

UDK 81:165.7  
81'37  
811.163.42'373  
Izvorni znanstveni članak  
Prihvaćeno za tisak 2. 10. 2001.

Sanja Fulgosi, Nina Tuđman Vuković  
Filozofski fakultet, Zagreb

## Relevantnost frekvencije jezične uporabe pri opisu strukture leksema<sup>1</sup>

Ovaj rad ispituje valjanost testova za određivanje prototipnih značenja polisemnih leksema koji uključuju jezične frekvencije i pretpostavljaju da najveća frekvencija uporabe odgovara središtu leksičke kategorije. Uz pomoć računalnog korpusa hrvatskoga jezika, Hrvatskoga nacionalnog korpusa (HNK) i metodologije korpusne lingvistike analiziraju se frekvencijski podaci za tri hrvatske imenice i uspoređuju s intuicijom izvornog govornika u određivanju prototipnog značenja istih leksema. Naglašava se da su razlozi postojanja prototipnosti u leksema kognitivne prirode te da postoje barem neke jezične kategorije u kojih frekvencija uporabe nije pokazatelj prototipnosti.

---

### Uvod: o prototipnosti u polisemnih leksema

Kognitivna semantika prihvatila je pojam **prototipa** kao temeljnog načela organizacije kategorija. Ova je pojava najprije uočena kod perceptivnih kategorija kao što su *boje* i *oblici* ili pak semantičkih<sup>2</sup> kategorija kao što su primjerice *voće*, *ptice* i *namještaj* (vidi npr. Rosch 1973, 1977, 1978; Rosch i Mervis 1975). Međutim, ubrzo se uvidjelo da je organizacija članova neke kategorije oko prototipa – najboljeg primjerka – po načelu veće ili manje sličnosti primjenjiva i šire u jeziku, budući da su sve jezične kategorije, uključujući i gramatičke, kognitivni konstrukti. Nadalje, i značenje leksema možemo promatrati kao kategoriju koja ima svoju unutrašnju organizaciju u kojoj postoji jezgra

---

1 Ovaj rad temelji se na izlaganju pod nazivom *A New View of Frequency and Prototypes*, održanom na skupu *New Theoretical Perspectives on Syntax and Semantics in Cognitive Science* od 8. do 10. rujna 2000. godine u Dubrovniku.

2 Nazive *perceptivna* i *semantička kategorija* rabi E. Rosch (1973).

(prototip) i periferni članovi koji su određeni većom ili manjom sličnošću s prototipom. Kao što je *voće*, primjerice, za govornike hrvatskog jezika vrlo dobar – prototipni – primjer kategorije *plod*, a plod svake nejestive biljke posve rubni član iste kategorije, analogijom bismo morali biti u mogućnosti odrediti jezgre i periferna značenja istoga leksema.

Ovdje valja naglasiti da želimo razlikovati kategorije pod čijim članovima u stvari razumijevamo entitete stvarnoga svijeta koje označujemo različitim leksemima, od leksema kao kategorije s više različitih, ali međusobno povezanih značenja. Među prve ubrajao bi se već spomenuti *plod*, s članovima kao što su *voće*, *povrće*, *žitarice* i drugi, a kod drugih se radi o značenjskim kategorijama gdje je leksem nadređena kategorija, a različita značenja njezini su članovi. Poslužimo li se istom kategorijom, leksemom *plod*, članove možemo označiti kao *plod*<sub>1</sub>, *plod*<sub>2</sub>, *plod*<sub>3</sub> i tako dalje, ovisno o broju značenja leksema. Pritom bi se *plod*<sub>1</sub> kao jezgra kategorije odnosio na 'jestivi dio biljke'<sup>3</sup>, dok su preostali članovi nastali različitim jezičnim mehanizmima proširenja značenja, pa tako imamo *plod*<sub>2</sub> = 'nerodeno dijete', *plod*<sub>3</sub> = 'rezultat djelatnosti' i tako dalje.

Jezične kategorije poput navedene imaju **polisemnu** strukturu. To znači da među značenjima postoji konceptualna povezanost čija je priroda takva da kategorija ostaje stabilan entitet, unatoč fleksibilnosti koja omogućuje nastanak novih značenja<sup>4</sup>. Značenja unutar polisemne kategorije organizirana su po načelu rodbinskih sličnosti (engl. *family resemblances*), a jedno od njih ima središnji status s obzirom na ostale članove. Središte kategorije ovdje nazivamo i **prototipnim članom**, odnosno prototipnim značenjem jezične kategorije, pri čemu naziv prototip rabimo u spomenutom smislu koji se odnosi isključivo na značenjski ustroj leksema. Predmet ovoga rada upravo je priroda prototipa jezične kategorije i način njegova određivanja te posredno i ustroj čitave polisemne kategorije koji odražava mehanizme nastanka značenja više ili manje udaljenih od središta. Vidjet ćemo da se jedan od čestih problema pri analizi polisemnih leksema koji je od velike važnosti za opis značenjskog ustroja odnosi upravo na određivanje statusa pojedinih značenja. Je li neko značenje utemeljeno kao zasebna **jezična jedinica**<sup>5</sup> pitanje je koje nije samo lingvističko već zadire i u područje psihologije, dok na leksikografskoj razini razgraničenje značenja, odnosno njihov broj često ovisi o subjektivnom odabiru leksikografa.

### Frekvencija uporabe i određivanje središta leksičke kategorije

Sve su kategorije kognitivni konstrukti zajednički govornicima nekog jezika. U različitim se pokusima pokazalo da ispitanici izuzetno dosljedno odabiru

3 Ovdje se radi o svakodnevnoj definiciji leksema *plod* koja se razlikuje od tehničke definicije kakva se navodi u rječniku, a plod opisuje kao 'dio biljke koji sadrži sjemenke'.

4 Usp. Taylor 1995: 53–54; također vidi Raffaelli 2000.

5 Status jezične jedinice imaju oni elementi jezika koje govornik ima na raspolaganju kao gotove i upotrebljava ih automatski, te za njih nije potreban napor koji se ulaže pri stvaranju novih struktura. (Langacker 1987: 57–76)

prototipne članove kategorije, odnosno da postoji vrlo visok stupanj suglasnosti pri određivanju rubnih članova (vidi npr. Rosch 1973). Isto vrijedi i za leksem kao značenjsku kategoriju, gdje izvorni govornik intuitivno može odrediti **prototipno značenje** leksema, kako kod konkretnih, tako i kod apstraktnih kategorija. Postojanje takvog središnjeg značenja oko kojega se organiziraju ostala značenja jedna je od posljedica ekonomičnosti u jeziku, odnosno kognitivnog ustroja u kojem su informacije povezane na takav način da govorniku budu najlakše dostupne.

Nije, dakle, sporno da je **prototipnost** načelo po kojem su ustrojene jezične kategorije. U načelu postoje tri načina za određivanje jezgrenog značenja leksičke kategorije (Dirven i Verspoor 1998: 31). Prvo, oslanjajući se na intuiciju izvornog govornika, možemo ispitati koje se od značenja kod najvećeg broja ispitanika pojavljuje kao prva asocijacija na spomen određene kategorije. Primjerice, kod leksema *plod* najčešća će biti uporaba sa značenjem 'dio biljke', dok je malo vjerojatno da će se pojaviti odgovor 'rezultat djelatnosti'. Drugo, pri određivanju prototipnog značenja možemo se poslužiti statističkim podacima o uporabi, odnosno frekvenciji pojavljivanja. Što je neka uporaba frekventnija, to je značenje bliže središtu kategorije. Treće, prototip kategorije predstavlja ono značenje koje služi kao objasnidbeno ishodište ostalih uporaba.

Prve dvije od navedenih metoda za određivanje središta leksičke kategorije služe se **frekvencijom** kao pokazateljem odnosa prototipnog i perifernih značenja. U prvom slučaju riječ je o frekvenciji odgovora ispitanika, gdje najveća frekvencija pojavljivanja nekog odgovora upućuje na prototipno značenje. U drugom slučaju promatra se broj pojavljivanja pojedinih značenja na jezičnom korpusu te se smatra da je ono najfrekventnije ujedno i prototipno značenje. U skladu s time u lingvističkim se radovima frekvencija uporabe često spominje kao dokaz prototipnih uporaba. Tvrdi se da jezična frekvencija upućuje na prototipno značenje te nam služi kao korisno oruđe u njegovu određivanju (npr. Geeraerts 1988: 222). Međutim, kasnije ćemo pokazati da postoje neke leksičke kategorije kod kojih takav odnos frekvencije uporabe i prototipnosti ne vrijedi. Vidjet ćemo da kod takvih kategorija prototipno značenje nije uvijek i najfrekventnije, a vjerujemo da razlog tome valja potražiti u prirodi unutrašnjeg ustroja kategorije, pri čemu je prototip njezin kognitivno najistaknutiji član.

Zbog svoje **kognitivne istaknutosti** (engl. *cognitive salience*), ali i drugih jezično-povijesnih razloga te prirode povezanosti središnjeg i ostalih značenja, prototipno značenje često jest, ali nije i uvijek, najfrekventnije. Drugim riječima, kod onih leksičkih kategorija gdje najfrekventnija uporaba odgovara prototipnom značenju, frekvencija jezične uporabe posljedica je postojanja prototipne strukture, a ne uzrok prototipnosti. Kako kaže Geeraerts, neke jezične uporabe nisu prototipne jer su frekventnije, već su frekventnije zato što su prototipne (Geeraerts 1988: 222). S druge strane, frekvencija je jedan od mogućih, kulturnološki uvjetovanih izvora prototipnosti izvanjezičnih kategorija. Primjerice, jedan od razloga, ako ne i jedini, zašto je *jabuka* u kulturi govornika hrvatskog jezika prototipno *voće* jest činjenica da je najčešće percipiramo kao pripadnika spomenute kategorije. Nadalje, na osnovi **referencijalne frekvencije** ne možemo donositi sudove o prototipnosti značenja u okviru leksema. Kod jezičnih

kategorija, dakle, ne radi se o referencijalnoj frekvenciji, već o **frekvenciji uporabe** prema kojoj možemo donositi sudove isključivo o prototipnosti jezične kategorije, a ne izvanjezične stvarnosti na koju se odnosi.

Treća tvrdnja o određivanju središnjeg značenja leksičke kategorije pomoću prototipa kao objasnidbenog ishodišta ostalih značenja ključna je te predstavlja općenit izraz načela prototipnosti u leksema i postojanja središnjeg i od njega više ili manje udaljenih perifernih članova. Možemo je dovesti u vezu sa **psihološkom hipotezom** za objašnjenje pojavnosti prototipa kod semantičkih kategorija, definiranom na temelju radova E. Rosch o načelima kategorizacije:

*“It states that it is cognitively advantageous to maximize the conceptual richness of each category through the incorporation of closely related nuances into a single concept because this makes the conceptual system more economic. Because of the maximal conceptual density of each category, the most information can be provided with the least cognitive effort.”* (Geeraerts 1988: 208)

Navodeći četiri hipoteze o mogućim uzrocima postojanja prototipa, koje sažima iz radova Roscheve, Geeraerts tvrdi kako je psihološka hipoteza općenitija od ostalih jer je funkcionalne prirode (Geeraerts 1988). Naime, kod nekih kategorija poput *boja* i *oblika* pojavnost prototipne strukture možemo objasniti prirodom perceptivnog sustava (*fiziološka hipoteza*), kod drugih postojanjem većeg broja zajedničkih atributa kod središnjih članova kategorije nego kod perifernih (*referencijalna hipoteza*), a kod trećih pak većom referencijalnom frekvencijom prototipnih članova (*statistička hipoteza*). Četvrta, *psihološka hipoteza* općenitija je jer ne upućuje na mjerljive znakove prototipnosti, već se poziva na kognitivne prednosti unutrašnjeg ustroja kategorija, odnosno načela povezanosti njihovih članova. Koncepti pojedinih kategorija organizirani su tako da dohvat informacija o članovima iziskuje najmanji kognitivni napor.

Za jezične kategorije to također znači ustroj koji sustav čini ekonomičnijim. Među pojedinim značenjima polisemne leksičke jedinice postoji povezanost u okviru zajedničkog koncepta, takva da postoji značenje koje se postavlja kao ishodište i temelj ostalih uporaba. Pritom se jedan ili više karakterističnih atributa različitim mehanizmima, među kojima je i **metaforizacija**, prenosi na ostala značenja. Vidjet ćemo da se upravo kod nekih polisemnih leksema sa značenjima nastalim metaforičnim proširenjima središnjeg značenja prototip uvijek odnosi na konkretno značenje koje je izvor takvih proširenja. Čak i u slučaju da prototipno značenje nema materijalnu potvrdu u frekvenciji uporabe, činjenica da ono služi kao ishodište ostalih značenja je ključna.

### **Korpusna i kognitivna lingvistika – empirijski pristupi jeziku**

Provjera postavke drugog testa za određivanje prototipnog značenja leksičke kategorije prema frekvenciji uporabe (Dirven i Verspoor 1998: 31) pretpostavlja korištenje **računalnog korpusa**, najrasprostranjenijeg izvora za kvantitativna lingvistička istraživanja. U ovom ćemo istraživanju na primjerima iz stvarne jezične uporabe, iz jezičnog računalnog korpusa, pokazati da frekvencija nekih

konkretnih kategorija sa značenjima nastalim procesom metaforizacije ne upućuje uvijek na prototipno značenje, kako bi se moglo očekivati prema drugom »testu« za određivanje prototipnosti.

U ovom radu koristimo se teorijskim postavkama i metodologijama dvaju empirijskih pristupa jeziku, korpusne i kognitivne lingvistike. Ova se dva pristupa s jedne strane suprotstavljaju čomskijevskom gledištu prema kojem se jezična djelatnost odvija u zasebnom jezičnom modulu neovisno o ostalim mentalnim aktivnostima. Za razliku od takva shvaćanja, kognitivna lingvistika jezičnu djelatnost vidi kao dio ukupne mentalne strukture. Iz toga proizlazi da u razmatranjima o prirodi i funkcioniranju jezika moramo uzeti u obzir njegovu uporabu, a upravo je uporaba jezika temelj svakog korpusnog pristupa jeziku.

Suvremena korpusna lingvistika grana je lingvistike koja pretpostavlja jezičnu analizu na računalnom korpusu pisanog ili govorenog jezika (McEnery i Wilson 2001: 1) (Ooi 1998: 34). Korpusna istraživanja uza sam korpus pretpostavljaju i suvremene alate i metodološke postupke koji jezičnu građu čine preglednom i dostupnom. Metodologijom korpusne lingvistike moguće je nadopunjavanje teorijskih tvrdnji i intuitivnih pretpostavki podacima iz stvarne jezične uporabe u gotovo svim lingvističkim disciplinama i područjima.

Od prvog računalnog jezičnog korpusa, milijunskog korpusa engleskog jezika, poznatijeg kao Brown–korpus (Kučera i Francis 1967), suvremena je korpusna lingvistika vrlo brzo napredovala. Veliki zamah korpusnih istraživanja u 80–im i 90–im godinama 20. stoljeća omogućen je brzim razvojem memorijskih mogućnosti osobnih računala, što je rezultiralo izradom velikih višemilijunskih korpusa, koji se često nazivaju korpusima druge generacije (Kennedy 1998: 45). Budući da semantičke analize zahtijevaju sustavno selektiranu opsežnu jezičnu građu, najveću su potrebu za takvim korpusima pokazala leksikološka istraživanja koja su vrlo široko primijenjena prije svega u engleskoj leksikografiji<sup>6</sup>. Budući da za potrebe ovog istraživanja valja utvrditi frekvencije pojedinih značenja polisemnih leksema, metodologija slijedi uporabu korpusa kakva je razvijena za potrebe navedenih leksikoloških istraživanja.

Osnovna je ideja takvih istraživanja promatrati pojavnice leksema u njihovu neposrednu okruženju i širem jezičnom kontekstu, što omogućavaju konkordancijski računalni alati<sup>7</sup>. Na temelju tako uredene jezične okoline moguć je pregledan uvid u gramatičke i leksičke značajke leksema s jasnim pokazateljima ukupnog broja pojavnica pojedinog leksema te frekvencije pojedinih značenja. Navedene značajke korpusne lingvistike čine korpusnu jezičnu analizu primjerenom metodologijom za određivanje što točnijih frekvencijskih odnosa središnjeg značenja i metaforičnih značenja pojedinog leksema. Podaci o fre-

6 U sklopu projekta sveučilišta u Birminghamu i izdavačke kuće Collins Cobuild sastavljen je *Cobuild corpus* (Sinclair i ur. 1987) koji je poslužio kao glavni izvor za istoimenu jednojezični rječnik engleskoga jezika (Sinclair 1987). Na temelju korpusa engleskog jezika sastavljen je i drugi poznati engleski jednojezični rječnik (LDCE 1995)

7 Najpoznatiji je oblik konkordancije tzv. KWIC (engl. *Key Word in the Context*) koji je u korpusnim istraživanjima postulirao John Sinclair (Sinclair 1991: 32–33). KWIC je ogledni okvir za korpusne analize koji se navodi i u temeljnoj korpusnoj literaturi (Ooi 1998), (Kennedy 1998), (McEnery i Wilson 2001), (Tognini–Bonelli 2001)

kvenciji, odnosno o jezičnoj uporabi do kojih se sustavno može doći s pomoću računalnog korpusa mogu biti dobro uporište za razmatranje utjecaja frekvencije pojedinog leksema na opis strukture njegova značenja.

Mogućnosti iscrpnih semantičkih analiza povećavaju se pojavom velikih referentnih, odnosno za pojedini jezik reprezentativnih korpusa koji sadrže više stotina milijuna pojavnica. Najpoznatiji je takav korpus stomilijunski BNC (*British National Corpus*) koji je svojim sastavom reprezentativan za engleski jezik te služi kao uzor nacionalnim korpusima ostalih jezika. Postojanje tako opsežnih i uravnoteženih korpusa može umanjiti prigovore da su korpusne analize nedostatne jer su korpusi nerepresentativni za pojedini jezik kao cjelinu.

Ovo je istraživanje provedeno na probnoj inačici suvremenog dijela *Hrvatskoga nacionalnog korpusa*<sup>8</sup> (HNK), koja je iznosila oko 8 milijuna pojavnica. HNK ishodište je referentnog korpusa od 30 milijuna pojavnica (Tadić 1998: 340) koji će biti reprezentativan za suvremeni hrvatski jezik. Stoga se već pri sastavljanju probne inačice korpusa težilo što boljoj reprezentativnosti jezične grade, što potvrđuje pet potkorpusa HNK kojima su obuhvaćene najvažnije jezične domene pisanog jezika: *novine, časopisi, knjige, beletristika, te eseji i govori*. Smatramo stoga da nam HNK svojim sastavom i opsegom može poslužiti kao glavni izvor za ovo istraživanje.

### Hrvatski leksemi *grana, korijen i plod*

Ovaj rad razmatra tri **leksema**, tri imenice koje se odnose na dijelove biljke: *grana, korijen* i *plod*. Jezični smo korpus pretraživali uzimajući u obzir sve pojavnice, odnosno sve morfološke oblike navedenih imenica. Prema tome u *Hrvatskom nacionalnom korpusu* nalazimo 316 pojavnica leksema *grana*, 531 pojavnicu leksema *korijen* te 446 pojavnica leksema *plod*. Svaka od navedenih imenica ima jedno konkretno značenje koje se odnosi na jedan od dijelova biljke. To je ujedno prototipno značenje, ono za koje izvorni govornik 'osjeća' da je osnovno i da predstavlja središte kategorije. Sva ostala značenja metaforična su proširenja konkretnoga značenja pa zaključujemo da su leksemi po svojem ustroju polisemni.

Korpus omogućava pregledan ispis svih uporaba navedenih leksema a njihova se značenja mogu utvrditi na temelju njihove neposredne okoline i širega jezičnog konteksta. Takvom smo analizom utvrdili da svi navedeni leksemi osim konkretnih značenja koja se odnose na dio biljke imaju i nekonkretna, apstraktna značenja koja u većini slučajeva odgovaraju jezičnoj realizaciji **konceptualnih metafora**<sup>9</sup> *ljudi su biljke* i *ideje su biljke*. Korpusna se analiza

8 HNK je u cijelosti dostupan na adresi [www.hnk.ffzg.hr](http://www.hnk.ffzg.hr)

9 Za razliku od shvaćanja metafore kao isključivo pjesničke i dekorativne jezične figure kognitivna lingvistika metafori pristupa kao svakodnevnoj, konvencionalnoj pojavi koju čovjek u svom jezičnom izražavanju često rabi gotovo nesvjesno. Iako broj mogućih metaforičnih izraza može biti neograničen, oni se ne proizvode sasvim proizvoljno, već proizlaze iz ograničenog broja sustavno organiziranih osnovnih metaforičnih koncepata koje nazivamo konceptualnim metaforama. Više o konceptualnim metaforama vidi u Lakoff i Johnson 1980, Lakoff i Turner 1989.

sastoji od dva dijela. U prvom smo dijelu svakoj pojavnici odredili značenje s obzirom na to je li konkretno ili metaforično.

Drugi i složeniji dio korpusne analize sastoji se u razgraničavanju metaforičnih značenja. Preglednost jezične građe iznesene u konkordancijama omogućava sustavan uvid u kontekst koji pomaže u preciznom određivanju i razgraničavanju značenja. Iako izlazi iz okvira ovog rada, kratkom ćemo digresijom naglasiti da izostanak takve korpusne analize može rezultirati izostankom upravo metaforičnih značenja iz rječnika (Fulgosi i Tudman Vuković 2000) koja iako česta, nisu sustavno dohvatljiva intuicijom. O tome svjedoči jednojezični rječnik hrvatskoga jezika (Anić 1998: 437) prema kojem za 218 pojava leksema *korijen* u našem korpusu ne možemo odrediti značenje na temelju značenja navedenih u rječniku. Budući da se radi o konceptualnoj metafori koja na isti način funkcionira i u engleskom jeziku, engleski leksem *root* ('korijen') ima slična metaforična značenja, pa smo se u njihovu određivanju rukovodili već postojećim opisom u rječniku sastavljenom na temelju korpusa (LDCE 1995: 1237).

Ta činjenica upućuje na potrebu da se tri temeljna dokaza u jezičnom opisu: *postojeći jezični opisi*, *intuicija izvornih govornika* i *jezik u kontekstu* (Sinclair 1991: 37) nadopunjuju i prožimaju. Korpusni pristup, stoga, nije jedini odgovor na složena pitanja leksičkih značenja. U korpusu značenja nisu eksplicitna. Snaga je suvremenog računalnog korpusa, potpomognutog tehnologijom, u sustavnom usmjeravanju istraživačke intuicije prema pravoj slici funkcioniranja nekog leksema. Upravo zbog te činjenice naglašavamo da je nezaobilazan u leksikološkim istraživanjima. Nedostaci koje nalazimo u rječnicima upućuju na nedostatnost pojedinačne jezične intuicije, posebno kada se radi o metaforičnim značenjima.

Na tablicama 1, 2. i 3. izneseni su rezultati cjelovite korpusne analize značenja leksema *grana*, *korijen* i *plod*.

<i>grana</i>	
značenje	broj pojava
1. dio biljke	64
2. znanstvena disciplina	205
3. idiomi	16
4. dio organa	3
5. obitelj	16
6. organizacija	1
7. pjesnička metafora	11
UKUPNO	316

**Tablica 1.**

Osnovne značenjske odrednice leksema *grana* s frekvencijama iz korpusa

<i>korijen</i>	
značenje	broj pojavnica
1. dio biljke	195
2. dio organa	5
3. početak, izvor	103
4. a) rezultat korjenovanja	2
b) rješenje	0
5. korijen riječi	8
6. podrijetlo	55
7. temelj (ljudi)	31
8. temelj (ideje)	83
9. uzrok	49
UKUPNO	531

**Tablica 2.**

Osnovne značenjske odrednice leksema *korijen* s frekvencijama iz korpusa

<i>plod</i>	
značenje	broj pojavnica
1. dio biljke	124
2. fetus	18
3. rezultat aktivnosti (ideje)	289
4. rezultat (ljudi)	9
5. kolokacija 'plodovi mora'	6
UKUPNO	446

**Tablica 3.**

Osnovne značenjske odrednice leksema *plod* s frekvencijama iz korpusa

Rezultati pokazuju da leksemi *grana*, *korijen* i *plod* imaju gotovo isti omjer konkretnih i metaforičnih značenja koji u prosjeku iznosi 30% : 70% u korist metaforičnih. To znači da su metaforična značenja gotovo dva puta učestalija. Slične smo omjere (68,9% : 31,8%) metaforičnih i nemetaforičnih značenja engleskog leksema *root* (hrv. *korijen*) dobili pretragom i analizom već spomenutog stomilijunskog BNC-a. Rezultati dobiveni na zaokruženom referentnom korpusu upućuju na činjenicu da navedeni omjeri prema HNK-u nisu posljedica slučaja, te da bismo slične rezultate dobili na opsežnijem i potpuno urav-



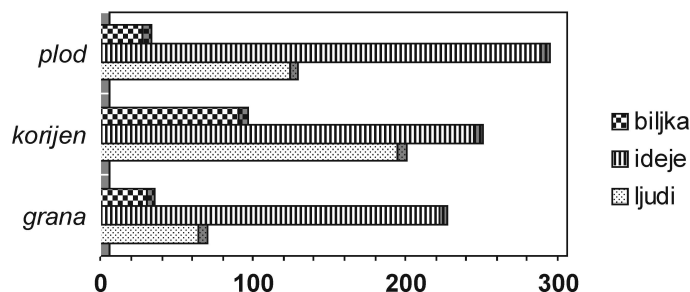
noteženom korpusu hrvatskoga jezika jer su takvi frekvencijski odnosi posljedica prirode značenjske strukture hrvatskog leksema *korijen*, odnosno engleskog leksema *root*. Stoga možemo ostati pri tvrdnji da su metaforična značenja ovih leksema u načelu češća od nemetaforičnih. *Tablica 1.* za leksem *grana* i *tablica 3.* za leksem *plod* jasno pokazuju da su metaforična značenja frekventnija. I konkretno značenje leksema *korijen* načelno je frekventnije (*tablica 2*) iako je na prvi pogled jedno od metaforičnih značenja ('početak, izvor') najfrekventnije. Razlog je tome što *korijen* ima najviše, čak osam metaforičnih značenja, pa podaci o frekvenciji pokazuju raspršenost njihove uporabe. Kod polisemnih leksema, vidjeli smo, često postoji problem razgraničenja značenja i određivanja statusa pojedinih značenja kao jezičnih jedinica. Rješavanje ovog problema često se svodi na subjektivnu prosudbu lingvista, odnosno leksikografa.

Napomenuli smo da gotovo sva metaforična značenja pripadaju konceptualnim metaforama *ljudi su biljke* i *ideje su biljke* kako ih definiraju Lakoff i Johnson (1980) te Lakoff i Turner (1989). Konkretno – prototipno – i metaforična značenja leksema *grana*, *korijen* i *plod* mogu se apstrahiranjem svesti na tri domene<sup>10</sup>; prototipno značenje pripada domeni *biljke*, a metaforična domena *ljudi* i *ideje*, kao što pokazuju primjeri iz korpusa:

- (1) koje mogu značajnije oštetiti *korijen* loze (**biljke**)
- (2) duboki hrvatski *korijeni* u BiH (**ljudi**)
- (3) Gdje bi bili *korijeni* slabe motiviranosti zaposlenih u građevinarstvu? (**ideje**)
- (4) *Grane* divljeg šipka što su se razrasle (**biljke**)
- (5) Kognitivna psihologija jest *grana* psihologije koja se bavi kognicijom (**ideje**)
- (6) svota koju će naslijediti hrvatska *grana* obitelji (**ljudi**)
- (7) Kako se zaraza proširi na *plodove* jabuke? (**biljke**)
- (8) Projekt je trebao biti *plod* multidisciplinarnog rada i temeljiti se na... (**ideje**)
- (9) Svako dijete je dar ljubavi, *plod* ljubavi (**ljudi**)

Svođenjem značenja na domene dokinuta je raspršenost frekvencija pojedinih metaforičnih značenja, što se posebno odnosi na leksem *korijen* koji ima najviše metaforičnih značenja. Podaci o frekvencijama na *slici 1.* pokazuju da su značenja koja pripadaju domeni *ideje* i domeni *ljudi*, tj. metaforična značenja svih triju leksema frekventnija od konkretnih (prototipnih) značenja. Iz svih se navedenih primjera vidi da je konkretno značenje izvorište ostalih značenja te je stoga prototipno.

10 Izraz *domena* ovdje rabimo u smislu u kojem se koristi u kognitivnoj gramatici. Domena je strukturirani dio konceptualizacije čovjekova iskustva u odnosu na koji se karakteriziraju semantičke jedinice (Langacker 1987: 147–154). Primjerice, koncepte *korijen*, *stabljika*, *list*, *plod* možemo razumjeti i njima baratati isključivo u odnosu na domenu *biljke* koja pruža svojevrsan konceptualni kontekst u koji se smještaju ovi leksemi.



Slika 1. Frekvencije značenja po domenama

Korijen je dio biljke koji osigurava postanak i opstanak pa se temeljno značenje prenosi s konkretnog na apstraktno na domenu *ljudi* (primjer 2) kao npr. *korijeni obitelji*, te na *ideje* (primjer 3) kao npr. uzrok ili početak, ili temelj. Prema istom načelu svoje konkretno značenje prenosi u metaforično leksem *grana* – dio biljke koji se odvaja od stabljike – pa se njegova metaforična značenja odnose na obitelj (primjer 5) te razradbu različitih područja ljudskog djelovanja (primjer 6). Plod je dio biljke koji nastaje kao posljedica biljnog reproduktivnog ciklusa, pa se u skladu s tim i njegova metaforična značenja preslikavaju kao rezultat i posljedica neke ljudske aktivnosti (primjer 8). Zanimljivo je vrlo frekventno metaforično značenje leksema *plod* koje pripada domeni *ljudi* i gotovo se uvijek odnosi na ljudski fetus. Ovo je značenje stupnjem metaforičnosti mnogo bliže središnjem značenju leksema *plod* od onog koje pripada domeni *ideje*. I druga dva leksema imaju metaforična značenja koja se s domene *biljke* preslikavaju na ljudsku fiziologiju kao 'dio organa'. Iako se ova značenja odnose na konkretno pa su time bliža prototipnom, možemo ih smatrati metaforičnim značenjima. Značenja iz domene *biljka* izvornom se govorniku pojavljuju kao prva reakcija na spomen riječi *grana*, *korijen* i *plod*, dok ostala značenja ostaju metaforičnima bez obzira na stupanj konkretnosti.

### Zaključak

U ovome smo radu na stvarnim primjerima iz hrvatskog jezika željeli pokazati da podaci o frekvenciji uporabe nekih polisemnih jezičnih kategorija s jednim konkretnim i više konceptualno povezanih metaforičnih značenja ne upućuju uvijek izravno na prototipno značenje, kako bi se to moglo očekivati prema pretpostavkama nekih autora. Kao temeljna se ipak postavlja činjenica da je prototipno značenje izvorište ostalih značenja te zbog svoje funkcionalne naravi ima prednost pred ostalim dvama »testovima« prototipnosti. Naime, velika frekvencija istog značenja kao reakcije na spomen određene kategorije posljedica je upravo **kognitivne osnovnosti**, mogli bismo čak reći kognitivne jednostavnosti toga značenja. Drugim riječima, ako neko značenje služi kao objasnidbeno ishodište ostalih uporaba, to znači da je u odnosu na njih temeljno.

Takav je slučaj i s konkretnim značenjima leksema *grana*, *korijen* i *plod* – onima koja se odnose na dio biljke, usprkos činjenici da su metaforična značenja, osobito ona iz kognitivne domene *ideje* znatno frekventnija. Također, to su značenja koja se izvornom govorniku pojavljuju kao prva reakcija na spomen riječi *grana*, *korijen* i *plod*, upravo stoga što su osnovna. Kao i kod psihološke hipoteze o prototipnosti, postoje kognitivni razlozi zbog kojih je značenje 'dio biljke' istaknutije od ostalih. Činjenica da je biljka opipljiv, konkretan predmet vrlo blizak čovjeku zasigurno je jedan od važnijih.

Sva preostala značenja više su ili manje utemeljena kao zasebni članovi leksičke kategorije. Pitanje može li koje od njih s vremenom preuzeti ulogu središnjeg, prototipnog značenja predstavlja poseban problem koji izlazi iz okvira ovog rada. Pa ipak, smatramo da su konkretna značenja leksičkih kategorija koje smo istraživali iz već spomenutih kognitivnih razloga sa svojom iskustvenom podlogom duboko ukorijenjena u središte kategorije. Moguće je da su u nekih drugih gramatičkih kategorija pomaci jednog od značenja prema središtu kategorije uslijed velike frekvencije uporabe izgledniji. Pitanje je li frekvencija uporabe u tih leksema u većem razmjeru s udaljenošću od središta kategorije, to jest je li najveća frekvencija ujedno i pokazatelj središnjeg značenja te koje su to specifične značajke pojedinih kategorija ostaje potencijalnim predmetom budućih istraživanja. U svakom slučaju možemo reći da za sve vrste leksičkih kategorija vrijedi funkcionalno usmjerena tvrdnja da je prototip središnji među značenjima zbog svojih kognitivnih prednosti, među kojima je i činjenica da on nerijetko predstavlja neposredno ili posredno izvorište ostalih značenja.

Osvrnimo se na kraju i na ulogu računalnoga jezičnog korpusa u ovoj vrsti semantičke analize. Referentni jezični korpus pokazao se neizostavnim alatom pri analizi frekvencijskih odnosa među pojedinim značenjima polisemnih leksema. Uvid u neposredni i širi kontekst također omogućava grupiranje sličnih uporaba u pojedina značenja, no ovaj je postupak kod nejasnijih razgraničenja, tj. značenja udaljenijih od prototipa, podložan subjektivnosti lingvista, što se često odražava u leksikografskim natuknicama polisema. Sličan problem uočili smo kod leksema *korijen* gdje se veliki broj metaforičnih uporaba može svesti na nekoliko značenja grupiranih prema kognitivnoj domeni, odnosno konceptualnoj metafori na temelju koje su nastala.

Nadalje, dobivena kvantitativna analiza sama po sebi nema objasnidbenu vrijednost. Pojavnost prototipnosti može se objasniti samo introspektivnim uvidom u strukturu leksema, odnosno uz pomoć izvornoga govornika i njegove jezične intuicije. Dok nam je olakšao uvid u velik broj uporaba, sam korpus nam nije mogao rasvijetliti njihove međuodnose. Štoviše, pokazalo se da prosuđivanje o prototipnosti isključivo na temelju kvantitativnih frekvencijskih podataka može biti pogrešno. Dakle, računalni jezični korpus zajedno sa svojom metodologijom može poslužiti kao važan alat u semantičkoj analizi, i to u svrhu dobivanja što potpunijeg popisa pojedinih značenja nekog leksema, ali priroda njihova odnosa, u našem slučaju konkretnog/prototipnog i metaforičnih značenja, iziskuje opis kognitivnih mehanizama kojima ona nastaju.

## Literatura

- Anić, V. (1998.) *Rječnik hrvatskoga jezika*, Novi Liber, Zagreb.
- British National Corpus*. World Edition. 2000. Izdao The Humanities Computing Unit of Oxford University.
- Dirven, R. i M. Verspoor. 1998. *Cognitive Exploration of Language and Linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Fulgosi, S. i N. Tudman Vuković. 2000. »Metaforična značenja i korpus; neke implikacije za leksikografiju«. Zbornik radova Savjetovanja Hrvatskoga društva za primijenjenu lingvistiku *Primijenjena lingvistika u Hrvatskoj – Izazovi na početku 21. stoljeća*. ur. Stolac D, Ivanetić N., Pritchard, B., 137–148.
- Geeraerts, D. 1988. »Where does prototypicality come from?« U Rudzka-Ostyn, B. *Topics in Cognitive Linguistics*. John Benjamins, Amsterdam.
- Kennedy, G. 1998. *An introduction to Corpus Linguistics*, Longman, London i New York.
- Kučera, H. i W. N. Francis. 1967. *Computational analysis of present-day American English*, Brown University Press, Providence, Rhode Island.
- Lakoff, G. i M. Johnson. 1980. *Metaphors We Live By*, The University of Chicago Press.
- Lakoff, G. i M. Turner. 1989. *More than Cool Reason, A Field Guide to Poetic Metaphor*, The University of Chicago Press.
- Langacker, R. 1987. *Foundations of Cognitive Grammar I*, Stanford University Press.
- Longman Dictionary of Contemporary English* 1995. Longman, London.
- McEnery, T. i A. Wilson. 2001. *Corpus linguistics*, Edinburgh University Press, Edinburgh.
- Ooi, V. B. Y. 1998. *Computer Corpus Lexicography*, Edinburgh University Press, Edinburgh.
- Raffaelli, I. 2000. »Neki vidovi kognitivne semantike u rekonstrukciji semantičkih struktura.« *Suvremena lingvistika* 49/50, 125–141.
- Rosch, E. 1973. »On the internal structure of perceptual and semantic categories.« U Moore, Timothy E. (ur.). *Cognitive Development and the Acquisition of Language*. Academic Press, New York.
- Rosch, E. 1977. »Human categorization.« U Warren, Neil. (ur.). *Studies in Cross-Cultural Psychology*. Academic Press, London.
- Rosch, E. 1978. »Principles of categorization.« U Rosch, Eleanor i Barbara B. Lloyd (ur.). *Cognition and Categorization*. Lawrence Erlbaum, Hillsdale.
- Rosch, E. i C. B. Mervis. 1975. »Family resemblances: Studies in the internal structure of categories.« *Cognitive Psychology* 7: 573–605.
- Sinclair, J. (ur.). 1987. *Collins Cobuild English Language Dictionary*. Collins, London i Glasgow.
- Sinclair, J. 1991. *Corpus, concordance, collocation*, Oxford University Press, Oxford.
- Sinclair, J. (ur.) 1987. *Looking Up, An account of the COBUILD Project*, Collins ELT, London i Glasgow.
- Tadić, M. 1998. »Raspon, opseg i sastav korpusa suvremenoga hrvatskoga jezika.« *Filologija* 30/31, 337–347.
- Taylor, J. R. 1995. *Linguistic Categorization*, Oxford University Press, Oxford.
- Tognini-Bonelli, E. 2001. *Corpus linguistics at work*. Benjamins, London.

*The relevance of linguistic frequency in the description of lexeme structure*

This paper questions the validity of tests considered to determine the prototypical senses of polysemous lexemes which involve linguistic frequencies and assume that the highest frequency of use corresponds to the center of the lexical category. Using a Croatian computer corpus and corpus linguistics methodology frequency data for three Croatian nouns are analyzed and compared to native speaker intuition in the establishment of prototypical meanings of these lexemes. It is pointed out that reasons for the existence of prototypicality in lexemes are cognitive in nature and that there are at least some linguistic categories in which frequency of use is not an indicator of prototypicality.

Ključne riječi: polisemni leksemi, korpusna lingvistika, kognitivna semantika, prototipnost (kognitivizam), hrvatski jezik, frekvencija uporabe

Key words: polysemous lexemes, corpus linguistics, cognitive semantics, prototypicality (cognitivism), Croatian language, frequency of usage