Računalno prepoznavanje i označavanje negacije u hrvatskom

Žanpera, Natalija

Master's thesis / Diplomski rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:131:050307

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-28**

Filozofski fakultet
Sveučilišta u Zagrebu

Repository / Repozitorij:

ODRAZ - open repository of the University of Zagreb
Faculty of Humanities and Social Sciences

zir.nsk.hr

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

Sveučilište u Zagrebu

Filozofski fakultet

Odsjek za lingvistiku

Natalija Žanpera

RAČUNALNO PREPOZNAVANJE I OZNAČAVANJE NEGACIJE U HRVATSKOM

(RECOGNIZING AND ANNOTATING NEGATION IN CROATIAN)

Diplomski rad

Zagreb, rujan 2020.

Sveučilište u Zagrebu

Filozofski fakultet

Odsjek za lingvistiku

Natalija Žanpera

RAČUNALNO PREPOZNAVANJE I OZNAČAVANJE NEGACIJE U HRVATSKOM

(RECOGNIZING AND ANNOTATING NEGATION IN CROATIAN)

Diplomski rad

Mentor: Doc. dr. sc. Krešimir Šojat

Komentorica: Izv. prof. dr. sc. Kristina Kocijan

Zagreb, rujan 2020.

PODACI O AUTORU

Ime i prezime: Natalija Žanpera

Studijske grupe: Lingvistika (smjer Računalna lingvistika) i Fonetika (smjer Govorništvo)

PODACI O RADU

Naslov rada na hrvatskome jeziku: Računalno prepoznavanje i označavanje negacije u hrvatskom

Naslov rada na engleskome jeziku: Recognizing and Annotating Negation in Croatian

Datum predaje rada: 26. rujna 2020.

IZJAVA O AUTORSTVU DIPLOMSKOGA RADA

Ovim potvrđujem da sam osobno napisala diplomski rad pod naslovom

Računalno prepoznavanje i označavanje negacije u hrvatskom

i da sam njegova autorica.

Svi dijelovi rada, podaci ili ideje koje su u radu citirane ili se temelje na drugim izvorima (mrežni izvori, udžbenici, knjige, znanstveni, stručni članci i sl.) u radu su jasno označeni kao takvi te su navedeni u popisu literature.

Natalija Žanpera

Zagreb, _____

Zahvala

Hvala mentoru profesoru dr. sc. Krešimiru Šojatu koji je rado pristao mentorirati ovaj diplomski rad i koji mi je kroz svoje kolegije približio i učinio zanimljivima različita područja lingvistike. Hvala što ste uvijek imali razumijevanja za studente.

Hvala mojoj komentorici profesorici dr. sc. Kristini Kocijan na svemu. Na tome što se moj odabir Vašeg izbornog kolegija s drugog odsjeka pokazao kao jedna od boljih odluka i što ste me na tom kolegiju toliko toga naučili. Možda niste ni svjesni, ali naučili ste me i da se bolje nosim s izazovima i više si vjerujem. Hvala Vam što ste mi pokazali cijeli prozor mogućnosti oko pisanja ovoga rada. Neizmjerno Vam hvala što ste na tom putu uvijek bili dostupni, što ste na svaki moj upit odgovarali, čak i najgluplji. Hvala Vam na ažurnosti oko povratne informacije, na automatskom rješavanju svih prepreka i problema, na usmjeravanju, savjetima i praktičnoj pomoći. Bez Vas ovo ne bi bilo moguće.

Hvala mom mužu Filipu na podršci, strpljenju, hrabrenju. Hvala na trpljenju mog paničarenja i jadikovanja. Hvala na slušanju o stvarima o kojima ništa nisi razumio, a o kojima se nisam gasila. Ne bih mogla bez tebe.

Hvala mojoj obitelji na razumijevanju i strpljenju s mojim podužim studiranjem.

Hvala svim mojim prijateljima što su se veselili zbog mene, bodrili me i navijali za mene.

Svi ste doprinijeli.

Sadržaj

1. Uvod ... 1

2. Određenje i vrste negacije ... 3

2.1. Jezične razine ... 6

2.2. Negacija u hrvatskim gramatikama ... 7

3. Prefiksalna negacija ... 9

3.1. Prefiksalna negacija pridjeva ... 11

3.2. Negacijski prefiksi ... 16

3.2.1. U hrvatskim gramatikama ... 16

3.2.2. U ostaloj hrvatskoj literaturi ... 23

3.2.3. U stranoj literaturi ... 26

4. Kognitivna obrada negacije ... 30

5. Računalna obrada negacije ... 32

6. Prikupljanje korpusa i analiza prefiksalne negacije pridjeva iz hrWaC-a ... 39

Prefiks a- ... 40

Prefiks ab- ... 41

Prefiks an- ... 42

Prefiks anti- ... 43

Prefiks bez- ... 44

(Inačica obez-) ... 45

Prefiks de- ... 46

Prefiks dez- ... 47

Prefiks dis- ... 47

Prefiks i- ... 48

Prefiks im- ... 51

Prefiks in- ... 53

Prefiks izvan- / van- ... 56

Prefiks kontra- ... 58

(Prefiks mal-) ... 58
Prefiks ne- ... 59
(Inačice prefiksa ne-) ... 60
(Prefiks one-) ... 60
(Prefiks une-) ... 61
(Prefiks izne-) ... 61
(Prefiks uzne-) ... 61
(Prefiks od-) ... 62
Prefiks protu- ... 63
Zamjena prefiksa protu- s anti- i kontra- ... 64
Podaci o prikupljenom korpusu ... 66
7. Metodologija ... 70
8. Računalno prepoznavanje i označavanje negacije hrvatskih pridjeva u NooJ-u ... 73
8.1. Rječnik ... 73
8.2. Gramatika ... 77
8.3. Izračun preciznosti, odziva i F-mjere ... 81
9. Zaključak ... 84
Literatura ... 88
Sažetak ... 92
Abstract ... 93
Životopis ... 94
Prilozi ... 95
Popis slika ... 95
Popis tablica ... 96
Popis grafikona ... 96

1. Uvod

Negacija je fenomen prisutan u svim prirodnim jezicima, a o njegovoj raširenosti i sveprisutnosti govori činjenica da se negaciju ne svrstava samo u područje jezika. Prema *Cambridge Encyclopedia of the Language Sciences,* „negacija je jezični, kognitivni i intelektualni fenomen“[1] (Hogan, 2011:554), dok Israel (2004:701) smatra da „negacija može biti logički operator, vrsta govornog čina, osnovni element semantičke reprezentacije ili pragmatički oblik komunikacijske interakcije.“[2] Na razini jezika, *A Dictionary of Linguistics and Phonetics* navodi da je negacija „proces ili tvorba u gramatičkoj i semantičkoj analizi kojim se uglavnom izražava proturječnost nekih dijelova rečenice ili cijelog njenog značenja“[3] (Crystal, 2008:323).

S obzirom na činjenicu da se izvori koji proučavaju negaciju na jezičnoj razini uglavnom odnose na engleski jezik, broj izvora za hrvatski jezik vrlo je siromašan. Stoga će se ovaj rad baviti negacijom u hrvatskom jeziku, odnosno negacijom na morfološkoj razini. Specifično, bavit će se negacijom morfološki povezanih ili istokorijenskih pridjeva, koji dodavanjem prefiksa postaju antonimi – riječi suprotnoga značenja. Točnije, negacija će se u ovom radu obrađivati u okviru definicije „bez prefiksa nešto *jest,* dodavanjem prefiksa to više *nije* ili dobiva suprotno značenje“, kako je određena i u radu koji se bavi negacijom imenica hrvatskoga jezika (Žanpera i sur., 2020:10).

Prvi dio rada uključivat će teorijski pregled negacije u gramatikama i ostalim jezičnim izvorima, kao i specifično pregled negacije pridjeva te pregled hrvatskih negacijskih prefikasa, ali i onih stranoga podrijetla. U ovome dijelu rada, po pitanju određenih tvorbenih pojava pokušat će se usporediti hrvatski i engleski jezik. Naposljetku, teorijski dio rada prikazat će i neke zaključke istraživanja kognitivne obrade negacije, ali i prikaz dosadašnjih istraživanja računalne obrade negacije.

Drugi dio ovoga rada potaknut je potrebom da se proučavanje ovog jezičnog fenomena odmakne od teorije. Kako navode Jimenez-Zafra i sur. (2019) negacija se premalo proučava na

[1] Vlastiti prijevod
[2] Vlastiti prijevod
[3] Vlastiti prijevod

primijenjenoj razini, a potreba za računalnim ophođenjem ovim fenomenom sve više raste. Primjena znanja iz ovog područja u obradi prirodnoga jezika može biti korisna u analizi sentimenta i izvlačenju informacija iz tekstova, a većina dosadašnjih istraživanja iz ovog područja također se, kao i teorija, odnosi na engleski jezik. Stoga ovaj rad donosi i analizu prefiksalne negacije pridjeva hrvatskoga jezika u programu za računalnu obradu prirodnoga jezika – NooJ (Silberztein, 2003). Obrada prefiksalne negacije pridjeva odnosi se na prikupljanje korpusa pomoću alata Sketch Engine[4], koristeći kao izvor hrvatski korpus hrWaC (Ljubešić i Erjavec, 2011). Prilikom prikupljanja korpusa pojedini prefiksi koji tvore negaciju pridjeva također će se i analizirati prema podacima iz navedenog korpusa, odnosno osvrnut će se na njihovu plodnost, podrijetlo i značenje. Sama obrada prikupljenog korpusa u NooJ-u odnosi se na pisanje rječnika prikupljenih pridjeva i izradu algoritma kojim bi morfološka gramatika uspješno prepoznala i označila pridjeve kao negirane. Također, iznijet će se i zaključci o učestalosti, odnosno plodnosti pojedinih prefikasa te će se izračunati točnost i sveobuhvatnost gramatike, a cjelokupni drugi dio rada teži doprinijeti istraživačkom pristupu ovome fenomenu.

[4] https://www.sketchengine.eu/

2. Određenje i vrste negacije

Zbog relativno oskudnog broja izvora o ovom fenomenu, opus koji se pretraživao za relevantne informacije proširivan je s knjiga i članaka vezanih uz ovo područje na pretragu gramatika i naposljetku, enciklopedija. Tako primjerice *Opća i nacionalna enciklopedija* negaciju smješta u kontekst formalne logike i definira ju kao „osporavanje, nijekanje nekog suda ili iskaza“ te kao „suprotnost afirmaciji, potvrđivanju ili tvrdnji“. Dodatno, definira ju kao „čin kojim se jedan pojam zamjenjuje njemu protuslovnim ili odbacuje kao neistinit“ (Vujić, 2007:220).

S druge strane, *Hrvatski jezični portal*[5] navodi iznenađujuće šturu definiciju, s obzirom na to da se radi o jezičnom izvoru. Navodi da je negacija poricanje, odbijanje ili nijekanje, a u lingvističkom smislu spominje se samo da je riječ za odbijanje tzv. niječnica *ne* te da joj je oprečni pojam afirmacija.

Zovko Dinković (2013) također ističe da je negacija u hrvatskom jeziku slabo i nesustavno obrađena te da se uglavnom spominje u okviru gramatika i jezičnih priručnika, a definira se kao poricanje, odsutnost ili proturječnost odnosno suprotnost nečemu potvrdnom te kao nepostojanje.

Kovačević (2016) navodi da je negacija univerzalna i jeziku inherentna pojava te da se eksplicitno izražava niječnim prefiksima, česticama, veznicima, prilozima, prijedlozima itd. Dodaje da je negacija sastavni dio gramatičke kategorije jesnost/niječnost te da se temeljno niječno značenje negacije označava pojmovima opreke, nijekanja ili poricanja.

Strani izvori plodniji su po pitanju određenja ovoga pojma, posebno u jezičnom kontekstu. Osim definicija već navedenih u uvodu, vrijedi istaknuti i da Horn i Kato (2000) ističu da je negacija u srži svakog sustava ljudske komunikacije, a zajedno sa svojim korelatima – istinite vrijednosti, pogrešne poruke, proturječnosti i ironija, čini jednu od temeljnih karakteristika ljudske vrste. Nadalje, autori navode da u srži jezične sposobnosti negacija ima značajnog dodira i s načelima morfosintakse, logike, jezičnog usvajanja te rečenične obrade,

[5] http://hjp.znanje.hr/

gdje se autori dotiču i nekih od jezičnih razina na kojima negacija može djelovati, o čemu će biti riječi kasnije u radu.

Kroz brojne izvore u kontekstu negacije pojavljuju se i ostali pojmovi koje je nužno razlikovati, a stoga ih je neophodno i spomenuti. Najčešće se spominju pojmovi suprotnost (engl. *contrary*) i proturječnost (engl. *contradictory*), a iz perspektive logike, prema Johnson (2007:63) „tvrdnje su u odnosu suprotnosti kada ne mogu obje biti istinite, ali mogu obje biti pogrešne“.[6] Ovu vrstu negacije objašnjavaju i De Clercq i Vanden Wyngaerd (2018:1) navodeći primjere tvrdnji „*Linus je visok.“* i *„Linus je nizak.“,* međutim osoba ne mora biti ni visoka, ni niska, već prosječne visine, što obje ove tvrdnje čini netočnima. S druge strane, Johnson proturječne pojmove objašnjava kao pojmove koji imaju suprotne istinite vrijednosti, odnosno ako je A točan, tada B mora biti pogrešan i obrnuto (Johnson, 2007).

Međutim, važno je spomenuti da ova podjela postoji još od Aristotela, koji je prvi negaciju prebacio iz područja ontologije u polje logike i lingvistike (Zovko Dinković, 2013). Horn (2005) i Zovko Dinković (2013) navode da je Aristotel u svom djelu *Kategorije* razlikovao četiri vrste oprečnosti: korelacija (odnosna opreka), suprotnost (engl. *contrary*), lišenost (engl. *privation*) i proturječnost (engl. *contradictory*). Chapman i sur. (2001b) navode primjere za vrste opreka prema Aristotelu, gdje se odnos između pojmova *duplo/pola* odnosi na korelaciju, *dobar/loš* na suprotnost, *slijep/sposoban da vidi* (engl. *blind/sighted*) na lišenost, a „*On sjedi.“ / „On ne sjedi.“* na proturječnost. Horn (2005) dodatno za proturječje navodi da se radi o odnosu jesnosti prema niječnosti, za korelaciju još navodi i primjer *roditelj/dijete,* dok za lišenost ističe da označava razliku između postojanja i odsutnosti nečega.

De Clercq i Vanden Wyngaerd (2018) odnose suprotnosti i proturječja objašnjavaju na primjeru istih pridjeva, odnosno *visok* i *nizak* su suprotni pojmovi, a u tvrdnjama „*On je visok.“ / „On nije visok.“* radi se o proturječnosti. Autori stoga ističu da se proturječna negacija ostvaruje niječnom riječju *not* (engl.), odnosno autori ju nazivaju rečeničnim negatorom, što je na primjeru hrvatskih rečenica ostvareno pomoću *nije.* Nazivajući *not* rečeničnim negatorom može se zaključiti da se odnos proturječnosti ostvaruje na razini rečenice, a to potvrđuju i ostali autori. Csapo (1997) govoreći o proturječnosti ističe da se radi o negaciji prema obrascu *„nije X“.* Nadalje, Jespersen (1917) tvrdi da negacija u lingvističkom smislu mijenja pojam u ono što

[6] Vlastiti prijevod

se u logici naziva proturječnim pojmom, odnosno *„A i nije-A“* (Jespersen, 1917). Naposljetku, Zovko Dinković (2013) ističe da se odnos proturječnosti tiče isključivo iskaza ili tvrdnji, dok se odnosi korelacije, suprotnosti i lišenosti uspostavljaju između pojmova.

Međutim, proturječnost prema obrascu *„X / nije X“* ne odgovara samo rečeničnoj razini. Naime, Jespersen (1917) navodi primjere *dostojan/nedostojan* (engl. *worty/unworthy)* koji odgovaraju i formuli *„X / nije X“,* odnosno može se reći i *dostojan/nije dostojan,* a jednako proturječno značenje ostvareno je prefiksom *un-* u slučaju engleskog jezika, to jest prefiksom *ne-* u slučaju hrvatskog jezika. Točnije, proturječna negacija osim na razini rečenice formulom *„nije X“* (engl. *not X*), može se ostvariti i na razini riječi formulom *ne-X,* odnosno *un-X* (engl.). Dodatno, Horn (2001) spominje prefiks *non-* kao prefiks koji također tvori proturječnu negaciju.

Već spomenuti autor Israel (2004:701) u kontekstu negacije navodi i pojam polarnosti koji razlikuje od negacije određujući ga kao „odnos između semantičkih suprotnosti – između značenja koja su u temelju nekonzistentna jedno s drugim“[7] te navodi da polarnost ne obuhvaća samo logički odnos između niječnih i potvrdnih tvrdnji, već i konceptualan odnos između suprotnih parova poput *vruće/hladno, dug/kratak, dobro/loše.*

O kompleksnosti ovog jezičnog fenomena govori činjenica da je ponekad teško povući granicu između ovih vrsta negacije, jer jezik je prepun prijelaznih razdoblja i stalno se mijenja, što uključuje i konstantu pojavu novih oblika koji ne odgovaraju uvijek u potpunosti ovim kategorijama (Csapo, 1977).

Iz prethodno navedene Israelove definicije kojom razlikuje negaciju i polarnost, vidi se prisutnost ovog fenomena na razini semantike, dok se kao što je već istaknuto, proturječnost pojavljuje na rečeničnoj razini, stoga će iduće potpoglavlje ukratko prikazati jezične razine na kojima se negacija može ostvariti.

[7] Vlastiti prijevod

2.1. Jezične razine

Hrvatski jezični savjetnik (Barić i dr., 1999, prema Kovačević, 2016:25) navodi da se nijekanje ostvaruje na nekoliko jezičnih razina:

- „na leksičkoj razini: *vidljiv – nevidljiv, prilika – neprilika*
- zanijekan može biti bilo koji rečenični dio – *Ivanu je ne vjerovati.*
- na razini rečenice – *Ivan nije došao.*
- na razini teksta – *Jesi li umoran? Nisam*."

Navedene jezične razine Kovačević (2016) obrađuje kao morfološku, sintaktičku, rečeničnu te negaciju u diskursu. Međutim, iako autorica za negaciju navodi da se radi o univerzalnoj jezičnoj pojavi, ističe da morfološka negacija nije univerzalna pojava, odnosno da ne izražavaju svi jezici negaciju i na morfološkoj razini. Zovko Dinković (2013) također spominje ove četiri jezične razine na kojima se može ostvariti negacija, nazivajući ih leksičkom, sastavničkom, rečeničkom razinom i razinom teksta.

Dodatno, Kovačević (2016) navodi da je u onim jezicima koji imaju morfološku negaciju, sintaktička negacija učestalija i uočljivija. Shodno tomu, literatura koja i obrađuje negaciju, najčešće ju obrađuje na rečeničnoj razini, primjerice Horn (2001) i Jespersen (1917), kao i na sintaktičkoj, primjerice Zovko Dinković (2013), dok radova koji obrađuju negaciju za hrvatski jezik na morfološkoj razini gotovo i nema. Međutim, vrijedi spomenuti Šarić (1992), rad o hrvatskom jeziku koji spominje negaciju na morfološkoj razini, točnije bavi se antonimijom, gdje autorica razlikuje morfološki povezane i nepovezane antonime, o čemu će biti riječi u poglavlju o negaciji pridjeva.

Joshi (2012) obrađuje afiksalnu negaciju u engleskom jeziku, ali također primjećuje da se većina literature po pitanju negacije bavi negacijom na rečeničnoj razini. Autor navodi da se antonimija, kao dio leksičke negacije, odnosno negacije na razini riječi uglavnom obrađuje u području leksikologije i donekle semantike. S druge strane, za afiksalnu negaciju ističe da kao i antonimija pripada negaciji na razini riječi, međutim njome se gotovo nitko ne bavi, čak ni u području morfologije (Joshi, 2012).

Ovaj rad, kao što je navedeno u uvodu, bavi se negacijom na morfološkoj, odnosno leksičkoj razini – na razini riječi. Točnije, bavi se negacijom koja se ostvaruje dodavanjem prefikasa, o čemu će biti riječi u narednim poglavljima.

2.2. Negacija u hrvatskim gramatikama

Negacija je u hrvatskim gramatikama nesustavno obrađena. Naime, neke gramatike svrstavaju je u preoblike gramatičkog ustrojstva rečenice, dok ju druge gramatike siromašno spominju u okviru prefiksalne tvorbe, odnosno kod navođenja prefikasa spominju se i oni koji tvore negaciju, ali ni na koji način se ne odvajaju i obrađuju kao zasebna jezična pojava. Međutim, određene gramatike negaciju ne svrstavaju niti definiraju, a neke je ni ne spominju kao pojavnost (Zovko Dinković, 2013).

Gramatika Barić i sur. (2005) negaciju smješta u preoblike gramatičkog ustrojstva rečenice. Navedeni autori iznose da postoje dvije vrste preoblika rečeničnog ustrojstva, preoblike koje gramatičke odnose u rečeničnom ustrojstvu ostavljaju nepromijenjenima i mijenjaju mu samo značenje te tu spadaju nijekanje, pitanje, usklik i zahtjev. S druge strane, postoje i preoblike koje mijenjaju gramatičko ustrojstvo rečenice, a značenje mu ostavljaju nepromijenjenim, gdje spadaju pasiv i obezličenje. Nijekanje opisuju kao preobliku kojom se čitav sadržaj rečeničnog ustrojstva niječe pa se otuda rečenici mijenja značenje. Rečenica preoblikovana nijekanjem po značenju je niječna, a ona koja to nije je potvrdna.

Katičić u svojoj *Sintaksi hrvatskog književnog jezika* (2002) negaciju postavlja unutar istog poglavlja kao Barić i sur. (2005). Spominje je u poglavlju o preoblikama gramatičkog ustrojstva rečenice gdje navodi da su to preoblike koje mijenjaju rečenično značenje, a odnose u ustrojstvu ostavljaju nepromijenjenima. Međutim, ovdje postoji sitna razlika u nazivlju između Katičića (2002) i Barić i sur. (2005), jer Katičić takve preoblike naziva nijekanje, pitanje, usklik i poticaj – koji Barić i sur. nazivaju zahtjev. Prema Katičiću, nijekanje ili negacija preoblika je koja zahvaća čitavo rečenično ustrojstvo. Njome se otklanja i poriče rečenični izričaj kako je sadržan u tom ustrojstvu. Ostatak poglavlja bavi se negacijom na razini rečenice, što nije predmet ovoga rada.

Takozvana *Velika hrvatska gramatika* (Babić, 2002), kao i njena prethodnica (Babić, 1991) ne sadrže negaciju kao posebno poglavlje, već ju spominju u poglavljima o prefiksalnoj tvorbi riječi, navodeći među ostalim prefiksima i prefikse kojima se tvori negacija.

Težak i Babić (1992) negaciju spominju u poglavlju o vrstama rečenica, gdje uz jesnu spominju i niječnu rečenicu navodeći da se prepoznaje niječnicom *ne* ili zanijekanim glagolskim oblikom *nisam, nisi* (1992). Međutim, kao i Babić (1991, 2002) spominju negacijske prefikse u poglavlju o tvorbi pridjeva. Autori su u navedenim izdanjima tvorbu

pridjeva podijelili na tvorbu opisnih i tvorbu odnosnih pridjeva, a ono što je zanimljivo i zbog čega je uspoređivano više izdanja Babićeve gramatike je da Težak i Babić (1992) spominju negaciju u poglavlju o tvorbi opisnih pridjeva s posebnim značenjima, točnije imaju posebno potpoglavlje o tvorbi pridjeva koji znače odsutnost ili nijekanje osobine. S druge strane, oba navedena izdanja Babićevih gramatika ne sadrže to potpoglavlje, iako sadrže poglavlje jednakoga naziva i unutar njega sva ostala potpoglavlja. Više o sadržaju tih poglavlja bit će izneseno u poglavlju o prefiksalnoj negaciji pridjeva, na čemu i je naglasak u ovome radu.

3. Prefiksalna negacija

Prema Subandowo (2014:137) „afiksalna negacija je proces tvorbe riječi u kojem nova riječ nastaje dodavanjem afiksa na korijen riječi“[8]. U slučaju ovoga rada, proces tvorbe nove riječi odnosi se specifično na dodavanje prefiksa, jer negacija se može tvoriti i dodavanjem sufiksa, što će biti spomenuto kasnije u ovom poglavlju.

Slično o ovoj vrsti negacije navodi i Zovko Dinković (2013), spominjući da se niječne riječi mogu podijeliti u dvije skupine. Prvu čini afiksalna negacija, odnosno niječne riječi koje se tvore dodavanjem afiksa na pozitivni oblik pridjeva, imenice ili priloga. Drugu skupinu čine riječi koje ne sadrže u sebi niječni morfem, ali imaju negativno značenje.

S obzirom na to da postoje riječi koje imaju negativno značenje, ali nisu negacija, važno ih je razlikovati. Primjerice, riječ *zločin,* gdje riječ *čin* postoji i samostalno, ali dodavanjem *zlo* ne negira se *čin* ili izražava nešto suprotno od čina, već *čin* samo dobiva negativno značenje. Također, riječ *nedjelo*, ne negira *djelo*, kao što riječ *nevolja* ne negira riječ *volja*, već samo riječi s pozitivnom ili neutralnom osnovnom daje negativno značenje. Stoga se u ovom radu negacija odnosi samo na slučajeve „bez prefiksa nešto *jest,* dodavanjem prefiksa to više *nije* ili dobiva suprotno značenje“, poput *normalno/abnormalno, produktivno/kontraproduktivno* i sl.

Afiksalnu negaciju Joshi (2012) dijeli na izravnu i neizravnu, a autor navodi da one ne isključuju jedna drugu te se njihov odnos ne može unaprijed odrediti. Autor za objašnjenje tih vrsta negacije upotrebljava primjere *happy/unhappy* (engl.) gdje se *unhappy* odnosi na osobu koja nije sretna, što je u izravnoj suprotnosti s osobom koja je sretna. S druge strane, pojam *infamous,* odnosno *ozloglašen,* nije u izravnoj suprotnosti s pojmom *famous,* odnosno *poznat.* Točnije, kod pojma *infamous* ne radi se o nekome tko nije poznat, nego o nekome tko je poznat iz loših razloga. Stoga *infamous* ima negativnu konotaciju i iz tog razloga odnosi se na neizravnu negaciju.

Takva podjela može se primijeniti i na hrvatski jezik, primjerice korištenjem istog primjera *sretan/nesretan,* radi se o izravnoj negaciji. S druge strane, primjer s imenicama naveden u prethodnom odlomku – *volja/nevolja,* ne odnosi se na izravnu negaciju, jer *nevolja*

[8] Vlastiti prijevod

dodavanjem negacijskog prefiksa nije u suprotnosti s riječju *volja,* već samo dobiva negativno značenje, odnosno negativnu konotaciju, ali kao što je već navedeno, ovaj rad ne bavi se takvom vrstom negacije. Stoga se može reći da bi negacija u ovom radu određena prema prethodno navedenoj Joshijevoj podjeli pripadala izravnoj vrsti negacije. Međutim, za razliku od imenica *volja/nevolja* gdje se radi o neizravnoj negaciji, vrijedi istaknuti da je s pridjevima *voljan/nevoljan* situacija drugačija, odnosno u slučaju ovih pridjeva radi se o izravnoj vrsti negacije. Točnije, kod ovih imenica prefiks je proizveo negativno značenje, ali ne i suprotni pojam, dok je kod pridjeva dodavanje prefiksa rezultiralo pojmovima suprotnoga značenja.

Pojava da se prefiksalnom tvorbom negacije ne tvore uvijek riječi istoga značenja nije neuobičajena. Kovačević (2016) napominje da tvorba riječi dodavanjem prefiksa na osnovu ne mora uvijek rezultirati time da osnova i nova riječ imaju isto značenje, ali među njima će postojati očita semantička veza. Odnosno, tvorbom riječi nastaju novi oblici, ali i nova značenja, gdje u primjeru već navedenih pojmova *volja* i *nevolja,* iako je *volja* osnova novonastale riječi, one ne dijele značenje, već se radi o imenici koja je dobila drugo značenje.

Kovačević (2016) afiksalnu negaciju naziva još i leksička ili morfološka te navodi da ona upućuje na tvorbu riječi. Kao što je već spomenuto, afiksalna negacija uključuje tvorbu riječi dodavanjem prefiksa, ali i sufiksa. Zimmer (1964, prema Kovačević, 2016) ističe da iako se afiksalna negacija najčešće ostvaruje prefiksima, u određenim jezicima moguće je ostvarenje afiksalne negacije i sufiksom. Kovačević za primjere navodi engleski sufiks -*less,* mađarski -*talan* te finski -*ton.* Govoreći o negaciji pridjeva, de Clercq i Vanden Wyngaerd (2017) također ističu da je pridjev moguće negirati negacijskim prefiksom, ali i sufiksom te navode primjere iz nizozemskog jezika s prefiksom *on-* i sufiksom -*loos.*

Babić (2002:48) prefiksalnu tvorbu definira kao „slaganje riječi kojem je u prvom dijelu složenice prijedlog ili niječnica *ne-*„„ primjerice *neumoljiv* te se takve riječi nazivaju prefiksima ili predmecima. Govoreći o prefiksima, Babić razlikuje samostalne i nesamostalne riječi, gdje su samostalne riječi upravo prijedlozi i niječnica *ne*, a o nesamostalnim prefiksima bit će riječi kasnije u radu.

Međutim, Težak i Babić (1992) ističu da se upravo zbog toga što su prefiksi uglavnom samostalne riječi, prefiksalna tvorba ubraja u slaganje. U istoj gramatici autori također navode da se prefiksalna tvorba često udružuje sa sufiksalnom, odnosno na osnovu se istodobno dodaje i prefiks i sufiks te se to naziva prefiksalno-sufiksalna tvorba. Takvih slučajeva ima i u korpusu korištenom za ovaj rad te iako je primarno određenje negacije za potrebe ovoga rada prefiksalna

negacija pridjeva, bili su neizbježni i ne tako „čisti“ primjeri, odnosno primjeri koji bi se mogli svrstati u takav složeniji oblik tvorbe.

Prema Težak i Babić (1992) svaka riječ koja u svom sastavu ima i prefiks i sufiks ne ide automatski u prefiksalno-sufiksalnu tvorbu, odnosno može ići i u neku drugu prema načinu na koji je nastala. Primjerice, imenica *beznadnost* je izvedenica, jer je nastala od pridjeva *beznadan*, koji je već u sebi imao prefiks *bez-* pa bi se tako imenica *beznadnost* mogla svrstati u sufiksalnu tvorbu, jer je nastala od riječi koja je već imala prefiks, ali je u toj tvorbi dobila sufiks *-ost*. Stoga Babić (2002) takve imenice i svrstava u sufiksalnu tvorbu, odnosno spominje ih upravo u poglavlju o imenicama nastalima sufiksom *-ost.*

Isto tako, neki pridjevi nastali su od imenica koje su već negirane prefiksalnom tvorbom. Primjerice, pridjev *kontraceptivni* nastao je od imenice *kontracepcija,* odnosno ovaj pridjev nije nastao prefiksacijom, već sufiksacijom. Međutim, za računalnu analizu koja će se provesti u ovom radu važno je da bez prefiksa nešto *jest*, a dodavanjem prefiksa više *nije* ili dobiva suprotno značenje – takvi primjeri se gledaju i smatraju negacijom, bez obzira na to je li taj pridjev nastao prefiksacijom ili od neke druge riječi koja je već sadržavala taj prefiks. Ovaj specifičan primjer s pridjevom *kontraceptivni* nije uključen u korpus za analizu, jer u hrvatskom jeziku ne postoji inačica toga pridjeva bez prefiksa.

3.1. Prefiksalna negacija pridjeva

Osim što je ovaj rad motiviran činjenicom da je ovo područje za hrvatski jezik premalo pokriveno – kako u teoriji, tako još više i u praksi, motivirala ga je i činjenica da su pridjevi hrvatskoga jezika izrazito plodni za tvorbu negacije. Pridjevi su uz imenice i glagole najbrojnija vrsta riječi, a o prepoznavanju i označavanju negacije imenica hrvatskoga jezika već je napisan rad – Žanpera i sur. (2020). Međutim, osim što su pridjevi jedna od najbrojnijih vrsta riječi, oni su najbrojnija i najučestalija vrsta riječi s niječnim prefiksima (Kovačević, 2016). Navedena autorica to povezuje s prototipnim značenjem pridjeva, ističući da je *svojstvo* kao prototipno značenje pridjeva apstraktnije od *predmetnosti* i *radnje* kao prototipnih značenja imenica i glagola te je stoga najviše podudarno doslovnom značenju negacije, koje je također apstraktno (Kovačević, 2016).

S druge strane, čini se da bi situacija s negacijom pridjeva u engleskom jeziku mogla biti drugačija. Kjellmer (2005) analizom pridjeva engleskoga jezika na temelju CobuildDirect korpusa dolazi do zaključka da većina pridjeva nema svoj negirani oblik. Točnije, od gotovo

1000 pridjeva, samo 8 % pojavilo se negirano prefiksima *dis-, in-, non-* ili *un-,* stoga autor zaključuje da je negacija prilično rijedak fenomen među pridjevima (Kjellmer, 2005). Međutim, ove brojke mogu biti i varljive, s obzirom na činjenicu da su parovi poput *unwritten* i *written* različito označeni. Odnosno, *unwritten* je označen kao pridjev, dok je *written* označen kao glagolski oblik. No i Kjellmer (2005) i drugi autori na koje se poziva, poput Gleby (2002) ističu da takvi ne-negirano/negirani parovi trebaju biti predstavljeni kao pridjevski oblici. Stoga se može zaključiti da korpus zasigurno sadrži puno više ne-negiranih pridjeva nego što je označeno.

Slično je utvrđeno i jednim ranijim istraživanjem. Naime, istraživanje koje je proveo Sherman (1973) pokazalo je da se prefiksalno negirani pridjevi pojavljuju rjeđe nego njihove osnove – pridjevi bez negacijskog prefiksa, osim pridjeva *tidy,* koji se češće pojavljuje u obliku *untidy.*

Nadalje, ono što je Kjellmer (2005) također zaključio je da ulogu u negaciji pridjeva ima i sama učestalost pridjeva. Riječi koje su rijetke puno rjeđe imaju svoj negirani oblik nego učestale riječi. Stoga autor zaključuje da govornici i pisci zasigurno puno češće imaju potrebu za pronalaskom negiranog ekvivalenta poznatog pridjeva, nego onih rjeđih i manje poznatih (Kjellmer, 2005).

Horn (2001) navodi da su pridjevi s negativnim afiksima vrlo često omalovažavajući, ponižavajući ili su procijenjeni kao negativno denotirani ili konotirani. Slično ističe i Jespersen (1917) govoreći specifično o pridjevima nastalima od prefikasa *un-* i *in-* gdje ističe da ih većina ima značenje umanjujuće vrijednosti, poput *unworthy, undue, imperfect* (engl.).

De Clercq i Vanden Wyngaerd (2018) govoreći o prefiksu *un-* ističu da se negativni pridjevi sustavno suprotstavljaju prefiksaciji negacijskim prefiksom *un-*. Zovko Dinković (2013) ističe da u takvim primjerima samo emotivno pozitivna ili neutralna osnova može poslužiti za tvorbu pridjeva i imenica s niječnim prefiksom. Slično navodi i Horn (2001) koji to objašnjava preuzimajući od Cruse (1980) oznake *e-pos* i *e-neg* za riječi procijenjene kao pozitivne, odnosno negativne. Iz toga izvodi formulu:

- negativni afiks *(un-)* + e-pos osnova *(happy)* → e-neg izvedeni oblik *(unhappy)*.

Odnosno, postoji riječ *unhappy,* ali ne i *unsad.* S druge strane, u hrvatskom jeziku postoje takve tvorbe. Točnije, postoji riječ *nesretan,* ali i *netužan,* što je pokazala i analiza

prefiksalno negiranih pridjeva iz hrWaC-a, o čemu će biti riječi u poglavlju o prikupljanju korpusa.

Tvrdnju da se niječni afiksi najčešće nalaze uz pozitivne osnove pokušao je provjeriti i Zimmer još 1964. (prema Kovačević, 2016). Pokazao je da iako u pojedinim jezicima i pojedinim vrstama riječi niječni prefiksi imaju tendenciju pridodavanja pozitivnim osnovama, ta pojava nije univerzalija. Također je zaključio da postoji i problem kod određenja, odnosno nije svaku osnovu moguće odrediti prema pozitivno/negativno odnosu, jer postoje i neutralne osnove. Stoga u tu tvrdnju Horn (2001) i Zovko Dinković (2013) dodaju i pojam neutralne osnove. Ono što je Zimmer dodatno uočio je da se u afiksalnoj negaciji ne može govoriti o negramatičnosti ili neprihvatljivosti određenih tvorbi (Kovačević, 2016). Točnije, postoje primjeri negacije (prefiks+osnova) koji se ne upotrebljavaju, ali ih to ne čini negramatičnima, s obzirom na to da ne krše gramatička pravila, već jednostavno govornici s takvim primjerima nisu upoznati. Primjerice, pridjev *unsad* se u engleskom jeziku ne upotrebljava, ali ga takva konstrukcija ne čini negramatičnim. Odnosno, kako nadalje ističe Kovačević (2016), postoji puno više mogućih primjera i kombinacija danih na slobodu govornicima nekog jezika, koji se mogu stvarati ovisno o potrebama konteksta, međutim rječnici ne mogu sve sadržati jer se zbog otvorenosti i slobode takvih tvorbi mogu konstantno stvarati novi oblici.

Kjellmer (2005) gledajući samo pridjeve koje počinju s *a-,* iznosi da riječi koje imaju „negativni" korijen nemaju svoje negirane ekvivalente, poput *anxious, awkward,* i sl., dok *agreeable, attractive* i slične riječi koje bi bile označene kao „pozitivni" korijen, imaju svoj negativni ekvivalent, što je u skladu sa zaključcima prethodno navedenih autora Horn (2001) te Zovko Dinković (2013). Međutim, s obzirom da postoje i iznimke, ovo pokazuje samo tendenciju u tvorbi negacije pridjeva, ne i pravilo. Primjerice, riječi poput *corruptible* i *provocative* imaju svoje negirane ekvivalente *incorruptible, unprovocative,* iako imaju negativan korijen (Kjellmer, 2005).

O iznimkama ove tendencije u tvorbi pridjeva pišu i De Clercq i Vanden Wyngaerd (2017) koji postavljaju izazov tvrdnji da negacijski prefiks ne može ići na već negativnu pridjevsku osnovu. Autori ističu da to nije slučajnost te predlažu da je neslaganje negativnih prefikasa jedan na drugoga, kao i na negativnu osnovu, povezano s dometom negacije. Točnije, ako imaju različite domete, slaganje je moguće, odnosno ako imaju isti domet ono nije moguće. Stoga se prema tome može predvidjeti da sintaktička negacija koja ima mali domet, neće podnijeti slaganje s afiksima kratkog dometa ili s negativnim pridjevom.

Međutim, Kjellmer (2005) ovu tendenciju primjećuje i kod pridjeva kojima se iskazuju emocije, posebno onih koji se upotrebljavaju u uzvicima, a to je da uglavnom nemaju svoj negirani ekvivalent. Primjerice, ponovno analizirajući samo pridjeve koji počinju s *a-,* pridjevi poput *amazing, annoying, awesome* nemaju svoj negirani ekvivalent. Autor zaključuje da „kao da bi negacija isisala emocionalni sadržaj iz riječi pa ga ne bi ni ostalo“[9] (Kjellmer, 2005:160).

Vrijedi spomenuti da je Kjellmer (2005) primijetio da superlativi – pridjevi najvišeg stupnja u gradaciji pridjeva, također nisu negirani, što je zapažanje koje je iznio i Zimmer (1964, prema Kjellmer, 2005). Odnosno, na primjeru hrvatskog jezika, superlativi *najbolji, najviši, najpametniji* nemaju svoje negirane parove. Međutim, čini se da općenito pridjevi koji sudjeluju u gradaciji nemaju svoje negirane parove, ne samo oni na najvišem stupnju gradacije – superlativi. Primjerice, stupnjevanje u slučaju superlativa *najbolji* glasi *dobar – bolji – najbolji,* u slučaju superlativa *najviši* glasi *visok – viši – najviši* te *najpametniji* glasi *pametan – pametniji – najpametniji.* Dakle, ne postoji *nedobar, nevisok, nepametan,* već pridjevi koji sudjeluju u stupnjevanju imaju svoje ne-negirane antonime – *dobar/loš, visok/nizak, pametan/glup.*

To potvrđuje i Babić (1991) navodeći ograničenje u tvorbi prefiksom *ne-* ističući da do takve tvorbe ne dolazi ako se opreka izriče drugim pridjevom, primjerice *debeo/mršav.* Međutim, sam autor navodi da takva ograničenja nisu apsolutna, jer se u jeziku svejedno upotrebljavaju primjeri poput *nečist,* iako postoji i opreka izrečena drugim pridjevom, odnosno *prljav,* kao i pridjev *neveseo,* iako postoji i opreka s ne-negiranim parnjakom *veseo/tužan.* Slično je o negaciji skalarnih pridjeva primijetila i Kovačević (2016) navodeći da, s obzirom da negacija znači da nešto *nije* ili ga *nema,* logična je i karakteristika niječnih pridjeva da se ne stupnjuju. Prikupljanjem pridjeva za ovaj rad iz korpusa hrWaC (Ljubešić i Erjavec, 2011) pokazalo se da postoje i iznimke, a detaljnije o tome iznijet će se na kraju poglavlja o prikupljanju i analizi korpusa.

Naposljetku, Kjellmer ističe da općenito pridjevi koji imaju svoj ne-negirani antonim, nemaju svoj negirani ekvivalent, ne samo pridjevi koji prolaze kroz stupnjevanje. Za primjer navodi pridjeve poput *dead/alive, amateur/professional, abstract/concrete.* Odnosno, pridjev *alive* ima svoj ne-negirani antonim *dead,* stoga ne postoji negirani ekvivalent od *alive: *in-/*non-/*un-alive* (Kjellmer, 2005). Međutim, u hrvatskom jeziku situacija je drugačija – iako

[9] Vlastiti prijevod

postoji ne-negirani antonim pridjeva *živ,* a to je *mrtav,* postoji i njegov negirani par, a to je pridjev *neživ.* Tako je i s primjerom *amater/profesionalac*, gdje postoji i njegov negirani par *neprofesionalac*, zatim *prisutan/odsutan*, ali postoji i *neprisutan* te *apstraktno/konkretno*, ali postoji i *nekonkretno*.

Šarić (1992) ovakve negirane pridjevske parove dijeli na morfološki povezane i morfološki nepovezane. Odnosno, par iz navedenog primjera *amater/profesionalac* prema Šarić bi bio morfološki nepovezan par te autorica navodi da teoretičari takve antonime nazivaju pravim antonimima. S druge strane, pridjevski par *profesionalac/neprofesionalac* smatrao bi se morfološki povezanim, odnosno istokorijenskim antonimom. Autorica za morfološki povezane antonime navodi primjer *sretan/nesretan,* dok bi morfološki nepovezan ekvivalent bio *sretan/tužan.* Ističe da se morfološki povezani antonimi najčešće tvore niječnicom *ne* ili nekim drugim prefiksom koji ima odrično značenje (Šarić, 1992).

Međutim, negirani i ne-negirani pridjevski parovi ne mogu se uvijek smatrati značenjskim ekvivalentima. Paradis (2006) u sličnom kontekstu spominje antonimne pridjevske parove koje dijeli na značenjski vezane i nevezane (engl. *bounded/unbounded*). Nevezane parove povezuje sa skalarnim pridjevima, odnosno onima koji se mogu stupnjevati, poput *dobar, bolji, najbolji,* što se može povezati s prethodno navedenim postavkama iz Kjellmer (2005), Zimmer (1964) i Babić (1991) prema kojima se takvi pridjevi ne negiraju prefiksom *ne-,* već imaju svoj ne-negirani ekvivalent. Dodatna razlika kojom Paradis (2006) pojašnjava vezane i nevezane pridjevske parove je da se vezani nalaze uz apsolutne modifikatore poput *potpuno, gotovo* (engl. *totally/almost*) izražavajući ideju *„ili-ili“* (engl. *„either-or“*). Odnosno, pojavljuju se uz primjere vezanih pridjeva u kojima nema prostora za između, poput *živ/mrtav,* koji se dakle mogu izraziti modifikatorima *totally dead/almost dead.* S druge strane, nevezani pridjevski parovi pojavljuju se uz modifikatore poput *vrlo, prilično* (engl. *very/fairly*). Odnosno, moguće je stupnjevanje između dvaju pojmova, primjerice cesta ne mora biti ni uska ni široka, a tvrdnja se može izraziti pomoću navedenih modifikatora, poput *vrlo široka/prilično široka.* Hipoteza koju je autorica potvrdila svojim istraživanjem je da kad se *ne/nije* kombinira s vezanim pridjevskim parom, pomoću *ne* se izražava potpuna suprotnost značenja – primjerice *ne mrtav* ili *nije mrtav = živ.* S druge strane, kad se *ne* kombinira s nevezanim pridjevskim antonimskim parovima poput *širok/uzak,* njegova funkcija je umanjivanje značenja modificiranog pridjeva te se tada ne izjednačava značenje – odnosno *širok* nije isto što i *ne uzak* – jer može postojati i stanje između, koje nije ni široko ni usko. Svojim istraživanjem autorica je potvrdila da je interpretacija negiranog pridjevskog značenja

osjetljiva na činjenicu radi li se o vezanim ili nevezanim parovima (Paradis, 2006), a ovakvo istraživanje bilo bi dobro učiniti i isključivo za hrvatski jezik, iako se isti obrasci mogu primijetiti već na navedenim primjerima.

3.2. Negacijski prefiksi

Ovo poglavlje donosi pregled prefikasa kojima se tvori negacija i koliko je moguće odrediti, koji se kada pojavljuju – kojega podrijetla i s kakvim osnovama. Bit će prikazani i prefiksi specifično za hrvatski jezik, odnosno u hrvatskoj literaturi i specifičnije, u hrvatskim gramatikama, ali i literatura o engleskom jeziku, s prikazom onih koji tvore negaciju i u hrvatskom jeziku te usporedba engleskog i hrvatskog jezika. Dzuganova (2006, prema Subandowo, 2014:137) negacijski afiks definira kao „afiks koji tvori suprotno značenje ili negira riječ“[10], što odgovara i definiciji negacije određenoj unutar okvira ovoga rada.

Neke prefikse koji će biti navedeni u idućim potpoglavljima, obrađeni izvori spominju u kontekstu prefiksalne tvorbe imenica, a neki prefiksi spominju se u kontekstu prefiksalne tvorbe pridjeva i glagola. Iako će se analiza u ovome radu baviti pridjevima, bit će spomenuti i prefiksi iz poglavlja o imenicama i glagolima, jer od tih vrsta riječi također nastaju i pridjevi koje će korpus obuhvaćati.

3.2.1. U hrvatskim gramatikama

Pretraživanjem tvorbe negacije u hrvatskim gramatikama, i specifično, prefiksalne tvorbe pridjeva negacijskim prefiksima, uspoređivane su različite gramatike, ali i različita izdanja istog autora. Naime, gramatika autora Babića pretraživana je u izdanjima iz 1991. i 2002., s obzirom na to da starije izdanje nije sadržavalo poglavlje o tvorbi pridjeva kojima se tvori negacija. Međutim, s obzirom na činjenicu da se u izdanju iz 1991. radi o *Nacrtu za gramatiku,* pretpostavka je bila da će prošireno i takozvano *poboljšano* izdanje koje se još naziva i *Velika hrvatska gramatika* sadržavati to poglavlje, kako ga sadrži i gramatika autora Težak i Babić iz 1992., kao i novija gramatika (Barić i sur., 2005), ali to nije bio slučaj.

[10] Vlastiti prijevod

Babić (2002) govoreći o prefiksima ističe da su, osim prefiksa *bez-* koji može biti i samostalna riječ jer je prijedlog, prefiksalni morfemi kao nesamostalne riječi brojni, ali navodi samo neke od njih koji sudjeluju u tvorbi negacije: *a-, ab-, de-, dis-, in-, kontra-*. Što se tiče *protu-* koji nije samostalna riječ, ali je prefiks u ovakvom načinu tvorbe, autor ga izdvaja kao alomorf prijedloga *protiv.*

Govoreći o prefiksima kojima se tvori negacija imenica, *Tvorba riječi u hrvatskome književnome jeziku* (Babić, 2002) za prefiks *a-* navodi da je sinoniman s prefiksima *ne-* i *bez-* te da u nekoliko složenica sa stranim osnovama označuje lišenost, odnosno da nema onoga što znači osnovna riječ ili da je suprotno od nje. Autor navodi primjere *amoraliza, asepsa, asimetrija, ateizam* itd. (Babić, 2002).

Prefiks *anti-* dolazi od grčkog *anti*, što znači protiv. Složenice s tim prefiksom označuju pojavu suprotnu značenju osnovne imenice, poput: *anticiklona, antisvijest, antitalent, antiteza* itd. Dodatno, Babić ističe da je u nekim složenicama jače naglašena aktivna suprotnost pojavi koju označuje osnovna imenica – *antialkoholizam, antifašizam* itd. Babić nadalje navodi da ako je u osnovi domaća riječ, složenice s *anti-* blago su stilski obilježene te je za stilsku neutralnost bolje upotrebljavati *protu-* (Babić, 2002).

Za prefiks *bez-* Babić tvrdi da je u imeničkoj tvorbi slabo plodan jer dolazi uglavnom od nekoliko starijih složenica ili u individualnoj tvorbi. Označava lišenost, nijekanje osnovnoga pojma ili njegovu suprotnost te navodi primjere *beskraj, besvijest, bezbroj* itd. (Babić, 2002). Međutim, primjer *bezbroj,* iako tvoren prefiksom koji uglavnom tvori negaciju, ne može se smatrati klasičnim primjerom negacije. Naime, *bezbroj* ne označava niti suprotnost broja niti stanje u kojem broj ne postoji – *bez broja,* već se odnosi na nešto što se ne može izbrojati jer je beskonačno. Stoga se u analizi pridjeva za potrebe ovoga rada pridjev *bezbrojan* neće tretirati kao negacija. Babić (2002) također spominje i imenice tvorene prefiksom *bez-* nastale od pridjeva koje su izrazito brojne, a više o njima bit će izneseno u nastavku rada.

Babić za prefiks *dis-* navodi da dolazi od latinskog *dis,* što znači *raz-,* ili grčkog *dys* što znači *ne.* U nekoliko složenica sa značenjem *ne-*, negira značenje osnove. Autor navodi primjere *disharmonija, disproporcija,* ali ističe da brojne riječi tvorene ovim prefiksom čine uskostručni nazivi ili se radi o riječima na tvorbenoj granici (Babić, 2002). Od ovih prefikasa, iako tvoreni od imenice, nastaju pridjevi poput *disharmoničan* i *disproporcionalan,* koji će iako nije riječ o čistoj prefiksalnoj tvorbi biti uključeni pretragom korpusa, kao i računalnom analizom.

Prefiks *kontra-* dolazi od latinskog *contra,* što znači protiv. U više složenica sa stranim i domaćim osnovama zapravo znači *protiv,* primjerice *kontrademonstracija, kontraigra, kontraprijedlog, kontranapad* te autor, kao i za prefiks *anti-* ističe da je bolje upotrebljavati *protu-* gdje se može. Dodatno navodi da u nekoliko složenica prefiks *kontra-* nosi i značenje *ispod, niži,* primjerice: *kontraadmiral, kontrabas,* odnosno označava hijerarhijski odnos (Babić, 2002). Međutim, riječi takvoga značenja nisu u okviru definicije negacije određene za ovaj rad te stoga takvi primjeri neće biti analizirani.

Prema Babić (2002) prefiks *ne-* izriče lišenost, odnosno niječe svojstvo osnovne riječi. Autor navodi primjere poput *nebiće, nebriga, neborac, neistina, nepar, nesporazum* itd. te dodaje da je nijekanje svojstva obično u glagolskih imenica poput: *neispunjenje, nepovjerenje, neshvaćanje* itd. Uz nijekanje, složenice dobivaju i dodatna značenja koja proizlaze iz nijekanja osnovnoga značenja. Tako Babić navodi složenice koje označuju suprotno svojstvo ili stanje, poput *nemir, nepravda, nerad.* Također izdvaja riječi u kojima *ne-* daje značenje *loš, zao,* poput *nebrat, nečovjek, nedjelo, nezgoda, nesreća, neprijatelj* itd. Naposljetku, autor ističe još jedno značenje prefiksa *ne-,* a ono je *nezgodan, nepriličan,* poput *nedoba, nevrijeme, neprilika* (Babić, 2002). S druge strane, potrebno je primijetiti da iako se zaista radi o tvorbi nove riječi dodavanjem niječnoga prefiksa, ne tvori svaki niječni prefiks negaciju u smislu u kojem se ona tumači u ovome radu. Imenica *nevrijeme* dodavanjem prefiksa *ne-,* ne negira *vrijeme* niti označava pojam suprotan vremenu. Međutim, s obzirom na to da bi se riječ *vrijeme* mogla protumačiti kao pozitivna riječ i u ovom slučaju pozitivna (ili neutralna) osnova, dodavanjem prefiksa *ne-* ona ipak dobiva značenje koje se može protumačiti kao *negativno značenje,* bez obzira na to što se ne radi o negaciji u doslovnom smislu riječi, ili barem u smislu obuhvaćenom definicijom negacije za potrebe ovoga rada.

Posljednji negacijski prefiks kojeg ovaj autor navodi u poglavlju o tvorbi imenica je *protu-*, ističući da dolazi u složenicama sa značenjem suprotni ili protivni, poput *protuakcija, protudokaz, protuotrov* itd. (Babić, 2002). Ova gramatika spominjala je i neke druge prefikse u poglavlju o pridjevima, o čemu će biti riječi u odlomcima koji slijede.

Babić (1991. i 2002.) u svojoj gramatici pridjeve dijeli na opisne i odnosne, a svaku vrstu podijelio je i prema prefiksalnoj, sufiksalnoj i prefiksalno-sufiksalnoj tvorbi. Ova gramatika kod opisnih pridjeva sadrži i poglavlje *Tvorba opisnih pridjeva s posebnim značenjem,* gdje postoje potpoglavlja: tvorba pridjeva koji znače mogućnost, opskrbljenost i obilje te sličnost i tvorba pridjevnih umanjenica i pojačanih pridjeva. Međutim, kao što je već

navedeno, nema potpoglavlja *Tvorba pridjeva koji znače nijekanje, lišenost, suprotnost* – što Težak i Babić (1992) te Barić i sur. (2005) imaju.

Babić (1991. i 2002.) ističe da je u prefiksalnoj tvorbi pridjeva plodna samo tvorba prefiksom *ne-,* dok je tvorba s ostalim prefiksima manje plodna ili neplodna. Međutim, autor navodi i ostale prefikse koji sudjeluju u tvorbi pridjeva, a ovdje će biti izdvojeni prefiksi koji tvore negaciju. U poglavlju o prefiksalnoj tvorbi opisnih pridjeva, negacijski prefiksi koje autor navodi su *a-, ab-, anti-, i-, in-* i *protu-*. U poglavlju o prefiksalnoj tvorbi odnosnih pridjeva, negacijski prefiksi koje autor navodi su prefiksi *anti-, kontra-, ne-* i *protu-.* Prefikse *bez-* i *izvan-* autor navodi u poglavlju o prefiksalno-sufiksalnoj tvorbi odnosnih pridjeva (Babić, 2002). Naposljetku, u poglavlju o tvorbi glagola, u potpoglavlju naziva *Prefiksi pojedinačno,* od prefikasa kojima se tvori negacija spominju se prefiksi *de-* i *dis-* (Babić, 2002). Stoga se može primijetiti da negacija na morfološkoj razini nije obrađena zasebno, već se određeni prefiksi spominju na različitim mjestima, što je nepraktično za pronalazak informacija. Također, za velik broj tih pridjeva autor ne opisuje njihovo značenje i etimologiju, već samo navodi primjere.

Za prefiks *a-,* slično kao u poglavlju o imenicama, Babić ističe da označuje nijekanje ili odsutnost kakve osobine te je stoga sinoniman s prefiksima *bez-* i *ne-* te dolazi na pridjeve stranog podrijetla, uglavnom grčke i latinske, navodeći primjere *alogičan, amoralan, apolitičan, asocijalan* itd. (Babić, 1991 i 2002). Za prefiks *ab-* u tvorbi pridjeva Babić samo navodi primjere, poput *abnormalan* i *aboralan.* Autor nadalje navodi prefiks *anti-* za kojeg spominje da se često pojavljuje s odnosnim pridjevima, ali ponekad i s opisnima, navodeći primjere *antireligiozan, antisocijalan* itd. (Babić, 2002). Prefiks *bez-* autor spominje samo navođenjem primjera, poput *bezgotovinski, bezvlasnički, bezatomski,* a navodi i alomorfsku inačicu *bes-* u primjeru *besteţinski* (Babić, 2002).

U poglavlju o tvorbi glagola autor za prefiks *de-* ističe da se javlja u dva alomorfa: *de-* i *dez-,* ističući da ako riječ počinje suglasnikom, dolazi prefiks *de-,* a ako počinje samoglasnikom, dolazi prefiks *dez-.* Za prefiks *dis-* autor navodi da dolazi s nekoliko glagola stranoga podrijetla. Međutim, ovi prefiksi ne smatraju se prefiksima kojima se tvore pridjevi, jer pridjevi koji sadrže ovaj prefiks nastali su od glagola koji ih je već sadržavao, stoga nije riječ o prefiksalnoj, već o sufiksalnoj tvorbi. Bez obzira na to, iako se ne radi o „čistoj" prefiksalnoj tvorbi, za ovaj rad je važno da pridjev bez toga prefiksa nije negiran, a dodavanjem toga prefiksa jest. Postavljanjem upita prilikom pretraživanja korpusa, ovakvi pridjevi svakako bi bili obuhvaćeni – jer radi se o riječi koja započinje određenim prefiksom i koja je označena

kao pridjev. Primjeri pridjeva sa spomenutim prefiksima su *deblokiran* (nastao od glagola *deblokirati*), *dezinficiran* (od glagola *dezinficirati*) te *diskvalificiran* (od glagola *diskvalificirati*).

Za prefiks *i-* Babić spominje da dolazi kod pridjeva stranoga podrijetla, u značenju prefiksa *ne-* te navodi primjere poput *ilegalan, imoralan, iracionalan* itd. te kod prefiksa *in-* navodi primjer *intranzitivan* (Babić, 2002).

U poglavlju o prefiksalno-sufiksalnoj tvorbi odnosnih pridjeva, Babić spominje i prefiks *izvan-* navodeći pridjeve poput *izvanbračni, izvaneuropski, izvanstranački* itd. te ističući da se ponekad umjesto *izvan-* može upotrijebiti i prefiks *ne-* (Babić, 2002).

Za najčešći prefiks u prefiksalnoj tvorbi pridjeva, prefiks *ne-,* Babić navodi da taj prefiks negira značenje osnovnoga pridjeva. Nadalje, autor ističe da su takvi pridjevi često složeni od pridjeva koji završavaju sufiksom *–(a)n,* poput *nebitan, neiskusan* itd. iako ih ima i s drugim završecima – *nečitak, nečitljiv, neslikovit*, ali i s nultim sufiksom, poput *nedragØ, nevještØ* itd. Također navodi da je prefiks *ne-* čest i s glagolskim pridjevima – *neangažiran, neviđen* itd. koji će biti uključeni u analizu za ovaj rad jer bi ih se prikupljanjem korpusa ionako obuhvatilo. Međutim, autor ističe da prefiks *ne-* ne dolazi sa svakim pridjevom, odnosno navodi ograničenja u tvorbi:

a) „ako pridjev ima više opreka: *crven: plav, zelen, žut…*

b) ako se opreka izriče drugim pridjevom: *debeo-mršav, desni-lijevi, jeftin-skup, …*

c) ako pridjev ima negativan smisao: *srdit, zao, zloban, …*“ (Babić, 1991 i 2002), što je u skladu i s Horn (2001), Zovko Dinković (2013) te Kovačević (2016) koja se poziva na Zimmer (1964).

S druge strane, autor također ističe da ova ograničenja nisu apsolutna jer se i uz takve opreke može pojaviti negacija. Primjerice postoje pridjevi poput *nečist, nelijep, neveseo* itd. – iako ti primjeri imaju opreku izrečenu drugim pridjevom, odnosno *prljav, ružan, tužan,* što je već spomenuto u prethodnom poglavlju.

Naposljetku, za prefiks *protu-* navodi da se pojavljuje s nekoliko opisnih pridjeva, poput *protuprirodan, protuzakonit* itd. (Babić, 2002) te ističe da „označuje da je što protiv onoga što znači pridjevna osnova“ (Babić, 1991:412). Ovaj prefiks autor spominje i u kontekstu prefikasa *anti-* i *kontra-.* Naime, navodi da *anti-* ima značenje domaćega prefiksa *protu-* te iako dolazi s pridjevima i stranog i domaćeg podrijetla, ističe da je s pridjevima stranoga podrijetla bolje

upotrebljavati prefiks *protu-* (Babić, 1991). Također ističe da su složenice s domaćim pridjevima stilski obilježene, a da su složenice s prefiksom *protu-* stilski neutralne te je u neutralnom značenju također bolje upotrebljavati prefiks *protu-*. Navodeći primjer za prefiks *kontra-: kontraobavještajni*, također navodi da je bolje upotrebljavati pridjev *protuobavještajni*.

Barić i sur. (2005) negaciju pridjeva spominju u poglavlju o prefiksalnoj tvorbi, unutar potpoglavlja o odnosnim pridjevima i posebno o opisnim pridjevima. Točnije, potpoglavlje *Opisni pridjevi* sadrži dodatno potpoglavlje *Pridjevi sa suprotnim ili zanijekanim svojstvom.* Navode da se svojstvo suprotno onome izrečenom osnovnim pridjevom ili zanijekano svojstvo osnovnoga pridjeva izriče prefiksom *ne-*. Ističu da uz jednostavne pridjeve kao osnovnu riječ dolaze „izvedeni pridjevi i popridjevljeni glagolski pridjev trpni“ te navode primjere: *mio-nemio, sretan-nesretan, veseo-neveseo, neizdržljiv, neosvojiv, nepobjediv, nebran, neizrečen, neobrijan, neželjen* itd. (Barić i sur., 2005:369).

Autori spominju da pridjevi mogu označavati i nepotpuno zanijekano svojstvo prefiksima *kvazi-,* poput *kvazidječji*, *nazovi-,* poput *nazoviteorijski* te *pseudo-,* poput *pseudopučki*. Međutim, tvorba negacije prefiksima koji označavaju nepotpuno nijekanje nije predmet ovoga rada stoga se takvi pridjevi neće uključiti u korpus za prepoznavanje i označavanje negacije. Autori u kontekstu odnosnih pridjeva navode i prefikse *anti-, kontra-* i *protu-,* poput *antiratni, kontraobavještajni* i *proturatni* (Barić i sur., 2005).

Težak i Babić (1992) navode podjelu pridjeva na opisne pridjeve koje još nazivaju kvalitativnima i odnosne koje nazivaju relacijskima. S druge strane, postoje i druge podjele pridjeva, poput podjele na opisne, gradivne i posvojne, a razlike između pojedinih vrsta nisu uvijek sasvim jasne. Naime, sami autori navode da i odnosni (relacijski) pridjevi mogu po svom značenju biti i kvalitativni te da se gradivni također često upotrebljavaju u kvalitativnom značenju. Međutim, ovdje naglasak nije na tim specifičnim vrstama pridjeva, ali zbog činjenice da autori negaciju spominju u poglavlju o tvorbi opisnih pridjeva te da će se pridjevi za rječnik u NooJ-u klasificirati prema podjeli na opisne i odnosne, navest će se definicije opisnih i odnosnih pridjeva zbog samoga razlikovanja. Opisne pridjeve autori definiraju kao one koji „označuju različita tvarna (materijalna, prirodna, fizička) i netvarna (misaona, duhovna, psihička) svojstva predmeta.“ S druge strane, odnosni pridjevi „izriču svojstva koja predmet (biće, pojava) ima u odnosu na koji drugi predmet“ (Težak i Babić, 1992:99).

Navedeni autori negaciju pridjeva obrađuju u poglavlju *Tvorba opisnih pridjeva s posebnim značenjem,* unutar posebnoga potpoglavlja *Tvorba pridjeva koji znače odsutnost ili*

nijekanje osobine. Ono što se može primijetiti je da za razliku od Babić (1991. i 2002.) i Barić i sur. (2005), ovi autori ne spominju prefiksalnu tvorbu odnosnih pridjeva, već samo opisnih.

Za ovakvu vrstu tvorbe pridjeva spominju samo prefikse *bez-* i *ne-*. Ističu da prefiks *bez-* može biti sa ili bez sufiksa te navode primjere poput *bezbrižan, bešuman, besan, bezruk, bezub* itd. (Težak i Babić, 1992), gdje su vidljive i glasovne promjene o kojima su autori posebno pisali, a koje će biti navedene na kraju ovoga potpoglavlja.

Za tvorbu prefiksom *ne-* Težak i Babić (1992) ističu da takve složenice niječu svojstvo koje je izrečeno drugim dijelom složenice ili znače suprotno svojstvo – što se uklapa u definiciju negacije ovoga rada, te navode primjere poput *neiskren, nečist* itd. Posebno ističu da se prefiks *ne-* često slaže s glagolskim pridjevima trpnim, poput *neiskazan, neispavan, neobrijan, neograničen* itd. Kao što je već navedeno, analiza negacije pridjeva uključivat će i glagolske pridjeve, jer su u hrWaC-u označeni kao pridjevi pa su obuhvaćeni pretragom.

Vrijedne spomena su i glasovne promjene koje su neizbježne u prefiksalnoj tvorbi. O njima pišu gramatike Težak i Babić (1992) i Babić (2002). Babić (2002) spominje da prefiksi koji završavaju na šumnik (obstruent) dolaze u svojim alomorfima: *bez-, bes-, beš-, be-, beza-* (*bezbrad, bescjen, bešćutan, bespolan, bezazlen*). Nadalje navodi da za neke moguće alomorfe nema primjera te kao takav prefiks, odnosno alomorf izdvaja *bež-*, iako postoje primjeri poput *beživotan, bežičan.* Dodatno ističe da se prefiks s dodanim blagozvučnim *a* (*beza-*) naziva prošireni prefiks.

Starija gramatika, (Težak i Babić, 1992) daje detaljniji opis takvih glasovnih promjena. Autori tvrde da takve promjene nisu toliko složene kao u sufiksalnoj tvorbi, jer se promjene u prefiksalnoj tvorbi vrše po jasno utvrđenim pravilima o jednačenju i ispadanju frikativa. Autori ističu da u glasovnim promjenama sudjeluju samo prefiksi koji završavaju na suglasnik te iako navode primjere za više prefikasa, ovdje će biti navedeni samo primjeri za prefiks kojim se tvori negacija, prefiks *bez-*.

U slučaju prefiksa *bez-* koji završava zvučnim glasnikom, on se jednačenjem po zvučnosti zamjenjuje sa svojim bezvučnim parnjakom, ako i osnova počinje bezvučnikom:

- bez+kraj+an → beskrajan
- (dodatni primjer kojeg autor ne navodi: bez+šum+an → bešuman)

Po mjestu tvorbe, *bez-* se ispred osnova koje počinju sa č, ć zamjenjuje sa š:

- bez+ćut+an → bešćutan

Gubi se završni suglasnik u prefiksu ako osnova počinje istim šumnikom (obstruentom):

- bez+zvuk+an → bezvučan
- bez+zub → bezub

Završni suglasnik u prefiksu također se gubi kad bi jednačenjem postao isti s početnim frikativom osnove:

- bez+život+an → beživotan
- bez+žica+an → bežičan.

3.2.2. U ostaloj hrvatskoj literaturi

Šarić (1992) ističe da u prefiksalnoj tvorbi sudjeluje niz prefikasa koji izražavaju odrično značenje, poput *ne-, a-, dis-, ab-, sub-, bez-, kontra-, protu-, anti-* itd. te da riječi kojima je osnova stranoga podrijetla mogu svoj antonimni parnjak tvoriti stranim prefiksom, poput *antifašist, ateist, deblokirati, demobilizirati* itd. Što se tiče prefiksa *a-* navodi da izriče djelomičnu ili potpunu suprotnost te navodi primjere poput *asimetrija, asepsa.* Prefiks *a-* niječe ili izriče odsutnost neke osobine te da je srodan prefiksu *ne-*, a uglavnom dolazi na riječi stranog podrijetla, primjerice grčke i latinske (*alogičan, apolitičan, amoralan*). Za prefiks *anti-* navodi da imenuje pojavu suprotnu značenju osnovne imenice *antitalent, antirodoljub.* Dodatno ističe da prefiks *anti-* najčešće dolazi u odnosnim i opisnim pridjevima od kojih se izvode imenice, poput *antireligioznost, antisocijalnost* itd. Prefiksom *bez-* izriče se krajnost, nijekanje pojma ili njegova suprotnost, poput imenice *bezuspješnost.* Prefiks *kontra-* izriče krajnju suprotnost te se može zamijeniti i prefiksom *protu-* (*kontranapad/protunapad*). Prefiks *protu-*, poput prefiksa *anti-* iskazuje protivnost imenice osnovnoj riječi (*protuteža*). Za prefiks *dis-* također navodi da izriče aktivnu suprotnost (*disharmonija*). O prefiksu *ne-* piše da u glagolskih imenica ukazuje na lišenost, nijekanje svojstva (*nepostojanje, neraspoloženje*) te da iskazuje suprotno svojstvo u imenica poput *nesloga, nesvijest, neprijatelj*. Nadalje, za prefiks *ab-* samo navodi primjere i to pridjeva poput *abnormalan, aboralan*. Za prefiks *i-* kaže da dolazi u stranih riječi u značenju *ne*, poput *ilegalan, imoralan, iracionalan* te uz njega navodi i inačice *in-* i *im-* u riječima poput *intranzitivnost*, *imperfektnost* itd. koje su tvorene od pridjeva *intranzitivan, imperfektan*. Autorica naposljetku navodi da je u latinskom osnovni prefiks *in-*, a da su *i-* i *im-* njegove inačice (Šarić,1992). Vrijedi spomenuti da analizirane gramatike prefiks *im-* ne spominju.

Zovko Dinković u svojoj knjizi *Negacija u jeziku* koja se bavi kontrastivnom analizom negacije u hrvatskom i engleskom jeziku, spominje prefikse *ne-, bez-* (*bes-*) i *in-* (*im-*) u kontekstu hrvatskoga jezika te prefikse *un-, in-, dis-, an-* i *a-* u engleskom jeziku. Ističe da se negacija pridjeva, priloga i imenica u hrvatskom jeziku uglavnom tvori prefiksom *ne-*, primjerice *prijatelj*/*neprijatelj*, a u nešto manjoj mjeri upotrebljavaju se ostali spomenuti prefiksi i njihovi morfofonološki ostvaraji (Zovko Dinković, 2013).

Naposljetku, Kovačević (2016) u analizi hrvatskog crkvenoslavenskog jezika govoreći o tvorbi negacije na morfološkoj razini spominje prefikse *ne-, bez-* i *ni-,* umjesto prefiks upotrebljavajući naziv predmetak, kojeg spominje i Babić (2002). Za prefiks *ne-* ističe da je nositelj općeg niječnog značenja, dok *bez-* i *ni-* imaju posebno niječno značenje. Odnosno, prefiks *ne-* tiče se označavanja referenta s obzirom na ono što nije te autorica navodi da su značenjske sastavnice poput suprotnosti, proturječnosti i negativnosti izvedene iz takvog osnovnog značenja. S druge strane, prefiks *bez-* niječnost ostvaruje kao odsutnost, a prefiks *ni-* kao nepostojanje (Kovačević, 2016). Autorica nadalje navodi da specijalizacijom značenja dolazi i do pada broja primjera u korpusu pa je tako najviše primjera s prefiksom *ne-,* a broj značajno pada za prefikse *bez-* i *ni-,* što je potvrđeno i prikupljenim korpusom za ovaj rad te će učestalost, odnosno plodnost svakog pojedinog prefiksa biti prikazana u nastavku ovoga rada.

Kovačević (2016) također ističe da su leksemi s prefiksom *ne-* pridjevi, imenice, prilozi i glagoli, a da su osnove imenske i glagolske. S druge strane, Horvat i Štebih Golub (2012) navode da je prefiks *ne-* tvorbeno najčešći kod pridjeva, zatim imenica, priloga i zamjenica. Nadalje, Kovačević (2016) za prefiks *bez-* navodi da su leksemi tvoreni njime također pridjevi, imenice, prilozi i glagoli, ali su im osnove samo imenske.

Autorica ističe da se po pitanju nastanka imenica s prefiksom *bez-* prednost u literaturi daje sufiksalnoj, odnosno prefiksalno-sufiksalnoj tvorbi. Napominje i da pri utvrđivanju tvorbenog načina treba uzeti u obzir smjer tvorbe i motiviranosti, jer primjerice *bezosjećajnost* nije izvedenica prefiksalno-sufiksalne tvorbe, već sufiksalne s obzirom da joj je polazna riječ *bezosjećajan,* riječ koja u sebi već sadrži prefiks *bez-* (Kovačević, 2016). Primjerice, pridjev

bezidejni dolazi od *bez+ideja* (imenica), ali imenica *bezidejnost* i takve tvorene sufiksom *-ost* nastaju od pridjeva, što potvrđuje i etimologija prema *Hrvatskom jezičnom portalu*[11].

Naime, Kovačević navodi da su pridjevi najbrojniji članovi leksika s prefiksom *bez-* te da upravo iz tog razloga postoji mogućnost da su imenske izvedenice motivirane pridjevima. U prilog tomu govori i činjenica da su imenice s prefiksom *bez-* uglavnom apstraktne imenice koje označavaju *pojavu* obilježenu nekim *svojstvom,* odnosno odsutnošću toga svojstva (Kovačević, 2016).

Ista autorica također navodi da se u niječne riječi s prefiksom *bez-* ubrajaju i alomorfi toga prefiksa, koji su svi uzeti u obzir i prilikom prikupljanja korpusa za potrebe ovoga rada, a to su: *be-, bes-, beš-, bež-* (Kovačević, 2016). Prefiks *bez-* izražava odsutnost, nepostojanje onoga imenovanoga osnovnom riječju. Izvedenicama s tim prefiksom najčešće se prevode grčke i latinske posvojne složenice s prefiksom *in-*. U prilog tomu govori i činjenica da se one katkada doista i prevode pomoću prijedložnoga izraza, poput *lat. insomnis – bez sna.* (Kovačević, 2016).

Naposljetku, riječi s prefiksom *ni-* mogu biti zamjenice, prilozi i pridjevi, ali osnove su im isključivo zamjeničke. Odnosno, autorica navodi da se prefiksom *ni-* tvore niječne zamjenice, zamjenički prilozi i zamjenički pridjevi (Kovačević, 2016). Primjerice, niječne zamjenice poput *nitko, ništa, ničiji,* zamjenički prilozi poput *nikako, nikada, nigdje* i zamjenički pridjevi, poput *nijedan, nikakav, nikoji.* Zamjenički pridjevi neće biti obuhvaćeni korpusom za računalnu analizu negacije pridjeva.

Pregled prefikasa koje spominju analizirani autori nalazi se u Tablici 1, a u nju nisu uključeni prefiksi kojima se ne tvori ova vrsta negacije, iako ih određeni autori spominju, poput *sub-* (Šarić, 1992) i *kvazi-, nazovi-, pseudo-* (Barić i sur., 2005).

[11] http://hjp.znanje.hr/

Tablica 1 Pregled negacijskih prefikasa prema analiziranim autorima

PREFIKSI: AUTORI:	A-	AB-	AN-	ANTI-	BEZ-	DE-	DEZ-	DIS-	I-	IM-	IN-	IZVAN-	KONTRA-	NE-	NI-	PROTU-
Težak i Babić (1992.)					+									+		
Šarić (1992.)	+	+		+	+			+					+	+		+
Babić (2002.)	+	+		+	+	+	+	+	+		+	+	+	+		+
Barić i sur. (2005.)				+									+	+		+
Zovko Dinković (2013.)					+					+	+			+		
Kovačević (2016.)					+									+	+	

U poglavlju o prikupljanju korpusa, gdje će se analizirati prefiksalna negacija pridjeva *Hrvatskog mrežnog korpusa – hrWaC* (Ljubešić i Erjavec, 2011), iznijet će se informacije o pojedinom prefiksu – njegovu podrijetlu i značenju i prema *Hrvatskom jezičnom portalu*[12].

3.2.3. U stranoj literaturi

Govoreći o negacijskim prefiksima, De Clercq (2013, prema De Clercq i Vanden Wyngaerd, 2017) navodi da jezici imaju raznolike označivače negacije, a dijeli ih u četiri kategorije – prema funkciji, semantici, dometu i razlikama u slaganju. Autorica nadalje objašnjava da se ne mogu sve vrste negacijskih prefikasa slagati jedan uz drugoga, kao niti, primjerice, na riječi već negativnog značenja, što u istom radu autorica potvrđuje svojom analizom.

Nadalje, Horn (2001:280) komentirajući prefikse engleskoga jezika *non-* i *un-* ističe da „afiks *non-* negira 'vidljivo', denotativno značenje, a afiks *un-* negira 'pozadinsko', konotativno značenje riječi“.[13] Horn za ovu vrstu negacije i s ovim prefiksom navodi da po formuli *nonX*, X može biti pridjev ili imenica. Dodatno, jedna značajka koju prefiks *non-* dijeli s ostalim prefiksima je mogućnost promjene kategorije, odnosno vrste riječi svoje osnove. Primjerice kako prefiksi *un-, de-* i *dis-* mogu promijeniti imenice u glagol (*unearth, unnerve, deflea, deplane, disbar, dismember*), tako u primjeru *nonskid,* prefiks pretvara glagol u pridjev (Horn, 2001). Konačno, što se tiče prefiksa *non-*, Horn ističe da se *non-* uglavnom pridaje uz

[12] http://hjp.znanje.hr/

[13] Vlastiti prijevod

neoznačene članove opozicije, primjerice u demografskoj taksonimiji postoje *nonwhites* (engl.), ali ne i *nonblacks* ili *noncoloreds* (Horn, 2001).

Kjellmer (2005) navodi da se ponekad upotrebljavaju različiti prefiksi za značenjski slične riječi. Primjerice navodi parove poput *disconnected/unplugged, incomplete/unfinished, inedible/uneatable* te *non-Ibearian/un-American.* Kao moguće objašnjenje navodi etimologiju, ističući za odnos između prefikasa *un-* i *in-* da se *in-* upotrebljava kod riječi latinskoga korijena, a *un-* u ostalim slučajevima.

Slično ističe i Joshi (2012), koji u poglavlju *Synonymy of negative affixes* piše o tome da i prefiks *dis-* i prefiks *mis-*, primjerice u pojmovima *displace, misplace* (engl.) u oba termina znače isto – „ne imati nešto na željenom mjestu", ali razlika je u dobrovoljnoj, svjesnoj kontroli nad akcijom. Odnosno, *displace* označava namjernu radnju, a *misplace* slučajnu. Također spominje upotrebu različitih afikasa na istoj osnovi, kao što su *a-, in-, non-* u riječima *amoral, immoral, non-moral*, s tim da postoje razlike u značenju tih riječi. Uporaba različitih prefikasa na istoj osnovi događa se i u hrvatskom jeziku, primjerice *imoralan/nemoralan, kontraobavještajni/protuobavještajni* itd., gdje u navedenim primjerima nema razlike u značenju, ali postoje i primjeri gdje uporaba različitih prefikasa doprinosi i različitom značenju. Primjerice, prema *Hrvatskom jezičnom portalu*[14]*, asocijalan/antisocijalan* te *anormalan/abnormalan* nemaju u potpunosti isto značenje.

Joshi (2012) kao što je već navedeno, za prefiks *un-* navodi da tvori izravnu negaciju (engl. *happy/unhappy*), a prefiks *in-* neizravnu (engl. *famous/infamous*). Za hrvatski jezik se može primijetiti da isti prefiks može tvoriti i izravnu i neizravnu negaciju, poput već navedenih primjera *sretan/nesretan* koji se tiču izravne negacije te *volja/nevolja* koji se tiču neizravne negacije. Također, spominjući upravo pojmove *famous/infamous,* može se primijetiti da su u hrvatskom jeziku moguće dvije inačice prijevoda, čiji negacijski parnjaci nemaju jednako značenje. Naime, prvi mogući prijevod je *poznat/ozloglašen,* a ne *nepoznat*, jer *nepoznat* nema isto značenje kao *ozloglašen.* Međutim, u drugom mogućem prijevodu *slavan/neslavan,* negativni parnjak *neslavan* koristi se upravo u kontekstu nekoga tko je *ozloglašen,* odnosno *infamous.* Neki od primjera iz hrWaC-a su *neslavna scena, neslavno gostovanje, neslavni kraj,*

[14] http://hjp.znanje.hr/

gdje pridjev *neslavan* nije u istom značenju kao *nepoznat,* već u značenju nečega „poznatoga po lošem“, odnosno *infamous.*

O negacijskim prefiksima pisao je i Jespersen (1917). Prefikse *un-* i *in-* naziva najvažnijim negacijskim prefiksima te *un-* navodi kao izvorni engleski oblik, dok je *in-* latinski oblik koji u engleski dolazi preko brojnih francuskih i engleskih riječi. Navodi da *in-* prema pravilima latinske fonologije ima alternativne oblike *ig-* (*ignoble*), *im-* (*impossible*), *il-* (*illiterate*) i *ir-* (*irreligious*). Ovdje se može primijetiti da se u hrvatskom jeziku s određenim navedenim prefiksima događa stapanje ukoliko prefiks završava konsonantom kojim osnova započinje. Primjerice pridjev *irelevantan* je stoga tvoren prefiksom *i-,* a ne *ir-,* dok je u engleskom jeziku tvoren prefiksom *ir-,* gdje se ne događa stapanje pa pridjev glasi *irrelevant.* Također, hrvatski primjer *iregularan,* na engleskom je *irregular* itd.

Jespersen nadalje ističe da se *un-* često preferira ispred kraće riječi, a *in-* ispred duže, izvedene od nje. Primjerice *unable/inability*, *unjust/injustice*, *unequal/inequality*. Ono što se ovdje može primijetiti je da se u ovim pridjevima radi o razlici pridjev/imenica, gdje je pridjev tvoren od prefiksa *un-* i čini kraću riječ, a prefiks *in-* daje imenicu koja je duža riječ. Međutim, vrijedi spomenuti da u prefiksalnoj tvorbi negacije hrvatskoga jezika situacija nije takva. Naime, isti primjeri u hrvatskom jeziku negirani su jednakim prefiksom, prefiksom *ne-,* radilo se o dužoj ili kraćoj riječi, odnosno imenici ili pridjevu: *nesposoban/nesposobnost, nepravedan/nepravda, nejednak/nejednakost.*

Nadalje, za prefiks *un-* Jespersen piše da se u tvorbi imenice mijenja i nekim drugim prefiksima, poput primjera *unfortunate/misfortune, unsatisfactory/dissatisfaction, uncomfortable/discomfort*. U hrvatskom jeziku situacija je ponovno drugačija, odnosno nevezano za promjenu vrste riječi, prefiks u ovim primjerima ostaje isti: *nesretan/nesreća, nezadovoljan/nezadovoljstvo, neudoban/neudobnost.* To je vidljivo i na već spomenutom prefisku *bez-* kojim velik broj imenica nastaje od pridjeva, zadržavajući isti prefiks: *bezazlen/bezazlenost, bezopasan/bezopasnost* itd.

Za prefiks *dis-* Jespersen navodi da je prefiks iz latinskoga jezika te da može imati različita značenja, ali i da označava čistu negaciju. Navodi primjere *dissimilar, dishonest, disagree, disuse, dislike* itd. te ističe da prefiks *dis-* „uglavnom implicira suprotnu, više nego

proturječnu negaciju“[15] (Jespersen, 1917:146). O prefiksu *non-* piše da se radi o latinskome prefiksu kojim se oblikuje velik broj riječi kada *un-* i *in-* nisu dostupni. Ističe da su za njegovu uporabu najodgovorniji pravni pojmovi gdje se on najviše i koristi, primjerice *non-payment, non-regardance*, *non-performance* itd. (Jespersen, 1917). Za *an-* i *a-* navodi da se radi o grčkim prefiksima, gdje se *an-* koristi ispred vokala, a *a-* ispred konsonanata te su etimološki identični latinskima *un-* i *in-*. Najčešći su u grčkim riječima, poput *anarchy, amorphous, achromatic*, ali također se u rjeđim slučajevima koriste i u engleskom jeziku za tvorbu novih riječi od latinskih korijena, poput *amoral, asexual* itd. Konačno, prefiks *no-* Jespersen (1917) spominje kao zamjenicu koja se ponekad koristi kao vrsta prefiksa te navodi primjere poput *no-education*, *no-thoroughfare, no-ball* itd.

Naposljetku, Csapo (1977) navodi da postoji pet negacijskih prefikasa u engleskom jeziku *un-, in-, non-, dis-* i *a-* te jedan sufiks *-less.* Autor nadalje navodi da se prefiks *a-* pojavljuje u riječima grčkog podrijetla koje se uglavnom upotrebljavaju u polju prirodoslovlja i medicine, odnosno radi se o različitim tehničkim pojmovima. Također, što se tiče njegove produktivnosti, prefiks *a-* stvara i proturječne i suprotne pojmove. Za prefiks *un-* autor ističe da je najproduktivniji s pridjevima i prilozima te odgovara *jednostavnoj negaciji: „*X dodavanjem *un-* postaje ne-X“. Tvori proturječnu negaciju, posebno s engleskim pridjevima koji završavaju na *-able* i *-ible,* kao što je *unadaptable*, a tvori i suprotnosti, primjerice *unjust* implicira suprotno od *just* itd. (Csapo, 1977). Međutim, ponekad je upotreba negacije s ovim prefiksom ograničena, odnosno *unfair* se odnosi na moralni smisao, ali ne i na vanjski izgled. Za prefiks *non-* autor navodi da tvori samo proturječne pojmove, ne i suprotne, a najčešće se upotrebljava s vlastitim imenicama i pridjevima, poput *non-English, non-American,* međutim nema pogrdno značenje, poput *un-English* (Csapo, 1977).

[15] Vlastiti prijevod

4. Kognitivna obrada negacije

Hoosain (1973) je svojim istraživanjem pokušao utvrditi obrasce i poteškoće u obradi negacije, uključujući negaciju na razini rečenice česticom *ne,* ali i na razini riječi kod pridjeva negiranih prefiksom *un-,* primjerice *„He is not happy.“* u usporedbi s *„He is unhappy.“* Svojim istraživanjem pokazao je da pridjevi s prefiksom *un-* nisu značajno teži za obradu nego isti pridjevi bez tog negacijskog markera. Odnosno, u navedenom primjeru *unhappy* nije značajno teži za obradu od *happy,* što bi moglo biti iznenađujuće s obzirom da su pridjevi s prefiksom *un-* rjeđi nego njihovi antonimski parovi koji ne sadrže taj prefiks, a sami negirani pridjevi također dodaju i slog više cjelokupnoj dužini rečenice. Međutim, pridjevi negirani s *not* značajno sporije su obrađeni nego isti ti pridjevi negirani prefiksom *un-*. Ovi rezultati impliciraju da, psihološki gledano, pridjevi poput *unhappy* i ostali tvoreni na isti način funkcioniraju više kao jedinstvena riječ, nego kao pridjev i negacija, odnosno rezultati sugeriraju da se radi o punoznačnim jedinicama (Hoosain, 1973).

Slično je utvrdio i Sherman u istraživanju objavljenom iste godine, kojim je pokušao utvrditi učinke negacije na obradu rečenice (Sherman, 1973). Istraživanje je pokazalo da negacijski prefiks otežava obradu, ali da je prefiksalna negacija lakša za obradu nego rečenična, ostvarena rečeničnim negatorom *not.* Sherman ističe da *un-* zapravo ima isto značenje kao *ne,* primjerice *unwilling* je isto što i *not willing* (1973), međutim *not* proizvodi rečeničnu negaciju, dok prefiks *un-* ne proizvodi (Klima, 1964, prema Sherman, 1973). Također, kako navodi autor, dok se *unwilling* može činiti sinonimnim konstrukciji *not willing,* rečenica *„John doesn't want to go and Frank is unwilling to go either“* u najmanju ruku smatra se čudnom, ako ne i negramatičnom (Sherman, 1973). Autor tvrdi da negacija bilo kakve vrste rečenicu čini težom za obradu (Just i Carpenter, 1971, prema Sherman, 1973) pa bi stoga i prefiksalna negacija pridjeva trebala otežati obradu rečenice, u usporedbi s njihovim antonimima. Iz tog razloga je u svoje istraživanje obrade negacije uključio antonimne pridjeve poput *correct-incorrect, possible-impossible* i slične.

Svojim istraživanjem potvrdio je da negacijski prefiks uzrokuje manje poteškoća pri obradi rečenice nego rečenična negacija česticom *ne,* odnosno engleskim rečeničnim negatorom *not,* a potkrjepljuje to ističući da je, povijesno gledano, dodavanje prefiksa *un-* motivirano kognitivnom ekonomičnošću (Sherman, 1973). Međutim, ovaj rezultat iznenađuje s obzirom na činjenicu da je *not* puno češći u engleskom jeziku nego negacijski prefiksi, a svejedno je teži za obradu. Autor to objašnjava pozivajući se na Miller i McNeill (1969, prema

Sherman, 1973), koji su istaknuli da je obrtanje značenja riječi koje uzrokuje *un-* psihološki manje kompleksno nego obrtanje cjelokupnog značenja rečenice, koje uzrokuje *not.* Nadalje, Sherman ističe da slušač čim čuje *un-* zna puno više o strukturi rečenice nego kad se pojavi *not.* Naposljetku, kada bi korištenje negacijskog prefiksa uzrokovalo jednaku poteškoću prilikom obrade kao i negacijska čestica, tada bi samo postojanje negacijskog prefiksa bilo nemotivirano (Sherman, 1973).

S ovim zaključkom može se povezati i Joshi (2012) koji ističe da je pozitivna strana afiksalne negacije pružanje učinkovitog načina da se stvore semantički negativni pojmovi, dok se u isto vrijeme dopušta stvaranje pozitivne rečenice. Primjerice, rečenica *„On nije sretan.“* može se izreći i *„On je nesretan.“* Također, Joshi (2012) navodeći primjer rečenice *„Krist nije smrtan.“* u usporedbi s *„Krist je besmrtan.“* navodi da bi se prva rečenica označila kao negativna rečenica, dok bi konstrukcija druge rečenice dala dojam govorniku da ne koristi negativnu izjavu. Osim toga, dodatna prednost afiksalne negacije je ekonomičnost, koja se odnosi na ekonomičnost riječi. Točnije, afiksalna negacija dopušta da se misao izrazi s manje riječi, bez dodatnog opisivanja, jer je jasno na što se misli samim dodavanjem prefiksa. Upravo ekonomičnost afiksalne negacije o kojoj govori Joshi (2012) može se povezati s prednošću prefiksalne negacije o kojoj piše Sherman (1973) kao one koja se lakše, odnosno brže obrađuje.

Naposljetku, Kovačević (2016) također ističe da su psiholingvistička i psihološka istraživanja pokazala da se negacija obrađuje sporije i teže od afirmacije. Navodi da ne samo da se niječna čestica sporije razumijeva od njenog ne-negiranog ekvivalenta, već se sporije razumijeva i od istoznačne jesne konstrukcije. Primjerice, ne samo da se *neprisutan* teže obrađuje od *prisutan,* već i od ne-negirane konstrukcije istoga značenja – *odsutan.*

5. Računalna obrada negacije

Iako postoje računalni programi koji funkcioniraju i za obradu hrvatskog jezika, negacija je ionako zanemareno područje već i na području same teorije, a posebno u području računalne obrade. Za engleski jezik primjerice postoji alat za automatsko prepoznavanje negacije – ConText (Harkema i sur., 2009, prema Mowery i sur., 2012) te NegEx – sustav za prepoznavanje negacije u medicinskim nalazima i otpusnim pismima temeljen na regularnim izrazima (Goldin i Chapman, 2003). Računalna obrada negacije uglavnom je rađena za engleski jezik, a u ovom poglavlju spomenut će se i radovi koji obrađuju negaciju u medicinskim nalazima za španjolski i švedski jezik. Prema Tanushi i sur. (2013) prepoznavanje negacije u kliničkom okruženju nužno je za sažimanje kliničkih tekstova, ali i za izvlačenje informacija o simptomima i dijagnozama, a sustavi koji to omogućuju mogu biti korisni kao potpora liječničkim odlukama. Neki autori koji će biti spomenuti u ovom poglavlju upotrebljavali su NooJ, alat koji je korišten i u ovom radu, dok su neki upotrebljavali sustave za parsanje, sustave za prepoznavanje negacije NegEx i ConText, zatim obradu negacije u analizi sentimenta, hibridne sustave te usporedbu NegExa s ostalim alatima, poput NooJ-a, a više o tim istraživanjima bit će navedeno u idućim odlomcima.

Pri računalnom prepoznavanju negacije važno je poznavati dva pojma koja se pojavljuju u gotovo svim istraživanjima negacije, a to su negacijski znakovi ili signali i negacijski domet. Sergeeva i sur. (2019) navode da je zadatak prepoznavanja negacije uglavnom podijeljen na dva pod-zadatka: prepoznavanje negacijskih znakova (engl. *negation cue*) i prepoznavanje negacijskoga dometa (engl. *negation scope*). Navedeni autori negacijski znak definiraju kao „riječ ili niz riječi koje ukazuju na postojanje negacije u tekstu“[16], a domet kao „jednu ili više riječi koje su negirane određenim negacijskim znakom“[17] (Sergeeva i sur., 2019:212). Odnosno, domet se može objasniti kao doseg djelovanja negacije, točnije koliki dio rečenice negacija obuhvaća i na koliki dio utječe. Pojednostavljeno, Kurtz i sur. (2020) negacijske znakove opisuju kao jezične izraze negacije, a domet negacije opisuju kao dio iskaza koji je negiran.

[16] Vlastiti prijevod
[17] Vlastiti prijevod

Prema Koza i suradnici (2018b), proces prepoznavanja i određivanja negacijskog dometa uglavnom uključuje upotrebu algoritama iz područja regularnih izraza, poput algoritma NegEx koji su napravili Chapman i suradnici (2001). Međutim, Koza i suradnici proveli su usporedbu NegEx i NooJ sustava na medicinskim nalazima te su zaključili da NooJ ima tri prednosti. Prva prednost odnosi se na činjenicu da u NooJ-u nije bilo potrebno prethodno označavanje nalaza, već se to ostvarilo jednim jednostavnim korakom i simultano s prepoznavanjem negacije. Druga prednost NooJ-a je činjenica da NooJ omogućava prepoznavanje materijala iz nalaza koji nisu uključeni u rječniku, kao i materijale koji sadrže tipfelere, služeći se sintaktičkim kontekstom. Naposljetku, NooJ prepoznaje više traženog sadržaja u jednoj rečenici, dok je NegEx-ov algoritam prepoznao samo u kojim rečenicama ima traženog materijala vezanog za negaciju u medicinskim nalazima, ali ne i koliko ih je. Odnosno, u NegEx-u nisu prepoznati koordinirani ili složeni rezultati – primjerice više njih odvojeno zarezom i sl. (Koza i sur., 2018b).

Koza i sur. (2018a) za medicinske nalaze španjolskoga jezika napravili su rječnik i sintaktičku gramatiku u NooJ-u kako bi olakšali proces automatskog prepoznavanja negacije. Rječnik je uključivao niječne čestice kojima bi se detektirala rečenična negacija, a riječi iz medicinskih nalaza odnosile su se na imenice i pridjeve. Međutim, ni rječnik ni gramatika nisu obuhvatili prefiksalnu negaciju, već samo rečeničnu. Na kraju su Koza i sur. (2018a) usporedili rezultate NooJ-a i NegEx-a koji su već prikazani prema Koza i sur. (2018b), gdje se vidi da oba sustava rade dobro, ali NooJ ima prednosti. Naposljetku, i prilikom izračuna same preciznosti gramatike NooJ je pokazao nešto bolje rezultate nego NegEx (Koza i sur., 2018a).

Nadalje, Kurtz i sur. (2020) pokušali su prikazati korisnost sintaktičkih informacija u rješavanju negacije, pojam kojim obuhvaćaju zadatak prepoznavanja negacije i određivanja njezinog dometa. Sustav u kojem su provodili ovaj zadatak je Sherlock (Lapponi i sur., 2012 i 2017, prema Kurtz i sur., 2020), sustav koji rješavanju negacije pristupa prepoznavanjem i obilježavanjem pojavnica te upotrebljava Conditional Random Field (CRF) klasifikator. Prije nego su oznake pridodali CRF-u, Sherlock je „izravnao“ sve pojave negacije u rečenici, dodjeljujući im sustav označavanja sa šest oznaka. Zatim klasifikator čini prepoznavanje na razini pojavnica – takozvani *POS tagging,* odnosno prepoznavanje vrsta riječi, nakon čega radi lematizaciju i dodjeljuje ovisnosne oznake. Autori su umjesto obilježavanja svake pojavnice oznakom *znak, domet* ili *događaj,* prepoznavanje negacije oblikovali kao zadatak parsinga, kreirajući ovisnosni negacijski graf (Kurtz i sur., 2020). Upotrijebili su neuralni parser (Dozat i Manning, 2018, prema Kurtz i sur., 2020) da bi preveli negacijske oznake u graf s ovisnosnim

odnosima te da bi parseri naučili kako predvidjeti negacijske znakove i njihov domet. Ovakvi grafovi izravno označavaju odnos negacijskog znaka i negacijskog dometa, čime omogućuju neuralnim mrežama da te odnose lakše nauče (Kurtz i sur., 2020), što se može primijeniti i na drugim vrstama zadataka i doprinijeti učinkovitijem rješavanju.

Mitchell i sur. (2004) kako bi prepoznali i označili negaciju u patološkim izvješćima, upotrijebili su sustav temeljen na NegEx algoritmu, čija je arhitektura napravljena u GATE-u – sustavu otvorenog tipa za razvoj jezično inženjerskih aplikacija. NegEx je algoritam prvotno opisan za upotrebu na otpusnim pismima i radiološkim nalazima (Mitchell i sur., 2004). Njihov označivač negacije rezultirao je s 84 %-tnom preciznošću i 80 %-tnim odzivom, međutim za bolje rezultate preciznosti i odziva u budućnosti, autori navode da bi trebalo proširiti broj negacijskih izraza na kojima je sustav testiran.

Chapman i sur. (2001a) pomoću NegEx sustava napravili su algoritam za prepoznavanje negacije kako bi analizirali negacijske izraze u medicinskim nalazima. Pokazalo se da je većina negacije izražena pomoću sedam izraza: *no, denies, without, not, no evidence, with no* i *negative for,* poredanih po frekventnosti – od najučestalijeg do najrjeđeg (Chapman i sur., 2001a). Osim kojim izrazima se postiže negacija, autori su se osvrnuli i na preciznost samih izraza pa se tako *not* pokazao kao najmanje precizan izraz. Primjerice u rečenici *„Ovo 'nije' izvor infekcije.“* (engl. *„This is 'not' the source of the infection.“),* izraz *not* negira izvor, a ne samu infekciju, međutim jednostavan algoritam poput NegEx sustava ne uzima u obzir sintaktičke znakove i leksičke informacije koje pomažu u određenju dometa negacijskog izraza. Stoga autori zaključuju da bi negacijski algoritam za prepoznavanje negacijskih izraza, kako bi bio otporniji na ovakve probleme, trebao obuhvatiti i sintaktičko i leksičko znanje (Chapman i sur., 2001a). Međutim, bez obzira na probleme s NegEx sustavom objavljene od istih autora iste godine, Chapman i sur. (2001b) NegEx nazivaju jednostavnim algoritmom temeljenim na regularnim izrazima koji objedinjuje različite indikatore negacijskih izraza, filtrira rečenice koje sadrže lažno negativne izraze te određuje domenu negacijskih izraza.

Mowery i sur. (2015) svojim istraživanjem potvrđuju jednu karakteristiku označavanja negacije koja se može pronaći u literaturi i u bazama podataka – postojanje leksičkog znaka koji ukazuje na negaciju, a može se pojaviti prije, usred i nakon spominjanja poremećaja ili bolesti u tekstu. Primjerice, kada leksički negacijski znak prethodi poremećaju *„Patient denies chest pain.“,* kada se nalazi usred imena poremećaja *„X-ray…possibly be indicative of pneumonia.“* te poslije *„abdominal cramping is unlikely.“* (Mowery i sur. 2012), a ovakvo istraživanje bilo bi korisno učiniti i za hrvatski jezik.

Harkema i sur. (2009) upotrijebili su algoritam ConText, koji je napravljen prema već spomenutom algoritmu za prepoznavanje negacije NegEx, temeljenom na regularnim izrazima. Autori su se također bavili prepoznavanjem negacije u medicinskim nalazima, ali su ovaj prilagođeni algoritam napravili kako bi pokušali poboljšati preciznost izvlačenja informacija iz medicinskih podataka te kako bi osim negacije prepoznali i hipotetsku situaciju, povijest bolesti te osobu koja ju je doživjela – osim samoga pacijenta (Harkema i sur., 2009). Slično kao NegEx, ConText je upotrijebio regularne izraze na unaprijed označenim medicinskim stanjima i specifičnim riječima u tekstu kako bi prepoznao i druge navedene potencijalne situacije, osim same negacije, s ciljem procjene učinkovitosti na ostalim kontekstualnim čimbenicima (Harkema i sur., 2009). Algoritam je napravljen tako da se medicinsko stanje navodi unutar domene takozvanog negacijskog okidača, a ConText mijenja vrijednost pojma u negirano – slično kao u prethodnim istraživanjima gdje se dodjeljivala vrijednost polariteta – pozitivno ili negativno. Pojmovi okidači za promjenu vrijednosti u negativnu su pojmovi poput *„ne“* i *„niječe“.* Autori su uključili ConText u sustav za obradu prirodnog jezika Topaz, a iako nije pokazao savršenu izvedbu, radi se o jednostavnom algoritmu kojim se lako dolazi do informacija i kojemu se s vremenom može poboljšati preciznost (Harkema i sur., 2009).

Nadalje, Huang i Lowe (2007) osmislili su hibridni pristup automatskom prepoznavanju negacije u medicinskim nalazima, s obzirom na to da su prethodna istraživanja pokazala da se negirani pojmovi teže prepoznaju ako su riječi koje ukazuju na negaciju – takozvani negacijski signali, više od nekoliko riječi udaljeni od njih, odnosno izvan negacijskoga dometa. Stoga ovaj hibridni pristup objedinjuje regularne izraze s parsanjem, kako bi se zaobišla ograničenja dosadašnjih modela za automatsko prepoznavanje negacije (Huand i Lowe, 2007). Autori navode da u većini slučajeva negacija uključuje negacijski signal, negirani izraz i mogući dodatni obrazac koji pomaže u prepoznavanju negiranog izraza (Huang i Lowe, 2007), a kako bi napravili gramatiku za prepoznavanje negacije, osim uključivanja negacijskih signala, negiranih izraza i ostalih negacijskih obrazaca, pretražili su i ostalu objavljenu literaturu kako bi stvorili opsežnu listu mogućih negacijskih obrazaca iz lingvističke perspektive (Huand i Lowe, 2007). Autori su negirane izraze izvukli uz pomoć sintaktičke strukture rečenice unutar parsnog stabla, međutim zbog brojnih pogrešaka unutar samih stabala, gramatika koja je kombinirala stabla i regularne izraze također je davala pogrešne rezultate. Bez obzira na to, autori smatraju da njihova gramatika strukturirana prema pravilima koristeći razna lingvistička načela funkcionira bolje nego prethodni pristupi prepoznavanju negacije (Huang i Lowe, 2007).

Tanushi i sur. (2013) bavili su se analizom negacijskoga dometa u kliničkim tekstovima švedskoga jezika koristeći tri pristupa – NegEx, PyConTextNLO i SynNeg. Autori ističu da se brojni sustavi za prepoznavanje negacije oslanjaju samo na negacijske znakove, a vrlo malen broj istražuje i sintaktičku strukturu rečenice koja se može upotrijebiti za određenje dometa negacijskih znakova. Sva tri pristupa su temeljena na pravilima, ali svaki ima različit način određivanja dometa negacijskih znakova – NegEx upotrebljava udaljenost između znaka i bolesti u tekstu, PyConTextNLP se oslanja na listu veznika koji ograničavaju domet znaka, a SynNeg koristi granice rečenica dobivene sintaktičkim parserom kako bi ograničio domet znakova (Tanushi i sur., 2013).

NegEx (Chapman i sur., 2001, prema Tanushi i sur., 2013) je široko upotrebljavan sustav za prepoznavanje negacije u kliničkim tekstovima koji se temelji na različitim listama negacijskih izraza koji služe kao negacijski znakovi (engl. *cue phrases*). Ovi negacijski znakovi mogu se pojaviti prije navođenja bolesti ili simptoma, poslije bolesti ili simptoma te mogu biti i pseudo-negacijski izrazi, odnosno izrazi koji nisu okidači negacije kliničkoga stanja, iako sadrže negacijski znak, primjerice *ne samo* (engl. *not only*). PyConTextNLP (Chapman i sur., 2011, prema Tanushi i sur., 2013) je prošireni NegEx algoritam koji uključuje modifikaciju pravila koja se odnose na domet negacije i povezivanje leksičkih okidača, a ima više mogućnosti prilikom definiranja pravila te radi na razini cijele rečenice. Naposljetku, SynNeg je alat za negacijski domet koji upotrebljava morfološke i sintaktičke informacije sustava MaltParser (Nivre i sur., 2007, prema Tanushi i sur., 2013). MaltParser je sustav ovisnosnog parsanja koji se za analizu novih rečenica koristi podacima banaka stabala. Autori su analizu ovih sustava proveli na švedskom kliničkom korpusu, a rezultati su za sva tri sustava vrlo slični – NegEx ima nešto veću preciznost, a SynNeg nešto viši odziv i F-mjeru. S obzirom na to da je SynNeg novi sustav koji se tek razvija, a pokazao je dobre rezultate upotrebljavajući sintaktičke oznake, autori zaključuju da upotreba sintaktičkih informacija može biti od koristi i za druge domene, osim medicinske, kao i za druge jezike (Tanushi i sur., 2013).

Goldin i Chapman (2003) smatraju da je negacija niječnom česticom *ne,* odnosno engl. *not,* teža za automatsko prepoznavanje nego ostale vrste negacije, jer ponekad *not* uopće ne negira rečenično značenje. Stoga autori primjenjuju sustave za strojno učenje kako bi razlikovali rečenice medicinskoga teksta u kojem *not* služi kao sredstvo negacije od onih u kojima ne negira. Točnije, autori upotrebljavaju dva različita algoritma za strojno učenje – Naive Bayes i Decision Trees. U prilog činjenici da je *not* problematičan za pravilnu detekciju negacije autori navode statističke podatke prema kojima je NegEx postigao 78 %-tni odziv i 84

%-tnu preciznost, a kada je negacijski izraz za prepoznavanje *not,* tada preciznost u jednom istraživanju pada na 58 %, a u drugom na 63 % (Goldin i Chapman, 2003). Stoga autori pokušavaju povećati sposobnosti NegEx sustava da ispravno odredi treba li kliničko mišljenje unutar domene od šest riječi od negacijskog izraza *not* biti negirano ili ne. Rezultati na sustavima Naive Bayes i Decision Trees pokazuju poboljšanje u preciznosti pronalazaka negacije s izrazom *not*, u usporedbi s prijašnjim istraživanjima na osnovnom NegEx algoritmu (Chapman i sur., 2001, prema Goldin i Chapman, 2003).

S druge strane, kompromis povećane preciznosti je smanjenje odziva, ali taj pad je zanemariv, dok vrijednosti F-mjere za Naive Bayes i Decision Trees nisu značajno drugačije od rezultata osnovnog NegEx algoritma. Međutim, autori su otkrili jednostavno pravilo za koje smatraju da ga treba primijeniti u temeljni NegEx algoritam – kada se negacija pojmova iz UMLS-a (*Unified Medical Language System*) ostvaruje negacijskom česticom *not,* ako pojmu prethodi član *the,* tada pojam ne treba označiti kao negiran (Goldin i Chapman, 2003). Bez obzira na doprinos navedenoga rada, autori smatraju da je njihovo istraživanje ograničeno iz nekoliko razloga – radi se o jednoj specifičnoj negacijskoj frazi te smatraju da bi se u budućim istraživanjima trebalo voditi intuitivnom idejom da bilo koji medicinski pojam u nalazu kojem prethodi član *the* ne bi trebao biti označen kao negiran o kojoj god negacijskoj frazi bila riječ. Također, smatraju da bi im trebao robusniji parser napravljen za rečenice iz medicinskih tekstova te da bi za veću vjerodostojnost trebalo upotrijebiti mišljenja više liječnika prilikom procjene tekstova za analizu, umjesto samo jednog (Goldin i Chapman, 2003).

Jimenez-Zafra i sur. (2019) bavili su se automatskim prepoznavanjem negacije španjolskoga jezika, s obzirom na to da se radi o drugom najrasprostranjenijem jeziku na svijetu. Grupa znanstvenika koja se bavi negacijom u španjolskom jeziku iz računalne perspektive naziva se NEGES, a nastala je 2017. godine (NEGES Task, n.d.). U radu koji su predstavili Jimenez-Zafra i sur. (2019), zadatak analize negacije u španjolskom jeziku podijeljen je na dva pod-zadatka – prvi se odnosi na prepoznavanje negacijskih znakova, a drugi na ulogu negacije u analizi sentimenta. Što se tiče prvog zadatka, cilj tima znanstvenika je promovirati razvoj i evaluaciju postojećih sustava za prepoznavanje negacijskih znakova u španjolskom. Drugi zadatak odnosi se na evaluaciju utjecaja pravilnog prepoznavanja negacije u analizi sentimenta, a trebalo se razviti sustav koji upotrebljava informacije o negaciji iz korpusa recenzija filmova, knjiga i proizvoda, sa svrhom poboljšanja zadatka određivanja polariteta.

Za evaluaciju sustava upotrijebljene su mjere preciznost, odziv i F-mjera, a različiti timovi na kraju su imali različite probleme sa svojim sustavima. Jedan tim imao je vrlo niske rezultate svoga sustava, jer nisu upotrijebili dodatna obilježja za tekstove iz svog korpusa, poput lematizacije i označavanja vrsta riječi, što su ostali timovi učinili i dobili bolje rezultate. Međutim, drugi timovi imali su problema prilikom prepoznavanja negacijskih znakova, odnosno sustav je pogrešno prepoznao znakove koji zapravo ne izražavaju negaciju. Također, imali su i probleme u prepoznavanju diskontinuiranih znakova – razdvojenih negacijskih znakova koji se sastoje od više riječi te općenito probleme u prepoznavanju višerječnih negacijskih znakova (Jimenez-Zafra, 2019).

Naposljetku, također se baveći analizom sentimenta, Ljajić i sur. (2017) istraživale su obradu negacije u analizi sentimenta za srpski jezik. Autorice ističu da analiza sentimenta nije lagan zadatak, s obzirom na to da postoje brojni problemi koje je teško razriješiti, poput negacije, ironije i sarkazma. Negacija mijenja značenje rečenice i njezin sentiment, ali postoje i slučajevi kada negacija ne mijenja sentiment rečenice, već samo intenzitet tog sentimenta. Autorice navode da je cilj analize sentimenta otkriti osjećaje, emocije i stavove u pisanom tekstu, a upotrijebile su algoritme strojnog učenja kako bi u označenom korpusu uvježbale klasifikator rečeničnog polariteta. Kiritchenko i sur. (2014, prema Ljajić i sur., 2017) pokazali su da negacija pozitivnih pojmova implicira negativni sentiment, dok u slučaju negativnih pojmova sentiment ostaje negativan. Relevantni radovi o analizi sentimenta za hrvatski jezik iz lingvističke perspektive, a specifično iz područja negacije nisu pronađeni, a radi se o sve potrebnijem i raširenijem području. Analiza sentimenta ključan je dio softvera za mrežno praćenje mišljenja o brendovima i tvrtkama, kao i velik dio programiranja virtualnih asistenata, stoga je to još jedno područje u kojem postoji potreba za primjenom lingvističkih znanja i vještina.

6. Prikupljanje korpusa i analiza prefiksalne negacije pridjeva iz hrWaC-a

Korpus je prikupljen uz pomoć alata Sketch Engine[18], koji uključuje brojne već prikupljene i učitane korpuse spremne za upotrebu, ali i alate za stvaranje, instaliranje i ophođenje vlastitim korpusom (Kilgarriff i sur., 2014). Kao izvor je korišten već dostupan i označen korpus hrvatskoga jezika hrWaC (Ljubešić i Erjavec, 2011). HrWaC je hrvatski mrežni korpus koji sadrži tekstove prikupljene s interneta, točnije s '.hr' domene. Korpus je napravljen u siječnju 2014. godine i ima preko 1.2 milijarde riječi.

Korpus koji se prikuplja za ovaj rad sadržavat će konkordancije iz hrWaC-a, koje uključuju pridjeve koji odgovaraju definiciji negacije u ovom radu – bez prefiksa nešto *jest,* dodavanjem prefiksa više *nije* ili dobiva suprotno značenje. Pridjevi su u tim konkordancijama prikazani kao *KWIC,* odnosno *keyword in context,* stoga korpus obuhvaća i njihovu ko-tekstnu okolinu. Prilikom prikupljanja korpusa pridjevi će se odmah i analizirati te će se u potpoglavljima s pojedinim prefiksima iznijeti zaključci o njihovoj tvorbi, učestalosti, podrijetlu i slično, gdje će se iznijeti informacije i definicije prema *Hrvatskom jezičnom portalu*[19]. Neki prefiksi ne tvore negaciju u smislu određenom ovim radom, ali će takvi prefiksi i pridjevi svejedno biti spomenuti, iako neće biti uključeni u računalnu analizu. U takvim slučajevima riječ je o prefiksima koji tvore neizravnu negaciju, odnosno izriču negativno značenje, ali ne i izravno negiranje te se stoga potpoglavlja s takvim prefiksima nalaze u zagradama.

Međutim, glavna svrha prikupljanja korpusa za ovaj rad je njegova primjena u NooJ alatu kako bi se napravila gramatika koja prepoznaje i označava negaciju pridjeva hrvatskoga jezika, o čemu će biti riječi u idućim poglavljima. Pridjevi uključeni u korpus prikupljani su uz pomoć gramatika (Babić, 1991; Težak i Babić, 1992; Barić i sur., 2005), pravopisa (Babić i sur., 1971) i rječnika (ur. Petrač i Miščin, 2004). Za prikupljanje korpusa upotrebljavani su regularni izrazi u CQL-u s negacijskim prefiksom i pridjevom kao glavnim uvjetima pretrage. CQL je *corpus query language* (engl.) te se upotrebljava kako bi se postavili uvjeti pretrage.

[18] https://www.sketchengine.eu/
[19] http://hjp.znanje.hr/

Regularni izrazi upotrebljavaju se unutar CQL pretrage i ograničavaju pretragu na određeni niz znakova.

Primjerice:

[lemma="ne.*" & tag="A.*"] filtrira riječi koje počinju prefiksom 'ne-' i koje su označene kao pridjev.

Također, upotrebljavala se i preciznija pretraga kako bi se izbjegli rezultati koji nisu negacija, poput:

[lemma="bez.*" & lemma=".*an" & tag="A.*"] filtrira riječi koje počinju s ‘bez-‘, završavaju sufiksom ‘-an’ i označene su kao pridjev, poput *bezosjećajan.*

Oznaka *lemma* označava da se pretražuju svi oblici neke riječi, a ne samo jedan njen određeni oblik. Oznaka *A* odnosi se na morfo-sintaktičku oznaku za vrstu riječi pridjev. Oznaka .* označava da se prije ili nakon postavljenog uvjeta može pojaviti bilo koji znak, nula ili više puta.

Prefiks a-

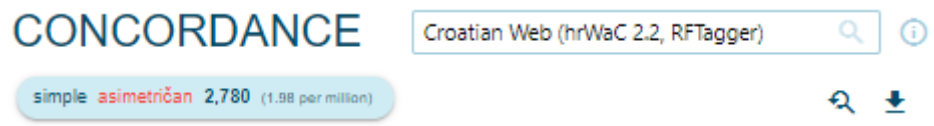
Prema *Hrvatskom jezičnom portalu*[20]*,* ovaj prefiks izražava suprotnost ili izostanak onoga što označava osnova – prema čemu odgovara definiciji negacije određenoj u ovom radu. Za negaciju prefiksom *a-* nije bilo moguće pretraživati na isti način CQL izrazom jer bi rezultat dao sve riječi koje počinju na *a-* i koje su pridjev, a takvih je puno koje nisu negacija. Iz tog razloga, za ovaj prefiks pretraživani su specifični primjeri pronađeni u rječnicima, gramatikama i pravopisu, u kojima prefiks *a-* tvori negaciju. Primjerice:

[lemma="adinamičan" | lemma="alogičan" | lemma="amoralan" | lemma="anormalan" | lemma="apolitičan" | lemma="asimetričan" | lemma="asocijalan" | lemma="aseksualan" & tag="A.*"]

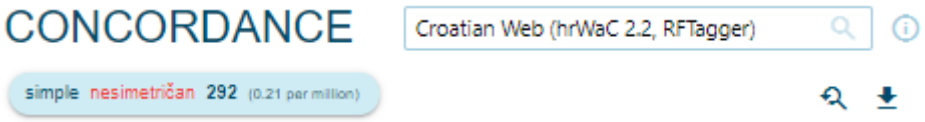
Upit je sadržavao i znak | koji označava *ili,* odnosno disjunkciju, kako bi pretraga rezultirala bilo kojim od traženih upita, odnosno prikazala nasumični uzorak.

[20] http://hjp.znanje.hr/

Prefiks *a-* često se može zamijeniti hrvatskom inačicom – prefiksom *ne-,* primjerice postoji i *nenormalan, nedinamičan, nelogičan, nesimetričan,* itd. Međutim, prema analizi u hrWaC-u, nije uvijek učestaliji pridjev s prefiksom *ne-*. Naime, primjer *asimetričan,* kao što prikazuju Slika 1 i 2, češći je od primjera *nesimetričan:*

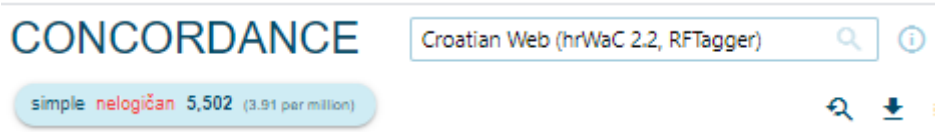


Slika 1 Učestalost pridjeva 'asimetričan' u hrWaC-u

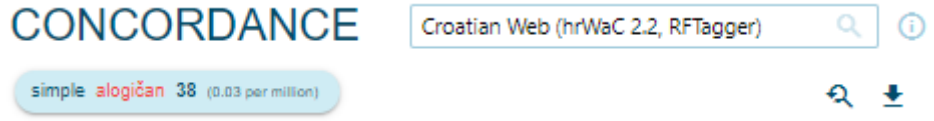


Slika 2 Učestalost pridjeva 'nesimetričan' u hrWaC-u

S druge strane, kod pridjeva *alogičan* suprotna je situacija, odnosno češći je pridjev *nelogičan,* što prikazuju Slika 3 i 4:



Slika 3 Učestalost pridjeva 'nelogičan' u hrWaC-u



Slika 4 Učestalost pridjeva 'alogičan' u hrWaC-u

Prefiks ab-

Pretraga korpusa upitom [lemma="ab.*" & tag="A.*"] obuhvaća i primjere *abortivan, aboridžinski,* koji ili nisu negacija ili nisu negacija u okviru definicije određene u ovome radu.

Odnosno, prema *Hrvatskom jezičnom portalu*[21], pridjev *abortivan* dolazi od *abortus,* gdje *oriri* znači roditi se, odnosno *ab+oriri* je suprotno od rođenja, ali ne radi se o primjeru koji bez prefiksa *ab* postoji u hrvatskom kao samostalan pridjev, što je uvjet.

Stoga, s obzirom na to da je pridjeva negiranih prefiksom *ab-* u hrvatskom jeziku vrlo malo, odnosno u literaturi su pronađeni samo primjeri *aboralan* i *abnormalan,* korpus je za te pridjeve pretražen specifičnim upitom: [lemma="aboralan" | lemma="abnormalan" & tag="A.*"].

Prefiks an-

Za ovaj prefiks nije se moglo pretraživati upitom [lemma="an.*" & tag="A.*"] jer takav upit obuhvaća i primjere sa *anti-,* obuhvaćene drugačijim upitom, ali i primjere koji nisu negacija, poput *anoniman, anksiozan, analitički, Anin, antropološki, angažiran, analiziran, anarhičan* itd.

Pridjev *anoreksičan,* ne odnosi se na negaciju obuhvaćenu definicijom ovoga rada iz istog razloga kao pridjev *abortivan,* jer anoreksičan se prema *Hrvatskom jezičnom portalu*[22] tvori od grčke riječi *oreksis* koja znači apetit, stoga je *an+oreksis* suprotno od apetita. Međutim, s obzirom na to da se ne radi se o tvorbi u okvirima hrvatskoga jezika, odnosno *oreksis* ne postoji kao osnova u hrvatskom jeziku, ovakvi primjeri nisu obuhvaćeni jer je uvjet da pridjev postoji i bez prefiksa.

Pridjeva koji u hrvatskom jeziku s prefiksom *an-* tvore negaciju nema puno, a nije ga spomenuo nijedan autor obrađen u teorijskom dijelu. U radu Žanpera i sur. (2020) koji se bavio prefiksalnom negacijom imenica hrvatskoga jezika, upotrijebljeni su samo primjeri *anorgazam* i *anovulacija.* Međutim, iako *Hrvatski jezični portal*[23] sadrži imenicu *anorgazam,* ne spominje i pridjev *anorgazmičan.* S druge strane, za imenicu *anovulacija,* spominje se i pridjev *anovulacijski.* Hrvatski mrežni korpus (hrWaC) pronašao je primjere pridjeva *anorgazmičan,* stoga su i oni uključeni u korpus ovoga rada, a osim toga obuhvaćen je i pridjev *anorganski.* S

[21] http://hjp.znanje.hr/
[22] http://hjp.znanje.hr/
[23] http://hjp.znanje.hr/

obzirom na to da nema puno primjera koji s ovim prefiksom tvore negaciju, upit postavljen u Sketch Engineu postavljen je prema specifičnim riječima:

[lemma="anorgazmičan" | lemma="anorganski" | lemma="anovulacijski" & tag="A.*"]

Primijećeno je također i da pridjevi negirani prefiksom *a-* uglavnom započinju konsonantom – *adinamičan, alogičan*, a pridjevi negirani prefiksom *an-* započinju vokalom – *anorganski, anovulacijski.*

Prefiks anti-

Prilikom prikupljanja korpusa, nisu se mogli izbjeći i neki primjeri koji zapravo ne čine negaciju. Primjerice, CQL upitom za prefiks *anti-* koji je glasio:

[lemma="anti.*" & tag="A.*"] obuhvaćeni su i primjeri *antički, antiohijski, antikvaran.*

Korpus je također obuhvaćao i primjere istoznačnih riječi, negiranih istim prefiksom, ali na osnovi različite etimologije, poput *antifungalni* i *antigljivični,* gdje je *fungalni* latinskog podrijetla, a *gljivični* praslavenskog.

Također, ono što je primijećeno je da korpus obuhvaća i pridjev *antikarcinogen* i *antikancerogen,* dok pridjev *karcinogen* ne postoji na *Hrvatskom jezičnom portalu*[24]*,* a *kancerogen* postoji. Međutim, obje imenice – *karcinom* i *kancer* postoje na *Hrvatskom jezičnom portalu*[25] te su navedene kao istoznačnice, gdje je imenica *karcinom* grčkog podrijetla, a *kancer* latinskog.

Više o ovom prefiksu u kontekstu negacije pridjeva hrvatskoga jezika bit će izneseno u posljednjem potpoglavlju ovoga poglavlja, gdje će se prefiks *anti-* usporediti s prefiksom *protu-*, s kojim ga je ponekad moguće zamijeniti.

[24] http://hjp.znanje.hr/
[25] http://hjp.znanje.hr/

Pridjevi koji bez ovoga prefiksa ne postoje u hrvatskom jeziku, poput *antibiotski,* gdje *biotski* ne postoji kao osnova, nisu uključeni u analizu jer je postavljen uvjet da riječ postoji i bez prefiksa.

Naposljetku, primijećeno je da se na istoj osnovi – *socijalan,* pojavljuju i prefiks *a-* i *anti-*. Prema *Hrvatskom jezičnom portalu*[26]*,* pridjev *asocijalan* odnosio bi se na osobu koja je nedruštvena te se više tiče samoga pojedinca i njegova karaktera. S druge strane, *antisocijalan* se odnosi na nekoga koji također nije društven, ali i koji je štetan društvu te ga *Hrvatski jezični portal*[27] navodi i u obliku s prefiksom *protu-,* odnosno *protudruštven.* Stoga pridjev negiran prefiksom *anti-* ima šire značenje i ipak nešto drugačije – onaj koji je više okrenut *protiv društva,* nego nedruštven kao dio osobnosti.

Prefiks bez-

S obzirom na moguće alomorfske inačice prefiksa *bez-,* kao što su *bes-, beš-, bež-,* regularni izraz u CQL upitu uključivao je upravo te inačice, jer pretraga sa *be-* kojom se obuhvaćaju i te alomorfske inačice, obuhvaća i riječi koje nisu negacija, poput *bečki, bećarski, betonski, beogradski, betlehemski.* Stoga je CQL izraz za ovaj prefiks izgledao:

[lemma="bes.*" | lemma="beš.*" | lemma="bez.*" | lemma="bež.*" & tag="A.*"]

Iako su ovim izrazom pročišćeni rezultati, odnosno puno je manje pogrešaka, i dalje su neizbježni primjeri koji nisu negacija u smislu određenom unutar ovoga rada. Primjerice *bezbrojan,* za koji se već pisalo da se ne odnosi na stanje *bez broja,* odnosno ne negira se postojanje broja, već bezbroj označava nešto toliko brojno da se ne može izbrojati. Prema *Hrvatskom jezičnom portalu*[28]*,* „ono što se ne može izbrojiti, velika množina, mnoštvo“. Međutim, u prenesenom značenju, odnoseći se na nešto *nebrojivo,* primjer bezbroj zanimljiv je primjer iz kognitivne perspektive i stoga bi se mogao smatrati rubnim primjerom negacije, ali u ovom radu neće se ulaziti u područje kognitivne lingvistike i ovaj primjer se iz tog razloga neće obuhvatiti.

[26] http://hjp.znanje.hr/
[27] http://hjp.znanje.hr/
[28] http://hjp.znanje.hr/

S druge strane, pridjevi *beskrajan* i *beskonačan* tumačeni su u drugačijem smislu. Naime, *beskraj* se odnosi na nešto *bez kraja,* gdje nema kraja, dok se *bezbroj* ne odnosi na stanje bez broja. Međutim, moglo bi se reći da se također radi o rubnom primjeru iz kognitivne perspektivne, jer nešto beskrajno je također neograničeno, ali primjer *beskrajan* zbog činjenice da se odnosi na nešto *bez kraja,* ipak će se obuhvatiti analizom. Pridjev *beskonačan,* prema *Hrvatskom jezičnom portalu*[29] u izravnoj je opreci s *konačan* te je stoga obuhvaćen.

Također, pridjev *bezazlen,* negiran alomorfskom inačicom *beza-,* za kojeg Babić (2002) ističe da se radi o prefiksu s dodanim blagozvučnim *a-* te ga naziva proširenim prefiksom, obuhvaćen je ovakvom pretragom, ali ne radi se o negaciji u potrebnom smislu, jer bez prefiksa, riječ *zazlen* ne postoji. Međutim, radi se o tvorbi *bez* + *zlo* koja se odnosi na nekog dobronamjernog i bezopasnog, ali s obzirom na to da nije tvoren na način „bez prefiksa nešto *jest,* dodavanjem prefiksa *nije* ili dobiva suprotno značenje", ovaj pridjev neće biti tretiran kao negacija u računalnoj analizi. Točnije, u navedenom primjeru radi se o neizravnoj negaciji, a područje ovoga rada je izravna negacija.

(Inačica obez-)

Još jedna moguća inačica prefiksa bez- je obez-. Točnije, od glagola poput obezglaviti, obezvrijediti, obezvučiti, obeščastiti tvore se glagolski pridjevi trpni obezglavljen, obezvrijeđen, obezvučen, obeščašćen. Pretraga upitom:

[lemma="obez.*" & tag="A.*"] generira i pridjev *obezbijeđen* – od glagola *obezbijediti*, što nije negacija, jer znači *osigurati.* Stoga je za ovakve primjere korpus pretraživan specifičnim upitom s navedenim pridjevima. Međutim, prefiks *obez-* iako tvori izravnu negaciju (označava da je netko ili nešto *bez glave, bez vrijednosti, bez zvuka, bez časti*), nije uključen u gramatiku za prepoznavanje negacije jer bez toga prefiksa pridjevi ne postoje kao samostalne osnove, odnosno takvi primjeri nisu obuhvaćeni definicijom određenom u ovome radu.

[29] http://hjp.znanje.hr/

Prefiks de-

Pretraga korpusa s prefiksom *de-* obuhvaćala je uglavnom pridjeve koji nisu negacija, poput *detaljan, definiran, depresivan, demokratski, debeli, desni* itd. Stoga su u upit postavljene specifične riječi, preuzete od imenica iz Žanpera i sur. (2020), kao i neke pronađene dodatno:

[lemma="deratiziran" | lemma="decentraliziran" | lemma="destabiliziran" | lemma="dehidriran" | lemma= "dehidratiziran" | lemma="dekomponiran" | lemma="dekontaminiran" | lemma="deblokiran" | lemma="dekoncentriran" & tag="A.*"]

Međutim, s obzirom na činjenicu da su neki od ovih pridjeva puno zastupljeniji u korpusu od drugih, ovakva pretraga koja generira nasumični uzorak nije dobra. Primjerice, pridjev *decentraliziran* u hrWaC-u pojavljuje se 2165 puta, dok se ostali pojavljuju otprilike 10-30 puta. Stoga rezultati generirani ovakvim upitom čine 2218 konkordancija, od kojih je većinu primjera u nasumičnom uzorku čini pridjev *decentraliziran* pa ako se iz toga za korpus ovoga rada želi prikupiti stotinjak primjera, vrlo teško je u takvom nasumičnom uzorku pronaći i ostale pridjeve. Iz tog razloga, ipak je stotinu primjera s ovim prefiksom prikupljeno pretraživanjem svakog pridjeva zasebno.

Također, kako bi pretraga obuhvatila oba oblika pridjeva, odnosno i *dehidriran* i *dehidracijski;* i *dekontaminiran* i *dekontaminacijski,* postavljeni upiti sadržavali su samo početak tih riječi, primjerice:

[lemma="decentraliz.*" | lemma="destabiliz.*" | lemma="dehidr.*" | lemma= "dehidratiz.*" | lemma="dekomponiran" | lemma="dekontamin.*" | lemma="deblokiran" | lemma="dekoncentriran" & tag="A.*"]

Međutim, s obzirom na to da ovakav upit generira još više konkordancija (5155) s neravnomjerno raspoređenim pridjevima, postavljen je upit u ovakvom obliku ali za svaki pridjev zasebno, kako bi se prikupilo stotinu pridjeva s ovim prefiksom, koji obuhvaćaju svaki od navedenih primjera.

S druge strane, pridjevi poput *degeneriran, defibriliran, dekompresijski, dekalcificiran* prema značenju na *Hrvatskom jezičnom portalu*[30] nisu negacija, stoga nisu uključeni u korpus.

[30] http://hjp.znanje.hr/

Prefiks dez-

Pretragom korpusa putem upita:

[lemma= "dez.*" & tag= "A.*"]

generiraju se i rezultati koji ne uključuju negaciju, poput *dezerterski, dežurni, dezen* (što je imenica, ali očito pogrešno označena u korpusu), *dežulovićev* itd. Pretraga nije generirala „čišće“ rezultate ni kada su uključeni sufiksi poput *-ski, -an* i slični, stoga je i prilikom prikupljanja primjera s ovim prefiksom korištena pretraga po specifičnim riječima.

Pridjevi koji su uključeni u korpus za ovaj rad su: *dezintegriran, dezintegracijski, dezorijentiran, dezorijentacijski, dezorganiziran, dezinformiran, dezinformacijski, dezinformativan, dezinsekcijski, dezinfekcijski, dezinficiran*, *deziluzioniran, dezideologiziran.*

Prikupljanjem korpusa primijećeno je da se ovaj prefiks vrlo često pojavljuje u značenju prefiksa *bez-*: *deziluzioniran* je *bez iluzija, dezorijentiran* je *bez orijentacije, dezideologiziran* je *bez ideologije.* Međutim, pojavljuje se i drugim značenjima, poput *dezinformiran*, što znači *pogrešno informiran*, a ne *bez informacija.*

Prefiks dis-

Ponovno, pretraga jednostavnim upitom:

[lemma="dis.*" & tag="A.*"] rezultira pogrešnim primjerima – onima koji nisu negacija, poput *disciplinski, diskrecijski, dispečerski, distanciran, distribuiran,* ali i *disco* i *disko,* što su pogrešno označene imenice.

Primjeri koji se odnose na negaciju su *disharmoničan, disfunkcionalan, disbalansiran, diskvalificiran, disproporcionalan* te su uključeni u korpus iako su nastali od imenica.

Pridjevi nastali od imenice *dismenoreja,* koji se u korpusu nalaze u oblicima *dismenoreičan, dismenorejičan* i *dismenoroičan* nisu uključeni u korpus ovoga rada jer *dismenoreja* nije nedostatak menstruacije, odnosno suprotnost od menstruacije što je izraženo imenicom *amenoreja,* već se odnosi na smetnje u menstruaciji, što ne odgovara negaciji u potrebnom smislu.

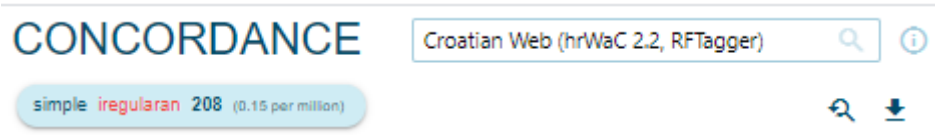
Također, pridjev *dispozicijski* ne odnosi se na negaciju, jer imenica *dispozicija* od koje nastaje nije negacija nijednom definicijom *Hrvatskog jezičnog portala.*[31]

Pridjev *disgenetički*, iako se pojavljuje u hrWaC-u, ne odnosi se na izravnu negaciju, nego onu koja je u teorijskom dijelu ovoga rada određena kao neizravna – ima negativno značenje, ali prefiksom se ne negira ili tvori neko suprotno značenje. Ovaj pridjev pojavljuje se u kontekstu s pridjevima poput *malformiran* i *abnormalan*, od kojih *malformiran* također samo ima negativno značenje, ali ne odgovara negaciji određenoj u ovom radu, dok je *abnormalan* izravna negacija, jer ima svoju jesnu/afirmativnu inačicu *normalan.*

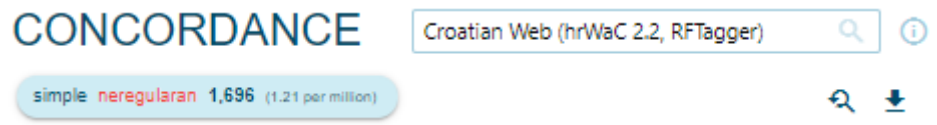
Prefiks i-

Prefiks *i-* je prefiks latinskog podrijetla i u hrvatskom jeziku slabo je plodan. Dodaje se osnovama koje su također stranog, uglavnom latinskog podrijetla. Primjeri u kojima tvori izravnu negaciju su pridjevi *iregularan, ilegalan, ilojalan, imoralan, imaterijalan, imobilan, iracionalan, irealan.* Ima značenje prefiksa *ne-,* kojim se negira značenje osnove, a u nekim primjerima može se i zamijeniti tim prefiksom, ostajući na istoj osnovi. Primjerice, u hrvatskom jeziku upotrebljava se i *neregularan, nemoralan, nematerijalan, nerealan* itd. Štoviše, primjeri s prefiksom *ne-* u hrWaC-u su puno češći.

Učestalost pridjeva *iregularan/neregularan* prikazana je slikama 5 i 6:



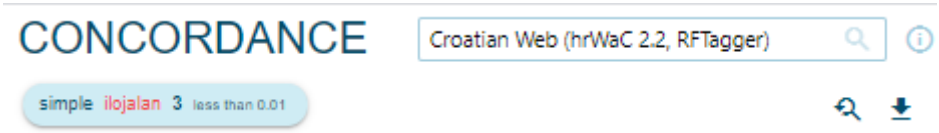
Slika 5 Učestalost pridjeva 'iregularan' u hrWaC-u



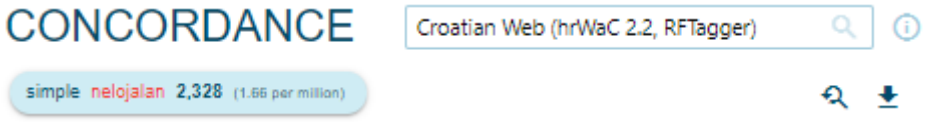
Slika 6 Učestalost pridjeva 'neregularan' u hrWaC-u

[31] http://hjp.znanje.hr/

Učestalost pridjeva *ilojalan/nelojalan* prikazana je slikama 7 i 8:

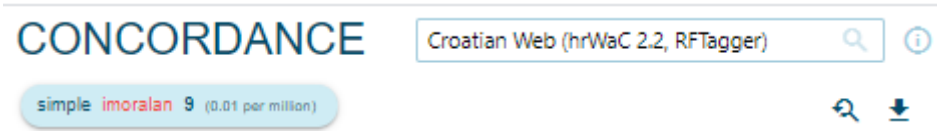


Slika 7 Učestalost pridjeva 'ilojalan' u hrWaC-u

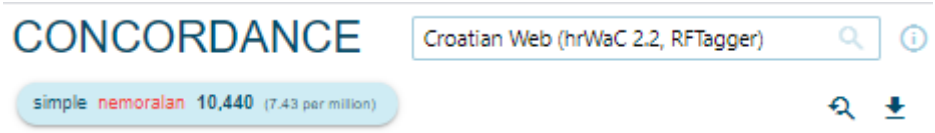


Slika 8 Učestalost pridjeva 'nelojalan' u hrWaC-u

Učestalost pridjeva *imoralan/nemoralan* prikazana je slikama 9 i 10:

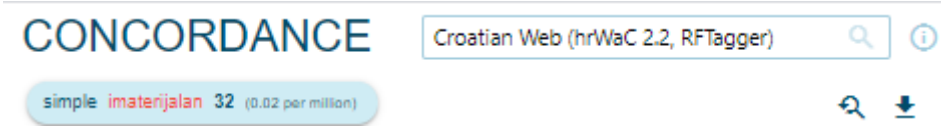


Slika 9 Učestalost pridjeva 'imoralan' u hrWaC-u

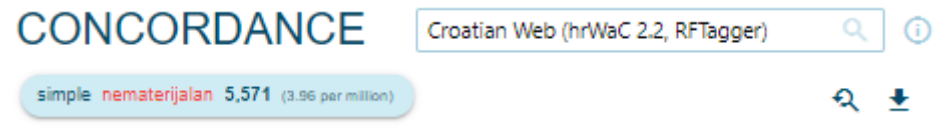


Slika 10 Učestalost pridjeva 'nemoralan' u hrWaC-u

Učestalost pridjeva *imaterijalan/nematerijalan* prikazana je slikama 11 i 12:

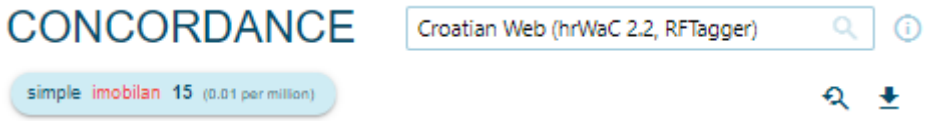


Slika 11 Učestalost pridjeva 'imaterijalan' u hrWaC-u

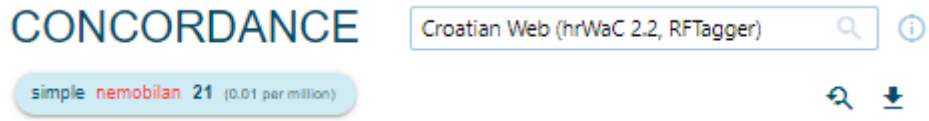


Slika 12 Učestalost pridjeva 'nematerijalan' u hrWaC-u

Učestalost pridjeva *imobilan/nemobilan* prikazana je slikama 13 i 14:

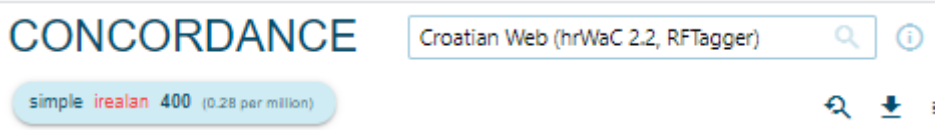


Slika 13 Učestalost pridjeva 'imobilan' u hrWaC-u

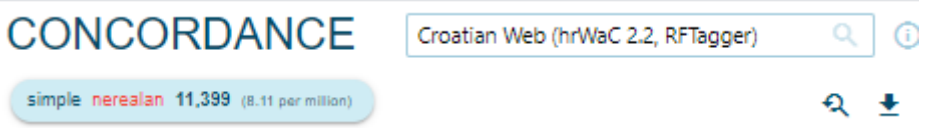


Slika 14 Učestalost pridjeva 'nemobilan' u hrWaC-u

Učestalost pridjeva *irealan/nerealan* prikazana je slikama 15 i 16:



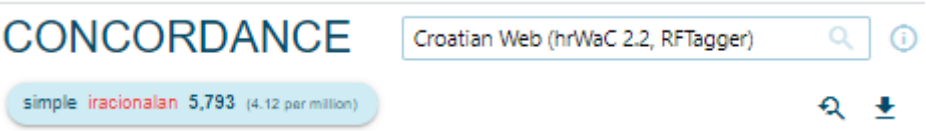
Slika 15 Učestalost pridjeva 'irealan' u hrWaC-u



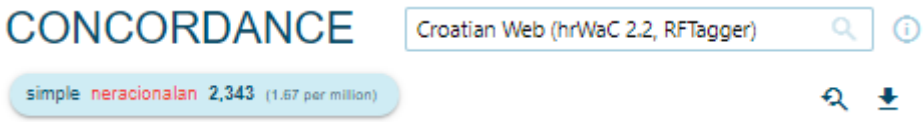
Slika 16 Učestalost pridjeva 'nerealan' u hrWaC-u

Međutim, iznimke čine pridjevi *iracionalan* i *ilegalan,* koji su u hrWaC-u zastupljeniji upravo u tim inačicama, umjesto s prefiksom *ne-*.

Usporedba učestalosti pridjeva *iracionalan/neracionalan* prikazana je slikama 17 i 18:

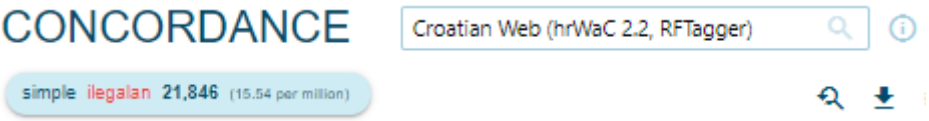


Slika 17 Učestalost pridjeva 'iracionalan' u hrWaC-u

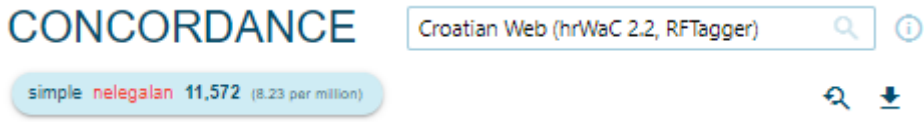


Slika 18 Učestalost pridjeva 'neracionalan' u hrWaC-u

Usporedba učestalosti pridjeva *ilegalan/nelegalan* prikazana je slikama 19 i 20:



Slika 19 Učestalost pridjeva 'legalan' u hrWaC-u

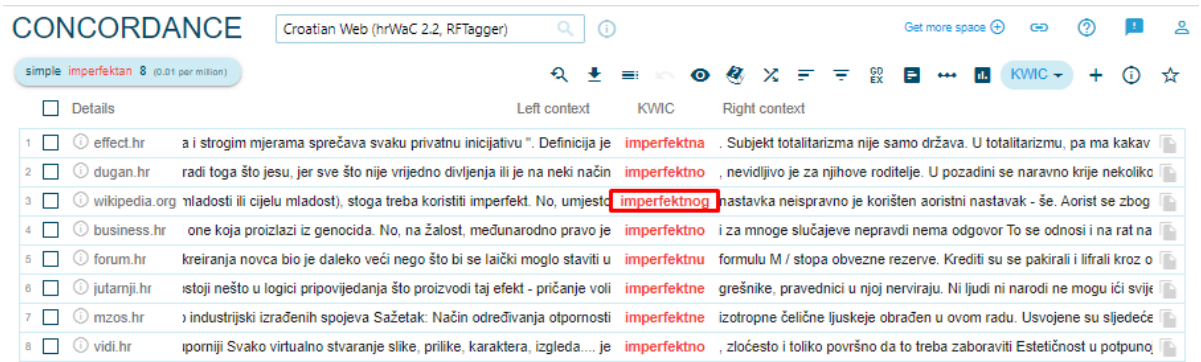


Slika 20 Učestalost pridjeva 'nelegalan' u hrWaC-u

Prefiks im-

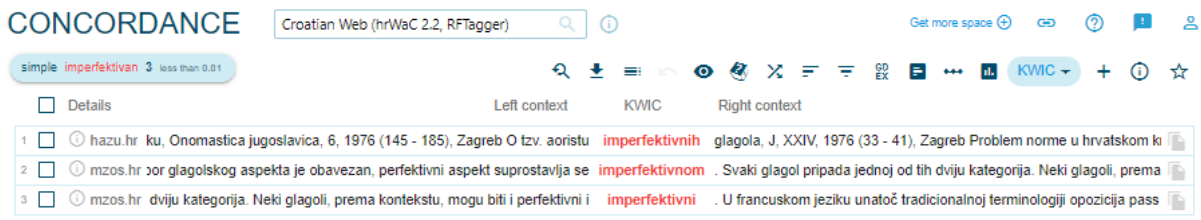
Prefiks *im-* latinskog je podrijetla i dolazi s osnovama stranog, uglavnom također latinskog podrijetla. Međutim, u hrvatskom jeziku vrlo je slabo plodan i definiciji negacije određenoj u ovom radu odgovaraju primjeri pridjeva *imperfektan, imperfektivan* i *impotentan.*

Pridjev *imperfektan* se odnosi na *nesavršen,* a *imperfektan* na glagolski oblik imperfekt. U hrWaC-u se pridjev *imperfektan* pojavljuje samo osam puta, od čega je jedna pojava u pogrešnom obliku – trebao je biti pridjev *imperfektivan* (Slika 21).



Slika 21 Učestalost i kontekst pridjeva 'imperfektan' u hrWaC-u

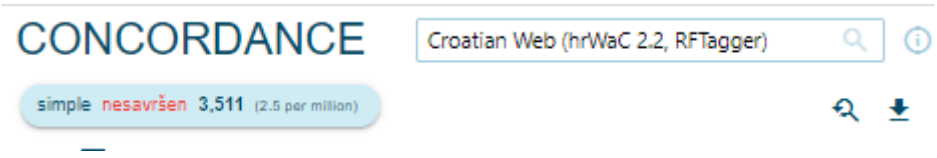
U kontekstu glagolskoga oblika, ispravan pojam je *imperfektivan*, a u takvom obliku postoje samo tri primjera u hrWaC-u (Slika 22), od kojih se sva tri odnose na njegovo ispravno značenje, odnosno prema *Hrvatskom jezičnom portalu*[32] – „ima svojstva imperfekta“, dok je imperfekt definiran kao „glagolski oblik, prošlo nesvršeno vrijeme“.



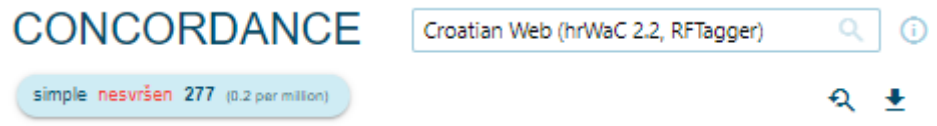
Slika 22 Učestalost i kontekst pridjeva 'imperfektivan' u hrWaC-u

Prefiks *im-* ponekad se može zamijeniti prefiksom *ne-,* a uz promjenu osnove istoga značenja su inačice *nesavršen* umjesto *imperfektan* i *nesvršen* umjesto *imperfektan.* Takvih oblika puno je više u hrWaC-u, međutim *nesvršen* ima puno šire značenje, odnosno *imperfektan* se odnosi samo na glagolski oblik. Iz toga se može zaključiti da promjena prefiksa ipak utječe i na značenje, odnosno s prefiksom *im-* u ovom primjeru ono je vrlo jasno i vrlo usko određeno.

Slika 23 prikazuje učestalost pridjeva *nesavršen,* a Slika 24 učestalost pridjeva *nesvršen*:



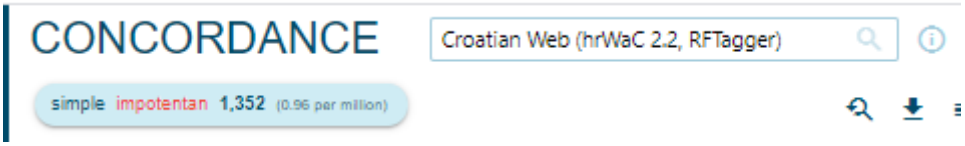
Slika 23 Učestalost pridjeva 'nesavršen' u hrWac-u



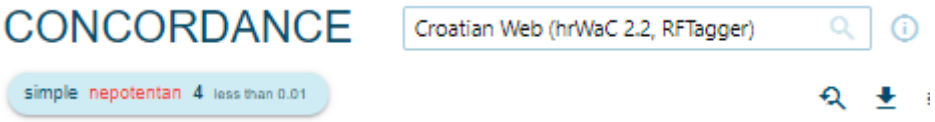
Slika 24 Učestalost pridjeva 'nesvršen' u hrWaC-u

[32] http://hjp.znanje.hr/

S druge strane, pridjev *impotentan* češći je od pridjeva *nepotentan,* što prikazuje Slika 25 i 26:



Slika 25 Učestalost pridjeva 'impotentan' u hrWaC-u



Slika 26 Učestalost pridjeva 'nepotentan' u hrWaC-u

Kratkom analizom ovih pridjeva u hrWaC-u, može se zaključiti da se pridjev *impotentan* upotrebljava u oba značenja definirana prema prema *Hrvatskom jezičnom portalu*[33]. Naime, pridjev *impotentan* dolazi od imenice *impotencija* koja ima dva moguća značenja – općenito „nemoć, nesposobnost" i drugo, specifičnije značenje – „spolna nemoć muškarca, nesposobnost postizanja ili održavanja erekcije". S druge strane, pridjev *nepotentan* upotrebljava se samo u prvom, općenitijem značenju, a ne u značenju spolne impotencije.

Prefiks in-

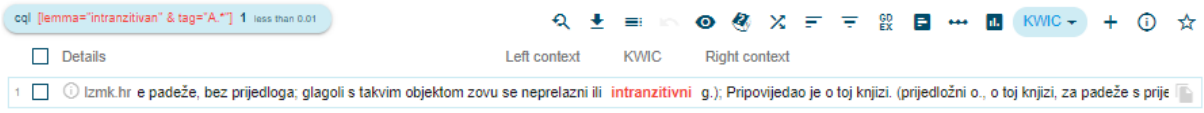
Pridjeva s ovim prefiksom je malo, stoga pretraga korpusa pomoću klasičnog upita kojim se uvjetuje početak riječi s *in-* i oznakom za pridjev nije dovoljno učinkovit jer rezultira s puno pogrešnih rezultata. Ovaj prefiks je latinskog podrijetla i uglavnom se pojavljuje s osnovama latinskog podrijetla. Primjeri u kojima se odnosi na negaciju određenu ovim radom su *intranzitivan, inkontinentan, inkompetentan, intolerantan.* Na ovim primjerima vidljivo je da je ovaj prefiks u značenju prefiksa *ne-,* jer negira značenje osnove i time tvori izravnu negaciju, a u nekim primjerima može se i zamijeniti prefiksom *ne-*. Pridjev *intranzitivan* ima i

[33] http://hjp.znanje.hr/

hrvatsku inačicu *neprijelazan; inkompetentan* se može zamijeniti s *nekompetentan,* a *intolerantan* se upotrebljava i kao *netolerantan.*

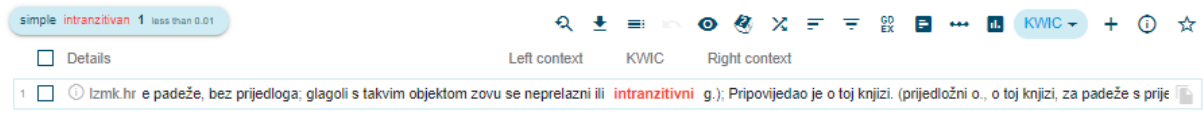
Upit postavljen za ove pridjeve sastavljen je ovako:

[lemma="intranzitivan" | lemma="inkontinentan" | lemma="inkompetentan" | lemma="intolerantan" & tag="A.*"], a pretraga je pokazala da je pridjev *intranzitivan* u hrWacu hapaks (Slika 27).



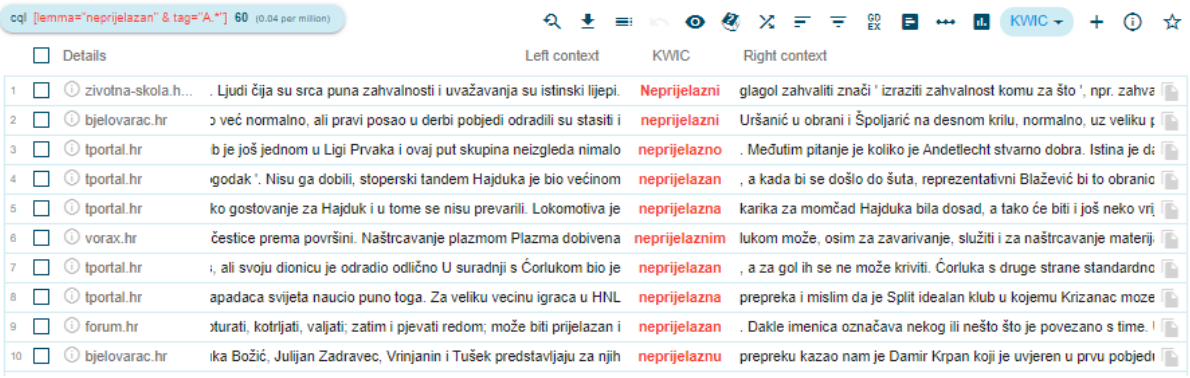
Slika 27 Pridjev 'intranzitivan' kao hapaks u hrWaC-u

Čak i kada se makne oznaka za pridjev, kako bi se obuhvatili mogući primjeri u kojima je pogrešno označen kao neka druga vrsta riječi, i dalje ima samo jednu pojavu u ovom korpusu (Slika 28).



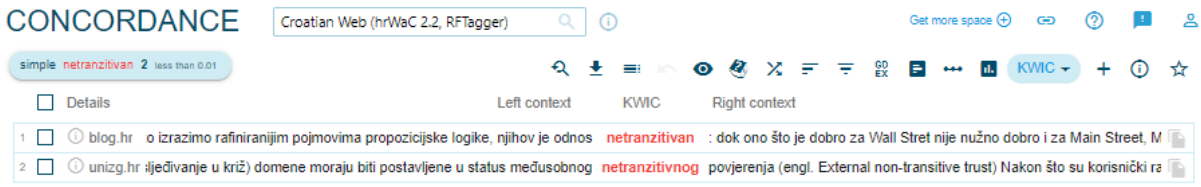
Slika 28 Učestalost primjera 'intranzitivan' bez pridjevske oznake

Pretraga s hrvatskom inačicom *neprijelazan,* generira 60 rezultata, od kojih nije svaki primjer upravo u ovom značenju. Stoga se može zaključiti da iako se *intranzitivan* i *neprijelazan* mogu i jedan i drugi upotrebljavati u kontekstu glagola – koji mogu biti neprijelazni ili intranzitivni, ipak pridjev *intranzitivan* ima uže značenje, odnosno u hrvatskom jeziku pojavljuje se samo u tom značenju. S druge strane, inačica *neprijelazan* odnosi se na više značenja (Slika 29):



Slika 29 Kontekst pridjeva 'neprijelazan' u hrWaC-u

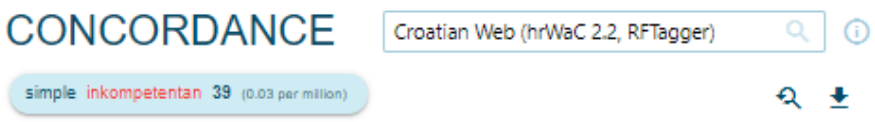
S druge strane kada se na pridjev iste osnove doda prefiks *ne-* umjesto *in-*, čineći pridjev *netranzitivan,* pretraga generira samo dva rezultata (Slika 30), od kojih nijedan nije u istom značenju kao pridjev *intranzitivan,* za koji je istaknuto da se pojavljuje samo u značenju koje se tiče glagola.



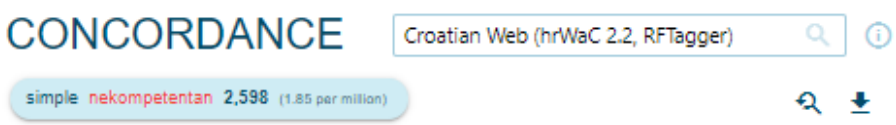
Slika 30 Učestalost pridjeva 'netranzitivan' u hrWaC-u

Pridjev *inkontinentan,* zamijenjen prefiksom *ne-* na istoj osnovi, čineći pridjev *nekontinentan,* ne pojavljuje se u hrWaC-u.

Međutim, pridjev *inkompetentan* (Slika 31) rjeđi je u hrWaC-u od pridjeva *nekompetentan* (Slika 32):

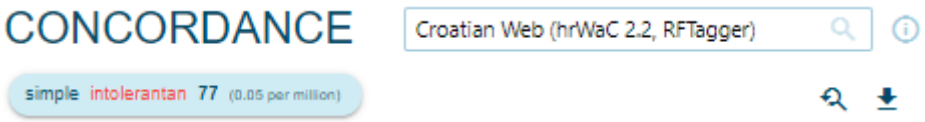


Slika 31 Učestalost pridjeva 'inkompetentan' u hrWaC-u

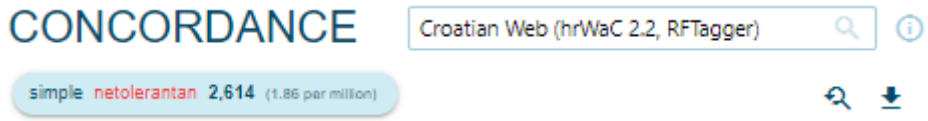


Slika 32 Učestalost pridjeva 'nekompetentan' u hrWaC-u

Isti je slučaj s pridjevima *intolerantan/netolerantan* u hrWaC-u, inačica s prefiksom *ne-* značajno je češća (Slika 33 i 34).



Slika 33 Učestalost pridjeva 'intolerantan' u hrWaC-u



Slika 34 Učestalost pridjeva 'netolerantan' u hrWaC-u

Prefiks izvan- / van-

Prefiks *izvan- / van-* prefiks je koji nije spomenut u radu o negaciji imenica (Žanpera i sur., 2020), jer se odnosi isključivo na tvorbu složenih pridjeva, ali se općenito i u literaturi o tvorbi pridjeva slabo spominje. Točnije, ovaj prefiks pojavljuje se uz imenice samo u obliku prijedloga kao samostalne riječi – *izvan prostora, izvan vremena,* od čega nastaju složeni pridjevi *izvanprostorni, izvanvremenski* itd. Međutim, kao što je već navedeno u pregledu hrvatskih gramatika, spominje ga Babić (2002) te ga vrijedi spomenuti i u ovom poglavlju jer sudjeluje u tvorbi negacije. Kako je istaknuo i Babić (2002), prefiks *izvan-* ponekad se može zamijeniti prefiksom *ne-*.

Ovaj rad bavi se prefiksalnom tvorbom negacije, a prefiks *izvan-* odnosi se na prefiksalno-sufiksalnu tvorbu, jer se osim prefiksom *izvan-*, pridjevi tvore i dodavanjem sufikasa poput *-ski, -čki, -ni.* Bez obzira na uglavnom ne-čistu prefiksalnu tvorbu, plodan je u tvorbi složenih pridjeva te ponekad tvori i izravnu negaciju, određenu u kontekstu ovoga rada, stoga će biti uključen u korpus za računalnu analizu.

Međutim, njegova klasifikacija isključivo unutar prefiksalno-sufiksalne tvorbe je diskutabilna. Naime, prema etimologiji *Hrvatskog jezičnog portala*[34], pridjev *izvanvremenski*

[34] http://hjp.znanje.hr/

nastao je od pridjeva *izvanvremen,* koji je nastao prefiksalnom tvorbom, odnosno dodavanjem prefiksa na pridjev *vremen.*

Ovaj prefiks pojavljuje se i u skraćenoj inačici *van-,* odnosno osim *izvanbračni* može se reći i *vanbračni,* umjesto *izvanjezični – vanjezični* i slično. *Hrvatski jezični portal*[35] navodi obje inačice i definira ga „kao prvi dio složenih pridjeva označuje da je nešto izvan onoga što znači pridjev u drugom dijelu složenice".

Označuje negaciju u kontekstu da se nešto *ne odvija* u okvirima označenima osnovom. Primjerice *izvanvremenski* se prema *Hrvatskom jezičnom portalu*[36] odnosi na „koji nije u vremenu, koji se ne događa u odsjeku vremena", *izvanbračni* se ne odvija u braku, *izvansudski* se ne odvija na sudu, i sl. Upravo u navedenom primjeru *izvanvremenski* može se zamijeniti i prefiksom *ne-,* jer prema *Hrvatskom jezičnom portalu*[37] pridjevi *nevremenski* i *izvanvremenski* su istoznačnice.

S druge strane, ne tvori uvijek negaciju u izravnom smislu, odnosno ne može se uvijek zamijeniti prefiksom *ne-.* Primjerice, pridjev *izvanškolski,* ne znači isto što i *neškolski,* jer neke aktivnosti mogu biti *izvanškolske,* ali i dalje se radi o *školskim* aktivnostima.

Vrijedi spomenuti i da je primijećeno da ovaj prefiks kada tvori negaciju, uglavnom tvori antonimno, odnosno suprotno značenje. Točnije, najčešće ne tvori negaciju u smislu da „dodavanjem prefiksa nešto što *jest,* više *nije*", već upravo određenu drugim dijelom definicije negacije u okvirima ovoga rada „dodavanjem prefiksa nešto dobiva suprotno značenje". Stoga je *izvanškolski* suprotan pojam od *školski,* ali ne negira *školski.* Neki pridjevi koje korpus obuhvaća odnose se na izravno negiranje – *izvanmaternični* je ujedno i *nematernični,* kao što je *izvanblokovski* ujedno i *neblokovski,* a neki primjeri mogu se činiti diskutabilni – radi li se o tvorbi suprotnoga pojma ili samo o pojmu koji označava nešto *van* osnove, bez da tvori negaciju. Međutim, s obzirom da definicija negacije u ovom radu obuhvaća i negiranje osnove i tvorbu suprotnoga pojma, primjeri s ovim prefiksom obuhvaćat će i neke diskutabilnog značenja.

[35] http://hjp.znanje.hr/
[36] http://hjp.znanje.hr/
[37] http://hjp.znanje.hr/

Prefiks kontra-

Pretraga korpusa putem upita:

[lemma="kontra.*" & tag="A.*"] generirala je i pogrešne rezultate, poput *kontrastna, kontraverzna* (umjesto *kontroverzna*, ali nijedan od tih oblika nije negacija), *kontraadimiralski* koji također nije negacija već se prema *Hrvatskom jezičnom portalu*[38] odnosi na *kontraadmirala,* najniži admiralski čin.

Pridjevi koji s ovim prefiksom tvore negaciju su *kontrarevolucionaran, kontraobavještajan, kontraproduktivan* i slični, a sveukupno su u hrvatskom jeziku malobrojni. Ovaj prefiks je prefiks kojim se tvori suprotno značenje osnove, latinskoga je podrijetla, a u hrvatskom se ponekad može zamijeniti prefiksom *protu-,* o čemu će se više pisati u posljednjem odlomku ovoga poglavlja.

Pridjevi poput *kontraceptivni/kontracepcijski* nisu uključeni u analizu jer u hrvatskom jeziku ne postoje bez prefiksa *kontra-* kao samostalni pridjevi, odnosno osnove.

(Prefiks mal-)

Ovaj prefiks naveden je u zagradi jer ne tvori negaciju u okviru definicije ovoga rada, stoga nije uključen u korpus za analizu. Pridjevi koji sadrže ovaj prefiks nastaju od imenica, ali nema svaka imenica i svoju pridjevsku inačicu. Međutim, tvorba ovim prefiksom u hrvatskom jeziku općenito je slabo plodna, a pretraga hrWaC-a pokazala je da dolazi s riječima stranoga podrijetla, tvoreći neizravnu negaciju. Riječ je o prefiksu latinskoga podrijetla, jer prema *Hrvatskom jezičnom portalu*[39] dolazi od latinske riječi *malus* – što znači loš te riječi koje tvori imaju značenje upravo nečeg lošeg, odnosno negativnog, ali ne radi se o izravnoj negaciji prilikom koje bi taj prefiks negirao osnovu ili stvorio suprotni pojam.

Primjerice, neke od imenica s ovim prefiksom su *malformacija, malapsorpcija, malnutricija, malrotacija, maldigestija,* dok je u hrWaC-u kao pridjev tvoren od tih imenica prisutan jedino pridjev *malformiran.* Međutim, upravo kako je navedeno, radi se o neizravnoj negaciji i taj pridjev stoga nije uključen u analizu. U značenju ovoga pridjeva, radi se o *lošoj*

[38] http://hjp.znanje.hr/
[39] http://hjp.znanje.hr/

formaciji, o odstupanju od normalne građe, ali ovaj pridjev samo ima negativno značenje i nema svoju jesnu/afirmativnu inačicu, kao *abnormalan/normalan.*

Ekvivalent ovoga prefiksa u nekim slučajevima u hrvatskom jeziku mogao bi biti *pod-/(pot-),* primjerice *pothranjen* (engl. *malnurished*)*,* međutim također tvori neizravnu negaciju, kao i ostali primjeri s prefiksom *mal-,* stoga nije uključen u korpus i računalnu analizu.

Prefiks ne-

Prefiks *ne-* najplodniji je prefiks u tvorbi negacije – bilo pridjeva, bilo imenica, o čemu su pisali i brojni drugi autori izneseni u teorijskom dijelu rada. To se potvrdilo i prikupljanjem ovoga korpusa, što će biti prikazano u grafikonu na kraju poglavlja o prikupljanju korpusa. Ovaj prefiks prema *Hrvatskom jezičnom portalu*[40] negira značenje osnove, znači opreku osnove – odnosno tvori suprotno značenje osnove ili govori da je nešto bez onoga što sadrži ne-negirana riječ. Prefiks *ne-* hrvatska je inačica koja također ponekad bez promjene značenja može zamijeniti i prefikse stranoga podrijetla poput *a-, an, i-, im-* i *in-* što je već spomenuto kod navedenih prefikasa. Također, može se pojaviti s istom osnovom kao i drugi prefiks domaćega podrijetla (*neuspješan/bezuspješan*).

Pretraga hrWaC korpusa regularnim izrazom:

[lemma="ne.*" & tag="A.*"] generira puno ispravnih rezultata, poput *neočekivan, nezakonit, nezamisliv, nečovječan* itd., ali generira i rezultate koji se ne odnose na prefiks *ne-,* poput *nebeski,* koji jest pridjev, ali *ne-* nije prefiks već je u osnovi riječi *nebo.* S druge strane, generiraju se i rezultati koji su ishod pogrešno označenih vrsta riječi, poput *neki,* što je zamjenica pogrešno označena kao pridjev. Na temelju primjera prikupljenih iz gramatika, pravopisa i rječnika, ali i putem vlastitog poznavanja jezika, prikupljeni su pridjevi koji zaista s prefiksom *ne-* čine negaciju. Međutim s obzirom na to da prikupljeni korpus obuhvaća i kotekst, odnosno ne samo ključne riječi, već i okolinu, zasigurno sadrži i koju zamjenicu koja započinje s *ne-* ili imenicu koja započinje s *ne-,* bez da se radi o negaciji i prefiksu. Prema pogreškama s kojima su se susreli Žanpera i sur. (2020) prilikom prepoznavanja prefiksalne negacije imenica – gdje je jedna od njih bila upravo pogreška zbog homografije – nešto se piše

[40] http://hjp.znanje.hr/

isto/započinje istim fonemima te ga stoga gramatika prepoznaje kao traženi izraz, takve pogreške mogu se očekivati i ovom računalnom analizom, stoga je gramatiku potrebno ograničiti.

(Inačice prefiksa ne-)

U inačicama prefiksa *ne-,* kao što su *one-, une-, izne-* i *uzne-* uglavnom je riječ o glagolskim pridjevima trpnim (primjerice *unereñen*) te glagolskim pridjevima radnim (primjerice *iznemogao*). Međutim, za razliku od prefiksa *ne-* koji je najplodniji prefiks u tvorbi negacije, njegove inačice vrlo slabo su plodne. Također, uglavnom se radi sufiksalnoj, a ne prefiksalnoj tvorbi – glagolski pridjev *unereñen* dolazi od nesvršenog glagola *unereñivati,* nastalog od svršenog glagola *unerediti,* koji je nastao od imenice *nered.* Odnosno, za tvorbu glagolskog pridjeva *unereñen* od glagola *unereñivati,* samo se odvija zamjena jednog sufiksa drugim. Negacija se stoga odvija još na razini imenice: *red/nered*, iako i glagol i pridjev imaju svoje jesne inačice – svršeni glagol: *unerediti / urediti,* nesvršeni glagol: *unereñivati / ureñivati,* glagolski pridjev trpni: *unereñen / ureñen.* Usporedbom jesnih i niječnih inačica na razini pridjeva, tvorba njihovih negacijskih parnjaka zapravo nastaje umetanjem *ne-* iza već postojećeg *u-,* stoga se ni s tog gledišta ne radi o prefiksalnoj tvorbi, odnosno prefiks nije *une-*, već se umeće *ne-* iza već postojećeg prefiksa. Ovakvi primjeri u NooJ-u se obrađuju u kategoriji glagola i zahtijevaju posebnu obradu, zbog čega ih gramatika napravljena za svrhu ovoga rada neće obuhvaćati.

(Prefiks one-)

Pretraga korpusa putem upita:

[lemma="one.*" & tag="A.*"]

generira većinom ispravne rezultate, poput *onečišćen, onesposobljen, onesviješten, oneraspoložen, onemogućen*, od kojih svi imaju svoje jesne inačice *očišćen, osposobljen, osviješten, oraspoložen, omogućen*. Rezultat uključuje i primjer *onemoćao*, koji nema svoj potvrdni parnjak jer ne postoji *omoćao* niti *moćao*, stoga se ne radi o izravnoj negaciji, već samo o riječi negativnoga značenja, odnosno neizravnoj negaciji, zbog čega taj primjer nije uključen u korpus za analizu. Primjeri *moćan/nemoćan* činili bi izravnu negaciju, ali *onemoćao*

se samo odnosi na nekoga bez snage, slaboga – dakle ima negativno značenje, ali nema svog parnjaka suprotnoga značenja.

(Prefiks une-)

Glagolski pridjevi koji sadrže ovaj prefiks i tvore negaciju definiranu ovim radom su *unereden, unesrećen, uneravnotežen,* s afirmativnim ekvivalentima *uređen, usrećen, uravnotežen.* Pretragom:

[lemma="une.*" & tag="A.*"] generiraju se i primjeri *unesen/unešen,* međutim oni nastaju od glagola *unijeti* i ne odnose se na negaciju. Kao i prethodno spomenuti primjeri glagolskih pridjeva trpnih, i ovi bi se trebali obraditi zasebno.

(Prefiks izne-)

Ovaj prefiks također je naveden u zagradi, jer osim što nije riječ o pridjevima, već o glagolskim pridjevima trpnim, nastalima od glagola, ovim prefiksom se, kao i prefiksom *mal-,* ne tvori izravna negacija. Radi se o primjerima poput *iznemogao* i *iznemoćao* (koji je u hrWaC-u hapaks), a prema *Hrvatskom jezičnom portalu*[41] dolaze od glagola *iznemoći* koji znači „klonuti, malaksati“, pri čemu ne tvore izravnu negaciju i nemaju svoj suprotni parnjak, stoga nisu uključeni u korpus ovoga rada. Pretraga hrWaC-a generira i rezultate poput *iznerviran, iznenađen/iznenadni* koji također nisu negacija te *iznesen/iznešen* koji nastaju od glagola *iznijeti* i također se ne odnose na negaciju.

(Prefiks uzne-)

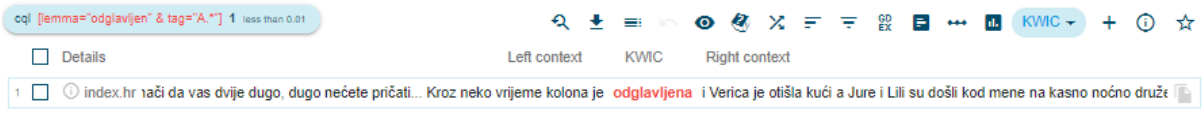
Primjer s „prefiksom“ *uzne-* je diskutabilan, jer glagolski pridjev *uznemiren* dolazi od glagola *uznemiriti,* koji nastaje od imenice *nemir* te iako postoji parnjak pridjeva *nemiran,* a to je *miran,* parnjak glagolskog pridjeva *uznemiren* bio bi *umiren,* pri čemu *uznemiren* nastaje dodavanjem *-zne,* a ne *uzne-*. Stoga, iako je riječ o negaciji u postavljenim okvirima, odnosno

[41] http://hjp.znanje.hr/

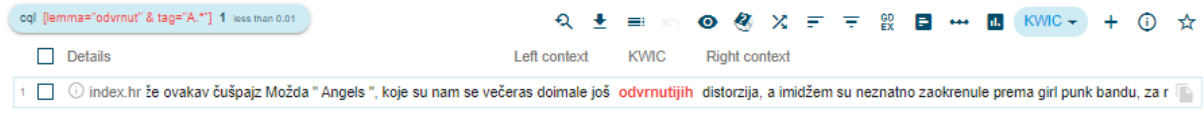
postoji i suprotni ekvivalent riječi *uznemiren*, ovaj primjer neće biti uključen u gramatiku za računalno prepoznavanje prefiksalne negacije pridjeva.

(Prefiks od-)

Iako ovi primjeri također neće biti uključeni u računalnu analizu jer nije riječ ni o čistoj prefiksalnoj tvorbi ni o tvorbi nastaloj od drugoga pridjeva, već od glagola, vrijedi spomenuti i ovakve slučajeve. Radi se o glagolima poput *odspojiti, odvrnuti, odglaviti* od kojih nastaju pridjevi *odspojen* (kao suprotnost od *spojen*), *odvrnut* (kao suprotnost od *zavrnut*), *odglavljen*, (kao suprotnost od *zaglavljen*). *Odvrnut* i *odglavljen* prikazani na slikama 35 i 36 u hrWaC-u su hapaksi, odnosno prema *Hrvatskom jezičnom portalu*[42] – riječ koja se u korpusu pojavljuje samo jedanput.

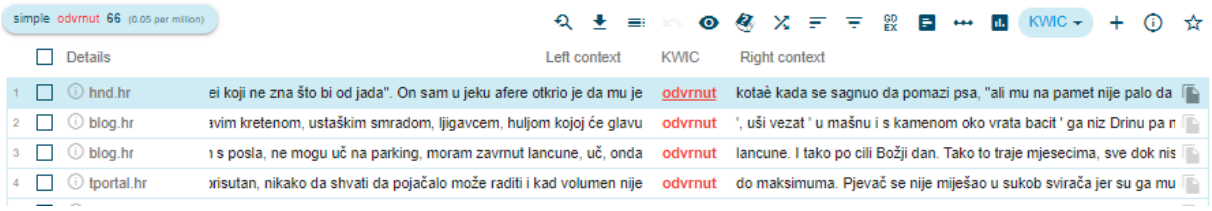


Slika 35 Pridjev 'odglavljen' kao hapaks u hrWaC-u



Slika 36 Pridjev 'odvrnut' kao hapaks u hrWaC-u

Međutim, kada se riječi *odvrnut* makne pridjevska oznaka, korpus sadrži 66 primjera (Slika 37) koji se u tom slučaju odnose na glagole, s redukcijom infinitiva *odvrnut'* umjesto *odvrnuti,* što se odnosi na radnju, primjerice *odvrnuti uši.*

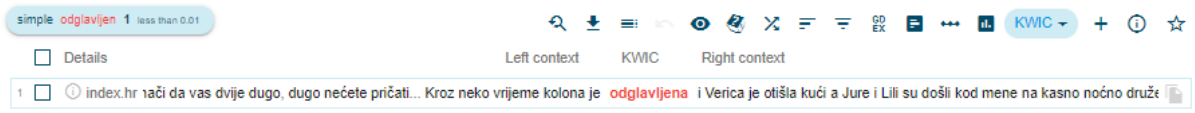


Slika 37 Učestalost pridjeva 'odvrnut' bez pridjevske oznake

[42] http://hjp.znanje.hr/

Također, u prvom primjeru *„(...)otkrio je da mu je odvrnut kotač(...)“,* radi se o pridjevu, ali je očito pogrešno označen prilikom *POS tagginga,* odnosno označavanja vrsta riječi. Stoga treba uzeti u obzir da se ne bi radilo o hapaksu kada bi oznake bile pravilno dodijeljene.

S druge strane, riječ *odglavljen* koja je s oznakom za pridjev hapaks, također je hapaks i kada se ta oznaka isključi u pretrazi (Slika 38):

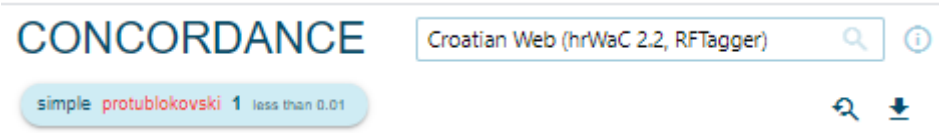


Slika 38 Učestalost pridjeva 'odglavljen' bez pridjevske oznake

Prefiks protu-

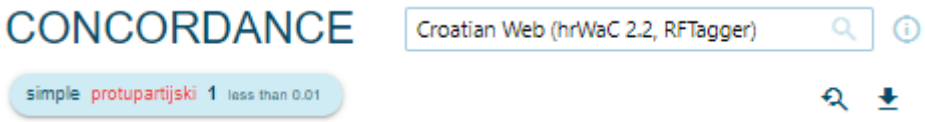
Prema *Hrvatskom jezičnom portalu*[43]*,* prefiks *protu-* prefiks je praslavenskog korijena koji dolazi od prijedloga *protiv.* Analizom pridjeva iz hrWaC-a koji su negirani tim prefiksom, primijećeno je da se ovaj prefiks najčešće pojavljuje u značenju „djeluje protiv čega“ – *protudržavni, protuzakonit, protuavionski, protutenkovski, protureformacijski* itd. Na temelju navedenih primjera također se može primijetiti i učestala pojava ovog prefiksa uz ideološke i vojne pojmove.

Neki pridjevi negirani ovim prefiksom su zastarjeli pojmovi specifičnoga značenja, poput *protublokovski* i *protupartijski,* pronađeni u *Hrvatskom pravopisu* iz 1971. godine te se oba navedena pojma u hrWaC-u pojavljuju samo jednom (Slika 39 i 40).



Slika 39 Pridjev 'protublokovski' kao hapaks u hrWaC-u

[43] http://hjp.znanje.hr/



Slika 40 Pridjev 'protupartijski' kao hapaks u hrWaC-u

Pridjev *protublokovski* odnosi se na politiku bivše Jugoslavije prema kojoj su bili *protiv* blokovske politike (primjerice američke i sovjetske). *Blok* se prema *Hrvatskom jezičnom portalu*[44] odnosi na „savez država ili političkih stranaka povezanih zajedničkim političkim, ideološkim, gospodarskim i drugim interesima“. Ovo se spominje kako bi se u istom kontekstu naveo još jedan mogući prefiks – *izvan,* odnoseći se na *izvanblokovsku* politiku, koja se za razliku od *protublokovske* ne odnosi na politiku *protiv* nekog od političkih blokova, već jednostavno na države koje ne pripadaju nijednom političkom bloku, odnosno *neblokovske.*

Pridjev *protunaravan* također se može smatrati zastarjelom riječju, jer danas se u tom kontekstu više upotrebljava *prirodan,* umjesto *naravan* te to potvrđuje i učestalost u hrWaC-u, gdje se pridjev *protunaravan* pojavljuje 124 puta, a protuprirodan 1008 puta.

Zamjena prefiksa protu- s anti- i kontra-

Kako je navedeno i u teorijskom dijelu rada gdje su prikazani prefiksi u hrvatskim gramatikama, prefiks *protu-* ponekad se može zamijeniti prefiksima *anti-* i *kontra-*. Odnosno, prefiks *protu-* češće se upotrebljava s riječima domaćeg podrijetla, a *anti-* i *kontra-* s riječima stranoga podrijetla. Stoga se ponekad zamjenom ovih prefikasa događa i zamjena osnove, ali ponekad se mijenjaju samo prefiksi, a osnova ostaje ista. Iz tog razloga napravljena je kratka analiza ovih prefikasa u hrWaC-u.

Primjerice, uz pridjev latinskog podrijetla – *religiozan,* pojavljuje se prefiks *anti-* tvoreći pridjev *antireligiozan* koji se u hrWaC-u pojavljuje 61 put. S druge strane, istoznačna osnova praslavenskog podrijetla – *vjerski,* pojavljuje se češće uz prefiks *protu-* tvoreći pridjev *protuvjerski* koji se u hrWaC-u pojavljuje 116 puta. Moguće su i kombinacije prefikasa s obje

[44] http://hjp.znanje.hr/

osnove, ali one su rjeđe – primjerice *antivjerski* se pojavljuje 43 puta, a *protureligiozan* samo 3 puta.

Isto je s pridjevom *antivladin* i *protuvladin,* gdje je *vladin* praslavenska osnova i pojavljuje se češće s prefiksom *protu-,* točnije 568 puta, a s *anti-* samo 32 puta. Prema tome se može potvrditi i ono što literatura navodi – *protu-* i *anti-* mogu se u nekim kontekstima zamjenjivati, ali *protu-* je češći s domaćim, a *anti-* sa stranim riječima.

Slično potvrđuju i brojni drugi primjeri. Primjerice *infektivan* se pojavljuje 51 put s prefiksom *anti-,* a samo 6 puta s prefiksom *protu-.* Inačica *zarazan* češća je uz *protu-,* iako je ta kombinacija vrlo rijetka jer se pojavljuje samo tri puta, dok se *antizarazan* pojavljuje samo jednom.

Nadalje, riječ stranog podrijetla – *korupcija,* u obliku pridjeva češća je s prefiksom *antikoruptivan/antikorupcijski,* s kojim se pojavljuje 120, odnosno 3324 puta. S druge strane, *protukoruptivni* se pojavljuje samo dvaput, a *protukorupcijski* 238 puta, i dalje znatno rjeđe od *antikorupcijski.* Pridjev *antifašistički* pojavljuje se 10 338 puta, a *protufašistički* samo 13 puta itd.

Međutim, „pravilo“ o podrijetlu riječi nije uvijek pravilo, odnosno postoje i iznimke. Naime, pridjev *antifungalni,* koji se sastoji i od stranog prefiksa i od strane osnove, pojavljuje se u hrWaC-u 10 puta, a *antigljivični* – kombinacija stranog prefiksa i domaće osnove pojavljuje se 115 puta. S druge strane, iako je *gljiva* osnova praslavenskog podrijetla, rjeđe se pojavljuje s prefiksom *protu-,* jer pridjev *protugljivični* pojavljuje se samo 28 puta.

Usporedba s prefiksom *kontra-* također potvrđuje iznimke od pravila, odnosno da „pravila“ zapravo nema. Primjerice, iako pridjev *obavještajni* dolazi od *vijest* koja potječe od praslavenskog korijena, češći je u kombinaciji s *kontra-,* prefiksom stranoga podrijetla s kojim se pojavljuje 463 puta, nego s domaćim prefiksom *protu-* s kojim se pojavljuje samo 26 puta.

Ono što se ovom analizom također primijetilo je da dok su ponekad osnove domaćega podrijetla češće sa stranim prefiksom, nego prefiksom *protu-,* neke riječi „domaćeg“, odnosno praslavenskog podrijetla pojavljuju se samo s prefiksom *protu-.* Primjerice, pridjev *protupožarni* pojavljuje se isključivo s tim prefiksom, odnosno nema nijedne pojave s prefiksima *anti-* i *kontra-.*

S druge strane, ponekad je moguća pojava sva tri prefiksa s istom osnovom. Primjerice, pridjev *revolucionaran* pojavljuje se sa sva tri prefiksa, od čega najčešće s prefiksom *kontra-.*

Točnije, pridjev *kontrarevolucionaran* pojavljuje se 296 puta, a pridjevi *proturevolucionaran* i *antirevolucionaran* imaju jednak broj pojava u hrWaC-u, gdje se svaki pojavljuje 30 puta.

Ono što je također zanimljivo i stoga vrijedi spomenuti, je da imenice *revolucija* i *reformacija,* obje latinskog podrijetla „preferiraju" različite prefikse. Točnije, *kontrarevolucija* se u hrWaC-u pojavljuje 484 puta, a *proturevolucija* samo 6 puta. S druge strane, *kontrareformacija* se pojavljuje samo 3 puta, a *protureformacija* 198 puta.

Naposljetku, vrijedi spomenuti i odnos između prefikasa *anti-* i *kontra-,* nevezano za prefiks *protu-.* Naime, riječ *antibaby/antibejbi,* u kontekstu *antibaby pilula* je engleski izraz koji se u hrvatskom jeziku upotrebljava u istom značenju kao *kontraceptivni/kontracepcijski.* Međutim, s obzirom na to da je kod primjera *antibaby* riječ isključivo o frazi uz *pilula/pilule* te nije sasvim jasno kako označiti *antibaby,* jer ponekad je u korpusu označen kao A – odnosno pridjev, a ponekad kao N – imenica ili kao X – ostalo, uz oznaku f za strano. Dodatno, dok *antibaby* ima samo usko značenje uz pojavu s *pilula/pilule,* pridjev *kontraceptivni/kontracepcijski* ima šire značenje, pojavljujući se uz *pilula, tableta, metoda, sredstvo* i sl.

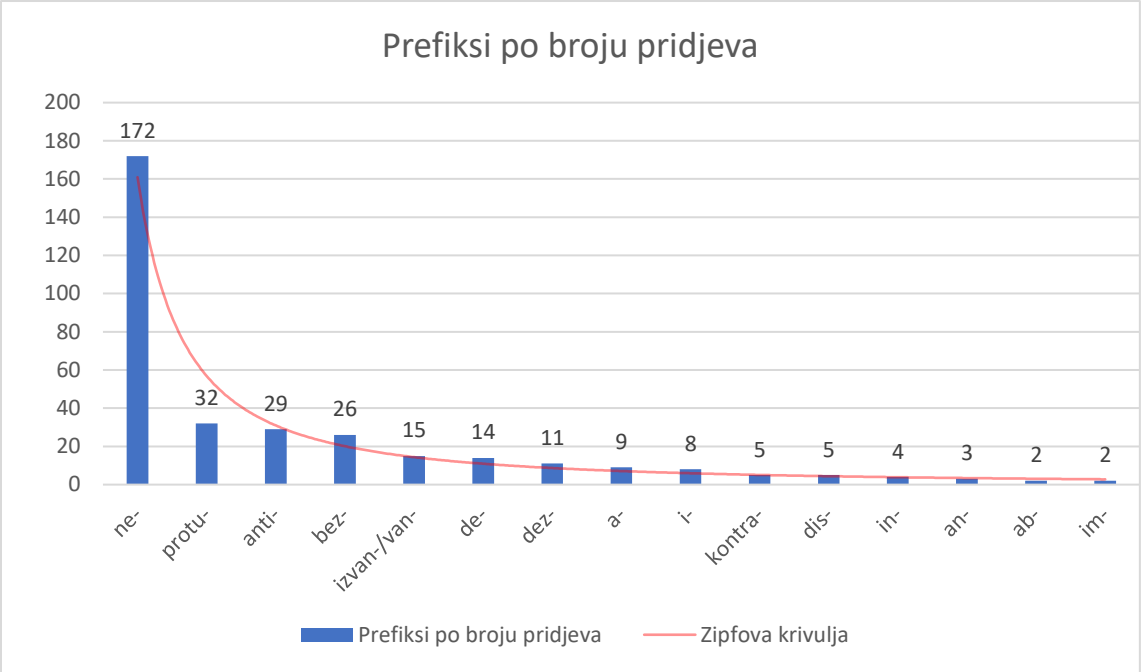
Podaci o prikupljenom korpusu

Prikupljeni korpus sastoji se od 2024 konkordancije, prikupljene sa 322 pridjeva koji s prefiksom tvore izravnu negaciju. Popis svih pridjeva i prefikasa koje korpus obuhvaća nalazi se u Tablici 2.

Tablica 2 Popis pridjeva po prefiksima za prepoznavanje u NooJ-u

a	dinamičan	bes	cementan		inficiran		indikativan	katolički	pobjediv	selektivan		zdrav
	logičan		ciljan		informacijski		obavještajni	kontaktibilan	podnošljiv	slavan		željen
	menoroičan		kompromisan		informativan		produktivan	koristan	poduzetan	spojiv		znanstven
	moralan		konačan		informiran		revolucionaran	korišten	pokoriv	sporan		zreo
	normalan		pomoćan		integracijski	ne	adekvatan	kršćanski	pokretan	sretan	protu	avionski
	političan		poštedan		integriran		aktivan	kršten	pomičan	stabilan		blokovski
	seksualan		povratan		organiziran		artikuliran	kvalitetan	popravljiv	steroidan		domovinski
	simetričan		prijekoran		orijentiran		bitan	legalan	poreciv	strpljiv		državni
	socijalan		principijelan	dis	balansiran		blag	lijep	posredan	stručan		eksplozijski
ab	normalan		težinski		funkcionalan		cijenjen	ljudski	pošten	stvaran		fašistički
	oralan	be	skrupulozan		harmoničan		civiliziran	logičan	potpisan	sumjerljiv		gljivični
an	organski		smislen		kvalificiran		častan	maran	potpun	supstancijalan		gradni
	orgazmičan		šuman		proporcionalan		čestit	mastan	potreban	svakidašnji		hrvatski
	ovulacijski		zvučan	i	legalan		čist	milosrdan	poznat	svjestan		infektivan
anti	alergijski		žičan		lojalan		čujan	miran	pravedan	svrstan		kršćanski
	bakterijski		životan		materijalan		davni	mjerljiv	pravilan	taknut		naravan
	celulitni	bez	bolan		mobilan		disciplиniran	moćan	pravodoban	talijanski		narodni
	civilizacijski		brižan		moralan		djelotvoran	moguć	prekidan	točan		oklopni
	crkveni		grešan		racionalan		dokazan	moralan	prepoznat	tolerantan		osovinski
	diskriminacijski		obziran		realan		dokučiv	nadahnut	prepoznatljiv	tradicionalan		partijski
	fašistički		opasan		regularan		doličan	običan	prerađen	udata		požarni
	fungalni		pridržajan	im	perfektan		donesen	objašnjiv	prihvatljiv	ugodan		pravni
	gljivični		rezervni		potentan		dopustiv	obrativ	prijateljski	ujednačen		prepadni
	hrvatski		uspješan	in	kompetentan		dosljedan	očekivan	prikladan	umjestan		prirodan
	infektivan		uvjetan		kontinentan		dostatan	odgodiv	priličan	umjetnički		propisan
	kancerogen		voljan		tolerantan		dostižan	odgovoren	primjeren	upotrebljiv		provalni
	karcinogen		blokiran		tranzitivan		dovoljan	odlučan	pristojan	uravnotežen		ratni
	korupcijski	de	centralizacijski	izvan/van	bolnički		dovršen	ograničen	privilegiran	uređen		reformacijski
	koruptivni		centraliziran		bračni		dvojben	okaljan	priznat	uspješan		tenkovski
	lipidni		hidratiziran		institucionalni		elastičan	opran	profitan	ustrašiv		upalni
	mikrobni		hidriran		jezični		goriv	opravdan	promijenjen	uvjerljiv		ustavni
	muslimanski		kompenziran		maternični		institucionaliziran	opterećen	prospavan	veseo		zakonit
	oksidativni		komponiran		nastavni		iscrpan	organiziran	proveden	vidljiv		zarazni
	parazitni		koncentriran		plućni		iskusan	organski	ranjiv	voljen		zračni
	septičan		kontaminacijski		proračunski		istinit	osjetljiv	razborit	zaboravan		zrakoplovni
	septički		kontaminiran		prostorni		izbježan	oštećen	razjašnjen	zaobilazan		životan
	stresni		moraliziran		redni		izlječiv	osvijetljen	razuman	zapaljiv		
	teroristički		reguliran		sudski		izmjenjiv	ovisan	razumljiv	zapamćen		
	tetički		stabilizacijski		školski		izreciv	ovlašten	razvijen	zaposlen		
	tuberkulozan		stabiliziran		tjelesni		izvjestan	ozbiljan	realan	zasićen		
	tumorski		ideologiziran		vremen		izvršen	oženjen	salomljiv	zasitan		
	virusni	dez	iluzioniran		vremenski		jasan	paran	samoupravan	zavisan		
	vladin		infekcijski	kontra	indiciran		karakteran	plodan	savršen	završen		

Prefiksi za tvorbu negacije pridjeva koje korpus obuhvaća su: *a-, ab-, an-, anti-, bez- (be, bes-), de-, dez-, dis-, i-, im-, in-, izvan- (van-), kontra-, ne-, protu-*. Učestalost pojedinog prefiksa u prikupljenom korpusu prikazana je u Grafikonu 1, a može se primijetiti da je, kako je navedeno i u teorijskom dijelu rada, prefiks *ne-* najplodniji.



Grafikon 1 Učestalost negacijskih prefikasa prema broju pridjeva u korpusu[45]

Također, ovom analizom primijećene su neke pojave o kojima se pisalo i u teorijskom dijelu, ali i neke suprotne njima. Naime, pronađeni su pridjevi koji imaju svoju prefiksalno negiranu inačicu, iako imaju i svoj ne-negirani parnjak. Primjerice, hrWaC sadrži pridjev *neveseo,* iako postoji *tužan; nelijep* iako postoji *ružan; nepametan,* iako postoji *glup* te *nečist* iako postoji *prljav.* S ovim pridjevskim parovima napravljena je kratka analiza učestalosti u hrWaC-u te iako se zaista svaki od navedenih pridjeva pojavljuje u obje inačice, ne-negirani pridjevski parnjaci puno su češći od negiranih prefiksom. Točnije, *nelijep* se pojavljuje 585 puta, a *ružan* 38 649 puta; *nepametan* se pojavljuje 351 puta, a *glup* 108 627, *neveseo* se pojavljuje 693 puta, a *tužan* 41 909 puta te se *nečist* pojavljuje 4 682 puta, a *prljav* 33 283 puta.

[45] Krivulja dobivena ovim grafikonom podsjeća na Zipfovu krivulju. George Zipf američki je lingvist koji je primijetio da se učestalost riječi može pratiti formulom, odnosno učestalost riječi proporcionalna je njezinom rangu (Zipf, 1932). Ne može se tvrditi da je ovdje slučaj u potpunosti takav, ali oblik krivulje zasigurno prati Zipfovu krivulju, gdje je *ne-* daleko najučestaliji prefiks, a prvi idući je puno niže rangiran, odnosno puno manje plodan, a ostali još manje.

Osim toga, pronađeni su i negirani pridjevi na višim razinama stupnjevanja. Primjerice pridjevi *nesretan, neizravan, nepoznat, bezbolan, dekoncentriran* i *dezorganiziran* u hrWaC-u se pojavljuju i kao stupnjevani pridjevi, iako su negirani prefiksom. Odnosno, osim pozitiva pojavljuje se i komparativ u primjerima poput *neizravniji, dekoncentriraniji, dezorganiziraniji* te superlativ u primjerima *najnesretniji, najnepoznatiji, najbezbolniji.*

7. Metodologija

Kao što je već navedeno u prethodnom poglavlju, svrha prikupljanja korpusa je njegova primjena u NooJ alatu, kako bi se napravila gramatika za prepoznavanje i označavanje prefiksalne negacije pridjeva. NooJ je razvojno okruženje koje se upotrebljava za pisanje formaliziranih opisa prirodnog jezika, kako bi ih se u realnom vremenu upotrijebilo na korpusu. Opisi prirodnog jezika koje pruža NooJ formalizirani su kao digitalni rječnici i gramatike predstavljene organiziranim setovima grafova. NooJ pruža alate za opis flektivne i derivacijske morfologije, varijacija u terminologiji i pismu, rječnika, sintakse i semantike. Također, upotrebljava se i kao alat za obradu korpusa, dopuštajući korisnicima obradu tisuća tekstualnih datoteka, kako bi primjerice mogli označiti morfo-sintaktičke obrasce, lematizirane konkordancije te na rezultatima provesti različite statističke izračune (Silberztein, 2003).

Stoga su, kako bi se računalno analizirala prefiksalna negacija pridjeva hrvatskoga jezika, osim prikupljanja korpusa potrebni i ostali zadaci. Na temelju prikupljenog korpusa napravit će se rječnik u NooJ-u, a svakom pridjevu u rječniku dodijelit će se paradigma po kojoj se on deklinira, stoga taj korak uključuje i izradu deklinacijskih obrazaca u obliku flektivnih gramatika. Potom će se izraditi morfološka gramatika, kako bi primjeri iz rječnika bili uspješno pronađeni i pravilno označeni. Odnosno, onima koji su tvoreni zadanim prefiksima potrebno je promijeniti polaritet – iz pozitivnog u negativni. Naposljetku, kako bi se izračunala točnost i sveobuhvatnost gramatike, izračunat će se njezina preciznost, odziv i F-mjera.

Iako ih korpus sadrži, u računalnu analizu (prepoznavanje i označavanje u NooJ-u) neće biti uključeni neki primjeri. U korpusu su se na višim razinama stupnjevanja (komparativ i superlativ) pojavili pridjevi koji su prefiksalno negirani, opovrgavajući teorijsku postavku da takvi pridjevi ne sudjeluju u stupnjevanju. Međutim, oni se dekliniraju po posebnim deklinacijskim obrascima te se stoga ne tvore samo dodavanjem negacijskoga prefiksa, već i dodatnom prefiksalno-sufiksalnom tvorbom (*bolan – bezbolan – bezbolniji – najbezbolniji*), zbog čega neće biti uključeni u gramatiku za prepoznavanje prefiksalne negacije pridjeva.

U računalno prepoznavanje i označavanje nisu uključeni ni pridjevi tvoreni negacijskim prefiksima koji bez njih ne postoje u hrvatskom jeziku, jer su oni izvan okvira definicije ovoga rada prema kojoj je određeno da „bez prefiksa nešto *jest*“. Točnije, pridjevi poput *kontraceptivni/kontracepcijski, antibiotski, deratizacijski, besposlen, besplatan, bezazlen* i sl. nisu uključeni jer ne postoji *ceptivni/cepcijski, biotski, ratizacijski, poslen, platan, zazlen*. Stoga

se pridjevi bez negacijskoga prefiksa ne mogu dodati u rječnik pa se, iako ih korpus sadrži, ne očekuje da ih gramatika pronađe. Nadalje, iako postoji pridjev *bèzub,* u rječnik se neće dodati njegova suprotna inačica bez prefiksa, jer se u tom slučaju mijenja vrsta riječi, odnosno *zûb* je u tom slučaju imenica, a ne pridjev. Isto tako, u korpusu postoje i pridjevi za koje se čini da su tvoreni negacijskim prefiksom, ali on je dio njihove osnove, poput *nestao, neophodan, neutralan, nebeski, antički* i sl. te stoga takvi pridjevi nisu uneseni u rječnik i ne želi ih se obuhvatiti gramatikom.

Kao što je također određeno na početku rada – računalno prepoznavanje i označavanje prefiksalne negacije pridjeva odnosit će se na pridjeve koji negacijskim prefiksom tvore izravnu negaciju, stoga primjeri koji tvore neizravnu negaciju neće biti uključeni u analizu. To su primjerice svi pridjevi tvoreni prefiksom *mal-* koji je iz tog razloga u analizi prefikasa naveden u zagradi.

Naposljetku, glagolski pridjevi trpni, iako su obuhvaćeni korpusom jer su u hrWaC-u označeni kao pridjevi, u NooJ-u se tretiraju kao glagolska kategorija. Oni su, kao što je već navedeno, tvoreni na drugačiji način, kompliciraniji za obuhvaćanje gramatikom na razini ovoga diplomskoga rada. Primjerice, *unesrećen* se ne može tretirati kao *une+srećen* jer *srećen* ne postoji i stoga se ne može dodati u rječnik u takvom obliku. Ovakvi primjeri nastaju od glagola, stoga se u NooJ-u tako i označavaju, a ako bi se tražio njihov ne-negirani pridjevski ekvivalent, u rječnik bi trebalo dodati *usrećen,* a gramatikom ubaciti prefiks *ne-* između *u-* i *-srećen.* Slično je i za *uznemiren,* gdje bi trebalo dodati *zne-,* zatim *onesviješten, onemogućen, onesposobljen, obezglavljen, obeščašćen* (gdje postoji *čašćen* ali u drugom značenju, a ovaj glagolski pridjev tvoren je od glagola *obeščastiti*). Ovaj rad pokušao je pružiti uvid i u takve, kompliciranije tvorbene slučajeve te su stoga takvi primjeri i spomenuti u dijelu analize prefiksalne negacije u hrWaC-u, međutim ovakve glagolske pridjeve trpne potrebno je posebno obraditi. Prefiks *od-* kojim se tvore isključivo glagolski pridjevi trpni također nije uključen.

S druge strane, glagolski pridjevi trpni koji na isti način kao obični pridjevi tvore izravnu negaciju dodavanjem prefiksa, poput *prepoznat-neprepoznat, taknut-netaknut, razvijen-nerazvijen, donesen-nedonesen,* uključeni su u analizu. Riječ je o tome da su obuhvaćeni prikupljanjem korpusa jer su u hrWaC-u označeni kao pridjevi, a za razliku od primjera spomenutih u prethodnom odlomku, njihov negirani ekvivalent tvori se isključivo dodavanjem negacijskoga prefiksa (a ne umetanjem u već postojeći prefiks). Osim toga, postoje kao samostalna riječ i bez prefiksa, stoga ih se može unijeti u rječnik te s prefiksom tvore izravnu

negaciju. Svi takvi primjeri tvoreni su prefiksom *ne-,* koji je ionako uključen u gramatiku zbog ostalih pridjeva, jer je najplodniji negacijski prefiks.

8. Računalno prepoznavanje i označavanje negacije hrvatskih pridjeva u NooJ-u

8.1. Rječnik

Za 322 pridjeva iz korpusa napisan je rječnik, a svakom pridjevu dodijeljene su odgovarajuće oznake i flektivni obrazac. S obzirom na to da Nooj za označavanje pridjeva sadrži kategorije *opisni, gradivni* i *posvojni,* dodana je kategorija *odnosni* te su pridjevi označeni prema dvjema kategorijama navedenima i u teorijskom dijelu rada – opisni i odnosni, jer posvojni i gradivni se također mogu opisati kategorijom odnosnih. Ova podjela upotrijebljena je jer većina pretražene literature upotrebljava upravo nju. Težak i Babić (1992) navode da se odnosni pridjevi tvore sufiksima *-ov, -ev, -in, -ski, -ni, -ji, -nji, -ći,* a opisni sufiksima *-(a)n, -en, -ljiv, -ast, -av, -it, -at* i *-ak.*

Potom je svaki pridjev označen i kao *određeni* ili *neodređeni.* Barić i sur. (2005) navode da neodređeni pridjevi odgovaraju na pitanje *kakav,* a određeni na pitanje *koji.* S druge strane, sami autori navode da ni ova podjela nije uvijek jasno razgraničena, jer se vrlo često može upotrijebiti i jedan i drugi vid. Barić i sur. ističu da samo opisni pridjevi mogu imati i određeni i neodređeni vid, a posvojni (koji su u NooJ rječniku napravljenom za ovaj rad označeni kao odnosni), mogu biti samo ili neodređeni ili određeni. Stoga autori navode da su posvojni pridjevi koji imaju neodređeni vid tvoreni sufiksima *-ov, -ev-, -ljev* i *-in,* a određeni sufiksima *-ji, -nji, -šnji, -ski/-ki* (Barić i sur., 2005).

Svakom pridjevu dodijeljena je i oznaka polariteta, što je semantička kategorija koja ne postoji u NooJ-u stoga je za ovu svrhu posebno dodana (Slika 41). Prije dodjeljivanja ove oznake, bilo je potrebno utvrditi što će se smatrati pozitivnim, a što negativnim polaritetom. Naime, oznaka polariteta odnosi se na značenje riječi, koje uvelike ovisi o kontekstu u kojem se riječ nalazi. Primjerice, za pridjev *tenkovski,* koji se pojavljuje s negacijskim prefiksom *protu-,* ne može se jasno odrediti je li polaritet pozitivan ili negativan, zbog činjenice da se ne zna kontekst. Također, problematične su i riječi kojima se bez obzira na poznavanje konteksta ne može dodijeliti isključivo pozitivan ili negativan polaritet, odnosno može postojati i neutralna osnova. Riječi koja bi se u rječniku označila kao neutralna nije moguće dodavanjem prefiksa odrediti polaritet isključivo kao negativan. Međutim, potreba poznavanja konteksta zahtijevala bi analizu negacije na široj razini nego je određena u ovome radu, odnosno tada bi analiza negacije izlazila van postavljenih okvira – okvira riječi. Stoga je određeno da će se dodjeljivanje oznake polariteta u NooJ-u zadržati na razini same riječi, prema čemu „bez

prefiksa nešto *jest,* dodavanjem prefiksa više *nije* ili dobiva suprotno značenje“ , što odgovara definiciji određenoj u ovome radu. Prema tome, svaka osnova na koju se dodaje negacijski prefiks smatrana je pozitivnom osnovom, a gramatika svakom pridjevu koji ima takvu oznaku treba nakon dodavanja negacijskog prefiksa promijeniti polaritet u negativan.

Na slici 41 prikazane su NooJ oznake za opis pridjeva, uz napomenu da se oznake za stupanj i kategoriju nisu upotrebljavale, a rod, broj i padež označeni su flektivnim gramatikama.

```
# Pridjevi - Adjectives: padež | rod | broj | komparacija (np|cp|sp)
A_Case = Nom | G | D | Acc | Voc | L | I;
A_Gender = m | f | n;                    #muški | ženski | srednji rod
A_Nb = s | p;                            # jednina | množina
A_Degree = np | cp | sp;                    #pozitiv | komparativ | superlativ
A_Type = op | gr | po | odn;                   #opisni | gradivni | posvojni | odnosni
A_Form = odr | nodr;                     #određeni | neodređeni
A_Category = boja | vrijeme | jez | mjera | obt;
A_Pol = pos | neg;                          #polaritet - pozitivan i negativan
```

Slika 41 Prikaz oznaka za opis pridjeva u NooJ-u

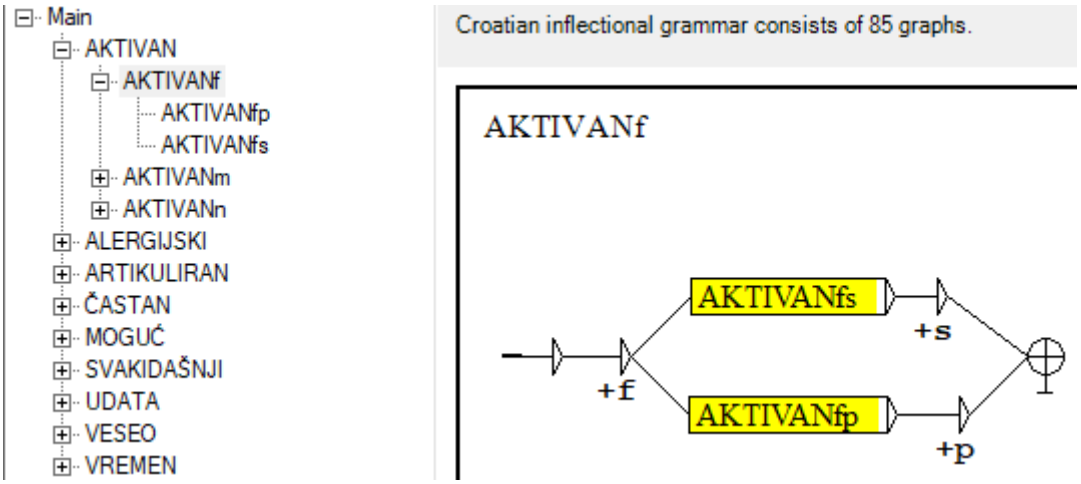
Naposljetku, svakom pridjevu u rječniku dodijeljen je i flektivni obrazac. Prikaz rječničkog unosa s dodijeljenim oznakama prikazuje Slika 42.

```
#use paradigme.nof
adekvatan,A+op+nodr+Pol=pos+FLX=AKTIVAN
aktivan,A+op+nodr+Pol=pos+FLX=AKTIVAN
alergijski,A+odn+odr+Pol=pos+FLX=ALERGIJSKI
artikuliran,A+op+nodr+Pol=pos+FLX=ARTIKULIRAN
```

Slika 42 Primjer rječničkog unosa s pridjevskim oznakama

Na slici 42 vidi se i da je naredbom '#use paradigme.nof' pozvana flektivna gramatika koja sadrži deklinacijski obrazac svih pridjeva iz rječnika. 322 pridjeva iz rječnika može se svrstati u 9 deklinacijskih obrazaca, stoga je napisano 50 flektivnih gramatika – za svaki pridjev muški, ženski i srednji rod te jednina i množina (osim za pridjev *udata* koji ima samo ženski rod jednine i množine).

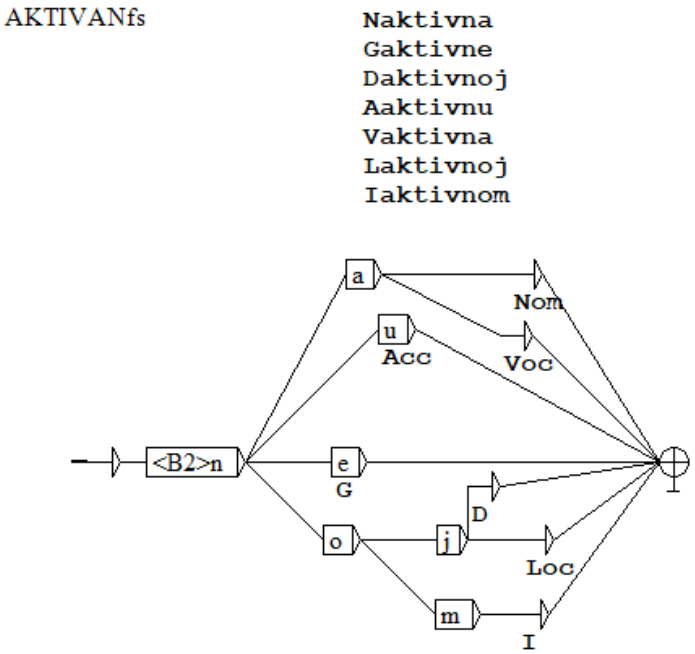
Primjer strukture flektivnih gramatika prikazuje Slika 43.



Slika 43 Primjer strukture flektivnih gramatika

Flektivni obrasci provjeravani su i prema *Hrvatskom jezičnom portalu,*[46] međutim s obzirom na to da on ne sadrži sve oblike koji se danas upotrebljavaju, neki oblici dodani su zasebno. Primjerice, za pridjeve koji se dekliniraju po flektivnom obrascu kao pridjev *aktivan, Hrvatski jezični portal*[47] za genitiv muškog roda jednine navodi samo: *koga? – aktivna (čovjeka).* Međutim, u govoru, kao i u korpusu prikupljenom iz korpusa hrWaC koji sadrži tekstove s internetskih portala i foruma, češće bi genitiv jednine muškoga roda glasio *aktivnog/aktivnoga,* stoga su takvi obrasci također dodani u gramatike. Slika 44 prikazuje primjer jednog deklinacijskog obrasca.

[46] http://hjp.znanje.hr/
[47] http://hjp.znanje.hr/



Slika 44 Primjer deklinacijskog obrasca za ženski rod jednine pridjeva 'aktivan'

Pisanjem rječnika primijećeno je da se 15 pridjeva pojavljuje uz više prefikasa, što prikazuje Tablica 3.

Tablica 3 Pridjevi iz korpusa koji se pojavljuju uz različite prefikse

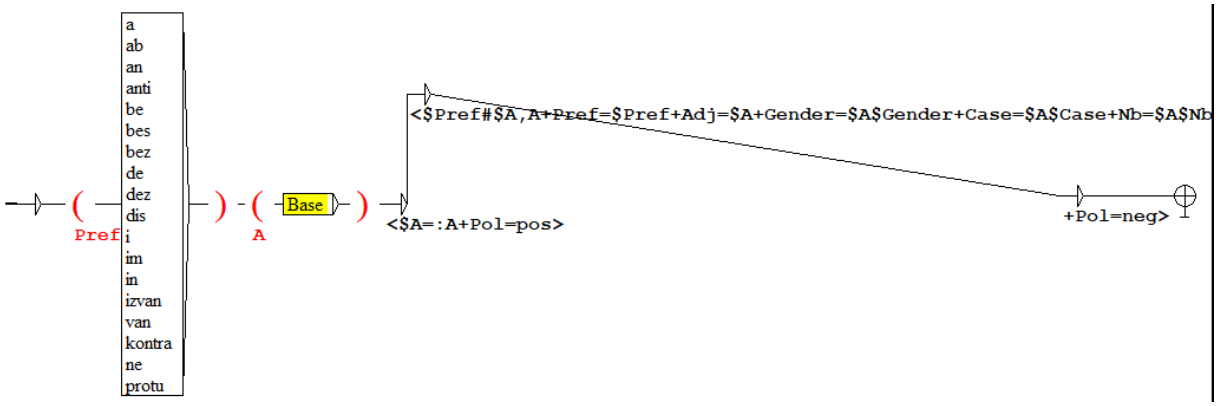
ANTI-/PROTU-	fašistički
	gljivični
	hrvatski
	infektivan
NE-/PROTU-	kršćanski
I-/NE-	legalan
	realan
	moralan
A-/NE-	logičan
	moralan
AN-/NE-	organski
A-/AB-	normalan
DEZ-/NE-	organiziran
IN-/NE-	tolerantan
BEZ-/NE-	uspješan
BEZ-/PROTU-	životan

Kod nekih pridjeva zamjenom prefiksa dolazi do promjene značenja, poput *beživotan/protuživotan,* dok kod nekih značenje ostaje isto, ali sadrže prefiks i stranog i domaćeg podrijetla, poput *antigljivični/protugljivični.* Ova pojava govori u prilog činjenici spomenutoj u poglavlju o prikupljanju korpusa – da se ovakve tvorbe ne mogu uvijek svrstati pod pravilo i ne slijede uvijek isti obrazac. Odnosno, riječi i stranog i domaćeg podrijetla mogu se pronaći i sa stranim i s domaćim prefiksima.

8.2. Gramatika

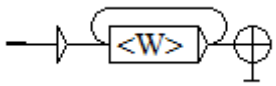
Cilj je napraviti gramatiku u NooJ-u koja bi se smatrala uspješnom kada bi prepoznala da je pridjev tvoren negacijskim prefiksom te bi mu shodno tome dodijelila oznake pridjevske osnove od koje je nastao te bi mu promijenila polaritet iz pozitivnog u negativni.

Prvu morfološku gramatiku za prepoznavanje prefikasa i pridjevskih osnova uz promjenu polariteta prikazuje Slika 45.



Slika 45 Prva morfološka gramatika za prepoznavanje prefiksalne negacije pridjeva

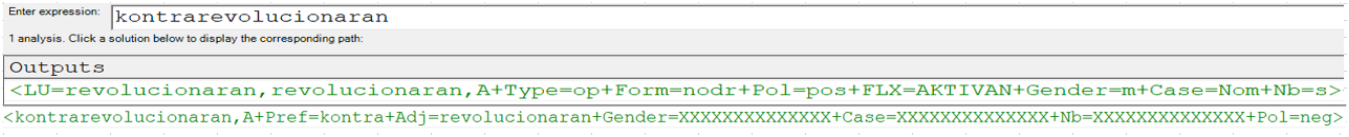
Svi negacijski prefiksi koje bi gramatika trebala prepoznati dodani su pod jednu varijablu (varijabla *Pref* na slici 45) te se izravno spajaju na čvor *Base,* koji se odnosi na ugrađeni graf (Slika 46) za prepoznavanje osnova na koje se dodaje prefiks. S obzirom na to da oznaka <W> predstavlja samo jedan znak, ovaj graf stavljen je u rekurziju kako bi se prepoznao niz znakova. Međutim, dodan je uvjet u obliku output oznake postavljene izvan te varijable (Slika 45) prema kojem gramatika treba prepoznati samo riječi koje su označene kao pridjev 'A' i koje imaju pozitivan polaritet.



Slika 46 Čvor 'Base' u morfološkoj gramatici koji 'poziva' osnove

Nakon spajanja prefiksa s osnovom označenom kao pridjev s pozitivnim polaritetom, novonastala riječ trebala bi poprimiti i ostale oznake toga pridjeva – rod, broj i padež te je na kraju dodana posljednja output oznaka gramatike (+Pol=neg) kojom se novonastaloj riječi mijenja oznaka polariteta.

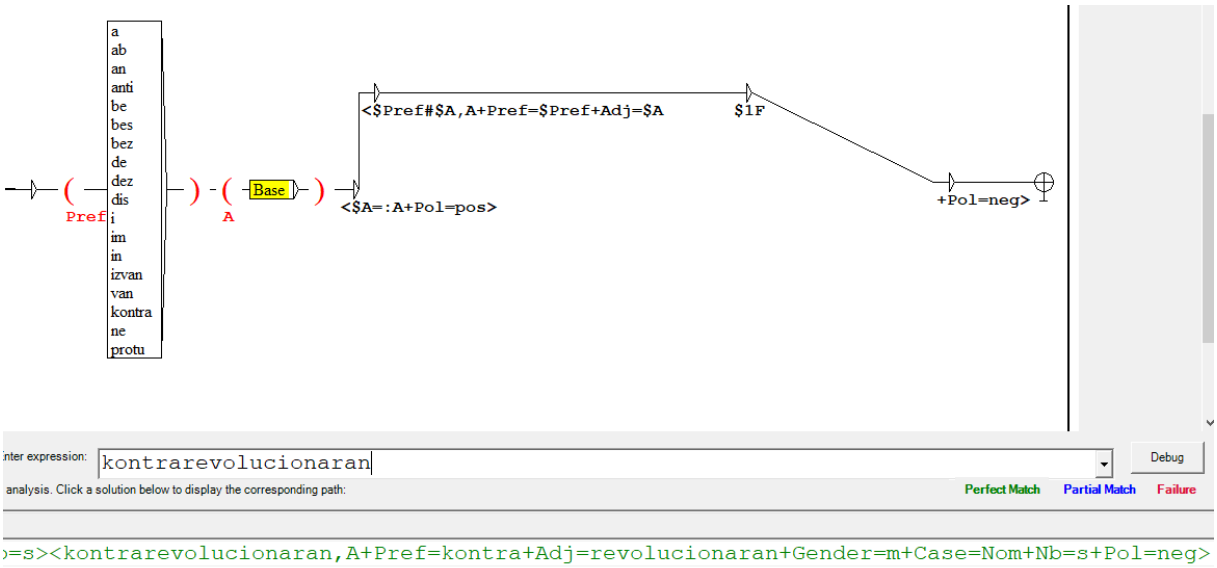
Međutim, na primjeru pridjeva *kontrarevolucionaran* može se vidjeti da ovakva gramatika nije dovoljno učinkovita (Slika 47).



Slika 47 Prikaz nedovoljno učinkovitog rezultata s prvom morfološkom gramatikom

Naime, ova gramatika prepoznala je pridjev *revolucionaran* i njemu dodijeljene oznake – da se radi o opisnom, neodređenom pridjevu kojemu je dodijeljen pozitivni polaritet i koji odgovara flektivnom obrascu pridjeva 'aktivan'. Za pridjev *revolucionaran* stoga je prema tom flektivnom obrascu prepoznato da se radi o pridjevu muškoga roda u nominativu jednine. Također je uspješno prepoznato da je *kontra* prefiks, a *revolucionaran* osnova, međutim prilikom spajanja toga prefiksa i osnove, gube se oznake za rod, broj i padež, odnosno ne dodjeljuju se iste oznake koje sadrži i pridjev *revolucionaran.* Naposljetku, polaritet je za pridjev koji sadrži negacijski prefiks uspješno promijenjen iz pozitivnoga u negativni.

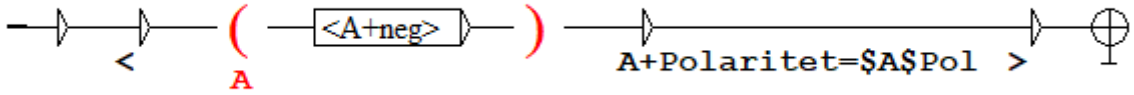
Stoga je gramatika ispravljena kako bi rezultat bio učinkovitiji, odnosno kako bi se oznake roda, broja i padeža dodijelile i prefiksalno negiranom pridjevu (Slika 48).



Slika 48 Ispravljena morfološka gramatika

Kako bi se oznake roda, broja i padeža ispravno pokazale, bilo je potrebno upotrijebiti varijablu $1F kojom se flektivne karakteristike pridjeva-osnove, dodjeljuju i novoj riječi nastaloj negacijskim prefiksom. S ovakvim oznakama moguće je napraviti podrobniju analizu prefiksalno negiranih pridjeva i pokušati utvrditi obrasce – primjerice koji prefiksi se pojavljuju s kakvim pridjevima, pojavljuje li se neki prefiks isključivo uz pridjeve muškoga roda i slično. Ovaj rad neće ulaziti u takvu analizu, ali pruža model za moguća buduća istraživanja.

Naposljetku, napravljena je sintaktička gramatika za primjenu na korpusu sa svrhom označavanja prefiksalno negiranih pridjeva negativnim polaritetom (Slika 49).



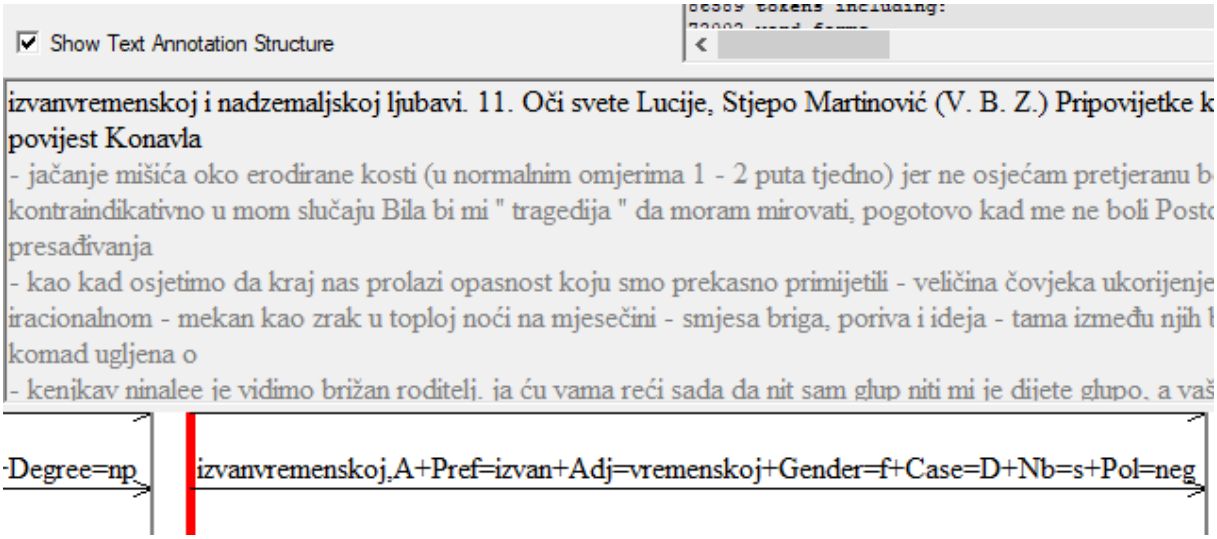
Slika 49 Sintaktička gramatika u NooJ-u za primjenu na korpusu

Ta gramatika primijenjena na korpus proizvodi označene konkordancije, prikazane na slici 50.

grandiozna ' Gojira ', revolucionarna ' Nećemo promjene ',	kontrarevolucionarna/<A+Polaritet=neg>	' Nisam vratio oružje u MUP
mafijaškim obračunima, ali i istinskoj,	izvanvremenskoj/<A+Polaritet=neg>	i nadzemaljskoj ljubavi. 11. Oči svete
pretjeranu bol ili je ono	kontraindikativno/<A+Polaritet=neg>	u mom slučaju Bila bi
veličina čovjeka ukorijenjena je u	iracionalnom/<A+Polaritet=neg>	- mekan kao zrak u toploj
Iran ili Pakistan vlast bude	destabilizirana/<A+Polaritet=neg>	i ne bude u mogucnosti
Petak, - ne kupujte knjigu od	nelegalnih/<A+Polaritet=neg>	distributera Uglavnom, Jergovićev ' Otac ' doživio
je ' šušur ' kakav je poprilično	kontraindiciran/<A+Polaritet=neg>	za jednu ozbiljnu i vrlo
hrtova na njegovoj farmi nađeno	dehidrirano/<A+Polaritet=neg>	i pothranjeno. Toplotni udar je

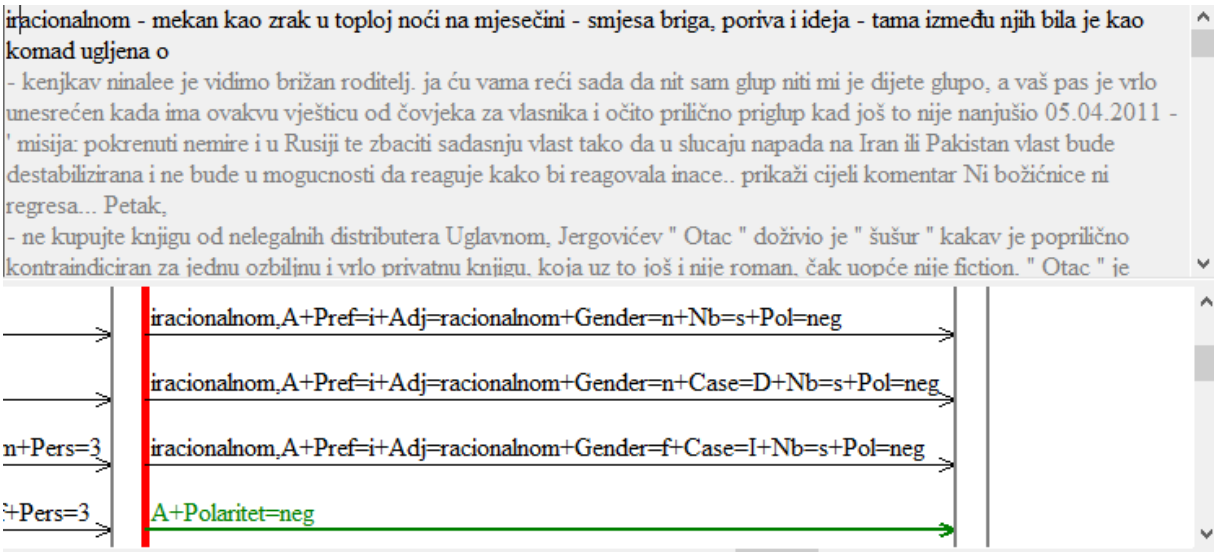
Slika 50 Označene konkordancije iz korpusa s primijenjenom sintaktičkom gramatikom

Također, uspješno je označen i tekst sa oznakama svega dodijeljenog morfološkom gramatikom – od kojeg prefiksa je nastala nova riječ, od koje osnove te koji je rod, broj, padež i polaritet (Slika 51).



Slika 51 Prikaz tekstualnih oznaka s kategorijama dodijeljenim morfološkom gramatikom

Odnosno, zelenom bojom uspješno se pokazuje oznaka polariteta dodijeljena sintaktičkom gramatikom (Slika 52).



Slika 52 Oznaka polariteta dobivena sintaktičkom gramatikom

8.3. Izračun preciznosti, odziva i F-mjere

Na manjem provjerenom uzorku čini se da su morfološka i sintaktička gramatika učinkovite, a za bolju provjeru učinkovitosti rezultati se provjeravaju na svim konkordancijama korpusa. Također, u izvore u NooJ-u dodani su i unaprijed napisani rječnici hrvatskoga jezika s ostalim vrstama riječi – imenicama, glagolima, zamjenicama i ostalim pridjevima, da bi se vidjelo kako ovaj model algoritma funkcionira izvan unaprijed pripremljenog i označenog rječnika. Točnije, obuhvaća li i pogrešne rezultate – nešto što nije pridjev, nešto čemu nije dodijeljena oznaka pozitivnog polariteta i sl.

Slika 53 prikazuje da je u 2024 konkordancije gramatika pronašla 2216 pridjeva kojima je nakon dodavanja negacijskog prefiksa polaritet promijenjen iz pozitivnoga u negativni.

Before	Seq.	After
istoka poput: blud, zguljen i	abnormalan/<+Pol=neg>	možda je najbitnije reći da
tom kriminalu mogu drogu postane	abnormalna/<+Pol=neg>	. Pošto je dopamin uključen u
i zajedničkom AV kanalu zbog	abnormalne/<+Pol=neg>	morfogeneze provodne srčane osovine. U
Papa test kojim se otkrivaju	abnormalne/<+Pol=neg>	promjene na stanicama koje bi
ranoga spontanog pobačaja. Najčešći uzrok	abnormalnih/<+Pol=neg>	jakih krvarenja jesu miomi i
u perimenopauzi se događa 50 % svih	abnormalnih/<+Pol=neg>	krvarenja, dok je ostalih 30 % rezervirano
ranoga spontanog pobačaja. Najčešći uzrok	abnormalnih/<+Pol=neg>	jakih krvarenja jesu miomi i
u perimenopauzi se događa 50 % svih	abnormalnih/<+Pol=neg>	krvarenja, dok je ostalih 30 % rezervirano
duljine 20 cm, na 6,5 cm od	aboralnog/<+Pol=neg>	ruba resekcije u lumenu debelog
znala biti uspavljujuća i relativno	adinamična/<+Pol=neg>	, dok je jazz, pop i
nižim od donje referentne vrijednosti.	Adinamična/<+Pol=neg>	koštana bolest može nastati ako
sloj, glasnice postaju krute i	adinamične/<+Pol=neg>	(ne vibrantne). Ovo rezultira asimetričnom
proljevom. Pacijenti su izuzetno dehidrirani,	adinamični/<+Pol=neg>	i pred kolapsom. Radna dijagnoza
Mario ' Studenti dolaze dosta dehidrirani,	adinamični/<+Pol=neg>	, hipotenzivni i pred kolapsom, ali

2216/2216

Slika 53 Broj dobivenih rezultata i oznaka polariteta uz prefiksalno negirane pridjeve

Točnost i sveobuhvatnost gramatike provjeravaju se izračunom preciznosti *(*engl. *precision),* odziva *(*engl. *recall)* i F-mjere.

Preciznost se izražava omjerom gramatikom pronađenih točnih rezultata i pronađenih rezultata ukupno. Odnosno:

$$P = \frac{\text{pronađeni sustavom točni}}{\text{pronađeni sustavom ukupno}}$$

Točnije, preciznost ukazuje na potencijalne netočno označene (engl. *false positive*) rezultate i upućuje na točnost sustava, dajući informaciju o tome koliko je alat s određenom gramatikom pronašao točnih, a koliko pogrešnih rezultata.

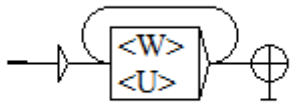
U 2216 rezultata dobivenih NooJ gramatikom, njih 103 su pogrešni rezultati, a 2113 je ispravno označeno, stoga je preciznost ove gramatike 95.3%. Sve pogreške odnose se na pogrešno označavanje priloga kao pridjeva zbog homografije. Primijećeno je da prilozi dijele oblik s pridjevima srednjeg roda u nominativu jednine. Primjerice, *nepokorivo, nezaobilazno, neobično, asocijalno* itd. može odgovarati na pitanje 'kakvo?', ali i 'kako?'. Stoga pridjeve nije bilo moguće analizirati same za sebe, već je bio potreban i kontekst, odnosno njihova ko-tekstna okolina. Ovaj problem mogao bi se razriješiti dodjeljivanjem dodatnih morfosintaktičkih oznaka. Primijećeno je da se najčešće ispred ili iza priloga nalazi glagol, primjerice *djeluje protuupalno, pate bezvučno, neprekidno raditi* i sl. te su vrlo često i modifikatori uz pridjev *disproporcionalno veći, bezrezervno prihvaćen* itd. S druge strane, iza pridjeva najčešće slijedi imenica ili glagolska imenica, primjerice *anormalan razvoj, neizlječiv poremećaj, izvanbračna zajednica, protuupalno djelovanje, neljudsko ponašanje* itd.

Odziv se izražava omjerom pronađenih točnih rezultata i svih točnih rezultata u korpusu, a upućuje na sveobuhvatnost sustava. Odnosno:

$$R = \frac{\text{pronađeni sustavom točni}}{\text{svi točni u tekstu}}$$

Prema tome, dok preciznost upućuje na netočno označene, odziv upućuje na neoznačene, uspoređujući točno pronađene sa stvarnim brojem rezultata koje je trebalo pronaći. Od 2113 točnih rezultata, gramatika je prvotno pronašla 2111 rezultata. Ustanovilo se da su propuštena dva pridjeva – *negoriv* i *nezapaljiv,* jer su u korpusu bili napisani velikim slovima (*'NEGORIV', 'NEZAPALJIV'*), a gramatika je, kako je prikazano u prethodnom poglavlju,

prvotno u *'Base'* čvoru sadržavala samo oznaku <W>. Ova oznaka odnosi se na znakove napisane malim slovom. Stoga je u taj čvor dodana i oznaka <U> koja se odnosi na znakove pisane velikim slovom (Slika 54), nakon čega odziv ove gramatike iznosi 100%, odnosno gramatika nije izostavila nijedan prefiksalno negirani pridjev koji je trebala pronaći.



Slika 54 Prepravljeni čvor 'Base' koji obuhvaća i niz znakova pisan velikim slovima

Naposljetku, prema podacima o preciznosti i odzivu, računa se F-mjera koja kombinira ova dva parametra. F-mjera sadrži i parametar β koji odražava relativnu važnost, odnosno težinsku vrijednost između preciznosti i odziva, prema čemu formula za izračun F-mjere glasi:

$$F - mjera = \frac{(\beta^2 + 1) * P * R}{\beta^2 * (P + R)}$$

S obzirom na to da preciznost i odziv imaju jednaku težinsku vrijednost, odnosno jednako su relevantni, β=1, prema čemu F-mjera iznosi:

$$F - mjera = \frac{2 * (P * R)}{P + R} = \frac{2 * (0.95 * 1)}{0.95 + 1} = \frac{1.9}{1.95} = 0.97$$

9. Zaključak

Negacija je fenomen koji se tiče logičke, kognitivne, jezične, ali i računalne domene, a ovaj rad pokušao ju je prikazati iz jezične i računalne perspektive. Međutim, negacija je kao jezični fenomen u literaturi koja se tiče hrvatskoga jezika vrlo nesustavno obrađena, stoga se u prvom dijelu rada pokušalo doprinijeti proučavanju negacije na teorijskoj razini, obuhvaćanjem hrvatskih gramatika, ali i ostalih relevantnih jezičnih izvora. S obzirom na to da se radovi koji spominju negaciju ili se bave njom, dotiču uglavnom sintaktičke i rečenične razine, ovaj rad pokušava doprinijeti proučavanju negacije na morfološkoj razini. Odnosno kako ju neki autori nazivaju – leksička razina ili razina riječi. Iako rad donosi pregled literature koja se tiče negacije i negacijskih prefikasa za više vrsta riječi (spominju se imenice, glagoli, pridjevi i zamjenice), naglasak je specifično na pridjevima. Proučavali su se pridjevi koji dodavanjem negacijskoga prefiksa tvore suprotni pojam – antonim, točnije riječ je o morfološki povezanim, odnosno istokorijenskim pridjevima.

U teorijskom dijelu rada upotrebljavali su se i izvori koji se tiču engleskoga jezika te su se gdje je bilo moguće usporedili hrvatski i engleski jezik. Sa svrhom provjere određenih teorijskih postavki i usporedbe hrvatskog i engleskog jezika te kako bi se utvrdilo odnose li se zaključci tih istraživanja i na hrvatski jezik, prilikom prikupljanja korpusa za računalnu obradu provedena je analiza prefiksalno negiranih pridjeva iz hrvatskog mrežnog korpusa hrWaC (Ljubešić i Erjavec, 2011).

Rad se s tom svrhom osvrće na istraživanja koja se bave upravo negacijom na morfološkoj razini u engleskom jeziku te se prema nekim navedenim teorijskim postavkama smatra da riječi koje imaju svoj ne-negirani ekvivalent, nemaju svoju negiranu inačicu. Primjerice, u engleskom jeziku postoje pridjevski parnjaci *dead/alive* te stoga ne postoji *unalive.* S druge strane, u hrvatskom jeziku postoje i pridjevski antonimi koji čine ne-negirani par, poput *živ/mrtav,* ali postoji i negirana inačica pridjeva *živ,* odnosno postoji i pridjev *neživ.*

Također, 'formula' za tvorbu negacije nalaže da se negacijski prefiks dodaje samo na pozitivnu ili neutralnu osnovu, stoga u engleskom jeziku postoji *unhappy* (nesretan), ali ne i *unsad* (netužan). Analiza pridjeva iz hrWaC-a pokazala je da u hrvatskom jeziku postoje upravo takvi pridjevi koji se smatraju semantički negativnom osnovom, a na njih je dodan i negacijski prefiks, poput upravo navedenog primjera *netužan.*

Nadalje, istraživanja negacije pridjeva engleskoga jezika (Kjellmer, 2005), ali i neka koja spominju negaciju pridjeva u hrvatskom jeziku (Kovačević, 2016), spominju da pridjevi koji sudjeluju u stupnjevanju nisu negirani na višim raznima stupnjevanja. Međutim, analizom hrvatskih pridjeva u hrWaC-u pronađeni su prefiksalno negirani pridjevi na svim razinama stupnjevanja, odnosno i komparativi i superlativi. Primjerice, *neizravniji, bezbolniji, dekoncentriraniji,* ali i *najbezbolniji, najneobičniji, najnepoznatiji* itd.

Drugi dio rada odnosio se na analizu prefiksalne negacije pridjeva u programu za računalnu obradu prirodnoga jezika – NooJ (Silberztein, 2003). Korpus je prikupljen pomoću alata Sketch Engine[48], a sadrži 2024 konkordancije prikupljene s 322 pridjeva. Pridjevi za analizu prefiksalno su negirani prefiksima *a-, ab-, an-, anti-, bez- (be, bes-), de-, dez-, dis-, i-, im-, in-, izvan- (van-), kontra-, ne-, protu-,* a uvjet je da se novonastali pridjev odnosi na izravnu negaciju. Odnosno, određenu prema definiciji „bez prefiksa nešto *jest,* dodavanjem prefiksa više *nije* ili dobiva suprotno značenje." Iz tog razloga, iz analize su isključeni prefiksi koji tvore neizravnu negaciju, poput *mal-* (*malformiran*), koji samo ima negativno značenje, ali ne negira osnovu. Također, isključeni su pridjevi koji u hrvatskom jeziku ne postoje kao samostalna osnova bez toga prefiksa, poput *kontraceptivni,* jer ne postoji *ceptivni,* zatim *antibiotski,* jer ne postoji *biotski* i sl. Prikupljanjem korpusa utvrđeno je i da je daleko najplodniji prefiks za tvorbu prefiksalne negacije, prefiks *ne-,* što navode i ostali autori.

U NooJ-u je napisan rječnik za svaki pridjev te su mu se dodijelile oznake *opisni* ili *odnosni* te *određeni* ili *neodređeni.* Svaki pridjev bez negacijskoga prefiksa tretiran je kao pozitivna osnova, jer pozitivni ili negativni polaritet vrlo često ovisi o kontekstu, a željelo se zadržati na razini same riječi. Stoga su u rječniku navedeni pridjevi bez negacijskoga prefiksa te im je dodijeljena oznaka pozitivnoga polariteta. Zatim je pridjevima dodijeljena i paradigma po kojoj se deklliniraju, a za tu svrhu napisane su flektivne gramatike prema devet deklinacijskih obrazaca.

Korpus je prikupljen i rječnik napisan sa svrhom da se u NooJ-u napiše morfološka gramatika, koja bi prepoznala svaki prefiks i svaku osnovu, spojila ih u novu riječ te novonastaloj riječi dodijelila sve morfološke karakteristike pridjeva-osnove, poput roda, broja i padeža. Naposljetku, gramatika je svakom prefiksalno negiranom pridjevu trebala promijeniti

[48] https://www.sketchengine.eu/

oznaku polariteta iz pozitivnoga u negativni. Zatim je napisana sintaktička gramatika koja se može primijeniti na bilo koji korpus te je u 2024 konkordancije generirala 2216 rezultata.

Naposljetku, provjerena je učinkovitost gramatike, točnije njezina točnost i sveobuhvatnost, koji se računaju parametrima preciznosti i odziva. Analizom rezultata utvrđeno je da je gramatika točno prepoznala 2113 rezultata, odnosno 103 su pogrešno označena. Sve pogreške tiču se homografije, jer pridjev srednjeg roda u nominativu jednine istog je oblika kao prilog. Stoga je preciznost ove gramatike 95.3%. Zatim se računala sveobuhvatnost (odziv), odnosno je li gramatika označila sve što je trebala označiti te se pokazalo da je gramatika preskočila dva pridjeva – *nezapaljiv* i *negoriv,* jer su u korpusu bila napisani velikim slovima. Nakon jednostavne preinake u gramatici, kojom se dodala opcija i za prepoznavanje velikih slova, odziv gramatike iznosi 100%.

Ovakva računalna obrada, s obzirom na to da je prva koja se bavi prefiksalnom negacijom pridjeva hrvatskoga jezika, pruža neka rješenja, ali i uvid u neke probleme te otvara prostor za buduća istraživanja. Ovim radom proširena je rječnička baza s pridjevima koji se mogu prefiksalno negirati, tvoreći izravnu negaciju, a s obzirom na to da već postoji i rad koji se bavi prefiksalnom negacijom imenica na morfološkoj razini (Žanpera i sur., 2020), u budućnosti bi se dodavanjem glagola mogao napraviti potpuni rječnik hrvatskih antonima.

Također, ovaj gramatički model koji već uključuje i ostale morfološke oznake, može poslužiti i utvrđivanju obrazaca u tvorbi ove vrste negacije. Primjerice, uključivši u analizu i oznake za vrstu pridjeva te rod, broj i padež, može se utvrditi pojavljuju li se neki prefiksi uz specifičnu vrstu pridjeva, uz pridjeve specifičnoga roda i slično. Osim toga, može se utvrditi pojavljuju li se neki pridjevi češće u svom prefiksalno negiranom ili osnovnom obliku.

Međutim, iako se ovaj rad bavi negacijom na morfološkoj razini te se pokušalo ostati unutar granica riječi, prilikom određivanja točnosti rezultata gramatike, nije se mogao zanemariti kontekst. Odnosno, upravo zbog homografije, trebala se uključiti i ko-tekstna okolina. Stoga je jedan od budućih zadataka pokušati razriješiti morfosintaktičku višeznačnost utvrđivanjem potencijalnih obrazaca za što trenutni model također može poslužiti. Ovaj model može poslužiti i u analizi sentimenta, za što bi se svakako trebalo odmaknuti od morfološke razine i uključiti ostale jezične razine u analizu.

Naposljetku, za buduću analizu ostaju i pridjevi koji bez negacijskoga prefiksa ne postoje kao samostalne riječi u hrvatskom jeziku. Za analizu ostaju i drugi fenomeni utvrđeni ovim radom – pridjevi koji su prefiksalno negirani iako im je osnova semantički negativna te

prefiksalno negirani pridjevi na višim razinama stupnjevanja. Također, kao što je već spomenuto u radu, za analizu ostaju i glagolski pridjevi trpni poput *uznemiren, onemogućen, unesrećen,* ali i sami glagoli.

Iz svega navedenog, može se potvrditi izneseno na samom početku rada – negacija je kompleksan, široko rasprostranjen fenomen koji seže i van granica samoga jezika. Važnost ovoga fenomena potvrđuje i nizozemski filozof Spinoza svojom poznatom izjavom *„Omnia determinatio est negatio“*, odnosno – ne možemo u potpunosti znati što nešto *jest,* bez da znamo što to *nije.*

Literatura

1. Babić, S., Finka, B., Moguš, M. (1971). *Hrvatski pravopis.* Zagreb: Školska knjiga.
2. Babić. S. (1991). Tvorba riječi u hrvatskom književnom jeziku – Nacrt za gramatiku, drugo izdanje. Zagreb: HAZU, Nakladni zavod Globus.
3. Babić, S. (2002). Tvorba riječi u hrvatskome književnome jeziku – Velika hrvatska gramatika, treće, poboljšano izdanje. Zagreb: HAZU, Nakladni zavod Globus.
4. Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V., Znika, M. (2005). *Hrvatska gramatika.* Zagreb: Školska knjiga.
5. Chapman, W. W., Bridewell, W., Hanbury, P, Cooper, G., F., Buchanan, B. G. (2001a). Evaluation of Negation Phrases in Narrative Clinical Reports. *Proceedings of the AMIA Symposium,* 105-109.
6. Chapman, W. W., Bridewell, W., Hanbury, P., Cooper, G. F., Buchanan, B. G. (2001b). A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries. *Journal of Biomedical Informatics,* 34, 301-310.
7. Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics.* Oxford: Blackwell Publishing.
8. Csapo, J. (1977). Adjectives With Negative Affixes in Modern English and Their Hungarian Equivalents – a Contrastive Study. *Hungarian Studies in English,* 11, 197-210.
9. De Clercq, K., Vanden Wyngaerd, G. (2017). Why Affical Negation is Syntactic. *Proceedings of the 34th West Coast Conference on Formal Linguistics,* 151-158.
10. De Clercq, K., Vanden Wyngaerd, G. (2018). Adjectives and Negation: Deriving Contarriety from Contradiction. *Studies van de Belgische Kring voor Linguistiek,* 12, 1-18.
11. *Englesko-hrvatski i hrvatsko-engleski rječnik* (2004). Gl. ur. B. Petrač, E. Miščin. Zagreb: Alfa.
12. Goldin, I., Chapman, W., W. (2003). Learning to Detect Negation With 'Not' in Medical Texts. *Proc Workshop on Text Analysis and Search for Bioinformatics.*
13. Harkema, H., Dowling, J. N., Thornblade, T., Chapman, W. W. (2009). ConText: An Algorithm for Determining Negation, Experiencer, and Temporal Status From Clinical Reports. *Journal of Biomedical Informatics,* 42, 839-851.
14. Hogan, P. C. (2011). *The Cambridge Encyclopedia of the Language Sciences.* Cambridge: Cambridge University Press.

15. Hoosain, R. (1973). The Processing of Negation. *Journal of Verbal Learning and Verbal Behavior,* 12, 618-626.
16. Horn, L., Kato, Y. (2000). Introduction: Negation and Polarity at the Millennium. U L. Horn i Y. Kato (ur.), *Negation and Polarity – Syntactic and Semantic Perspectives,* 1-19. Oxford: Oxford University Press.
17. Horn, L. R. (2001). *A Natural History of Negation.* Chicago: CSLI Publications.
18. Horn, L.R. (2005). An *Un-* Paper for the Unsyntactician. U S. S. Mufwene, E. J. Francis, R. S. Wheeler (ur.), *Polymorphous Linguistics – Jim McCawley's Legacy*, 329-365. Cambridge: MIT Press.
19. Horvat, M., Štebih Golub, B. (2012). Tvorba riječi u hrvatskim dopreporodnim gramatikama. *Rasprave instituta za hrvatski jezik i jezikoslovlje*, 38,2, 295-326.
20. Hrvatski jezični portal. URL: http://hjp.znanje.hr/ (pristupljeno 9. rujna 2020).
21. Huang, Y., Lowe, H., J. (2007). A Novel Hybrid Approach to Automated Negation Detection in Clinical Radiology Reports. *Journal of the American Medical Informatics Association,* 14,3, 304-311.
22. Israel, M. (2004). The Pragmatics of Polarity. U L. Horn, G. Ward (ur.), *The Handbook of Pragmatics,* 701-723. Oxford: Blackwell.
23. Jespersen, O. (1917). *Negation in English and Other Languages*. Kopenhagen: A. F. Host.
24. Jimenez–Zafra, S, M., Cruz Diaz, N., P., Morante, R., Martin–Valdivia, M., T. (2019). NEGES 2019 Task: Negation in Spanish. *Proceedings of the Iberian Languages Evaluation Forum.* 329-341.
25. Johnson, R. M. (2007). *A Logic Book.* Belmont: Thomas Wadsworth.
26. Joshi, S. (2012). Affixal Negation – Direct, Indirect and Their Subtypes. *Syntaxe et semantique.* 1, 49-63.
27. Katičić, R. (2002). *Sintaksa hrvatskoga književnog jezika.* Zagreb: HAZU, Nakladni zavod Globus.
28. Kilgarriff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovar, V., Michelfeit, J., Rychly, P., Suchomel, V. (2014). The Sketch Engine: Ten Years On. *Lexicography,* 1,1, 7-36.
29. Kjellmer, G. (2005). Negated Adjectives in Modern English. *Studia Neophilologica,* 77, 156-170.
30. Kovačević, A. (2016). Negacija od čestice do teksta: Usporedba i povijesna raščlamba negacije u hrvatskoglagoljskoj pismenosti. Zagreb: Staroslavenski institut.
31. Koza, W., Muñoz, M., Rivas, N., Godoy, N., Filippo, D., Cotik, V., Stricker, V., Martinez, R. (2018a). Automatic Detection of Negated Findings with NooJ: First Results. *Natural*

Language Procesing and Information Systems. NLBD 2018. Lecture Notes in Computer Science, vol. 10859, 298-302.

32. Koza, W., Filippo, D., Cotik, V., Stricker, V., Muñoz, M., Godoy, N., Rivas, N., Martinez-Gamboa, R. (2018b). Automatic Detection of Negated Findings in Radiological Reports for Spanish Language: Methodology Based on Lexicon-Grammatical Information Processing. *Journal of Digital Imaging,* 32, 19-29.
33. Kurtz, R., Oepen, S., Kuhlmann, M. (2020). End-to-End Negation Resolution as Graph Parsing. Proceedings of the 16th International Conference on Parsing Technologies and the IWPT 2020 Shared Task, 14-24.
34. Ljajić, A., Marovac, U., Avdić, A. (2017). Processing of Negation in Sentiment Analysis for the Serbian Language. *IcETRAN 2017 Conference Proceedings.*
35. Ljubešić. N., Erjavec, T. (2011). hrWaC and slWaC: Compiling Web Corpora for Croatian and Slovene*. Text, Speech and Dialogue. TSD 2011. Lecture Notes in Computer Science* (ur. I. Habrenal, V. Matoušek), vol. 6836, 395-402.
36. Mitchell, K. J., Becich, M. J., Berman, J. J., Chapman, W. W., Gilbertson, J., Gupta, D., Harrison, J., Legowski, E., Crowley, R. S. (2004). Implementation and Evaluation of a Negation Tagger in a Pipeline-based System for Information Extraction from Pathology Reports. *Medinfo,* 663-667.
37. Mowery, D. L., Velupillai S., Chapman, W. W. (2012). Medical diagnosis lost in translation – Analysis of uncertainty and negation expressions in English and Swedish clinical texts. *BioNLP: Proceedings of the 2012 Workshop on Biomedical Natural Language Processing,* 56-64.
38. NEGES Task (n.d.). *Negation in Spanish*. URL: http://www.sepln.org/workshops/neges/grupo.php?lang=en (pristupljeno 11. kolovoza 2020).
39. Paradis, C., Willners, C. (2006). Antonymy and Negation – The Boundedness Hypothesis. *Journal of Pragmatics,* 38, 1051-1080.
40. Sherman, M, A. (1973). Bound to be easier? The negative prefix and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* 12, 76-84.
41. Silberztein, M. (2003). *NooJ Manual.* URL: http://www.nooj-association.org/ (pristupljeno 10. kolovoza 2020).
42. Sketch Engine. URL: https://www.sketchengine.eu/ (pristupljeno 7. rujna 2020).
43. Subandowo, D. (2014). Negation Affixes in English. *Premise Journal,* 3,2, 135-144.

44. Šarić, Lj. (1992). Antonimija: Neke značenjske i tvorbene odrednice. *Rasprave ZHJ,* 18, 177-191.
45. Tanushi, H., Dalianis, H., Duneld, M., Kvist, M., Skeppstedt, M., Velupillai, S. (2013). Negation scope delimitation in clinical text using three approaches: NegEx, PyConTextNLP and SynNeg. *Proceedings of the 19th Nordic Conference of Computational Linguistics,* 387-397.
46. Težak, S., Babić, S. (1992). Gramatika hrvatskoga jezika: Priručnik za osnovno jezično obrazovanje. Zagreb: Školska knjiga.
47. Vujić, A. (2007). Opća i nacionalna enciklopedija u 20 knjiga (XIV. Mi-No). Zagreb: Pro leksis d.o.o.
48. Zipf, G. K. (1932). Selected Studies of the Principle of Relative Frequency in Language. Cambridge: Harvard University Press.
49. Zovko Dinković, I. (2013). Negacija u jeziku: Konstrativna analiza negacije u engleskome i hrvatskome jeziku. Zagreb: Hrvatska sveučilišna naknada.
50. Žanpera, N., Kocijan, K., Šojat, K. (2020). Negation of Croatian Nouns. Formalizing Natural Languages with NooJ 2019 and Its Natural Language Processing Applications. NooJ 2019. Communications in Computer and Information Science (ur. H. Fehri, S. Mesfar, M. Silbertzein), vol. 1153, 52-64.

Sažetak

Računalno prepoznavanje i označavanje negacije u hrvatskom

Ovaj rad bavi se prefiksalnom negacijom pridjeva hrvatskoga jezika, odnosno negacijom na morfološkoj razini. Definicija negacije unutar ovoga rada odnosi se na izravnu negaciju, prema kojoj „bez prefiksa nešto *jest,* dodavanjem prefiksa više *nije* ili dobiva suprotno značenje.“ Točnije, rad se bavi pridjevima koji dodavanjem negacijskoga prefiksa čine suprotni pojam, stoga se radi o morfološki povezanim ili istokorijenskim antonimima. Prvi dio rada donosi prikaz negacije i negacijskih prefikasa kroz hrvatske gramatike i ostale relevantne jezične izvore, a gdje je to moguće radi se i usporedba s engleskim jezikom. Također se prikazuju dosadašnja istraživanja ovog fenomena iz kognitivne i računalne domene. Prikupljanjem korpusa za računalnu obradu napravljena je i analiza prefiksalno negiranih pridjeva iz hrvatskog mrežnog korpusa „hrWaC“, pri čemu su opovrgnute neke dosadašnje teorijske postavke. Korpus sadrži 2024 konkordancije u kojima je bilo potrebno pronaći 2113 pridjeva. Drugi dio rada odnosi se na pisanje rječnika i gramatike u programu za računalnu obradu prirodnog jezika – NooJ, sa svrhom da gramatika prepozna kojim prefiksom i s kojom pridjevskom osnovom je stvorena nova riječ te novonastaloj riječi dodijeli morfološke oznake osnove i promijeni joj polaritet iz pozitivnoga u negativni. U rječnik su unesena 322 pridjeva – osnove, odnosno pridjevi bez negacijskoga prefiksa kojima je stoga dodijeljena oznaka pozitivnog polariteta te im je dodijeljena paradigma. Za tu svrhu napravljeno je 50 flektivnih gramatika s deklinacijskim obrascima. Zatim je napravljena morfološka gramatika za prepoznavanje i promjenu polariteta pridjeva i naposljetku sintaktička gramatika za primjenu na korpusu. Naposljetku se izračunala učinkovitost gramatike, prema parametrima preciznosti i odziva. Preciznost gramatike je 95.3%, ukazujući na probleme u označavanju zbog homografije, a odziv je 100%, odnosno nijedan prefiksalno negirani pridjev nije promaknuo morfološkoj gramatici u prepoznavanju i označavanju. Ova gramatika pruža model za daljnja istraživanja ovoga fenomena, poput nadograđivanja sa svrhom razrješenja morfosintaktičke višeznačnosti, ali i detaljnije analize sa svrhom prepoznavanja ostalih jezičnih obrazaca.

Ključne riječi: negacija, prefiksalna negacija, pridjevi, hrvatski jezik, obrada prirodnog jezika, NooJ

Abstract

Recognizing and Annotating Negation in Croatian

The paper studies the prefixal negation of Croatian adjectives, i.e., morphological negation. The paper narrows its scope to direct negation – 'before adding the prefix something *is,* after the prefix something *isn't,* or it gets the opposite meaning'. Specifically, this paper analyses morphologically related antonyms – adjectives that produce the opposite term by adding the negation prefix. The first part of the paper consists of a negation and negational prefixes overview in Croatian grammars and other relevant language sources. Where possible, a comparison between the Croatian and English language is made. Additionally, this paper overviews some of the cognitive and computational research about this phenomenon. While collecting the corpora for computational processing, a short analysis of the prefixal negation is made for adjectives from the Croatian web corpus 'hrWaC'. This analysis resulted in a rebuttal of some of the theoretical assumptions. The corpus consists of 2,024 concordances with 2,113 adjectives. The second part of the paper consists of writing a dictionary and a grammar in a program for natural language processing – NooJ. The purpose of the grammar is to recognize from which prefix and which base a new word is formed, assigning it the morphological annotations of the base and changing its polarity from positive to negative. The dictionary consists of 322 adjectives. These are the adjectives without the negation prefix that are assigned the positive polarity annotation and a paradigm. For this purpose, 50 inflectional grammars have been made, as well as the morphological grammar for recognizing and changing the adjective's polarity, and the syntactic grammar for applying on the corpus. Finally, the grammar's efficiency is calculated with precision and recall parameters. The precision of the grammar is 95.3% due to the annotation mistakes related to homography. The recall of this grammar is 100%, meaning the grammar has recognized and annotated all adjectives. This grammar provides a model for future research of this phenomenon, such as upgrading it with the purpose of morphosyntactic disambiguation, as well as further analysis to recognize other language patterns.

Keywords: negation, prefixal negation, adjectives, Croatian language, natural language processing, NooJ

Životopis

Natalija Žanpera rođena je 1993. godine u Zagrebu, gdje je završila Gornjogradsku gimnaziju. 2015. godine na Filozofskom fakultetu Sveučilišta u Zagrebu upisala je studij fonetike i lingvistike. Na diplomskom studiju fonetike 2018. godine upisala je smjer govorništvo, a na lingvistici računalni smjer. Akademske godine 2017/2018. bila je potpredsjednica Kluba studenata fonetike Eufonija. Dobitnica je Sveučilišne stipendije za izvrsnost za akademsku godinu 2016/2017., 2017/2018. te 2018/2019. Na Odsjeku za lingvistiku koautorica je članka *Negation of Croatian Nouns* koji je u lipnju 2019. godine predstavljen na konferenciji u Hammametu, a u siječnju 2020. objavljen u zborniku *Formalizing Natural Languages with NooJ 2019 and Its Natural Language Processing Applications* izdavača Springer. Na Odsjeku za fonetiku sudjelovala je u projektu *Mali i veliki čitači,* zbog čega je sudjelovala i u pisanju grupnog znanstvenog rada.

Prilozi

Popis slika

Slika 1 Učestalost pridjeva 'asimetričan' u hrWaC-u 41
Slika 2 Učestalost pridjeva 'nesimetričan' u hrWaC-u 41
Slika 3 Učestalost pridjeva 'nelogičan' u hrWaC-u 41
Slika 4 Učestalost pridjeva 'alogičan' u hrWaC-u 41
Slika 5 Učestalost pridjeva 'iregularan' u hrWaC-u 48
Slika 6 Učestalost pridjeva 'neregularan' u hrWaC-u 48
Slika 7 Učestalost pridjeva 'ilojalan' u hrWaC-u 49
Slika 8 Učestalost pridjeva 'nelojalan' u hrWaC-u 49
Slika 9 Učestalost pridjeva 'imoralan' u hrWaC-u 49
Slika 10 Učestalost pridjeva 'nemoralan' u hrWaC-u 49
Slika 11 Učestalost pridjeva 'imaterijalan' u hrWaC-u 49
Slika 12 Učestalost pridjeva 'nematerijalan' u hrWaC-u 49
Slika 13 Učestalost pridjeva 'imobilan' u hrWaC-u 50
Slika 14 Učestalost pridjeva 'nemobilan' u hrWaC-u 50
Slika 15 Učestalost pridjeva 'irealan' u hrWaC-u 50
Slika 16 Učestalost pridjeva 'nerealan' u hrWaC-u 50
Slika 17 Učestalost pridjeva 'iracionalan' u hrWaC-u 50
Slika 18 Učestalost pridjeva 'neracionalan' u hrWaC-u 51
Slika 19 Učestalost pridjeva 'legalan' u hrWaC-u 51
Slika 20 Učestalost pridjeva 'nelegalan' u hrWaC-u 51
Slika 21 Učestalost i kontekst pridjeva 'imperfektan' u hrWaC-u 51
Slika 22 Učestalost i kontekst pridjeva 'imperfektivan' u hrWaC-u 52
Slika 23 Učestalost pridjeva 'nesavršen' u hrWac-u 52
Slika 24 Učestalost pridjeva 'nesvršen' u hrWaC-u 52
Slika 25 Učestalost pridjeva 'impotentan' u hrWaC-u 53
Slika 26 Učestalost pridjeva 'nepotentan' u hrWaC-u 53
Slika 27 Pridjev 'intranzitivan' kao hapaks u hrWaC-u 54
Slika 28 Učestalost primjera 'intranzitivan' bez pridjevske oznake 54
Slika 29 Kontekst pridjeva 'neprijelazan' u hrWaC-u 55
Slika 30 Učestalost pridjeva 'netranzitivan' u hrWaC-u 55

Slika 31 Učestalost pridjeva 'inkompetentan' u hrWaC-u 55
Slika 32 Učestalost pridjeva 'nekompetentan' u hrWaC-u 55
Slika 33 Učestalost pridjeva 'intolerantan' u hrWaC-u 56
Slika 34 Učestalost pridjeva 'netolerantan' u hrWaC-u 56
Slika 35 Pridjev 'odglavljen' kao hapaks u hrWaC-u 62
Slika 36 Pridjev 'odvrnut' kao hapaks u hrWaC-u 62
Slika 37 Učestalost pridjeva 'odvrnut' bez pridjevske oznake 62
Slika 38 Učestalost pridjeva 'odglavljen' bez pridjevske oznake 63
Slika 39 Pridjev 'protublokovski' kao hapaks u hrWaC-u 63
Slika 40 Pridjev 'protupartijski' kao hapaks u hrWaC-u 64
Slika 41 Prikaz oznaka za opis pridjeva u NooJ-u 74
Slika 42 Primjer rječničkog unosa s pridjevskim oznakama 74
Slika 43 Primjer strukture flektivnih gramatika 75
Slika 44 Primjer deklinacijskog obrasca za ženski rod jednine pridjeva 'aktivan' 76
Slika 45 Prva morfološka gramatika za prepoznavanje prefiksalne negacije pridjeva 77
Slika 46 Čvor 'Base' u morfološkoj gramatici koji 'poziva' osnove 78
Slika 47 Prikaz nedovoljno učinkovitog rezultata s prvom morfološkom gramatikom 78
Slika 48 Ispravljena morfološka gramatika 79
Slika 49 Sintaktička gramatika u NooJ-u za primjenu na korpusu 79
Slika 50 Označene konkordancije iz korpusa s primijenjenom sintaktičkom gramatikom 80
Slika 51 Prikaz tekstualnih oznaka s kategorijama dodijeljenim morfološkom gramatikom .. 80
Slika 52 Oznaka polariteta dobivena sintaktičkom gramatikom 81
Slika 53 Broj dobivenih rezultata i oznaka polariteta uz prefiksalno negirane pridjeve 81
Slika 54 Prepravljeni čvor 'Base' koji obuhvaća i niz znakova pisan velikim slovima 83

Popis tablica

Tablica 1 Pregled negacijskih prefikasa prema analiziranim autorima 26
Tablica 2 Popis pridjeva po prefiksima za prepoznavanje u NooJ-u 67
Tablica 3 Pridjevi iz korpusa koji se pojavljuju uz različite prefikse 76

Popis grafikona

Grafikon 1 Učestalost negacijskih prefikasa prema broju pridjeva u korpusu 68