

MILAN MOGUŠ

O NASTANKU KOMPJUTORSKIH KONKORDANCIJA I NJIHOVOJ UPOTREBI PRI ANALIZI TEKSTA

Godina 1958. i 1959. bila je u nas ispunjena optimizmom da je kompjutor našao svoje mjesto i u lingvistici. Bilo je to u vezi sa strojnim prevođenjem. Val oduševljenja ponio je, s pravom, grupu mladih lingvista okupljenih u Hrvatskom filološkom društvu u Zagrebu koji su rezultate svojih istraživanja i razmišljanja prezentirali zbirkom rasprava pod zajedničkim nazivom *Strojno prevođenje i statistika u jeziku*¹). Nije mi namjera da ovdje iznosim sve što je tada bilo rečeno, koje su objekcije dane, koliko je ushićenja i sumnji bilo. Ali jedna se misao provlačila ipak: suradnja čovjeka i stroja moguća je pod uvjetom da svaka strana obavlja svoj dio posla — stroju pripada kvantiteta, čovjeku kvaliteta, stroju brzina, čovjeku analiza. Pa kako to da danas, nakon toliko godina od prvih vijesti o strojnom prevođenju, nema gotovo ni traga ni glasa pothvatu da za fantastično kratko vrijeme dobijemo prijevod nekog književnog teksta na jedan, dva ili deset jezika. Drugim riječima: kako to da je strojno prevođenje, nakon tako blistavog početka, ostalo danas samo uspomenom (ako se zanemare tzv. grubi prijevodi nekih tehničkih tekstova).

Možda je zakazao stroj koji ne može prenijeti u drugi jezik sve finese jezičnoga izraza, ili lingvist koji još nije pronikao u jezičnu strukturu i opisao tako njezine zakonitosti da bi to i stroju bilo jasno. Kao što je poznato, u suvremenim su jezicima nataložena stoljeća. Je li moguće tako brzo, kao što to stroj želi, razmrsiti sve probleme i propisati kada se što upotrebljava.

Razmotrimo, ilustracije radi, nekoliko problema.

Naš oblik riječi *Englez* odgovara poljskom obliku *Anglik*. Posebnim programom, gdje se nalaze ne samo upute za leksički dio riječi

¹ *Naše teme*, sv. 6, Zagreb, 1959, str. 105—298. — Pojedinačno su objavili: Bulcsú László: *Broj u jeziku*; Svetožar Petrović: *Može li stroj prevoditi poeziju*; Željko Bujas: *SP u SAD i Engleskoj od 1946. do 1954*; Malik Mulić: *SP u SSSR-u*; Breda Pogorelec: *Počeci rada na SP u ČSR*; Krunoslav Pranjić: *Suvremeno stanje SP*; Božidar Finka: *Odjeci SP u Jugoslaviji*; Stjepan Babić: *Bibliografija*.

nego i za gramatički, stroj može rečenicu *on je Englez* prevesti odmah *on jest Anglikem*. Međutim, ubaci li se u stroj izreka *pravi se Englez*, može se dobiti u tren oka poljski prijevod *udawać Anglika*, ali — to u poljskom jeziku ne znači ništa. U poljskom se, naime, netko može, ako želi ostaviti dojam naivna čovjeka, praviti Grk (a ne Englez) jer se kaže *udawać greka*. Uzmimo obratan slučaj. Poljsku riječ *taniec* stroj lako prevede sa *ples*, pridjev *święty* sa *sveti*, a ime *Wit* sa *Vid*. Poljski izraz *taniec świętego Wita* u hipu se prevodi *ples svetoga Vida*, ali takva nam je fraza potpuno nerazumljiva jer je *taniec świętego Wita* ostao u poljskom jeziku sinonim za padavicu, epilepsiju. I sad možemo nizati: *plakati kao godina* (gdje *godina* znači *kiša*) ili *śutjeti kao zaliven* (elipsa, od: *śutjeti kao zaliven voskom*, tj. *śutjeti kao mrtvac* jer su u pogansko doba mrtvacima zalijevali voskom sve otvore u glavi) ili *očitati komu bukvicu* (gdje je *čitati* sačuvalo prvotno značenje *brojiti*, tj. nabrojiti pogreške, mane) itd. Iz ovoga se može lako naslutiti koliko je mnogo predradnji učiniti, koliko savršeno točnih programa, da bi stroj znao kada prevesti *godina* u značenju '12 mjeseci', a kada 'kiša', kada *očitati (broj)* znači njemački *ablesen (die Ziffer)*, a kada *očitati bukvicu* znači *die Leviten lesen*. Ako se tomu dodaju rečenice, kao što je npr. *vidjela ga je pijana*, gdje ni sami ne znamo što znači (tj. *on je bio pijan i ona ga je vidjela* ili *ona je bila pijana i vidjela ga*), onda bi valjda moglo biti jasno pred kakvim se problemima našao čovjek koji bi sve to trebao prethodno najpreciznije programirati. Iz ovog se može razabrati da je strojno prevođenje napušteno zato što čovjek nije mogao izdržati konkurenciju.

Ljudski se duh, ipak, ne miri tako lako s porazom. Lingvisti nisu prestali ispitivati gdje bi mogli iskoristiti onu vrlinu stroja koja se zove — brzina. Brzinu ne treba miješati s površnošću. Riječ je o brzom dobivanju najtočnijih i najiscrpnijih podataka. Problem je u ovome: kako u istraživanju osloboditi lingviste onog mehaničkog posla, onog nezaobilazno dugog ispisivanja, razvrstavanja, podvrstavanja, abecediranja itd. koja ih, osobito kad se radi o analizi golemih tekstova, toliko iscrpe da nisu u punoj kreativnoj snazi kad treba povezivati, zaključivati, sintetizirati — ukoliko do toga uopće dođu. I sasvim je logično da se pomisao o stroju — u prvom redu o kompjutoru zbog njegove nevjerovatne brzine i točnosti — počela nametati kao moguće rješenje upravo tamo gdje je trebalo beskrajno dugo ispisivati kartice da bi se dobio uredno abecediran slijed potvrda za svaki oblik svake riječi u svim rečenicama nekoga korpusa, drugim riječima pri izradi konkordancija. Konkordancija je, dakle, strogim abecednim redom složena lista svih oblika svake riječi nekoga djela s kontekstom koji tu riječ okružuje i s točnom oznakom mjesta gdje se sve javlja u djelu. Dok nije bilo stroja, pogotovu elektroničkog stroja što ga zovemo kompjutor, čitav se posao morao odvijati ručno. To znači: svaku riječ ispisati ručno na kartici i uz nju svaki put navesti kontekst koji joj pripada, autora, djelo, stranu i redak (ili stih). Tko je god to pokušao raditi, zna koliko je strpljivosti, koncentracije i volje potrebno da se na kartice prenese sav (ponavljam: sav!) leksički inventar nekoga djela. Relativno je lako kad neki tekst

ima, na primjer, 10.000 riječi. Sasvim je drugačije kad počnemo ispisivati djelo koje sadrži 30.000—40.000 riječi ili, kao Gundulićev *Osman*, 55.400 riječi²) ili Krležin *Povratak Filipa Latinovicza* sa 65.159 riječi³) ili, prema mojoj procjeni, Krležin roman *Zastave* s preko 400.000 konkordiranih riječi (da i ne govorimo o broju riječi u svim Krležinim djelima).

A ipak su se ljudi upuštali bar u neke slične pothvate jer je vrijednost konkordancija, kao što ćemo vidjeti, zaista velika. Poznato je da su se opsežna i važna djela, kao npr. Biblija ili Shakespeareova djela, počela ručno konkordirati još u prošlom stoljeću⁴), ali je manje poznato da je već 1842. godine Gustav Flügel izdao u Leipzigu konkordanciju Kurana pod naslovom *Concordantiae Corani Arabicae*⁵). Koliko je na tom poslu Flügel radio, nije zapisano. Zna se, međutim, da je na konkordanciji Shakespeareovih djela radio John Barlett punih 16 godina, a »konkordancija Chaucerovih djela vukla se uz prekide preko 50 godina«⁶).

Može se, na sreću, tvrditi da izrada konkordancija na osnovi ručnog ispisivanja građe pripada prošlosti. Slijed je događaja zanimljiv: koliko je, nakon početnog uzleta, slabila ideja o strojnom prevodenju, toliko je jačala misao o upotrebi kompjutora za izradu konkordancija.

Prve su kompjutorske konkordancije izrađene u SAD 1959. godine, a sa strojnima se, iako ne elektroničkim, počelo već 1951. godine u Italiji⁷). Od 1960. godine počela su se takva djela pojavljivati u većem broju. Biblioteka kompjutorskih konkordancija broji sada vjerojatno nekoliko stotina primjeraka.

Početak rada na kompjutorskom konkordiranju u Jugoslaviji vezan je uz djelatnost prof. dr. Željka Bujasa (Odsjek za anglistiku Filozofskog fakulteta u Zagrebu). On je za vrijeme svoga boravka u Americi (1967—1968) uspio na univerzitetu u Texasu izraditi 5 konkordancija hrvatskih književnih tekstova, i to: Marulićeva *Susana*, isto djelo (*Susana*) odostražno konkordirano, Gundulićev *Osman*, Krležin *Povratak Filipa Latinovicza* i engleski prijevod istog djela (*The Return of Philip Latinovicz*).

U to, dakle, vrijeme dobio sam od kolege Ž. Bujasa prve informacije o kompjutorskim konkordancijama. Bilo je to u doba kad se pomišljalo na novo, kritičko izdanje Marulićevih djela jer se 1969.

² Željko Bujas: *Kompjutorska konkordanca Gundulićeva »Osmana«*, Filologija, sv. 7, Zagreb, 1973, str. 42.

³ Željko Bujas: *Navedeno djelo*, str. 42.

⁴ Željko Bujas: *Navedeno djelo*, str. 36.

⁵ Za taj podatak zahvaljujem kolegi Srđanu Jankoviću, koji mi je pokazao jedan primjerak Flügelove konkordancije što se nalazi u biblioteci Filozofskoga fakulteta u Sarajevu.

⁶ Željko Bujas: *Navedeno djelo*, str. 36—37.

⁷ Uspor. Roberto Busa: *S. Thomae Aq. Hymnorum Ritualium Varia Specimina Concordatarum*, Milano, 1951. (Citirano prema navedenom djelu Ž. Bujasa, str. 37)

navršavalo sto godina otkako je izašla 1. knjiga Akademijine edicije *Stari pisci hrvatski* (SPH) u kojoj je Ivan Kukuljević Sakcinski izdao Marulićeva djela. Novo je izdanje Marulićevih djela o toj stogodišnjici bilo motivirano uglavnom ovim razlozima: 1. pojavila se sumnja u Marulićevo autorstvo nekih pjesmotvora što ih je Kukuljević unio u SPH⁸), 2. u djelima koja su bez sumnje Marulićeva (a tu se u prvome redu misli na *Juditu* i *Susanu*) ima slabo ili pogrešno pročitanih mjesta⁹), 3. javljaju se glasovi da su pronađena nova Marulićeva djela¹⁰), pa bi ih — ako je to istina — trebalo u novom izdanju priključiti Marulićevoj opusu. Zbog svih triju razloga bila je potrebna uža suradnja književnih historičara i lingvista. Drugim riječima: jezična bi analiza pomogla razrješavanju tih pitanja. Zato je 1968. godine u Institutu za lingvistiku Filozofskoga fakulteta (Zagreb) koncipiran znanstveni projekat *Jezik Marka Marulića*¹¹).

Počelo se s ručnim ispisivanjem i konkordiranjem cjelokupne građe iz Marulićeve *Judite*¹²).

Budući da je Ž. Bujas — kao što je rečeno — radio u Americi na kompjutorskom konkordiranju nekih djela hrvatske književnosti, poslao sam mu prema ranijem dogovoru Marulićevoj *Susanu*. Na taj su se način dva Marulićeva djela počela priređivati za konkordiranje: *Judita* ručno u Zagrebu i *Susana* kompjutorski u Texasu.

Kad sam po povratku Ž. Bujasa u Zagreb doznao koje je on sve predradnje učinio za kompjutorsko konkordiranje i koliko je malo vremena za to utrošio, napustio sam ručno ispisivanje ostalih Marulićevih djela (npr. pjesama i poslanica) kao neekonomično i dogovorio se s njim da zajednički nastavimo kompjutorskim sistemom konkordirati ne samo Marulićeva djela nego i ostala djela starije hrvatske književnosti. Tako smo u okviru rada Instituta (sada: Zavoda) za lingvistiku predložili novi projekt pod naslovom *Kompjutorska analiza*

⁸ Antun Đamić: *Nekoliko priloga za poznavanje Marulićeve poezije*, Hrvatsko kolo, sv. 3, Zagreb, 1950.

⁹ Vjekoslav Štefanić: *O izdanjima Marulićeve Judite* (u knjizi: Marko Marulić, *Judita*, izd. Zora, Zagreb, 1950, str. 151—159).

¹⁰ Carlo Verdiani: *Il codice Dalmatico-Laurenziano*, Ricerche slavistiche, sv. 5, 1957; Isti: *Prose e versi inediti di Marco Marulo nel codice Dalmatico-Laurenziano*, Ricerche slavistiche, sv. 6, 1958.

¹¹ Nosilac projekta: Milan Moguš.

¹² Treba ipak reći da prvenstvo za ovakav postupak pripada Antunu Đamiću, bibliotekaru Odsjeka za klasičnu filologiju Filozofskoga fakulteta u Zagrebu. Naime, za vrijeme održavanja Znanstvenoga sastanka o Marku Maruliću u prosincu 1974. godine, kad se pojavila kompjutorska konkordancija Marulićevih djela, otkrio je kolega Đamić da je još prije 35 godina ispisivao na listiće i abecedirao svaku riječ iz djela koja je smatrao Marulićevim. Zanimljivo je da se A. Đamić prihvatio toga posla iz istih pobuda kao i ja, tj. da se totalnim ispisom onoga što jest Marulićevo pomogne odgonetati pripada li Maruliću i ono što mu se pripisivalo. Šteta što je taj Đamićev rad ostao u rukopisu i što se za nj nije znalo sve dosada.

tekstova stare hrvatske književnosti¹³). Projekt je prihvatio Savjet za naučni rad SR Hrvatske i uvrstio ga u svoj 5-godišnji plan (1971—1975). U tom je razdoblju trebalo elektronički obraditi određeni broj djela hrvatske pismenosti i književnosti od Bašćanske ploče do ilirskoga preporoda.

Iz prvog, hrvatskoglagoljskog razdoblja do 1300. godine konkordiran je najprije *Razvod istarski*, glagoljicom pisani tekst čiji je prvi sloj nastao oko godine 1275. U priređivanju glagoljaških tekstova za kompjutorsku obradu iskrsavaju dvije teškoće: jedna osnovna, druga posebna. Osnovna je transliteracija s gdjekad vrlo nečitke glagoljice u latinicu, a posebna s latinice u kompjutorsku abecedu, tj. u znakovlje koje ima tastatura pojedinog kompjutora, ali tako da svaki znak (pa i najmanji) označava uvijek samo jedno.

Kod ostalih, neglagoljskih, djela starije hrvatske književnosti bilo je drugih problema. Tako, na primjer, pri priređivanju Marulićevih djela za konkordiranje pojavile su se dvije teškoće. Jedna je već spomenuta: odrediti autentičan Marulićev opus. Na temelju dosadašnje stručne literature i vlastitog istraživanja došao sam do zaključka da se u Marulićev korpus mogu ubrojiti u prvome redu *Judita* i *Susana*, zatim dvije poslanice upućene benediktinki Katarini Obirtića i 12 pjesama¹⁴). Takvoj je strogosti u odabiru bila jedina svrha da stvori autentičnu riznicu Marulićeva jezika koja bi, s jedne strane, bila podlogom najraznovrsnijem lingvističkom i lingvostilističkom opisu, a, s druge, da posluži kao dobrodošla i sigurna pomoć onim stručnjacima koji žele pri utvrđivanju Marulićeva autorstva uzeti u obzir i jezične elemente.

Drugi je problem u vezi s priređivanjem Marulićeva korpusa za konkordiranje bio u tome da se utvrdi postoji li raskorak između jezika originala i drugih izdanja. Analize su pokazale da takva raskoraka ima zbog stare nesavršene grafije. Zato nije bilo druge nego na temelju izvornikâ prirediti novo svojevrsno kritično izdanje koje je otisnuto ispred konkordancije Marulićevih djela. Ali takav postupak nije proveden samo kod Marulića. Za potrebe kompjutorske analize priređivao sam nova izdanja i drugih starih pisaca, npr. Karnarutića, Zoranića, Barakovića, Pelegrinovića i Vitezovića, od kojih su neka čak posebno objavljena: Vitezovićeve *Senjčica* u »Senjskom zborniku« (Senj, 1973), trogirski verzija Pelegrinovićeve *Jejupke* u »Mogućnostima« (Split, 1974), a Karnarutićevo djelo *Piram* i *Tižba* u »Čakavskoj riči« (Split, 1976). S tako priređenim tekstovima započeta je kompjutorska obrada.

¹³ Nosioci projekta: Željko Bujas i Milan Moguš. Sve kompjutorske programe za naš projekt izradio je dipl. ing. Milutin Cihlar, rukovodilac sektora za programiranje u Elektroničkom računskom centru Centra za ekonomski razvoj grada Zagreba, komu dugujemo veliku zahvalnost. Zahvaljujemo na pomoći i prof. Maji Bratanić-Čimbur, asistentici projekta.

¹⁴ Tih je 12 pjesama uvrstio i Ivan Slamnig u svoje izdanje Marulićevih hrvatskih djela (»Pet stoljeća hrvatske književnosti«, sv. 4, Zagreb, 1970). To ne znači da se još pokoja pjesma ne bi mogla naći u sastavu Marulićevih hrvatskih stihova, ali »sumnjive« nisam uzimao u obzir.

Svaki se takav, filološki uređen, tekst najprije priprema za posebnu vrstu prepisivanja: programiranim se znakovima obilježava početak i kraj svakog dijela teksta, npr. svaki odlomak proznog teksta, svaki prizor ili čin drame, svaka pjesma u zbirci, ili se čak svaki naslov pjesme odvađa od stihova koji slijede. Osim toga, numerira se svaki redak, odnosno stih. Ovako opremljen tekst prepisuje se na posebnom pisačem stroju, na tzv. fleksorajteru (Flexowriter). Tastatura je toga stroja slična tastaturi običnog pisačkog stroja. Pritiskivanjem na tastere bilježi se tekst dvostruko: na papir običnoga formata koji se pokreće pomoću valjka da se vidi što je prepisano i, što je osobito važno; ubušava na specijalno montiranu papirnu traku tako da svakom slovu odgovara određeni broj ubušenih kružnica poredanih na određen način¹⁵).

Tek sada, kad je tekst ubušen na papirnoj traci, stvorena je mogućnost da bude primljen u kompjutor. Međutim, da bi se ubrzao taj postupak, i to fantastično ubrzao, prenosi se sadržaj papirne trake preko električnog čitača na magnetsku traku jer kompjutor očitava znakove s magnetske trake 200 puta brže nego s papirne (200.000 prema 1.000 znakova u sekundi). Tako očitani tekst kompjutor sprema u tzv. memoriju.

Ono što je kompjutor »zapamtio« u memoriji prenosi se pomoću štampača kao ispis na velike kompjutorske liste. U našem slučaju štampač je ispisivao kompjutorski formirane retke, i to tako da prozni tekstovi imaju najviše 65 znakova (stihovi obično nemaju više od 65 znakova, pa svaki stih stane u kompjutorski redak). Svaki je redak u ovoj fazi rada označen onim brojem koji je predvidio programer.

Takav se kompjutorski ispis uspoređuje s izvornim rukopisom, tj. onim tekstom koji se ubušavao na fleksorajteru. Korektor označava sve pogreške kao kod tiskarskoga sloga. Pogrešna se mjesta ponovo ubušavaju i preslikavaju na magnetsku traku i kao novi ispis vraćaju opet korektoru na novu korekturu. Na taj se način uklanjaju pogreške sve do potpuno pročišćenog teksta.

Kad je potpuno pročišćen tekst prenesen na magnetsku traku, nastupa veliki trenutak konkordiranja. Prema određenim programima kompjutor sâm izdvaja iz teksta — bez obzira na njegovu veličinu — nepogrešivo svaku riječ, dodaje joj s jedne i s druge strane otprilike deset riječi teksta (tako da čitav konkordirani redak ima najviše 132 znaka), slaže nepogrešivo u stupac svaki oblik svake riječi po strogoj abecedi i označava točno redak ili stih gdje se izdvojena riječ, tzv. stožernica, nalazi u tekstu. Vrijeme za taj posao ovisi o veličini teksta. Tako je npr. tekst *Razvoda istarskoga* sa 18.252 riječi konkordiran za 27 minuta, a sva Marulićeva djela sa 35.914 riječi za 83 minute. Isto je toliko trebalo da se konkordirani zapis na magnetskoj traci prenese u obliku otiska na kompjutorsku listu sa 54 retka na svakoj stranici liste. Tako smo za otprilike jedan sat dobili odštampanu knji-

¹⁵ Na jedan metar papirne trake ubuši se otprilike 75 riječi.

gu od 383 stranice konkordiranoga teksta *Razvoda istarskoga*, a za manje od tri sata konkordanciju Marulićevih djela koja ima čak 650 strana. Uz te podatke bilježim jednu napomenu: ni u jednom retku ni na jednoj stranici kompjutorski otisnutoga teksta konkordancije nema tiskarske pogreške (uz uvjet da je nije bilo prilikom ubušivanja na papirnu traku). To je jedan od razloga zašto se kompjutorske konkordancije, dobivene na velikim listama, ne štampaju na klasičan način, nego se umnožavaju preslikavanjem. Da bi se mogle lakše upotrebljavati, preslikavaju se i umnožavaju na posebnom stroju u manji format (tzv. A 4). U želji da se prigodom upotrebe dobije što veća slika podataka, konkordancije su uvezane po svom duljem bridu papira i listaju se odozdo prema gore (a ne zdesna nalijevo). Na taj se način pojavljuju konkordancije kao knjige nekonvencionalna izdanja i uveza, pa predstavljaju novost i u bibliotekarstvu.

Konkordancije mogu biti totalne, selektivne i odostražne. Kod totalnih se konkordira cjelokupan vokabular sa svim potvrđama (takve su sve konkordancije izrađene u Zavodu za lingvistiku Filozofskoga fakulteta u Zagrebu), kod selektivnih se prikazuju samo odabrani podaci (npr. samo leksičke riječi, tj. izostavljaju se kao stožernice veznici, glagolske enklitike i sl.)¹⁶, a kod odostražnih se riječi abecediraju prema slovima od desna na lijevo¹⁷). Pored spomenutih pet konkordancija što ih je Ž. Bujas izradio u Americi, dosada su izrađene u Zagrebu i mogu se koristiti ove kompjutorske konkordancije:

1. Željko Bujas: *Kompjutorska konkordancija Gundulićeva »Osmana«*, sv. 1 i 2, izd. Liber, Zagreb

2. Milan Moguš i Željko Bujas: *Kompjutorska konkordancija hrvatskih djela Marka Marulića*, izd. Instituta za lingvistiku Filozofskoga fakulteta, Zagreb, 1974, str. 1—650.

3. Milan Moguš i Željko Bujas: *Kompjutorska konkordancija Karnarutićevih djela*, izd. Institut za lingvistiku Filozofskoga fakulteta Zagreb, 1975, str. 1—321.

4. Milan Moguš i Željko Bujas: *Kompjutorska konkordancija »Razvoda istarskoga«*, izd. Institut za lingvistiku Filozofskoga fakulteta, Zagreb, 1976, str. I—XVIII i 1—383.

5. Milan Moguš i Željko Bujas: *Kompjutorska konkordancija Krležinih »Balada Petrice Kerempuha«*, Zagreb, 1977, str. 1—310¹⁸).

U koje se svrhe mogu upotrijebiti konkordancije?

To se pitanje često postavlja, ali na nj nije lako odgovoriti zato što nije tako jednostavno predvidjeti sve one neslućene zahvate i

¹⁶ Selektivne, tzv. male konkordancije, mogu poslužiti kao priručnik u svim srednjim školama.

¹⁷ Odostražne su konkordancije pogodne za morfološka i tvorbenia istraživanja. Takvu je konkordanciju Marulićeve *Susane* izradio Ž. Bujas u Americi.

¹⁸ Potpuno su dorađene i čekaju na sâm čin konkordiranja djela Zoranićeva, Pelegrinovićeve, Lucićeva, Hektorovićeve i Benetovićeve. Razlog je čekanju administrativne prirode: prijelaz sa sistema u ERC-u (Elektronički računski centar) na sistem u SRCU (Sveučilišni računski centar).

pothvate u koje se istraživači mogu upustiti sada kad postoje konkordancije. Zato je možda najbolje prikazati bar neke radove koji su izvršeni ili koji bi bogli biti izvršeni na osnovi konkordancija.

Upravo kad se dovršavala kompjutorska konkordancija Marulićevih djela, izašla je knjiga Carla Verdianija pod naslovom *O Marulićevu autorstvu Firentinskoga hrvatskog zbornika iz XV stoljeća* (Split, 1973), gdje je ugledni talijanski slavist podastro razloge koji su ga naveli da kodeksu, što ga je pronašao u biblioteci Mediceo-Laurenziana u Firenzi, pripiše autorstvo Marka Marulića. Stoga je razumljivo što sam s osobitim zanimanjem čitao Verdianijev tekst, pogotovu što mi se pružila izvanredna mogućnost da Verdianijeve ilustrativne primjere iz Firentinskoga zbornika odmah uspoređujem sa svim onim zakonitostima, dakle obrascima koje pružaju distribucije uočljive u okviru bogatstva od 35.914 konkordiranih Marulićevih potvrda. Bio sam, kako rekoh, u mogućnosti da podvrgnem analizi čvrstoću onih Verdianijevih dokaza o Marulićevu autorstvu spomenutoga zbornika koje on — proučivši tekst — temelji na jeziku.

U 3. poglavlju knjige pod naslovom »Neke osobitosti kodeksa« Verdiani ističe da je upotreba tzv. kontaktnih sinonima jedna od prvih odlika teksta zbornika. Taj postupak »postaje karakteristična crta našega pisca, jer se ponavlja od početka do kraja rukopisa« (str. 36). I nižu se primjeri: *devota i goruća molitva, imamo figuru ali priliku, febar ali ognjenica, fundament ali zastava, poklisara ali imbasadura, poklisar ali legat, ki nam tolika doktura i naučitelja poda, zdarži njegovo pristolje ali katridu, činiti mirakuli ali čudesa* itd. Budući da je takav postupak »gotovo stalan«, da se nalazi podjednako u svim dijelovima teksta, talijanski slavist smatra da »moramo pretpostaviti da su sve teme našega kodeksa djelo jednog jedinog autora« (str. 37). Postavlja se sada pitanje: je li ta stilska osobitost koja se provlači — kako kaže Verdiani — »od početka do kraja rukopisa«, karakteristična i za Marulićev jezik.

Vidjeli smo da se kontaktni sinonimi povezuju konjukcijama *i* i *ali* (= ili). Uvidom u konkordanciju možemo se osvjedočiti da Marulić

1. za veznik koji znači »ili« upotrebljava oblike *al*, *ali* i *aliti*, tj. uopće nema veznika u obliku *ili*;

2. da *al* ima 21 potvrdu, *ali* 43 potvrde, *aliti* 1 potvrdu;

3. da se u svih 65 potvrda (*al* + *ali* + *aliti*) nalazi samo jedan primjer upotrebe sinonima gdje je strana riječ objašnjena domaćom. To je bilješka uz 285. stih 6. libra *Judite* koja glasi: »*Kantika ali pisan Juditina*«. Svi ostali primjeri toga veznika imaju pravu rastavnu funkciju: *pakal al raj, u polju al u gori, glavu otkla kako svinji ali kako bravu, čuvši ali vidivši, koliko prude ali ude životu našem* itd.

U vezi s kontaktnom sinonimikom još je zanimljiviji primjer s veznikom *i*. Marulić je upotrijebio riječ *i* preko tisuću puta (točno 1007). Pregledao sam sve primjere (a bilo ih je lako pregledati jer su svi složeni uz stožernicu *i*) i mogu reći da su samo tri slučaja slična onima koje navodi Verdiani.

Po toj, dakle, glavnoj karakteristici, po čestoj i pravoj sinonimici, jezik Marulićev uvelike odudara od jezika Firentinskoga zbornika.

Druga je odlika kodeksa iz Firenze upotreba riječi *sestra* ili *sestrica*. Verdiani kaže: »Riječi *sestra*, *sestrica* dolaze gotovo na svakoj stranici, mnogo češće u jednini uz različite osjećajne dodatke« (str. 32). To je točno. U građi što je prezentirana u drugom dijelu Verdianijeve knjige riječ se *sestrica* nalazi gotovo na svakoj strani, na mnogim čak po dva puta. Istina je i to da se riječima *sestra* ili *sestrica* dodaju razni atributi. To su *poljubljena*, *bogoljubna*, *draga*, *pridraga*, *jedina*. Važno je istaknuti da od deset sintagmi tipa *sestrice* (*sestro*) + *atribut* osam ih ima pridjev *poljubljena*. Odnos je dakle 8 (*poljubljena*) prema 2 (svi ostali atributi). Prema tome: Firentinski se zbornik odlikuje vrlo frekventnom uporabom riječi *sestrica* ili *sestra* uz atribut *poljubljena*. Kompjutorski sređeni podaci omogućuju nam da ustvrdimo dvoje:

1. U Marulićevim djelima slabo su frekventne riječi *sestra* i *sestrica*. *Sestra* ima svega 12 potvrda, a *sestrica* samo dvije. To, dakle, nisu riječi koje Marulić često upotrebljava. Ako se koja riječ nalazi u Marulićevim djelima gotovo na svakoj strani, kao što se prema Verdianijevoj tvrdnji nalazi riječ *sestrica*, onda se njezina frekvencija kreće negdje oko 100 povrda¹⁹).

2. Sintagmu *poljubljena sestrice* upotrebljava autor kodeksa gotovo automatski, a u konkordiranim Marulićevim djelima sa preko 35.000 riječi uopće nema riječi *poljubljena*. I s drugim je atributima gotovo isto: riječi *bogoljubna* također nema, pridjev *jedini* nalazi se samo jedanput u muškom rodu i jedanput u srednjem, ali nijedanput u ženskom (nema eventualno *jedina sestrice* ili *sestro*), dok je za *pridraga* samo jedna potvrda uz imenicu Anka.

Analiza Verdianijevih argumenata nije ovime završena²⁰), ali je vjerojatno i ovo dovoljno da se pokaže kako konkordancije, oslobođene predrasuda i sentimentalnosti, mogu pomoći pri dokazivanju ili, obrnuto, opovrgavanju nečijega autorstva. Stariji period svake književnosti pun je rukopisâ pohranjenih u arhivima čiji su autori ostali nepoznati. Pomnim uspoređivanjem konkordiranih djela poznatih autora s konkordancijama djela čiji autori nisu poznati mogu se prikupiti dokazi koji u velikoj mjeri pomažu razrješavanju pojedinih zagonetki u vezi s autorstvom.

Potpuni leksički inventar u konkordancijama upravo nuka na leksikografske zahvate svake vrste. To se u prvom redu odnosi na izradu rječnika pojedinih autora. U Sovjetskom Savezu imaju, između

¹⁹ Usporedbe radi može se napomenuti da npr. riječ *riba*, koja ne predstavlja nikakvu posebnost u Marulićevim djelima, ima 48 potvrda, tj. 4 puta više nego *sestra*, odnosno 24 puta više nego *sestrica*.

²⁰ Uspor. Dragica Malić: *O Verdianijevu pristupu Firentinskom zborniku*, Forum, sv. 9, Zagreb, 1976, str. 401–424; Milan Moguš: *Je li Marulić autor Firentinskog zbornika*, Radovi Zavoda za slavensku filologiju, sv. 14, Zagreb, 1976, str. 45–51.

ostaloga, Puškinov rječnik²¹), u Italiji Danteov²²), u Engleskoj Shakespeareov²³), a u Poljskoj se izrađuje Mickiewiczev. Na osnovi konkordancija mogli bismo relativno lako izraditi npr. rječnik Marulićev, Gundulićev, Krležin. Ili rječnik samo jednog djela, npr. *Razvoda istarskoga*, Mažuranićeve *Smrti Smail-age Čengića*, Goranove *Jaime*, Krležinih *Balada Petrice Kerempuha* i sl.

Ali, ako momentano nije najpreči zadatak izrada rječnika nekog književnika, svakako je potrebno popunjavati i dorađivati postojeće rječnike. U tom su pogledu neophodne dopune velikog *Rječnika hrvatskoga ili srpskog jezika* Jugoslavenske akademije znanosti i umjetnosti u Zagrebu. Ne radi se samo o dopunama pisaca čija djela nisu iz ovih ili onih razloga do sada ekscerpirana. Riječ je o djelima na osnovi kojih je rađen Akademijin rječnik. Tako, na primjer, za taj su rječnik ispisivana ista Marulićeva djela koja su uzeta u obzir za konkordiranje. Marulićeve su potvrde vrlo često u Akademijinu rječniku ili najstarije ili jedine, radilo se o apelativima ili o vlastitim imenima, pa su, na primjer, da ostanemo na početnom slovu abecede, imena *Agnesina*, *Amonit* i *Atlanta* potkrijepljene samo Marulićevim potvrdama. Ali nasuprot tomu, u ARj. nema imena *Abra*, *Absalon*, *Agaba*, *Agnes*, *Akeloj*, *Akior*, *Alcest*, *Amalek*, *Amon*, *Amoreja*, *Antiok*, *Arceslav*, *Arfaksat*, *Arkadija*, *Astijag*, *Aseur* i *Atropos* iako se sva spominju u Marulićevim djelima. Razmjer je, dakle, uvrštenih i neuvrštenih imena samo kod slova A veoma velik. Istina je da se neka od neuvrštenih imena susreću u Marulićevim djelima jedanput, pa je kod ručnog ekscerpiranja moglo doći do propustâ koji se kompjutoru ne mogu dogoditi, ali je istina i to da su npr. imena *Asuer* i *Atropos* potvrđena u Marulićevim djelima 2 puta, *Amalek* 3 puta, *Amon* 4 puta, *Abra* 9 puta, a *Akior* čak 32 puta, i to sva u *Juditi*.

Nije Akademijin rječnik manjkav samo kad se radi o imenima. Uspoređivanjem s konkordancijama može se vidjeti da ili nema potvrda iz djela koja su za taj rječnik ekscerpirana bila, ili nema svih značenja koja se nalaze u građi, ili se ne donose najstarije potvrde. Tako, da spomenem nekoliko nasumce odabranih primjera, nema pridjeva *kostelski* iako se u *Razvodu istarskom* spominje 3 puta, nema riječi *poni* u značenju *barem*, odnosno *dakle*²⁴) iako je u Marulićevim djelima zabilježena ravno 47 puta (i to u *Juditi* 20 puta, u *Susani* 6 puta, u pjesmama 17 puta, u poslanicama 4 puta), nema prilogâ *vdaleč* i *većkrate* iako je prvi prilog upotrijebio Karnarutić u 228. stihu *Vazetja Sigeta grada*, a drugi u 825. stihu *Pirama i Tižbe*, tako je i s pridjevom *višnji* 'vječni' unatoč dvjema Karnarutićevim potvrdama, nema Karnarutićeva suponiranog infinitiva *većivati se* iako u

²¹ *Slovar jazyka Puškina*, izd. Akademija nauka SSSR-a, Moskva, 1957.

²² *Enciclopedia Dantesca*, tom. I—III, Roma, 1970.

²³ A. Schmidt: *Shakespeare-Lexicon*, vol. I—II, Berlin-Leipzig, 1923; *A Shakespeare Encyclopaedia*, London, 1974.

²⁴ Petar Skok: *Etimologijski rječnik hrvatskoga ili srpskog jezika*, knj. 2, Zagreb, 1972, str. 700.

Vazetju (159. stih) piše: Tri baše vićnici s kimi *se većuje*²⁵⁾, izostao je i suponirani infinitiv *obljetavati* od prezenta *obljetuje* koji se nalazi u 539. stihu 12. pjevanja Gundulićeva *Osmana*²⁶⁾, nema... itd. Za glagol *sitovati* kaže se u ARj. da znači *svjetovati* i navodi jedina potvrda iz Habelićeva rječnika. Međutim, Marulićeva rečenica: »Judita... leže u grob muža svoga; sedam dan ju puk *sitova*« očito pokazuje da glagol *sitovati* znači ovdje *oplakivati*²⁷⁾, a ne *svjetovati*. Za riječ *kuba* u značenju *kupula* navodi se da se javlja tek u 17. stoljeću, a 4 primjera iz *Razvoda istarskoga* pokazuju da je riječ *kuba* upotrijebljena u nas već u 13. stoljeću, dakle četiri stoljeća ranije. Zar je sve što navedosmo mala spoznaja?

Akademijin je rječnik sasvim sigurno »jedan od najvećih leksi-kografskih pothvata u slavenskom svijetu, djelo kakvih nemaju ni mnogo veći narodi i u koje je uloženo preko sto godina napornog i samoprijegornog rada mnogih znanstvenih radnika«²⁸⁾. Upravo je zato potrebno, sistemom dopunâ, uklanjati ono što je propušteno pri njegovoj obradi. U Institutu za jezik u Zagrebu već se ispusuju dopune koje bi imale obuhvatiti sve one novootkrivene nezaobilazne spomenike naše jezične i kulturne prošlosti, osobito starijih razdoblja, koji ne mogu izostati u djelu tako značajne koncepcije²⁹⁾. U tom će pogledu izrađene konkordancije djela starih hrvatskih pisaca biti prvo-razredni pomagač.

Do sada smo razmatrali korist konkordiranih tekstova u odnosu na nekonkordirane, a sada bismo se, također samo ilustrativno, mogli pozabaviti pomoći koju pružaju konkordancije pri analizi konkordiranih tekstova.

Možemo postaviti hipotetično pitanje: kojem dijalekatnom sustavu pripada npr. *Razvod istarski*. Ako postoje kriteriji za određivanje dijalekatne pripadnosti, onda se može provjeriti da li se osobine koje karakteriziraju neko narječje susreću u konkordiranom tekstu u svim primjerima ili samo u većini. Budući da postoje kriteriji kojima se određuje, na primjer, čakavsko narječje³⁰⁾, moguće je vrlo brzo, za svega nekoliko minuta, konstatirati da u korpusu *Razvoda istarskoga* sa 18.252 riječi postoje jedino oblici *ča* i *zač*, *malin* 'mlin', *kadi* 'gdje',

²⁵ Ovaj je primjer i inače zanimljiv zbog različitog refleksa jata (*vićnik*: *većivati se*) što je čest stilem u starijih hrvatskih pjesnika. Na to mi je skrenuo pažnju kolega Josip Vončina.

²⁶ Na ovaj je primjer upozorio Željko Bujas (v. Navedeno djelo, str. 53).

²⁷ To je značenje glagola *sitovati* dobro naveo I. Slamnig u izdanju Marulićevih djela (*»Pet stoljeća hrvatske književnosti«*, sv. 4, Zagreb, 1970, str. 235) upućujući na isti korijen koji se nalazi u riječi *sjeta*. Marulićevu potvrdu glagola *sitovati* u tom značenju navodi i Petar Skok u *Etimologijskom rječniku s. v. sjetiti se*.

²⁸ Dragica Malić: *Uz završetak jednog velikog posla*, Jezik, sv. 5, Zagreb, 1970, str. 145.

²⁹ Dragica Malić: *Navedeno djelo*, str. 151.

³⁰ Božidar Finka: *O čakavskom identitetu*, Suvremena lingvistika, sv. 7—8, Zagreb, 1973, str. 11—15; Milan Moguš: *Fonološki kriteriji za određivanje čakavskoga narječja*, Radovi Zavoda za slavensku filologiju, sv. 13, Zagreb, 1973, str. 23—36.

maša 'misa', va 'u', vazmite 'uzmite', jazikom 'jezikom'. Postoje dakle, prvo, primjeri u kojima se nalaze eminentno čakavske fonološke crte i, drugo, ti se primjeri javljaju jedino u takvu čakavskom obliku. Istraživanje drugih crta samo će potvrditi činjenicu da se tekst *Razvoda istarskoga* može smatrati čakavskim.

Unutar toga problema drugo je pitanje kojem dijalektu čakavskoga narječja pripada *Razvod istarski*. Poznato je, naime, da npr. refleks jata nije jedinstven u čakavaca: jedni su govori ekavski, drugi ikavski, treći ikavsko-ekavski po pravilu Jakubinskoga, a četvrti jekavski. Stanje se u *Razvodu istarskom* — strogo uzevši — ne poklapa ni s jednom od nabrojanih mogućnosti jer u istim riječima nalazimo dvostruki refleks. Svedemo li tu konstataciju, ilustracije radi, na oblike glagola *biti*, možemo na temelju konkordancije za doslovno minutu-dvije s potpunom sigurnošću utvrditi: a) da se u *Razvodu istarskomu* nalaze oblici *behu* i *bihu*, b) da je njihov odnos 30 (*behu*) prema 9 (*bihu*), c) da samo dva oblika *behu* imaju u kasnijem prijepisu varijantu *bihu*, a da čak sedam od devet oblika *bihu* imaju u kasnijem prijepisu varijantu *behu*. Prevaga je, prema tome, ekavske varijante očita u ova dva lika. Ako želimo utrošiti još nešto vremena listajući konkordanciju, možemo se osvjedočiti da postoje dublete *crekva-crikva*, *imehu-imihi*, *imeli-imili*, *imeše-imiše*, *napred-naprid*, *nemški-nimški*, *onde-ondi*, *potreba-potriba*, *stena-stina*, *zapovedi-zapovidati*, *živeli-živili*, *živeti-živiti* i da je u većini tih dubleta ekavska varijanta frekventnija. Supostojanje dubleta upućuje na heterogenost teksta u *Razvodu istarskomu*, na dva sloja od kojih je jedan osnovni. A osnovni je onaj koji je pretežit ili jedini. U dubletama je pretežit ekavski oblik, a jedini je u *beli*, *bežali*, *breg*, *brest*, *brežac*, *cesta*, *del*, *delati*, *deva*, *drevo*, *lepo*, *leto*, *mera*, *mesto*, *naposred*, *navestiti*, *obed*, *odnet*, *ponedela*, *pred*, *prek*, *reka*, *seć*, *sedeti*, *sediniti*, *smeti*, *sused*, *večni*, *verno*, *vera*, *videti*. Samo dvije imenice, *misec* i *vrime*, javljaju se isključivo u ikavskoj formi. Kao što vidimo, za nekoliko je minuta skica ove fonološke pojave bila gotova, a osnova te skice, tj. podaci, potpuni su iako korpus ima 18.252 riječi.

Naravno, dublete ne moraju pokazivati samo raznovremenske slojeve nekoga teksta. One mogu biti i namjerno unesene zbog stilističkih razloga, ali i zbog kojih drugih čije opravdanje danas, nakon nekoliko stoljeća, nije moguće sasvim dokučiti. Već smo kod Karnarutićeva primjera: Tri baše *vićnici* s kim *se većuje* (Vazetje Sigeta grada, 159. stih) vidjeli da oblike s različitim refleksom jata zna iskoristiti, kao i drugi stari hrvatski pjesnici, u stilističke svrhe; takvi su, naime, stilistički klišeji poznati. Nasuprot tomu, teže će biti bez dublje analize objasniti zašto Karnarutić upotrebljava *vira* i *vera*, *virna* i *verna*, *virno* i *verno*, *viruj* i *veruj*. Međutim, podatak da Karnarutić upotrebljava dublete samo tih oblika glagola *virovati/verovati* možemo doznati iz konkordancije za dvadesetak sekundi iako se radi o djelima čiji opseg prelazi 18.000 riječi.

Ali »igra dubletâ« nije specifična samo za pjesnike starijega razdoblja. Uzmemo li u ruke Krležine *Balade Petrice Kerempuha*, možemo se lako uvjeriti da čitav tekst upravo vrvi dubletama. O tome

sam uostalom i sâm nešto pisao³¹). Međutim, postojanje konkordancije toga djela omogućuje barem dvoje: prvo, da se izradi kompletna slika dubleta, tj. da se jednostavno utvrdi koji sve oblici imaju dubletni odnos i kakav je taj odnos i, drugo, stoje li dubletne riječi u istim kolokacijama ili ne jer okolina može utjecati na izbor dublete i onda je taj izbor stilogen. Tako se sada može s potpunom sigurnošću reći da npr. postoje dublete *dan* i *den*, odnosno *danas* i *denes* (nema hipotetičnog *danes*³²), da je *dan* frekventnije od *den* (34:5), ali da je *denes* frekventnije od *danas* (11:2). Slična je situacija s dubletom *maša* i *meša* 'misa' (7:1), ali je više izvedenica i češće su od manje frekventnog oblika *meša* negoli od frekventnijeg *maša*. Dakako, mnoge riječi nemaju ovakvih dubleta iako bi ih — teoretski — mogle imati, npr. *deska*, *melin* 'mlin', *pes*³³) i dr. Ali ako *pes* nema dubleta sa *pas*, ima sa — *cucek*. I tu bi već moglo početi ono stilističko »fino tkanje« jer su — zahvaljujući konkordanciji — svi *pesi* i *cucki* na okupu. Pokazuje se, naime, da su kolokacije tih dubleta ipak ponešto drugačije. *Cucek* je *lovni* i *vern*, a *pes* je *stekli*, *peklenski* i *kervavi*. Razlika je sasvim očita. A da se uoči ta razlika u masi od 2.825 stihova (ili 15.041 riječi), treba manje od pola minute.

Često se postavlja pitanje: ima li smisla navoditi sve potvrde svih veznika, enklitika, prijedloga i sl. Već smo se mogli osvjedočiti da je upravo totalni ispis veznika *i* u Marulićevim djelima (1007 potvrda) bio jedan od protudokaza za njegovo autorstvo Firentinskoga zbornika. Totalni su ispisi veoma pogodni za sintaktička istraživanja. Tako, na primjer, grupiran na jednome mjestu oblik *da* iz Krležinih Balada omogućava da se, pošto se odvoji prezent *da* od veznika *da*, odmah uoči različita sintaktička upotreba veznika *da* u rečenicama kao što su: tak je na svetu *da* za najvekšeg suca smert se s kosum vre oko smuca; — a zakaj *da* fkradne kozu kak potepuh i lola, ak more vernut za trispet ranjčkov vola?; a *da* bi ga oprali od smertnoga greha, z terbuha su mu zrezali za orgule meha; — dvojmba ga vu mozgu muči *da* Biškup je Sardanapal; ...znutra vočje lakom ftič, kak stvorili *da* ga je hudič; ... nekom rastolnačiti: *da* nemre kmet plaćati kad mu penez ni; — Čuli su se glasi kak z terbuha basi: Vivat Gašoar, v nebe dobre došel *da* si! Itd.

Slično je i s veznikom *i*. Dakako, kad bismo trebali rukom ispisivati svaki *i* koji se nalazi u obilju od 15.041 riječi da bismo tek onda vidjeli postoji li kakva zakonitost kod upotrebe te konjukcije, zacijelo bismo se teško latili takva posla. Međutim, kad je umjesto nas taj mukotrpan posao obavio kumpjutor, i to obavio savršeno točno ne preskočivši ni jedan veznik *i*, kad dakle imamo, kao na tanjuru, sve Krležine rečenice iz *Balada* u kojima se nalazi veznik *i*, možemo žrtvovati minutu-dvije za zavirimo u to vezivo. A ono između ostalog poka-

³¹ Milan Moguš: *Što pokazuju Krležine dublete u Baladama?* — Forum, sv. 9, Zagreb, 1973, str. 376—386.

³² Tako je i u organskim dijalektima kajkavskoga narječja: u jednim se dijalektima govori *den* i *denes*, u drugima *dan* i *danas*, ali ni u kojem *danes*.

³³ Nema *pas* nego samo *pasoglavci*.

zuje da postoje dvije strogo odvojene semantičke grupe i da je Krležino vezivanje, podređeno isticanju podvojenosti između tih dviju grupa, organizirano tako da se uvijek spajaju samo istoklasne jedinice, a ne raznoklasne. Tako se na jednoj strani nalaze veliki i zmožni, poglaviti i velemožni, veleučeni i poglaviti, špani i kapetani, biskupi i stožernici, a na drugoj pobožni i podložni, kokošar i tat, kramari i kukavci, grintavci i slinavci, cafute i fakini, fačuk i smujin sin.

Nema povezivanja predstavnika jedne grupe s predstavnicima druge. Takav podatak može u nekoj analizi nešto značiti³⁴). Do njega se može doći zahvaljujući upravo potpunom ispisu gdje momenat slučajnosti otpada.

Kad već imamo konkordanciju u rukama, lako je tragati za podacima koji nam svakoga časa mogu pasti na um. Na primjer: koje sve brojeve od 1 do 10 upotrebljava Krleža u *Baladama* i koja je njihova frekvencija. Vrlo se brzo mogu dobiti ovi precizni podaci:

- a) da su neparni brojevi frekventniji od parnih,
- b) da brojevi 6, 8 i 10 uopće nisu upotrijebljeni,
- c) da za broj 3 postoje 54 potvrde, za 1 (u značenju *jedan* i *neki*) ima 46 potvrda, za 5 ima 7 potvrda, za 7 ima 4 potvrde, za 2 su 2 potvrde, a za 4 i 9 jedna potvrda. Ovakav redoslijed, s brojem 3 na čelu kao prvim samostojećim brojem, odgovara onomu što susrećemo u našoj narodnoj i umjetničkoj književnosti³⁵).

Mogli bismo tako nabrajati i povezivati koliko je god potrebno, što god hoćemo, čime god hoćemo, do podataka iz konkordancije dolazimo uvijek vrlo, vrlo brzo i sasvim, sasvim točno. Važno je samo jedno: znalački povezivati podatke. Postojanje konkordancija omogućuje nam da svoje snage usmjerimo na stvaralački dio posla, a ne da ih iscrpimo u beskrajnim tehničkim pripremama. Od naše umješnosti, dakle, zavisi koliko ćemo ono što nam za filološka i druga istraživanja pruža sada u obliku kompjutorskih konkordancija suvremena tehnika — a to je fantastično brzo dobiven i nadasve točan podatak — znati upotrijebiti za znanstveni napredak. U tom kreativnom naporu stroj niti može niti treba da zamijeni čovjeka.

ON THE ORIGIN OF THE COMPUTER CONCORDANCES AND THEIR APPLICATION IN TEXT ANALYSIS

In this paper the author presents the historical background of the work on computer concordances produced in the Institute of Linguistics, Faculty of Philosophy, University of Zagreb, and the advantages of their application for the exact analysis of literary texts: ranging from proving the authorship of a certain work to detecting subtle stylistic variations.

³⁴ Zanimljivo je da su upravo takva vezivanja karakteristična za hrvatsku srednjovjekovnu književnost. Na »duboku srodnost, izraženu i sadržajnu, misaoni i osjećajni svijet, koji tu davnu glagoljašku poeziju prisno veže s Krležinim *Baladama* Petrice Kerempuha« upozorio je Radoslav Katičić u članku *Zapisi s izvorišta* (*»Slovo«*, br. 25—26, Zagreb, 1976, str. 393—406).

³⁵ Anglist Željko Bujas mi reče da u engleskom jeziku frekvencija brojeva od 1 do 10 odgovara aritmetičkom redu, tj. od 1 dalje.