućava šaroliku primjenu. Primjerice, može poslužiti sveučilišnom profesoru kao materijal u poučavanju o jezičnim poremećajima, kliničaru u pokušaju stvaranja jasnije slike o postignućima djeteta s jezičnim kašnjenjem ili kao alat za evaluaciju napretka u terapiji osoba s afazijom (MacWhinney i Fromm, 2022).

Unatoč tome što neke metode i alati za procjenu imaju odlične karakteristike u vidu velikog broja kvalitetnih znanstvenih istraživanja i dobrih psihometrijskih obilježja, to nije nužno kriterij po kojem će logopedi u praksi odabrati metodu ili alat, već je važno da provedba te metode ne predstavlja preveliki izazov u kliničkom radu (Betz i sur., 2013; Pavelko i sur., 2016; Denman i sur., 2017). Stoga nije neobično da unatoč svim prednostima, ova metoda nije popularna među logopedima (Overton i Wren, 2014; Pavelko i sur., 2016; Klatte i sur., 2022).

O najistaknutijim izazovima pri provedbi, koji će pobliže prikazati razloge rjeđe uporabe te moguća rješenja, bit će više rečeno u nadolazećim odjeljcima rada.

## **1.2. Izazovi analize jezičnog uzorka**

Kao što je ranije navedeno, metoda analize jezičnog uzorka, unatoč svim prednostima, nije popularna među logopedima. Značajan dio logopeda (33%) u praksi ne provodi ovu metodu (Pavelko i sur., 2016), pri čemu su ovo podatci za SAD.

Kao najveći izazov u literaturi izdvaja se nedostatak vremena za provedbu (Hux i sur., 1993; Kemp i Klee, 1997; Pavelko i sur., 2016; Klatte i sur., 2022). Upravo ova prepreka pokazala se zasad nepremostivom te se uz nju veže i izostanak financijske potpore od nadležnih tijela za dodatno vremensko opterećenje (Klatte i sur., 2022). Uz vremensku komponentu veže se i izazov dugotrajnosti samog procesa transkripcije (MacWhinney i Fromm, 2022).

Literatura pokazuje da postoji i potreba za boljim prijenosom znanja iz istraživanja u praksu jer se pokazalo da logopedi u analizi upotrebljavaju neprikladne zadatke (Nippold i sur., 2008; Pavelko i sur., 2016) ili neproverjene protokole (Pavelko i sur., 2016).

Postoji i potreba za djelovanjem na negativne stavove logopeda o vlastitom znanju o ovoj metodi (Klatte i sur., 2022).

Ne postoje istraživanja na ovu temu za Republiku Hrvatsku, no u ovom radu bit će prikazane informacije koje prikazuju djelomičan odgovor na predstavljene izazove.

Pri Edukacijsko-rehabilitacijskom fakultetu na Sveučilištu u Zagrebu u sklopu diplomskog studija Logopedije provodi se izborni kolegij pod nazivom “Računalna obrada dječjeg jezika” koji svojim silabusom pokriva teorijsku pozadinu i praktične korake za provedbu navedene metode (Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu, 2010). Zbog tehničkih i organizacijskih razloga kolegiji ima ograničen broj sudionika te nije dio obveznog kurikulumu. Osim ovog temeljnog kolegija, studenti u sklopu različitih obveznih i izbornih kolegija spominju navedenu metodologiju („Psiholingvistika“, „Neurolingvistika“ i „Razvojni govorno-jezični poremećaji I“ u sklopu preddiplomskog studija te „Metodologija jezičnih istraživanja“ i „Neurolingvistika II“ na diplomskom studiju).

Pri Filozofskom fakultetu na Sveučilištu u Rijeci u sklopu preddiplomskog studija Logopedije na kolegiju „Psiholingvistika“ studenti se uvode u navedenu temu tijekom predavanja o metodama istraživanja jezika. Na diplomskom studiju provodi se obvezni kolegij pod nazivom “Metodologija istraživanja jezika i govora” koji svojim silabusom pokriva teorijsku pozadinu ove metode, vježbe i samostalan rad u programima za obradu jezika: CHAT, CLAN, Sketch Engine (Sveučilišni diplomski studij Logopedija, 2023). Osim navedenog studenti logopedije imaju i mogućnost upisa mikro kvalifikacije koju izvodi više sastavnica Sveučilišta u Rijeci: „Jezične tehnologije i digitalna obrada teksta“ koja pruža interdisciplinarni pristup ovoj temi.

Naravno, samo postojanje kolegija u edukaciji hrvatskih logopeda ne garantira njihove pozitivne stavove prema vlastitom znanju i kompetentnosti za provedbu ove metode. Također je važno imati na umu da nisu svi logopedi, koji su zaposleni u struci, imali priliku pohađati ove kolegije.

U razdoblju pisanja ovog rada u Republici Hrvatskoj ne postoji usvojen Zakon o logopedskoj djelatnosti te se može pretpostaviti kako ranije naveden izazov izostanaka financijske potpore od nadležnih tijela za dodatno vremensko opterećenje nije riješen za hrvatske logopede.

Navedene prepreke su već dugi niz godina istaknute u literaturi te su za njihovo rješavanje predložene brojne ideje. Jedna od tih ideja, koja se realizirala u različitim formama je uporaba računalnih programa u procesu transkripcije i analize.

## **2. Programi za računalnu obradu jezika**

Kao što je u prethodnom odjeljku navedeno najveći izazov u provedbi analize jezičnog uzorka predstavlja upravo vremensko opterećenje. Proces transkripcije i analize danas izgleda svakako drugačije od početaka provedbe ove metode, umjesto olovke i papira, najčešće se rabe složeni računalni programi. U novijem istraživanju Pezold i sur. (2020) su na dvije studije slučaja pokazali da je za cijeli proces transkripcije, kodiranja i analize zapisa od 50 iskaza u konverzacijskom diskursu (uz uporabu računalnih programa CLAN, SALT i SUGAR) bilo potrebno između trideset i četrdeset i pet minuta. Autori navode kako je to relativno kratko vrijeme za dobiti kvantitativne podatke o djetetovom svakodnevnom jezičnom funkcioniranju. Možda uporaba ovih programa može dovesti do šire primjene ove metode. No, kako bi se taj cilj ostvario, program mora posjedovati određene karakteristike.

Prvotno, mora imati određenu razinu automatizirane ili poluautomatizirane transkripcije. Zatim, mora posjedovati sustav kodova koji omogućava različite analize te omogućiti izračun različitih objektivnih mjera.

Naravno, kako bi izračun mjera bio pouzdan i precizan nužno je zadovoljiti neke preduvjete. Prvi preduvjet je razdvajanje govorenog jezika na iskaze te bi bilo poželjno da je ovaj korak automatiziran. Ako se tekst automatski razdijeli na određene sintaktičke strukture, lakše je i precizno odrediti sintaktičku složenost. Drugi preduvjet je omogućiti automatiziran proces lematizacije s visokom razinom preciznosti. Treći važan preduvjet je omogućiti automatiziran proces morfološkog označavanja s visokom razinom preciznosti. Kako bi bila jasnija potreba za njihovom automatizacijom, objasnit će se temeljni koncepti koji su u pozadini mjera jezičnog razvoja.

Hržica i Trtanj (2021) navode “Pojavnica (eng. *token*) svaka je riječ u jezičnom uzorku, to jest svaki niz znakova okružen bjelinama. Natuknica (eng. *lemma*) citatni je oblik riječi kakav se pojavljuje u rječnicima” (p. 170). Treći temeljni pojam je obličnica (eng. *inflected form*) koja se odnosi na “sve glasovno i sadržajno jedinstvene oblike riječi u tekstu pridruženog gramatičkog sadržaja, bez ponavljanja” (Jelaska, 2005, p. 139). Na primjer rečenicu “Velika koza je priskočila u pomoć manjoj kozi dok je treća koza sve to gledala” čini 15 pojavnica, 12 natuknica (“velik”, “koza”, “biti”, “priskočiti”, “u”, “pomoć”, “manji”, “dok”, “treći”, “sve”, “to”, “gledati) te 12 obličnica (“koza” se ponavlja dva puta i u oba slučaja je u nominativu jednine, treće lice jednine prezenta pomoćnog glagola „biti“ “je” se također ponavlja dva puta).

Određivanje obličnica i natuknica može se provesti ili ručno ili s pomoću programa, no kako je cilj postići češću uporabu metode analize jezičnog uzorka smanjivanjem vremenskog opterećenja za provedbu, važno se usmjeriti na automatizaciju procesa lematizacije to jest određivanja natuknica. Automatizirani proces lematizacije omogućio bi precizniji izračun mjera rječničkog bogatstva. Također kako bi računalo ispravno podijelilo govoreni niz na sintaktičke cjeline, važno je da ispravno odredi temeljni oblik te riječi (na primjer: da u rečenici iz prošlog odjeljka „je“ odredi kao glagol, a ne zamjenicu). Dakle, veća preciznost u procesu lematizacije pripomaže i omogućuje i precizniju segmentaciju na sintaktičke cjeline.

Treći važan preduvjet je omogućiti automatiziran proces morfološkog označavanja s visokom razinom preciznosti. Istovremeno, uz proces lematizacije važno je osigurati i visoku preciznost pri određivanju morfološkog značenja. Morfologija nužno za sobom povlači i sintaksu na primjer: imenica u akuzativu unutar rečenice vrlo vjerojatno ima službu objekta te samim time njena morfološka označenost implicira njenu sintaktičku ulogu. Računalo će lakše uočiti i predvidjeti sintaktičke veze te samim time lakše odvojiti govoreni niz na ispravne sintaktičke cjeline. Ovaj proces je također značajniji u morfološki bogatim jezicima, poput hrvatskog, dok je manje važan u jezicima poput engleskog jezika.

U nastavku rada predstaviti će se nekoliko programa koji su najčešće u uporabi te odgovoriti na pitanje koliko oni posjeduju potrebnih karakteristika i preduvjeta za hrvatske logopede.

## 2.1. CHAT i CLAN

U sklopu TalkBank projekta razvijena su tri različita, ali međusobno integrirana alata (MacWhinney, 2000).

Prvi alat je CHAT (Codes for the Human Analysis of Transcripts, MacWhinney, 2000) koji predstavlja niz pravila za transkripciju i kodiranje. Transkripti koji poštuju pravila i kodiranje CHAT-a mogu se dalje obrađivati u programu CLAN.

Drugi alat CLAN (Computerized Language Analysis, MacWhinney, 2000) je paket programa za računalnu analizu jezičnog uzorka, koji se sastoji od različitih programa. CLAN omogućava računalni izračun jezičnih mjera na četrdeset i devet jezika (Pezold i sur., 2020).

Treći alat je TalkBank, baza podataka koju čine jezični korpusi govorenog jezika (MacWhinney, 2007). Prilikom korištenja CLAN-a moguće je dobivene rezultate usporediti s postojećim podacima unutar postojeće baze (djeca do 6 godina) u samom programu, no to je moguće samo za engleski jezik. Podaci iz Talkbank baze korišteni su u preko devet tisuća publikacija, također alati CHAT i CLAN dio su kurikuluma preko sto fakulteta te se ranije spomenute metode rabe u dvadeset i četiri zemlje (Liu i sur., 2023).

Ovaj paket programa rabi se i u velikom broju domaćih jezičnih istraživanja (Kelić i sur., 2012; Olujčić i Matić, 2017; Trtanj i Kuvač Kraljević, 2017; Matić i sur., 2018; Kuvač Kraljević i sur., 2020; Leko Krhen i sur., 2020; Bedeković i sur., 2021; Košutar i sur., 2022b, Leko Krhen i sur., 2023).

Jedna od prednosti ovog programa je da uz pomoć starijih jezičnih tehnologija može samostalno morfološki kodirati transkripte na engleskom jeziku s devedeset i četiri postotnom točnošću (Bernstein Ratner i MacWhinney, 2016). Za druge jezike, ako je označavanje i dostupno, preciznost je značajno niža.

Ova baza podataka i paket programa, osim što su omogućili povezivanje velikog broja istraživača u području jezika te značajno olakšali međujezična istraživanja, važan su primjer automatizacije procesa analize jezičnog uzorka. Autori kontinuirano razvijaju ovu bazu i programe te sve više umanjuju barijere za provedbu ove metode. No, Liu i sur. (2023) navode kako se preporučuje da se za automatiziranu transkripciju i analizu u CLAN-u s trenutnim ograničenjima upotrijebi razgovor dvaju odraslih govornika engleskog jezika bez artikulacijskih i fonoloških teškoća.

Iz navedenih preporuka jasno se može zaključiti da većinski automatiziran proces, koji omogućuje brzo provođenje analize jezičnog uzorka, nije pogodan za hrvatske logopede. Iz svega navedenog može se zaključiti da CLAN ne zadovoljava ranije navedene preduvjete za hrvatske logopede (poluautomatiziranu/automatiziranu razinu transkripcije, automatizirani proces lematizacije i morfološkog označavanja).

Ovaj paket alata dostupan je u otvorenom pristupu te nudi veliki broj resursa za veće jezike, poput engleskog jezika. No, manji jezici i manje istraživani jezici nažalost nemaju jednake dobrobiti od njihove uporabe. Stoga, je važno usmjeriti se na razvoj programa koji bi ponudili potrebne karakteristike i hrvatskim logopedima.

Unatoč značajnoj razini automatizacije procesa, Klatte i sur. (2022) tijekom su provedbe edukacije nizozemskih logopeda naišli na kritike pri uporabi računalnog programa CLAN. Logopedi su naveli kako im je rad u CLAN-u bio preizazovan te nedovoljno intuitivan. Nastavno na to odustali su od uporabe te su se prebacili na ručni izračun.

Ovo istraživanje nas upućuje da u razvoju programa za računalnu obradu jezika nisu važne isključivo napredne značajke te brzina procesa, već je važno voditi računa da ciljana skupina korisnika pristane rabiti alat. Važno je i da im sučelje programa bude intuitivno kako bi se smanjio otpor prema početnoj uporabi.

## **2.2. SALT**

SALT (Systematic Analysis of Language Transcripts) je skup računalnih programa za analizu i objašnjenje jezičnih uzoraka koji su razvili John F. Miller i Robin S. Chapman 1982. godine (Kuvač i Palmović, 2007). Oni omogućuju morfološku, pragmatičku i semantičku analizu jezika (Weston i sur., 1989). Osim analize djetetove pragmatike, omogućavaju i analizu roditeljske (Kuvač i Palmović, 2007). SALT je komercijalan program te nudi mogućnost plaćanja transkripcije na dva jezika (španjolski i engleski), koju provode njihovi zaposlenici.

Brojna strana istraživanja navode ovaj program u svojoj metodologiji (primjerice, Nippold i sur., 2005; Justice i sur., 2006; Nippold i sur., 2008; Heilman i sur., 2010; Terry i sur., 2013).



U usporedbi s druga dva programa (CLAN i SUGAR), SALT ima manji broj jezičnih uzoraka za usporedbu u svojoj bazi za predškolsku dob, no nudi i uzorke za stariju djecu (Pezold i sur., 2020). SALT nudi bolju podršku za neke jezike dok je CLAN orijentiran na razvoj alata za različite jezike. Također SALT je komercijalan program dok je CLAN dostupan u otvorenom pristupu.

Unatoč navedenim karakteristikama, ovaj program također nije rješenje za hrvatske logopede jer ne zadovoljava ranije navedene preduvjete (poluautomatiziranu/automatiziranu razinu transkripcije, automatizirani proces lematizacije i morfološkog označavanja). Kao i CLAN, nudi značajne pogodnosti i resurse za velike jezike, u ovom slučaju engleski i španjolski jezik, no za manje jezike, kao što je hrvatski, ne nudi mnogo.

### **2.3. SUGAR**

SUGAR (Sampling Utterances and Grammatical Analysis Revised) je alat koji je razvijen kao odgovor na ranije navedenu anketu (vidjeti: Pavelko i sur., 2016) provedenu na velikom broju američkih logopeda (*SUGAR Language*, n.d.). Razvili su ga Robert E. Owens i Stacey L. Pavelko. Autori navode kako se djetetov uzorak od pedeset iskaza može prikupiti, transkribirati i analizirati u svega dvadeset minuta. Pezold i sur. (2020) u svom istraživanju potvrđuju da u usporedbi s ranije navedenim programima, SUGAR omogućuje najbržu transkripciju i kodiranje. Dio brzine postiže se izostavljanjem ubacivanja, ponavljanja i ispravljanja u zapisu (Pezold i sur., 2020). Također se izostavljaju i komunikacijske izmjene drugih osoba tijekom konverzacije s djetetom (Guo i sur., 2018). SUGAR omogućava izračun četiriju mjera: prosječna duljina iskaza, ukupni broj riječi, ukupan broj klauza po rečenici i ukupan broj riječi po rečenici (Westby, 2021). SUGAR za razliku od drugih navedenih programa, nije samostalan softver, već je to skup protokola za uporabu u drugim softverima za obradu riječi (na primjer: Word).

Autori su također proveli više istraživanja rabeći ovaj alat (Pavelko i Owens, 2017; Pavelko i Owens, 2019; Pavelko i sur., 2020).

Guo i sur. (2018) navode dvije ključne kritike SUGAR-a: smanjen opseg analize (nudi samo četiri mjere) i ovisnost o ručnoj segmentaciji (na iskaze ili morfeme dok CLAN nudi računalnu segmentaciju na morfeme za engleski jezik, ali ne i na iskaze).

Unatoč prednostima u kliničkoj praksi koje nudi ovaj program – jednostavna i brza provedba, ovaj program također nije rješenje za hrvatske logopede. Razlozi tomu su: smanjen opseg same analize zbog malog broja mjera, izostanak poluautomatizirane/automatizirane razine transkripcije, automatiziranog procesa lematizacije i morfološkog označavanja. Također ovaj program nema podršku jezičnih tehnologija koje nude precizniju i pouzdaniju analizu. Zaključno, ovaj program unatoč navedenim prednostima, ne zadovoljava ranije navedene potrebne preduvjete.

### **3. Mrežne aplikacije utemeljene na jezičnim tehnologijama**

Osim ranije navedenih računalnih programa, u posljednje vrijeme došlo je do razvoja različitih mrežnih aplikacija temeljenih na jezičnim tehnologijama. Navedene tehnologije dio su ranije spomenutog širokog područja umjetne inteligencije. S obzirom na primjenu velikog broja različitih alata razvijenih unutar ovog područja, nije neočekivano da se njihova primjena pronalazi i u logopedskoj domeni. U različitim stranim istraživanjima navode se različiti alati u dijagnostici poremećaja logopedске domene i šire: dizatrije (Tao i sur., 2016), afazije (Le i sur., 2016), dječje govorne apraksije (Hosom i sur., 2004), Parkinsonove bolesti (Vásquez-Correa i sur., 2019) te veliki broj istraživanja vezan uz demencije i prepoznavanje govornih obrazaca kao ranog znaka Alzheimerera (Martinez de Lizarduy i sur., 2017; Balagopalan i sur., 2021; Guo i sur., 2021; Gagliardi, 2023). Također su primjenu pronašli i kod procjene kvalitete glasa laringektomiranih osoba (Clapham i sur., 2016). Osim u procjeni govornih karakteristika alati iz područja umjetne inteligencije svoju su ulogu pronašli i u jezičnoj procjeni i različitim drugim jezičnim aktivnostima. Neke od njih spomenute su ranije u radu (na primjer: uporaba jezičnih tehnologija za postizanje visoke razine automatiziranosti procesa analize jezičnog uzorka u CLAN-u za engleski jezik), a o novije razvijenim mrežnim aplikacijama temeljenim na jezičnim tehnologijama (na primjer: Text Inspector, Coh-Matrix) bit će više riječi u nastavku rada.

Ove bi aplikacije mogle dati odgovor na ranije predstavljene izazove logopeda. Iako obje navedene aplikacije služe za analizu pisanog teksta, njihove mogućnosti, svrha i primjena su različite. Text Inspector razvijen je kao pomoć u pripremi pisanih materijala i omogućava izračun mjera: broj rečenica, broj riječi, broj različitih riječi, omjer različenica i pojavnica, prosječna duljina rečenice (*Text Inspector*, n.d.). Ova mrežna aplikacija se plaća, ali ima i besplatnu verziju, no ona je značajno ograničena (za vrlo male količine teksta). Coh-Matrix je sofisticiraniji alat koji nudi

dublju analizu teksta, to jest analizu na razini kohezije, različite jezične mjere i informacije o razini čitljivosti teksta (Graesser i sur., 2004; McNamara i sur., 2014). Međutim, iako nude automatsku analizu (morfološku i sintaktičku) te izračun mjera za jezike na kojima su razvijene, ni ove aplikacije nisu rješenje za hrvatske logopede jer nisu namijenjene konkretno analizi jezičnih uzoraka s usmjerenjem ka kliničkom radu i trenirane su na pisanom tekstu.

#### **4. Aplikacija za obradu dječjeg govornog jezika (MultiDis)**

MultiDis je mrežna aplikacija namijenjena analizi jezičnog uzorka na razini mikrostrukture za hrvatski jezik (Košutar i sur., 2022a). Temelji se na jezičnim tehnologijama za hrvatski jezik. Jezična tehnologija u podlozi ove aplikacije je transformatorski model umjetne inteligencije BERTić (Ljubešić i Lauc, 2021). On je predtreniran na 8 milijardi pojava s hrvatskih, srpskih, bosanskih i crnogorskih mrežnih stranica. Aplikacija MultiDis počela se razvijati u sklopu projekta *Višerazinski pristup govorenom diskursu u jezičnom razvoju* (MultiDis, n.d.). Aplikacija se u procesu razvoja, trenira na postojećim pripremljenim transkriptima dječjeg govorenog jezika kako bi se prilagodila i pripremila za analizu novih transkripata hrvatskog dječjeg jezika. Izlaganje većem broju transkripata u fazi predtreniranja modela omogućava modelu višu razinu preciznosti i pouzdanosti pri izračunu jezičnih mjera.

Važan cilj ove aplikacije je omogućiti analizu govorenog jezika. Kako bi se on ostvario potrebno je provesti postupke validacije jer jezične tehnologije trenirane na pisanom jeziku (uključujući tu i transformatorski model BERTić) nisu jednako uspješne s govorenim jezikom.

Osnovni cilj ove aplikacije je korisnicima omogućiti automatski izračun mjera jezične produktivnosti (broj komunikacijskih jedinica, ukupni broj riječi, broj različitih riječi), leksičkog bogatstva (omjer različenica i pojava, indeks leksičke različitosti D, prosječan pomični omjer različenica i pojava), sintaktičke složenosti (prosječna duljina komunikacijske jedinice, prosječna duljina klauze i gustoća klauza) te kohezije diskursa (omjer konektora, omjer različitih konektora) govorenog i pisanog jezika. Dakle, unutar aplikacije MultiDis postoje različiti moduli, no u ovom radu fokus će biti stavljen na mjere sintaktičke složenosti.

#### **4.1. Izazovi odjeljivanja govorenog niza i morfološkog označavanja**

Preduvjet izračuna mjera sintaktičke složenosti, koje su fokus ovog rada, prikladno je odjeljivanje teksta na sintaktičke strukture, no automatizirano odjeljivanje teže je provesti u govorenom jeziku nego što je slučaj u pisanom. Govorenje i pisanje se temelje na različitim zakonitostima te predstavljaju dva modaliteta jezične proizvodnje (Hržica i sur., 2021). Prilikom pisanja osoba/dijete ima više vremena posložiti i isplanirati svoje ideje. Također postoji i veća mogućnost revidiranja. Govorenje, za razliku od toga, karakteriziraju česta nizanja te granice između riječi/sintaktičkih cjelina nisu jasne. Dakle, pisanje i govorenje predstavljaju inherentno drugačiji prijenos misli.

Neke od specifičnosti govorenog jezika koje Hržica i sur. (2021) navode su da sintaksu karakterizira necjelovitost i česta uporaba jednostavnijih struktura. S obzirom na sve navedeno, odjeljivanje govornog niza na sintaktičke cjeline predstavlja dodatni izazov u odnosu na pisani jezik. Automatizirano odjeljivanje na sintaktičke cjeline u govoru se ne može voditi interpunkcijskim znakovima kao što je to slučaj u pisanom tekstu. Iako je moguće donekle se voditi pauzama ili intonacijom, to nije najpouzdaniji pristup zbog visoke razine subjektivnosti (vidjeti: Stockman, 2010). Stoga, u literaturi se predlaže da temeljna sintaktička jedinica govorenog jezika bude klauza (vidjeti: Biber, 1995).

Osim specifičnosti sintakse, Hržica i sur. (2021) navode i druge karakteristike govorenog jezika: spontanost i nepripremljenost izvedbe koja rezultira oklijevanjima, prekidima i ponavljanjima, određeni dio značenja nose i izvanjezični elementi (na primjer: intonacija, mimika, geste) te na nadrečeničnoj razini značenje nose deikse i diskursne oznake. Navedene i ostale specifičnosti govorenog jezika mogu predstavljati izazove: na primjer, morfološko označavanje može biti teže jer u govorenom jeziku ima dijalektizama, ili u slučaju dječjeg jezika, neologizama. Kao što je ranije navedeno, visoko precizno automatizirano morfološko označavanje omogućava i olakšava određivanje sintaktičkih uloga pa tako i odjeljivanje govorenog niza na sintaktičke cjeline.

## 4.2. Mjere sintaktičke složenosti

Sintaktička složenost može se odrediti mjerama temeljenim na duljini jedinice govorenog jezika i mjerama temeljenima na vrsti sintaktičkih veza. Preduvjet je preciznog izračuna mjera sintaktičke složenosti ispravno odjeljivanje govorenog niza na sintaktičke cjeline. Kako bi računalo treniranjem naučilo ispravno odjeljivati govoreni niz na klauze, važno ga je izložiti većem broju transkripata odijeljenih na taj način. Također važno mu je omogućiti usvajanje pravilnosti pri procjeni vrsta veza između klauza.

Klauza (eng. *clause*) predstavlja temeljni sintaktički element govorenog jezika, a čine ju barem predikat i subjekt, koji može biti neizrečen. Klauze mogu biti nepovezane ili povezane na različite načine. Mogu biti povezane nezavisnom ili zavisnom vezom. Zavisna veza među klauzama smatra se višom razinom sintaktičke složenosti (Loban, 1976). Dakle, kao što je ranije navedeno najpouzdaniji pristup, koji najmanje podliježe subjektivnosti jest podjela govorenog niza na klauze.

C-jedinica (eng. *C-unit, communication unit*) temelji se na klauzama te predstavlja sintaktičku cjelinu koja se učestalo rabi u istraživanjima dječjeg jezika i istraživanjima konverzacija.

Može se sastojati od jedne klauze. Na primjer: “Lucija je kupila mašnu.” Također može imati i složeniju sintaktičku strukturu. Na primjer: rečenica “Lucija je kupila mašnu jer je bila tužna” bi činila jednu C-jedinicu dok bi rečenica “Lucija je kupila mašnu, a Klara je otišla u šetnju” predstavljala dvije C-jedinice (1.“Lucija je kupila mašnu” i 2.“Klara je otišla u šetnju”). Razlog tomu je da kada je klauza nezavisno povezana s drugom smatra se novom C-jedinicom.

Prosječna duljina komunikacijske jedinice (PDKJ, eng. *mean length of communication unit*) jedna je od mjera sintaktičke složenosti u jezičnoj proizvodnji. Vrijednosti ove mjere dobivaju se stavljanjem u omjer ukupnog broja riječi i ukupnog broja C-jedinica. Što je broj riječi unutar C-jedinice veći, to će i vrijednost ove mjere biti veća.

Ova je mjera odabrana jer njen porast odražava porast sintaktičke složenosti, to jest ona govori da osoba/dijete upotrebljava veći broj dopuna (na primjer: više objekata, priložnih oznaka, apozicija ili atributa), proizvodi veći broj zavisnosloženih veza među klauzama ili oboje (Bedeković i sur., 2021).

Računalo značajno može olakšati izračun mjere PDKJ, ali uz preduvjet preciznog odjeljivanja na klauze te određivanja vrste veze. Međutim, računalo može imati različite stupnjeve preciznosti u ovim zadacima. Zato se u razvoju jezičnih tehnologija, koje su u podlozi aplikacije MultiDis, primjenjuje predtreniranje i validacija kao korak razvoja, to jest, prati se koliko se u označavanju podudaraju aplikacija i stručnjak.

## **5. Problemska pitanja i ciljevi**

Iz ranije navedene literature ističu se jasne prepreke u provedbi analize jezičnog uzorka. Unatoč ekološkoj valjanosti i značajnoj informativnoj vrijednosti ove metode, vremensko opterećenje koje zahtijeva, umanjuje njenu popularnost među logopedima. Kako bi se doskočilo ovom i ostalim izazovima razvijeni su brojni programi koji omogućuju automatizaciju ovog procesa: od transkribiranja i kodiranja pa do izračuna mjera. Jedan od takvih programa, koji se razvija za hrvatski jezik, mrežna je aplikacija MultiDis. Ova aplikacija uči iz jezičnih materijala te u tom procesu može biti manje ili više precizna te je stoga potrebna intervencija čovjeka kako bi se utvrdila njena preciznost i razina podudarnosti prilikom segmentacije na C-jedinice. Iz svega navedenog proizlaze ciljevi i istraživačka pitanja ovog rada.

Cilj je ovog rada stvoriti temelje za podučavanje modela i istražiti razinu podudarnosti dviju metoda izračuna mjere prosječne duljine komunikacijske jedinice, računalnog računanja unutar aplikacije MultiDis i ručnog izračuna kao koraka u smjeru automatizacije ove mjere. Pretpostavlja se da neće postojati značajne razlike u izračunu mjere prosječne duljine komunikacijske jedinice između računalnog i ručnog računanja.

Drugi je cilj rada opisati dobivene nepodudarnosti te donijeti zaključke o mogućim razlozima istih.

Iz navedenih ciljeva slijede ova istraživačka pitanja:

1. Kolika je podudarnost u izračunu mjere prosječne duljine komunikacijske jedinice prilikom računalnog računanja unutar aplikacije MultiDis i ručnog izračuna?
2. Koje i kakve su razlike u izračunu mjere prosječne duljine komunikacijske jedinice prilikom računalnog računanja unutar aplikacije MultiDis i ručnog izračuna?

Validacija ove mjere važan je korak u procesu automatizacije i daljnjem unapređenju mrežne aplikacije MultiDis pa tako i jedan korak manje na dugom putu ka što bržem i bezbolnijem procesu analize jezičnog uzorka za hrvatske logopede. Također je uočen nedostatak ovakvih radova u hrvatskoj literaturi te zbog navedenih razloga je validacija ove mjere izabrana kao cilj ovog rada.

## **6. Metodologija istraživanja**

### **6.1. Materijal**

Jezični uzorci sudionika preuzeti su iz baze unutar aplikacije MultiDis. Pohranjeni transkripti u navedenoj bazi služe za treniranje aplikacije te su prikupljeni u sklopu projekta MultiDis. Uzorkovanje je provedeno snimanjem zvučnih zapisa pripovijedanja i prepričavanja prikupljenih pomoću instrumenta Multilingual Assessment Instrument for Narratives (MAIN, Gagarina i sur., 2012) prilagođenog hrvatskom jeziku Višejezični instrument za ispitivanje pripovijedanja (Hržica i Kuvač Kraljević, 2012). On se sastoji od četiri usporedne priče koje su jednake kognitivne i lingvističke složenosti. Namijenjen je za procjenu djece kronološke dobi od tri do deset godina. Za potrebe ovog istraživanja odabrano je 30 transkripata.

## 6.2. Sudionici

Odabrani transkripti prikupljeni su na uzorku od 20 djece, izvornih govornika hrvatskog jezika, od kojih je 6 ženskih, a 14 muških, prosječne dobi  $M = 9,91$  godina ( $SD = 0,35$ ;  $Min = 9,09$ ,  $Max = 10,11$ ). Odabrana je starija školska dob (najveća što dopušta instrument) jer u toj dobi možemo očekivati višu razinu sintaktičke složenosti u odnosu na predškolsku dob. Frizelle i sur. (2018) su potvrdili navedeni porast te navode kako su djeca u dobi od 7 do 12 godina pokazala gotovo jednaku razinu sintaktičke složenosti što ide u prilog odabiru ove dobi za istraživanje. Zvučni zapisi su transkribirani i kodirani prema sustavu CHAT-a (MacWhinney, 2000), dok je dio transkripata kodiran prema kodovima aplikacije MultiDis. Pri odabiru transkripata, naglasak je bio na postizanju što veće raznolikosti kako bi se procijenila učinkovitost aplikacije u različitim uvjetima. Osnovne informacije o transkriptima prikazane su u Tablici 1.

**Tablica 1.** Osnovne informacije o transkriptima

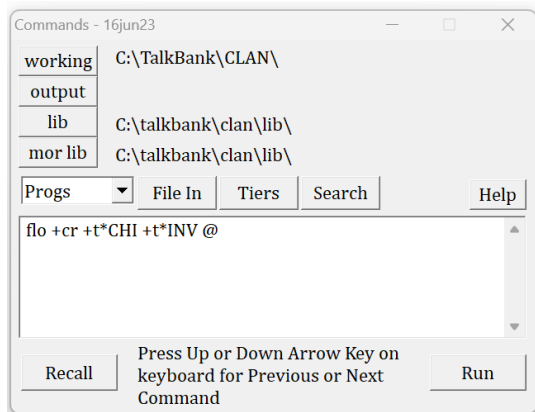
	Priča				Vrsta zadatka		Kodovi	
	Mačka	Pas	Ptice	Koze	Pripovijedanje	Prepričavanje	Bez	Multidis kodovi
Broj transkripata	10	4	7	9	16	14	17	13
Ukupno			30			30		30



## 6.3. Postupak

### 6.3.1. Priprema transkripata

Dio transkripata bio je kodiran kodovima za aplikaciju MultiDis, no dio njih, koji su bili označeni kodovima za sustav CHAT (MacWhinney, 2000), trebalo je pripremiti za daljnje označavanje. To je provedeno u programu CLAN (MacWhinney, 2000), naredbom flo uz kod +cr +t\* (Slika 1). Ova naredba uz dodatak +cr iz transkripta uklanja sve kodove za sustav CHAT.

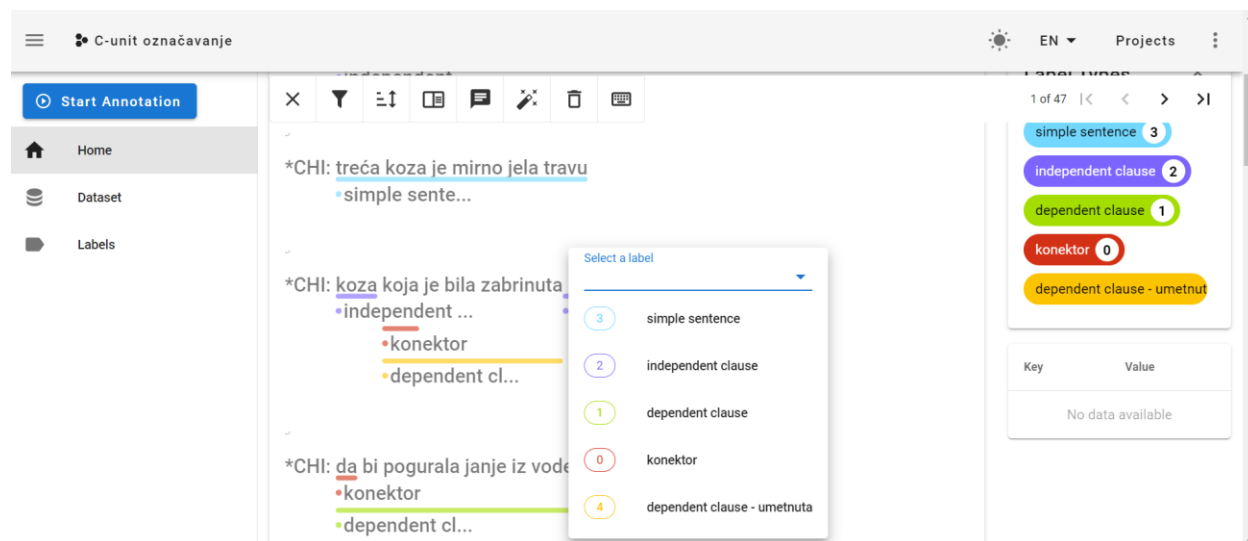


Slika 1. Prozor s naredbom u CLAN-u

### 6.3.2. Označavanje komunikacijskih jedinica

Transkripti su označeni u alatu Doccano (Nakayama i sur., 2018) putem mrežnog preglednika. Doccano je alat otvorenog koda za označavanje teksta (Doccano, n.d.). Nudi mogućnost označavanja u svrhu klasifikacije teksta (eng. *text classification*) i označavanje slijeda (eng. *sequence labeling*) te usluge označavanja slijeda za slijed (eng. *sequence to sequence task*). Za svaku kategoriju, moguće je postaviti različite boje prikaza i različite prečace što omogućuje brzo i jednostavno označavanje. Na Slici 2 prikazano je sučelje Doccana. Za ovaj rad upotrijebljeno je 5 kategorija. Kategorije su prikazane u Tablici 2. U ovom radu navedeni alat se koristio za zadatak označavanja slijeda.

Ovaj alat rabi se i u stranim istraživanjima u području umjetne inteligencije (Moussa i Mourhir, 2023; Li i sur., 2024).



**Slika 2.** Sučelje za označavanje u Doccanu. Kad se odabere granica teksta prikaže se popis kategorija te se odabire odgovarajuća kategorija s popisa.

**Tablica 2.** Kategorije označavanja u Doccanu. Dijelovi primjera na koje se odnosi kategorija su podebljani i napisani u kurzivu.

Naziv kategorije	Primjer
Jednostavna rečenica	Sudionik: <b><i>mačka je lovila leptira.</i></b>
Osnovna klauza	Sudionik: <b><i>koza je ugledala malu kozu</i></b> kako se utapa.
Zavisna klauza	Sudionik: <b><i>kad se mama vratila</i></b> , mačak se počeo penjati.
Zavisna klauza – umetnuta	Sudionik: koza, <b><i>koja je bila zabrinuta</i></b> , je ušla u vodu.
Konektor	Sudionik: koza je pomogla djetetu <b><i>da</i></b> izađe iz vode.

### 6.3.3. Petlja povratne veze

Nakon početnog označavanja u alatu Doccano korištena je petlja povratne veze (eng. *feedback loop*). Autorica ovog rada označila je sve zadane transkripte, zatim je dodatni procjenjivač provjerio te oznake te ukazao na moguće pogreške ili uzroke budućih nepodudarnosti. Nakon toga autorica je još jednom u Excel tablici provjerila sve transkripte te ispravila istaknute primjere i analizirala razloge nepodudarnosti. Detaljnija analiza i objašnjenja bit će prikazana u raspravi rada. Nakon toga je dodatni procjenjivač završno pregledao i još jednom ispravio izrađenu tablicu.

### 6.3.4. Obrada podataka

Označeni transkripti podignuti su u aplikaciju MultiDis. Dobiveni rezultati su uneseni i uspoređeni u programu za statističku obradu podataka „Statistical Package for Social Sciences“ (IBM SPSS Statistics) inačica 29. Usporedba dviju metoda testirana je Bland-Altmanovom analizom kako bi se odgovorilo na pitanje je li razlika između ovih metoda statistički značajna. Ova analiza je slikovna statistička metoda s pomoću koje zaključujemo i postoji li sistematska pristranost (eng. *systematic bias*) između dviju metoda, to jest daje li neka od metoda konzistentno manje ili veće vrijednosti. Rabi se za procjenu nove metode mjerenja s referentnom metodom kako bi se opravdala uporaba nove metode (Bland i Altman, 1986). U slučaju ovog rada referentnu metodu predstavlja ručni izračun jer se ljudski faktor smatra najpouzdanijim za ovakav tip mjera te kao takav čini zlatni standard. Nova metoda zasniva se na računalnom izračunu unutar aplikacije MultiDis.

## 7. Rezultati

U Tablici 3 prikazani su rezultati dobiveni računalnim i ručnim izračunom mjere prosječne duljine komunikacijske jedinice, njihova razlika te aritmetička sredina. Kako bi se saznala razina podudarnosti između ovih dviju metoda (ručni i računalni izračun PDKJ-a), za svaki transkript izračunato je koliko se razlikuju dobivene vrijednosti PDKJ-a te je izračunata i kolika im je srednja vrijednost (aritmetička sredina) obaju mjera. Raspon rezultata razlika je 0 – 2,69; no većina razlika je manja od 1. Postoji i nekoliko transkripata s većim razlikama (na primjer: 1,46; 2,02; 2,69).

Na temelju tih podataka je provedena Bland-Altman analiza.

**Tablica 3.** Rezultati izračuna prosječne duljine komunikacijske jedinice (PDKJ), razlika dvaju mjerenja te srednja vrijednost između mjerenja.

PDKJ (ručni izračun)	PDKJ (MultiDis)	Razlika	Aritmetička sredina
9,08	8,43	,65	8,75
7,77	8,42	–,65	8,09
10,27	8,81	1,46	9,54
8,94	8,56	,38	8,75
5,18	5,18	,00	5,18
8,33	7,89	,44	8,11
7,06	6,44	,62	6,75
7,15	6,70	,45	6,93
7,62	6,50	1,12	7,06
7,00	7,64	–,64	7,32
5,64	5,04	,60	5,34
11,05	11,88	–,83	11,47
4,33	5,33	–1,00	4,83
4,75	5,81	–1,06	5,28
4,40	4,10	,30	4,25
8,96	8,88	,08	8,92
7,71	8,73	–1,02	8,22
5,25	4,88	,37	5,06
7,00	7,00	,00	7,00
5,62	6,64	–1,02	6,13
5,50	5,90	–,40	5,70
6,60	9,29	–2,69	7,94

8,52	7,58	,94	8,05
7,06	6,38	,68	6,72
6,89	7,23	-,34	7,06
9,23	8,15	1,08	8,69
8,64	7,55	1,09	8,10
7,29	7,89	-,60	7,59
8,69	8,80	-,11	8,75
6,39	8,41	-2,02	7,40

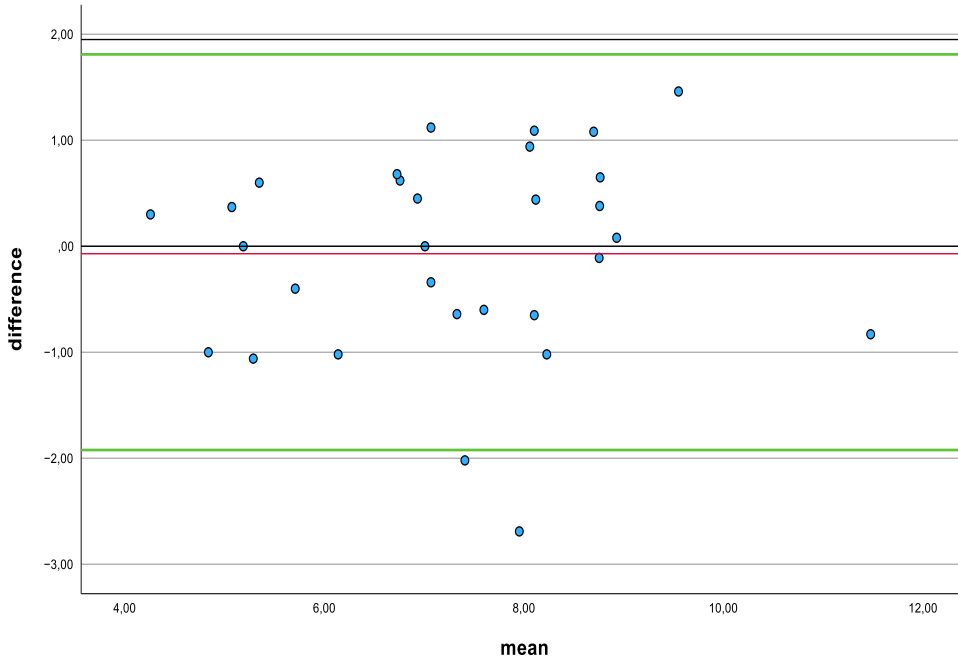
Na Slici 3 prikazani su rezultati Bland-Altman analize. Na x osi grafa prikazane su razlike izračuna PDKJ-a ručnom i računalnom metodom (vidjeti: Tablica 3, stupac *Razlika*). Na y osi grafa prikazane su aritmetičke sredine izračuna PDKJ-a za svaki transkript između ručne i računalne metode (vidjeti: Tablica 3, stupac *Aritmetička sredina*). Svaka od plavih točkica predstavlja jedan transkript s određenom vrijednosti razlike i aritmetičke sredine.

Zatim je izračunata aritmetička sredina svih vrijednosti razlika iz Tablice 3 te je ona na grafu predstavljena crvenom linijom, a iznosi -0,07.

Također su izračunate i granice podudarnosti koje su predstavljene zelenom linijom. One iznose -1,95 i 1,81 te predstavljaju predefiniran raspon od  $\pm 1,96 SD$ .

Kako bi se opravdala uporaba nove metode mjerenja s referentnom metodom važno je postići visoku podudarnost (približno 95%). Također ovom metodom provjerava se i pristranost metode to jest da ne postoji sistematska tendencija da jedna od metoda daje ili veće ili manje rezultate.

Iz navedenog grafa možemo vidjeti da se većina podataka (93 %) nalazi unutar intervala  $\pm 1,96 SD$  (između zelenih linija) to jest samo 2 rezultata su izvan tog intervala. Također nije prisutna pristranost niti jedne metode. Stoga možemo zaključiti da računalni izračun mjere PDKJ unutar aplikacije MultiDis pokazuje visoki stupanj podudarnosti sa zlatnim standardom to jest ručnim izračunom PDKJ-a. U budućim izračunima cilj je postići još veći stupanj podudarnosti između ovih dviju metoda, no važno je istaknuti da i ovi rezultati (93 % rezultata se nalazi unutar intervala  $\pm 1,96 SD$ ) opravdavaju uporabu nove metode za izračun PDKJ.



**Slika 3.** Bland-Altman analiza usporedbe dviju metoda izračuna. Na x osi prikazane su razlike izračuna PDKJ-a ručnom i računalnom metodom. Na y osi prikazane su srednje vrijednosti (aritmetička sredina) izračuna PDKJ-a između ručne i računalne metode.

Kako bi se model u pozadini aplikacije MultiDis mogao još preciznije istrenirati, analizirane su i opisane dobivene nepodudarnosti. Većina nepodudarnosti bila je uzrokovana pogreškama u procesu transkripcije. Sam transkript izgleda tako da transkriptor u retke ubaci određene riječi, a model iz takvog rascjepkanog teksta prepoznaje sintaktičke cjeline. Stoga, cilj je da pri automatskoj podjeli na C-jedinice model točno odjeli transkripte na sintaktičke cjeline bez obzira na to kako su te cjeline razdijeljene u transkriptima.

Kao najčešći uzrok neslaganja kod izračuna PDKJ bila je pogrešna segmentacija na komunikacijske jedinice. Modelu u podlozi aplikacije MultiDis najveći izazov u pravilnom izračunu predstavljala je netočna raspodjela C-jedinica po redovima unutar transkripata. No, model bi trebao prepoznavati i druge sintaktičke cjeline kako bi u budućnosti mogao samostalno segmentirati tekst.

Na primjer: „*kad je lisica išla hvatati drugu malu kozu okrenule su se velika*

*i prva  
mala koza.*“

Kad bi se model susreo s ovakvim primjerima, glavnu klauzu (u ovom slučaju: *okrenule su se velika i prva mala koza*) ne bi označio kao zasebnu komunikacijsku jedinicu zbog neispravne transkripcije. Dakle, model u ovom primjeru ne bi označio dvije komunikacijske jedinice, već samo jednu moguće zbog neispravnog formatiranja teksta.

Drugi je najzastupljeniji uzrok također bio povezan s netočnom raspodjelom po redovima, no u ovim slučajevima radilo se o pogrešnoj transkripciji u vidu odvajanja zavisne klauze u zaseban red. Ukoliko je ta zavisna klauza bila dio višestruko složene rečenice, što u govorenom jeziku nije rijetkost, utoliko nije bila prepoznata. Na primjer: „*i onda je mačka došla.*

*pa je vidjela te dvije male ptice.  
kad je mama otišla.*“

U ovom primjeru bi model zavisnu klauzu „*kad je mama otišla*“ označio kao zasebnu komunikacijsku jedinicu što ona nije.

Treći najčešći uzrok predstavljale su nepodudarnosti u vidu neispravnih ili izostalih kodova u izvornom transkriptu. Na primjer: „*i dot dotle je mama spašavala svoje to dijete.*“ Kad bi se model susreo s ovakvim primjerima pogrešaka u transkripciji – „*dot*“ nije označen kao fonološki fragment te kod za ponavljanje nije zabilježen – model bi takav fragment uvrstio u izračune jezičnih mjera, a to želimo izbjeći jer on ne predstavlja zasebnu riječ. Stoga, ako kodovi za ponavljanja, oklijevanja i različite vrste pogreški nisu ispravno upisani, u ovoj fazi model sve uključuje u izračun.

Dio nepodudarnosti ne možemo objasniti ranije navedenim uzrocima te ne znamo zašto je do njih došlo. Ova pojava nije nepoznanica u području istraživanja umjetne inteligencije te se naziva fenomen crne kutije (eng. *Black-Box Phenomenon*) (vidjeti: Holm, 2019). S obzirom na složenost arhitekture modela u podlozi ove aplikacije, razlozi nekih neslaganja ostaju nam nepoznanicom. No, to ne znači da iz ranijih primjera ne možemo pokušati izvući korisne zaključke za daljnje treniranje modela.

## 8. Rasprava

Mrežna aplikacija MultiDis uči iz jezičnih materijala te u tom procesu može biti manje ili više precizna. Zbog toga je potrebna intervencija čovjeka kako bi se utvrdila njena preciznost i razina podudarnosti prilikom segmentacije na C-jedinice. Kako bi se omogućio automatizirani proces provedbe analize jezičnog uzorka, potrebno je model u podlozi ove aplikacije podučiti i provjeriti. Cilj ovog rada bio je stvoriti temelje za podučavanje modela i istražiti razinu podudarnosti dviju metoda izračuna mjere prosječne duljine komunikacijske jedinice, računalnog računanja unutar aplikacije MultiDis i ručnog izračuna kao korak u smjeru automatizacije ove mjere.

Rezultati su pokazali da postoji visoki stupanj podudarnosti između računalnog izračuna mjere PDKJ unutar aplikacije MultiDis i zlatnog standarda, to jest ručnog izračuna PDKJ. Slijedom toga, prihvaćena je pretpostavka da neće postojati značajne razlike u izračunu PDKJ između računalnog i ručnog računanja. Također, rezultati su pokazali da ne postoji sistematska pristranost niti jedne metode to jest ne postoji sistematska tendencija da jedna od metoda daje ili veće ili manje rezultate. Kako ne postoje slična strana ili domaća istraživanja s ovom temom ne možemo usporediti rezultate s drugim istraživanjima. Preduvjet izračuna PDKJ i ostalih mjera sintaktičke složenosti temeljenih na duljini jedinice adekvatno je odjeljivanje na određene sintaktičke cjeline. Navedeni rezultati ovog rada upućuju na opravdanost uporabe nove i brže metode izračuna jezičnih mjera uz pomoć aplikacije MultiDis jer je pokazala visok stupanj podudarnosti u usporedbi sa stručnjakom u procesu segmentacije.



Drugi cilj ovog rada bio je opisati dobivene nepodudarnosti te donijeti zaključke o mogućim razlozima istih kako bi se pospješilo daljnje treniranje modela te tako omogućio precizniji i pouzdaniji izračun mjera. Rezultati detaljnije analize primjera neslaganja pokazali su da se kao najčešći uzroci ističu pogreške u procesu transkripcije (netočna raspodjela po redovima unutar transkripata, odvajanja zavisne klauze u zaseban red, neispravni ili izostali kodovi). Oni upućuju na važnost pripreme transkripata prije same analize. Kad su transkripti imali manje pogrešaka, razina slaganja je bila veća. Iako je jedan od krajnjih ciljeva razvoja ove aplikacije automatizacija cijelog procesa, u ovoj fazi razvoja važno je voditi računa o ispravnosti samih transkripata te njihovoj kompatibilnosti s modelom. MacWhinney i Fromm (2022) izvještavaju kako ključni izazov pri provedbi analize jezičnog uzorka čine dugotrajnost i monotonost procesa transkripcije te je svakako jedan od daljnjih ciljeva razvoja smanjivanje potrebne količine intervencije čovjeka tijekom procesa transkripcije.

U daljnjem treniranju ovog modela vodit će se računa o vrsti transkripata te njihovoj kvaliteti i točnosti. U kasnijim fazama razvoja model će se trenirati i validirati na transkriptima koji iziskuju nižu razinu pripreme kako bi se motiviralo veći broj logopeda da upotrijebe ovu metodu procjene za potpuniju sliku o jezičnom statusu osobe.

Iako nam interpretacija mogućih razloga neslaganja između ručnog i računalnog izračuna može poslužiti za daljnje treniranje modela, važno je uzeti u obzir da su navedeni uzroci pretpostavljeni te da ne možemo sa sigurnošću tvrditi da su ispravni. Bolukbasi i sur. (2021) u svom radu na jezičnom modelu BERT, na temelju kojeg je razvijen BERTi<sup>ć</sup>, model u pozadini aplikacije MultiDis, navode kako je upravo jedan od uzroka iluzije prilikom interpretacije rezultata jezičnih modela pogreška anotatora to jest procjenjivača. Ova pojava odnosi se na tendenciju ljudi da pronalaze uzorke tamo gdje ih nema te tako objašnjavaju rezultate do kojih su jezične tehnologije došle. Stoga, važno je pažljivo interpretirati drugi dio rezultata ovog rada.

S obzirom na sve navedeno, ovo je istraživanje pripomoglo u postavljanju temelja za treniranje te razvoj specijaliziranog modela za dječji jezik za hrvatski jezik koji se može primijeniti i za govoreni i pisani jezik. Aplikacija MultiDis primjer je programa razvijenog za analizu jezičnih uzoraka s usmjerenjem ka kliničkom radu te je model u njoj podlozi treniran na govorenim uzorcima. Kao takva predstavlja revoluciju u provedbi analize jezičnih uzoraka za hrvatske logopede te pokazuje potencijal da odgovori na većinu ranije iznesenih prepreka (na primjer: veća

preciznost mjera smanjuje utrošak vremena). Zasad nepremostiva prepreka ostaje vremenski utrošak potreban za provedbu, no s daljnjim razvojem te automatizacijom mjere PDKJ i ostalih jezičnih mjera očekuje se da neće ostati trajna. Kao što je ranije u radu spomenuto, model uči iz jezičnih materijala te u tom procesu može biti manje ili više precizan. Stoga, je važno provesti daljnja istraživanja za validaciju i automatizaciju ostalih mjera, kako bi procesi lematizacije i morfološkog označavanje te segmentacije na komunikacijske jedinice bili što precizniji kako bi omogućili uporabu ovog programa u kliničkom radu.

Prilikom tumačenja rezultata važno je istaknuti ograničenja ovog rada. Transkripti su dio već postojeće baze te se dovodi u pitanje ujednačenost procesa provedbe, broja ispitivača i uvjeta provedbe. Važno je navesti da su navedeni transkripti na kojima se provodila analiza donekle unificirani u vidu iste priče, slične strukture, ujednačene razine složenosti s obzirom na dob. Također analiziralo se 30 transkripata te bi možda rezultati bili drugačiji da je bio uključen veći broj transkripata, stoga treba biti oprezan pri generalizaciji ovih rezultata. U obzir treba uzeti i da su ispitanici bili djeca urednog jezično-govorno-komunikacijskog statusa te da bi možda model imao više izazova u segmentaciji transkripata djece s teškoćama u logopedskoj domeni. Subjektivnost procjenjivača također igra ulogu te bi bilo preporučljivo u budućim istraživanjima uključiti veći broj procjenjivača te tako dobiti pouzdanije podatke.

Razvoj ove aplikacije, osim implikacija za znanstvena istraživanja, ima svoje doprinose za stručnu primjenu. Često različiti razvijeni alati ostaju dostupni i korisni isključivo za istraživanja dok u kliničkom radu bivaju zanemareni. Cilj razvoja aplikacije MultiDis upravo je primjenjivost za logopede u struci. Naravno, osim razvoja samog programa važno je djelovati na stavove logopeda o uporabi tehnologije u procjeni. Za tu svrhu postoje brojna rješenja: od edukativnih radionica u sklopu stručnog usavršavanja do promjene paradigme poučavanja na studiju Logopedije.

Validacija ove mjere važan je korak u procesu automatizacije. Postojanje programa, koji zaobilazi izazove ranije navedene u radu (nedostatak vremena za provedbu, dugotrajnost samog procesa transkripcije, izostanak financijske potpore od nadležnih tijela za dodatno vremensko opterećenje, potreba za boljim prijenosom znanja iz istraživanja u praksu te potreba za djelovanjem na

negativne stavove logopeda o vlastitom znanju o ovoj metodi) svakako bi utjecalo na popularnost ove metode među logopedima.

## 9. Zaključak

Analiza jezičnog uzorka ekološki je valjana metoda procjene jezičnog statusa te se uz procjenu rabi za planiranje terapijskih ciljeva i praćenje napretka. Omogućuje dublju analizu morfosintaktičkih, semantičkih i pragmatičkih obilježja. Također se može rabiti s djecom i odraslima iz različitih jezičnih i kulturoloških sredina. Unatoč njenim brojnim prednostima, nije popularna među logopedima. Kao najveći izazovi u literaturi ističu se: nedostatak vremena za provedbu, dugotrajnost samog procesa transkripcije, izostanak financijske potpore od nadležnih tijela za dodatno vremensko opterećenje, potreba za boljim prijenosom znanja iz istraživanja u praksu te potreba za djelovanjem na negativne stavove logopeda o vlastitom znanju o ovoj metodi. Za hrvatske logopede uz navedene izazove ističe se i znakoviti izazov jezične barijere to jest da su dosadašnji programi većinom razvijeni za engleski jezik. Mrežna aplikacija MultiDis primjer je nove generacije programa koji u svojoj podlozi imaju složene jezične tehnologije. Aplikacija za cilj ima automatizirati proces analize jezičnog uzorka za hrvatski jezik i tako omogućiti primjenu u kliničkom radu. Aplikacija je još u razvoju pa je model u njenoj pozadini potrebno trenirati te provjeravati jezične mjere u različitim točkama. Osim, mjere PDKJ čijom automatizacijom se bavio ovaj rad, važno je automatizirati i ostale mjere sintaktičke složenosti, ali i druge mjere (na primjer: rječničkog bogatstva). Kontinuiranim treniranjem modela i provjeravanjem ovih mjera dobiva se povratna informacija o preciznosti izračuna, kao i mogućim razlozi nepodudarnosti. Upravo ova saznanja daju smjernice za budući razvoj.

## 10.Literatura

1. *ASHA: Assessment tools, techniques, and data sources.* (n.d.). American Speech-Language-Hearing Association. Preuzeto 1. ožujka 2024., s <https://www.asha.org/practice-portal/clinical-topics/late-language-emergence/assessment-tools-techniques-and-data-sources/>
2. Balagopalan, A., Eyre, B., Robin, J., Rudzicz, F. i Novikova, J. (2021). Comparing Pre-trained and Feature-Based Models for Prediction of Alzheimer’s Disease Based on Speech. *Frontiers in Aging Neuroscience*, 13. <https://doi.org/10.3389/fnagi.2021.635945>
3. Bedeković, M., Hržica, G. i Kramarić, M. (2021). ANALIZA SINTAKTIČKE SLOŽENOSTI DJEČJEG PRIPOVJEDNOG DISKURSA. *FLUMINENSIA*, 33 (2), 417-443. <https://doi.org/10.31820/f.33.2.8>
4. Bernstein Ratner, N. i MacWhinney, B. (2016). Your laptop to the rescue: Using the Child Language Data Exchange System Archive and CLAN utilities to improve child language sample analysis. *Seminars in Speech and Language*, 37(02), 074–084. <https://doi.org/10.1055/s-0036-1580742>
5. Betz, S. K., Eickhoff, J. R. i Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of specific language impairment. *Language, Speech, and Hearing Services in Schools*, 44(2), 133. [https://doi.org/10.1044/0161-1461\(2012\)12-0093](https://doi.org/10.1044/0161-1461(2012)12-0093)
6. Biber, D. (1995). *Variation across speech and writing*. Cambridge University Press.
7. Bland, J. M. i Altman, D. G. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *Lancet*, 327(8476), 307–310. [https://doi.org/10.1016/s0140-6736\(86\)90837-8](https://doi.org/10.1016/s0140-6736(86)90837-8)

8. Bolukbasi, T., Pearce, A., Yuan, A., Coenen, A., Reif, E., Viégas, F. i Wattenberg, M. (2021). An interpretability illusion for BERT. *arXiv preprint arXiv:2104.07143*.
9. Clapham, R., Martens, J., Van Son, R., Hilgers, F. J. M., Van Den Brekel, M. W. M. i Middag, C. (2016). Computing scores of voice quality and speech intelligibility in tracheoesophageal speech for speech stimuli of varying lengths. *Computer Speech & Language*, 37, 1–10. <https://doi.org/10.1016/j.csl.2015.10.001>
10. Denman, D., Speyer, R., Munro, N., Pearce, W. M., Chen, Y. W. i Cordier, R. (2017). Psychometric Properties of Language Assessments for children aged 4–12 years: A Systematic review. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01515>
11. Doccano. (n.d.). *GitHub - doccano/doccano: Open source annotation tool for machine learning practitioners*. GitHub. <https://github.com/doccano/doccano>
12. Edukacijsko-rehabilitacijski Fakultet Sveučilišta u Zagrebu. (2010). Obnovljeni program diplomskog studija Logopedija. Preuzeto 14. veljače 2024., s [https://www.erf.unizg.hr/studiji/diplomski\\_studiji/logopedija](https://www.erf.unizg.hr/studiji/diplomski_studiji/logopedija)
13. Erin, J. N. (1990). Language samples from visually impaired four- and five-year-olds. *Journal of Childhood Communication Disorders*, 13(2), 181-191. <https://doi.org/10.1177/152574019001300205>
14. Frizelle, P., Thompson, P. A., McDonald, D. i Bishop, D. V. M. (2018). Growth in syntactic complexity between four years and adulthood: evidence from a narrative task. *Journal of Child Language*, 45(5), 1174–1197. <https://doi.org/10.1017/s0305000918000144>

15. Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bognár-Kolbe, U. i Walters, J. (2012). MAIN Multilingual Assessment Instrument for Narratives (Manual). Berlin, Germany: ZASPiL.
16. Gagliardi, G. (2023). Natural language processing techniques for studying language in pathological ageing: A scoping review. *International Journal of Language & Communication Disorders*, 59(1), 110–122. <https://doi.org/10.1111/1460-6984.12870>
17. Gavin, W. J. i Giles, L. (1996). Sample size effects on temporal reliability of language sample measures of preschool children. *Journal of Speech Language and Hearing Research*, 39(6), 1258–1262. <https://doi.org/10.1044/jslr.3906.1258>
18. Graesser, A. C., McNamara, D. S., Louwerse, M. M. i Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods Instruments and Computers*, 36(2), 193–202. <https://doi.org/10.3758/bf03195564>
19. Guo, L., Eisenberg, S., Bernstein Ratner, N. i MacWhinney, B. (2018). Is putting SUGAR (Sampling Utterances of Grammatical Analysis revised) into language sample analysis a good thing? A response to Pavelko and Owens (2017). *Language Speech and Hearing Services in Schools*, 49(3), 622–627. [https://doi.org/10.1044/2018\\_lshss-17-0084](https://doi.org/10.1044/2018_lshss-17-0084)
20. Guo, Y., Li, C., Roan, C., Pakhomov, S. i Cohen, T. (2021). *Crossing the “Cookie Theft” Corpus Chasm: Applying What BERT Learns From Outside Data to the ADReSS Challenge Dementia Detection Task*. *Frontiers in Computer Science*. <https://doi.org/10.3389/fcomp.2021.642517>

21. Heilmann, J., Nockerts, A. i Miller, J. F. (2010). Language sampling: Does the length of the transcript matter? *Language, Speech, and Hearing Services in Schools*, 41(4), 393.  
[https://doi.org/10.1044/0161-1461\(2009/09-0023\)](https://doi.org/10.1044/0161-1461(2009/09-0023))
22. Holm, E. A. (2019). In defense of the black box. *Science*, 364(6435), 26–27.  
doi:10.1126/science.aax016
23. Hosom, J. P., Shriberg, L. i Green, J. R. (2004). Diagnostic Assessment of Childhood Apraxia of Speech Using Automatic Speech Recognition (ASR) Methods. *Journal of Medical Speech-Language Pathology*, 12(4), 167–171.
24. Hržica, G. i Kuvač Kraljević, J. (2012). MAIN - hrvatska inačica: Višejezični instrument za ispitivanje pripovijedanja [MAIN - Croatian version: Multilingual Assessment Instrument for Narratives]. *ZAS Papers in Linguistics*, 56, 323–323.
25. Hržica, G. i Trtanj, I. (2021). Mjere rječničke raznolikosti u pričama djece predškolske i rane školske dobi. *Rasprave Instituta Za Hrvatski Jezik I Jezikoslovlje*, 47(1), 105–126.  
<https://doi.org/10.31724/rihjj.47.1.3>
26. Hržica, G., Košutar, S. i Posavec, K. (2021). Konektori i druge diskursne oznake u pisanome i spontanome govorenom jeziku. *Fluminensia*, 33(1), 25–52.  
<https://doi.org/10.31820/f.33.1.12>
27. Hržica, G., Košutar, S., i Kramarić, M. (2019). Rječnička raznolikost pisanih tekstova osoba s razvojnim jezičnim poremećajem. *Hrvatska revija za rehabilitacijska istraživanja*.
28. Hux, K., Morris-Friehe, M. i Sanger, D. (1993). Language sampling practices. *Language Speech and Hearing Services in Schools*, 24(2), 84–91. <https://doi.org/10.1044/0161-1461.2402.84>

29. Jelaska, Z. (2005). Oblici hrvatskih riječi. U Z. Jelaska (Ur.), *Hrvatski kao drugi i strani jezik* (str. 136-143). Zagreb: Hrvatska sveučilišna naklada.
30. Kelić, M., Hržica, G. i Kuvač Kraljević, J. (2012). Mjere jezičnog razvoja kao pokazatelji posebnih jezičnih teškoća. *Hrvatska revija za rehabilitacijska istraživanja*, 48(2), 23-40.  
<https://hrcak.srce.hr/87793>
31. Kemp, K. i Klee, T. (1997). Clinical language sampling practices: results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy*, 13(2), 161–176. <https://doi.org/10.1177/026565909701300204>
32. Klatte, I., Van Heugten, V., Zwitserlood, R. i Gerrits, E. (2022). Language sample analysis in Clinical practice: Speech-Language Pathologists' Barriers, Facilitators, and Needs. *Language Speech and Hearing Services in Schools*, 53(1), 1–16.  
[https://doi.org/10.1044/2021\\_lshss-21-00026](https://doi.org/10.1044/2021_lshss-21-00026)
33. Košutar, S., Karl, D., Kramarić, M. i Hržica, G. (2022a). Automatic Text Analysis in Language Assessment: Developing a MultiDis Web Application. In *Proceedings of the Conference on Language Technologies and Digital Humanities* (pp. 93-99). Ljubljana: Institute of Contemporary History.
34. Košutar, S., Kramarić, M. i Hržica, G. (2022b). The relationship between narrative microstructure and macrostructure: Differences between six- and eight-year-olds. *Psychology of Language and Communication*, 26(1), 126–153.  
<https://doi.org/10.2478/plc-2022-0007>
35. Kuvač Kraljević, J., Hržica, G. i Gorup, I. V. (2020). A Comparative Macrostructural Analysis of Narrative Discourse in Children with Typical Language Development and



- Children with Developmental Language Disorder. *Drustvena Istrazivanja*, 29(3), 453–470. <https://doi.org/10.5559/di.29.3.06>
36. Kuvač, J. i Palmović, M. (2007). *Metodologija istraživanja dječjega jezika*. Jastrebarsko: Naklada Slap.
37. Le, D., Licata, K., Persad, C. i Provost, E. M. (2016). Automatic Assessment of speech intelligibility for individuals with Aphasia. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(11), 2187–2199. <https://doi.org/10.1109/taslp.2016.2598428>
38. Leadholm, B. J. i Miller, J. F. (1994). Language sample analysis: The Wisconsin guide (Bulletin 92424). Madison, WI: Wisconsin State Dept. of Public Instruction.
39. Leko Krhen , A., Hržica, G. i Kokot, N. (2020). Sintaktičke sposobnosti djece koja mucaju. *Govor*, 37(2), 179–204. <https://doi.org/10.22210/govor.2020.37.09>
40. Leko Krhen, A., Hržica, G. i Lasan-Grevelhörster, M. (2023). Netečnosti u pripovijedanju odraslih govornika hrvatskoga jezika. *Govor*, 39(2), 119–150. <https://doi.org/10.22210/govor.2022.39.07>
41. Li, P., Zhou, G., Guo, Y., Zhang, S., Jiang, Y. i Tang, Y. (2024). EPIC: An epidemiological investigation of COVID-19 dataset for Chinese named entity recognition. *Information Processing and Management*, 61(1), 103541. <https://doi.org/10.1016/j.ipm.2023.103541>
42. Liu, H., MacWhinney, B., Fromm, D. i Lanzi, A. (2023). Automation of language sample analysis. *Journal of Speech Language and Hearing Research*, 66(7), 2421–2433. [https://doi.org/10.1044/2023\\_jslhr-22-00642](https://doi.org/10.1044/2023_jslhr-22-00642)

43. Loban, W. (1976). Language development: Kindergarten through grade twelve [Izvrješće istraživanja]. National Council of Teachers of English.
44. Ljubešić, N. i Lauc, D. (2021). *BERTić - The Transformer Language model for Bosnian, Croatian, Montenegrin and Serbian*. ACL Anthology.  
<https://www.aclweb.org/anthology/2021.bsnlp-1.5>
45. MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk (3. izd.). Mahwah, NJ: Lawrence Erlbaum Associates.
46. MacWhinney, B. (2007). The TalkBank project. In *Palgrave Macmillan UK eBooks* (pp. 163–180). [https://doi.org/10.1057/9780230223936\\_7](https://doi.org/10.1057/9780230223936_7)
47. MacWhinney, B. i Fromm, D. (2022). Language sample analysis with TalkBank: an update and review. *Frontiers in Communication*, 7.  
<https://doi.org/10.3389/fcomm.2022.865498>
48. Mäkinen, L., Loukusa, S., Nieminen, L., Leinonen, E. i Kunnari, S. (2013). The development of narrative productivity, syntactic complexity, referential cohesion and event content in four- to eight-year-old Finnish children. *First Language*, 34(1), 24–42.  
<https://doi.org/10.1177/0142723713511000>
49. Martinez de Lizarduy, U., Calvo Salomón, P., Gómez Vilda, P., Ecay Torres, M. i López de Ipiña, K. (2017). ALZUMERIC: A decision support system for diagnosis and monitoring of cognitive impairment. *Loquens*, 4(1), e037.  
<https://doi.org/10.3989/loquens.2017.037>
50. Matic, A., Hržica, G., Kuvač Kraljević, J. i Olujić, M. (2018). Syntactic complexity of spontaneous spoken language of adult Croatian speakers. U. V. Brala Vukanović i A.

Memišević (Ur.), *Language in Research and Teaching: Proceedings from the CALS Conference 2016* (str. 143-156). Berlin: Peter Lang.

51. McNamara, D. S., Graesser, A. C., McCarthy, P. M. i Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*.  
<https://doi.org/10.1017/cbo9780511894664>
52. Moussa, H. N. i Mourhir, A. (2023). DarNERcorp: An annotated named entity recognition dataset in the Moroccan dialect. *Data in Brief*, 48, 109234.  
<https://doi.org/10.1016/j.dib.2023.109234>
53. MultiDis. (n.d.). Višerazinski pristup govornogovorenom diskursu u jezičnom razvoju MultiDis. Preuzeto 16. veljače 2024., s <http://www.multidis.erf.hr/index.html>
54. Nakayama, H., Kubo, T., Kamura, J., Taniguchi, Y. i Liang, X. (2018). *doccano: Text Annotation Tool for Human* [Software]. Dostupno na:  
<https://github.com/doccano/doccano>
55. Nippold, M. A., Hesketh, L. J., Duthie, J. K. i Mansfield, T. C. (2005). Conversational Versus Expository Discourse. *Journal of Speech Language and Hearing Research*, 48(5), 1048. [https://doi.org/10.1044/1092-4388\(2005/073\)](https://doi.org/10.1044/1092-4388(2005/073))
56. Nippold, M. A., Mansfield, T. C., Billow, J. L. i Tomblin, J. B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development. *American Journal of Speech-Language Pathology*, 17(4), 356–366.  
[https://doi.org/10.1044/1058-0360\(2008/07-0049\)](https://doi.org/10.1044/1058-0360(2008/07-0049))
57. Olujić, M. i Matić, A. (2017). Govorni i pisani jezik odraslih: koliko se razlikuju? *Govor*, 34(1), 33–52. <https://doi.org/10.22210/govor.2017.34.02>

58. Overton, S. i Wren, Y. (2014). Outcome measurement using naturalistic language samples: A feasibility pilot study using language transcription software and speech and language therapy assistants. *Child Language Teaching and Therapy*, 30(2), 221–229. <https://doi.org/10.1177/0265659013519251>
59. Paul, R. i Norbury, C. F. (2012). Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating (4. izd.). St. Louis, MO: Elsevier.
60. Pavelko, S. L. i Owens, R. E. (2017). Sampling Utterances and Grammatical Analysis Revised (SUGAR): New Normative Values for Language Sample Analysis Measures. *Language, speech, and hearing services in schools*, 48(3), 197–215. [https://doi.org/10.1044/2017\\_LSHSS-17-0022](https://doi.org/10.1044/2017_LSHSS-17-0022)
61. Pavelko, S. L. i Owens, R. E. (2019). Diagnostic Accuracy of the Sampling Utterances and Grammatical Analysis Revised (SUGAR) Measures for Identifying Children With Language Impairment. *Language, speech, and hearing services in schools*, 50(2), 211–223. [https://doi.org/10.1044/2018\\_LSHSS-18-0050](https://doi.org/10.1044/2018_LSHSS-18-0050)
62. Pavelko, S. L., Owens, R. E., Ireland, M. i Hahs-Vaughn, D. L. (2016). Use of language sample analysis by School-Based SLPs: results of a nationwide survey. *Language Speech and Hearing Services in Schools*, 47(3), 246–258. [https://doi.org/10.1044/2016\\_lshss-15-0044](https://doi.org/10.1044/2016_lshss-15-0044)
63. Pavelko, S. L., Price, L. R. i Owens, R. E. (2020). Revisiting reliability: Using sampling Utterances and Grammatical Analysis Revised (SUGAR) to compare 25- and 50- Utterance language samples. *Language Speech and Hearing Services in Schools*, 51(3), 778–794. [https://doi.org/10.1044/2020\\_lshss-19-00026](https://doi.org/10.1044/2020_lshss-19-00026)

64. Pezold, M. J., Imgrund, C. M. i Storkel, H. L. (2020). Using computer programs for language sample analysis. *Language Speech and Hearing Services in Schools*, 51(1), 103–114. [https://doi.org/10.1044/2019\\_lshss-18-0148](https://doi.org/10.1044/2019_lshss-18-0148)
65. Prath, S. P. (2018). *The How and Why of Collecting a Language Sample*. ASHA Wire. Preuzeto 02. veljače 2024., s <https://leader.pubs.asha.org/doi/10.1044/the-how-and-why-of-collecting-a-language-sample>.
66. Scott, C. M. (2020). Language sample analysis: New and neglected clinical applications. *Topics in Language Disorders*, 40(2), 132-134. <https://doi.org/10.1097/TLD.0000000000000215>
67. Stockman, I. J. (1996). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language Speech and Hearing Services in Schools*, 27(4), 355–366. <https://doi.org/10.1044/0161-1461.2704.355>
68. Stockman, I. J. (2010). Listener reliability in assigning utterance boundaries in children's spontaneous speech. *Applied Psycholinguistics*, 31(3), 363–395. <https://doi.org/10.1017/s0142716410000032>
69. *SUGAR language*. (n.d.). Owens, R.B i Pavelko, S.L. Preuzeto: 15. veljače 2024. s <https://www.sugarlanguage.org/>
70. Sveučilišni diplomski studij Logopedija. (2023). Izvedbeni plan nastave sveučilišnog diplomskog studija Logopedija u 2023./2024. akademskoj godini. Preuzeto 14. veljače 2024., s <https://logo.uniri.hr/studij/nastava/detaljni-izvedbeni-nastavni-plan/>
71. Tao, M., Wisler, A., Berisha, V. i Liss, J. (2016). The relationship between perceptual disturbances in dysarthric speech and automatic speech recognition performance. *Journal*

*of the Acoustical Society of America*, 140(5), EL416–EL422.

<https://doi.org/10.1121/1.4967208>

72. Terry, N. P., Mills, M. T., Bingham, G. E., Mansour, S. i Marencin, N. (2013). Oral narrative performance of African American prekindergartners who speak nonmainstream American English. *Language, Speech, and Hearing Services in Schools*, 44(3), 291–305. [https://doi.org/10.1044/0161-1461\(2013/12-0037\)](https://doi.org/10.1044/0161-1461(2013/12-0037))
73. *Text Inspector*. (n.d.). Preuzeto 16. veljače 2024., s <https://textinspector.com/>
74. Trtanj, I. i Kuvač Kraljević, J. (2017). Jezična i govorna obilježja dječjega pripovjednog diskursa: analiza na mikrostrukturnoj razini. *Govor*, 34(1), 53–69. <https://doi.org/10.22210/govor.2017.34.03>
75. Vásquez-Correa, J. C., Arias-Vergara, T., Orozco-Arroyave, J. R., Eskofier, B. M., Klucken, J. i Nöth, E. (2019). Multimodal Assessment of Parkinson’s Disease: a Deep learning approach. *IEEE Journal of Biomedical and Health Informatics*, 23(4), 1618–1630. <https://doi.org/10.1109/jbhi.2018.2866873>
76. Weston, A. D., Shriberg, L. D. i Miller, J. F. (1989). Analysis of Language-Speech Samples with Salt and Pepper. *Journal of Speech Language and Hearing Research*, 32(4), 755–766. <https://doi.org/10.1044/jshr.3204.755>