

## FREKVENCIJSKA ANALIZA ENGLESKOG JEZIKA ELEKTRONIČARSKE STRUKE

SANJA ČURKOVIĆ KALEBIĆ  
Filozofski fakultet u Zadru  
*Faculty of Philosophy in Zadar*

UDK/UDC: 802.0:801:510  
Stručni članak  
*Professional paper*

Primljeno  
: 1990-12-12  
*Received*

Nakon što su spomenuti osnovni principi lingvističke statistike i poznate frekvencijske analize u članku se dalje govori o rezultatima frekvencijske analize engleskog jezika elektroničarske struke. Posebno se obrađuju podaci koje daje frekvencijska lista, a odnose se na raspodjelu frekvencija, rang i kumulativnu frekvenciju te odnose između gramatičkih i leksičkih jedinica. Ovi podaci se zatim uspoređuju sa podacima dobivenim analizom takozvanog "Brown Corpora" i "A Word-Frequency List of Scientific English" ("Popis frekvencije riječi znanstvenog engleskog"). U članku se također ponovo ispituju neke od zakonitosti lingvističke statistike.

Prisutnost matematičkih metoda u lingvistici obično se naziva matematička lingvistika. Matematička lingvistika kvantitativnog tipa ili lingvistička statistika je jedan od postupaka matematičke lingvistike kojim se utvrđuju kvantitativni odnosi među jezičnim jedinicama s pomoću statističkih metoda. Pojmovi kojima operira lingvistička statistika preuzeti su iz opće statistike. Neki od tih pojmova su *frekvencija*, tj. broj ponavljanja nekog svojstva kod članova zadanog skupa, *distribucija* ili raspodjela pojedinih frekvencija te *rang* odnosno položaj pojedinog člana skupa s obzirom na njegovu frekvenciju u odnosu na druge članove toga skupa.

Lingvistička statistika formulirala je veći broj zakona od kojih su sigurno najpoznatiji *Zipfovi zakoni*, nazvani tako prema Georgeu Kingsleyju Zipfu, tvorcu ove metode u lingvistici. Jedan od Zipfovih zakona kaže da je odnos između ranga i

frekvencije riječi obrnuto proporcionalan tako da je njihov produkt otprilike konstantan. Prema drugom Zipfovom zakonu aproksimativno je konstantan produkt kvadrata frekvencije neke riječi i broja riječi s tom frekvencijom:  $f^2 \cdot n = c$ . Također je interesantan zaključak lingvističke statistike prema kojem prvih 15 riječi po rangu u nekom jeziku predstavlja otprilike 25% nekog ostvarenog teksta tog jezika, prvih 100 riječi pokriva 60%, prvih 1000 pokriva 85%, prvih 4000 - 97,5%, a sve ostale samo 2,5%.

Lingvistička statistika mnogo je doprinijela egzaktnosti pojedinih istraživanja. Utvrđivanje najfrekventnijih riječi omogućava stvaranje *bazičnih rječnika* u kojima su za svaki jezik skupljene najčešće riječi. Ovakvi rječnici mogu se upotrebljavati u učenju jezika.

Do sada je u svijetu i u nas izvršen veći broj analiza frekvencijskih svojstava vokabulara. Među onima u svijetu najpoznatija je ona pod nazivom "Standard Corpus of Present-day English" ili "Brown Corpus". On predstavlja najpoznatije istraživanje frekvencijskih svojstava vokabulara općeg engleskog jezika. Autori toga korpusa, profesori W.Nelson Francis i Henry Kučera sastavili su korpus od oko milijun riječi vezanog teksta američke proze na engleskom jeziku, objavljene u periodu od 1961. do 1963.god. Korpus je podijeljen u 15 područja i ima 500 uzoraka sa po oko 2000 riječi. Jedan dio Brown Corpusa (u daljnjem tekstu BC) poslužio je Arne Zettersenu da izradi "A Word-Frequency List of Scientific English"(WFLSE). U osnovi tog popisa je područje J.-Learned and Scientific Writings koje sadrži riječi općeg stručnog engleskog jezika.

U obje ove analize riječi su rangirane prema broju pojavljivanja u korpusu tako da su rezultati ovih radova zapravo rang-liste frekvencije riječi za pojedine oblasti. I u BC i u WFLSE, dakle i u općem i u stručnom tekstu uočavamo iste zakonitosti. Naime, na vrhu dvije rang-liste nalaze se funkcionalne ili gramatičke i neleksičke jedinice kao npr. članovi, prijedlozi, veznici, modalni glagoli, zamjenice i brojevi. Prva leksička jedinica javlja se u BC na 53. mjestu (said), a u WFLSE još kasnije tj. na 65. mjestu (used).

Posljednjih godina u našoj zemlji je izvršeno nekoliko frekvencijskih analiza jezika struke. Primjenom metoda lingvističke statistike na korpus stranog jezika struke ustanovljene su najfrekventnije jedinice unutar jezika određene struke. Na temelju tih jedinica sastavljeni su minimalni Vokabulari (MV) jezika struke. Tako je npr. Milica Gačić izradila MV engleskog jezika krivičnih disciplina, Svetlana Leković izradila je MV engleskog jezika prostornog planiranja, a Istočnica Babić je sastavila MV stručnog engleskog jezika građevinske struke. Valjanost ovih vokabulara provjerena je na tekstovima tipa određene struke. Rezultati provjere pokazali su da oni pokrivaju više od 90% teksta pripadajuće im struke.

U ovom radu bit će riječi o frekvencijskim analizama koje sam vršila na

korpusu engleskog jezika elektroničarske struke. Praktični lingvistički cilj vršenja ovih analiza i ovdje se ogledao u izradi MV jezika struke. Provjerom efikasnosti MV elektronike koji broji 1772 jedinice ustanovljeno je da u potpunosti zadovoljava zadane kriterije jer pokriva 96% teksta elektroničarske struke pisanog na engleskom jeziku. Međutim na ovom mjestu neće biti riječi o praktičnom lingvističkom cilju ovog rada, dakle o MV elektronike, već o teoretskom lingvističkom cilju koji je sadržan u otkrivanju zakonitosti leksičke strukture jezika elektronike. Usporedbom ovog jezika s općim i općim stručnim engleskim jezikom ustanovit će se da li on ima značajke prirodnih odnosno stručnih jezika. Također će se ispitati valjanost nekih zakona lingvističke statistike na korpusu elektronike.

Korpus engleskog jezika elektronike (u daljnjem tekstu KEL) sastavljen je odabirom uzoraka iz sedam reprezentativnih udžbenika američke provenijencije. Ti udžbenici sadržajem pokrivaju Nastavni plan studija elektronike u Splitu iz 1987/88. Broj i veličina uzoraka određen je na temelju učešća pojedine knjige u NPP studija. Korpus zadane veličine od 50.000 riječi teksta ukucan je na kompjutor, presnimljen na disketu i poslan na kompjutorsku obradu u SRCE u Zagreb gdje je za potrebe analize korpusa struke izrađen paket program pod nazivom SAT (Sistem za analizu teksta). Ovakva obrada korpusa daje tri vrste ispisa: konkordancijske liste, abecedni popis cijelog korpusa i frekvencijski popis, tj. rječnik cijelog korpusa. Pošto je za potrebe analiza o kojima će ovdje biti riječi najviše korišten frekvencijski rječnik potrebno je nešto reći o podacima koje on daje.

Frekvencijski rječnik cijelog korpusa daje 4 podatka za svaku različnicu. Podaci su izneseni u stupcima. Prvi stupac odnosi se na rang riječi, tj. na položaj određene različnice u ukupnom broju različenica s obzirom na njezinu frekvenciju. Npr. određeni član ima najveću apsolutnu frekvenciju (4695) i prvi je na rang-listi. Što je apsolutna frekvencija manja različnica se nalazi niže na rang listi i kažemo da ima veći rang. Različnice imaju isti rang ukoliko imaju istu frekvenciju. Npr. različnice *frequency* i *from* imaju rang 22 jer obje imaju frekvenciju 218. Međutim jedinica koja slijedi ima rang 24 (a ne 23) jer su na listi dvije jedinice s rangom 22. Podaci u drugom stupcu frekvencijskog rječnika odnose se na broj pojavljivanja određene različnice u ukupnom korpusu, dakle na njezinu apsolutnu frekvenciju. Tako npr. jedinica *the* ima najveću apsolutnu frekvenciju koja iznosi 4695. Jedinica *frequency* ima apsolutnu frekvenciju 218, a jedinica *resistance* ima apsolutnu frekvenciju 48. U trećoj koloni iznosi se relativna frekvencija različnica (na tri decimale). Npr. *the* pokriva 9,225 % teksta, *frequency* 0,428 %, a jedinica *resistance* 0,094 %. Kako opada apsolutna frekvencija različnica opada i njihova relativna frekvencija. Tako npr. jedinica *circuit*, čija je apsolutna frekvencija 133 ima relativnu frekvenciju 0,261 %, dok jedinica *message* čija je

apsolutna frekvencija 46 ima relativnu frekvenciju 0,090 %. U četvrtom, posljednjem stupcu nalazi se kumulativna relativna frekvencija. Ovaj podatak nam kazuje koliki postotak teksta je pokriven određenim brojem različenica. Tako npr. prvih 10 različenica pokriva 28,271 % teksta, prvih 100 različenica pokriva 50,668 % teksta, a prvih 1000 različenica pokriva 82,210 % teksta. Ako usporedimo dobivene podatke s naprijed iznesenim zaključcima lingvističke statistike u vezi pokrivenosti teksta određenim brojem različenica vidjet ćemo da su odstupanja neznatna. Podaci o kumulativnoj frekvenciji bili su dragocjeni kod odlučivanja o jedinicama koje je trebalo isključiti iz MV.

Koristeći podatke koje daje rang lista može se također izračunati da u korpusu od 50.889 pojava ima 5279 različenica. Dakle, na nešto manje od 10 pojava, točnije 9,66 dolazi 1 različenica.

Promatrajući podatke koji su izneseni u opisanim stupcima rang-liste javlja se potreba da se odgovori na nekoliko pitanja koje bismo mogli formulirati na sljedeći način:

1. Kakva je raspodjela apsolutnih frekvencija u korpusu?
2. Kakva je pokrivenost teksta raznim kumulativnim rangovima?
3. Kakav je odnos gramatičkih i leksičkih jedinica?

Da bi se odgovorilo na prvo pitanje bilo je potrebno načiniti pregled raspodjela frekvencija u korpusu. Ovaj pregled dobijen je tako da su prvo izbrojane sve različenice s istom frekvencijom, a potom se iznio njihov postotak i kumulativni postotak. Raspodjela određene frekvencije izračunata je tako da je određena frekvencija pomnožena s brojem različenica koje imaju istu frekvenciju. Npr. frekvenciju 218 imaju dvije različenice *frequency* i *from*. Raspodjela ove frekvencije iznosi  $218 \times 2 = 436$ . Pošto iz objektivnih razloga ovdje nije moguće iznijeti kompletan pregled raspodjela frekvencija u korpusu, tabelarno ćemo dati skraćeni prikaz odnosa frekvencija i broja različenica s istom frekvencijom. Slovom f označena je apsolutna frekvencija, a slovom H broj različenica s istom frekvencijom.

| f          | H        |
|------------|----------|
| 4695 - 234 | 1        |
| 218 - 17   | 1 - 27   |
| 16 - 8     | 33 - 100 |
| 7          | 115      |
| 6          | 177      |
| 5          | 211      |
| 4          | 294      |
| 3          | 470      |
| 2          | 812      |
| 1          | 2248     |

Iz skraćenog prikaza raspodjele frekvencija možemo zaključiti da uvijek postoji samo jedna različnica s određenom visokom frekvencijom. Opadanjem frekvencija povećava se broj različenica s istom frekvencijom.

Ovdje se također može provjeriti jedan od važnijih zaključaka lingvističke statistike koji kaže da je približno konstantan produkt kvadrata frekvencije nekih riječi i broja riječi s tom frekvencijom. Uvrstit ćemo podatke o raspodjeli frekvencija u formulu  $f^2 \cdot n = C$  i dobiti sljedeće rezultate:

$$7^2 \cdot 115 = 5635$$

$$6^2 \cdot 177 = 6272$$

$$5^2 \cdot 211 = 5275$$

$$4^2 \cdot 294 = 4704$$

U navedenim slučajevima produkt je prilično konstantan (varira od 4704 do 6272) pa je i ovo jedan dokaz da KEL ima značajke prirodnih jezika.

Podaci u raspodjeli frekvencija u KEL pokazuju da od 5280 različenica ima 2248 različenica s frekvencijom 1 / "hapax legomena" ili 42,575 % od ukupnog broja različenica.

U BC broj "hapax legomena" iznosi 22543 ili 44,72 % od 50.406 različenica.

Pošto je ovaj rezultat blizu rezultata dobivenog za KEL pokazuje nam da se i ovdje radi o stalnoj jezičnoj zakonitosti.

Podaci o raspodjeli frekvencija i zaključci koje smo na temelju njih izveli upućuju nas na drugo pitanje koje smo naprijed postavili, a u vezi pokrivenosti teksta raznim kumulativnim rangovima. Naime, ustanovili smo da različnice s frekvencijom 1 predstavljaju nešto manje od polovice od ukupnog broja različenica. Međutim, podaci o kumulativnoj frekvenciji koje nalazimo u rang-listi pokazuju nam da različnice s frekvencijom većom od 1 pokrivaju 95,574 % teksta, iz čega proizlazi da različnice s frekvencijom 1 pokrivaju samo 4,426 % teksta. Zaključujemo, dakle, da "hapax legomena" ne predstavljaju onaj dio vokabulara koji je bitan za razumijevanje elektroničarskog teksta, pošto pokriva mali dio korpusa.

Iz podataka o kumulativnoj frekvenciji možemo također vidjeti da prvih 100 različenica na rang listi ili 1,8943 % od ukupnog broja različenica pokriva otprilike polovinu teksta tj. 50,668 %. Postotak pokrivenosti iza prvih 100 riječi raste sporije, iza prvih 1000 riječi on iznosi 82,210 %, a iza prvih 2000 različenica on iznosi 91,382 %. Ovaj posljednji podatak bio je značajan za izradu MV pri kojoj se kao kriterij eliminacije jezičnih jedinica uzima postotak pokrivenosti veći od 90 %.

Podaci o kumulativnoj frekvenciji u KEL mogu se usporediti s odgovarajućima u BC. Usporedba je prikazana u sljedećoj tablici:

| KR  | Kf/BC  | Kf/KEL |
|-----|--------|--------|
| 1   | 6,899  | 9,225  |
| 10  | 24,256 | 28,271 |
| 500 | 61,927 | 72,109 |

Primjećujemo da postoji razlika u pokrivenosti teksta između BC i KEL, međutim ta razlika nije velika.

Postoji razlika u pokrivenosti prvom različnicom (određeni član *the*). Kumulativna frekvencija ove različnice u KEL veća je od one u BC. Razlog učestalije upotrebe člana u elektroničarskim tekstovima nalazimo u većoj egzaktosti stručnih tekstova u odnosu na opći jezik BC. Razlika u pokrivenosti postoji i kod drugih rangova zbog toga što KEL ima manje različnica od BC pa ih stoga lakše i pokrije.

Podatak da prvih 100 različnica pokriva oko 50 % teksta nameće nam potrebu da utvrdimo koje su to različnice. Pretpostavljalo se da će se, kao i kod dosadašnjih frekvencijskih analiza engleskog i drugih, na samom vrhu rang liste nalaziti neleksičke jedinice. Pretpostavka je potvrđena konstatacijom da su prvih 16 različnica u KEL gramatičke jedinice. To su:

| R  | Različnica | R   | Različnica |
|----|------------|-----|------------|
| 1. | the        | 9.  | be         |
| 2. | of         | 10. | that       |
| 3. | a          | 11. | are        |
| 4. | is         | 12. | as         |
| 5. | to         | 13. | this       |
| 6. | in         | 14. | by         |
| 7. | and        | 15. | an         |
| 8. | for        | 16. | with       |

Dakle, prvih 16 gramatičkih jedinica na vrhu rang-liste su još jedna potvrda da KEL ima osobine prirodnih jezika.

Pošto se i na vrhu rang-liste BC nalaze gramatičke jedinice izvršena je usporedba ovih rang-lista, odnosno njihovih prvih desetak različnica. Među prvih deset različnica u BC nalazi se 8 različnica koje se nalaze među prvih 10 različnica u KEL. Od ovih različnica postoji manje odstupanje u rang-u. Tako se, npr., *to* u

BC nalazi na 4. mjestu, a u KEL na 5. mjestu, dok se *a* u BC nalazi na 5. mjestu, a u KEL na 3. mjestu. Različnice sa rang-liste BC - *was* i *he* ne nalazimo među prvih deset različenica na rang-listi KEL, na kojoj se do broja 10 nalaze i različnice *for* i *be* kojih nema među prvih deset jedinica na rang-listi BC.

Prisutnost odnosno odsutnost nekih različenica na samom početku tih rang-lista mogla se i pretpostaviti. Logično je da se pri vrhu rang-liste stručnog jezika neće naći lična zamjenica, dok je ona na rang-listi BC na 10. mjestu (radi se o ličnoj zamjenici *he*). Prva zamjenica koja se javlja u KEL jest zamjenica *it* i to na 18. mjestu, dok se ostale zamjenice javljaju znatno kasnije.

Uspoređujući KEL s WFLSE utvrdit ćemo da li postoji, i u kolikoj mjeri, poklapanje između ova dva stručna teksta.

Prvih deset različenica KEL ujedno su i prvih deset različenica WFLSE, pri čemu različnice *the*, *of* i *to* imaju isti rang, dok se rangovi ostalih različenica neznatno razlikuju. Npr., *for* je u KEL na 8. mjestu, a u WFLSE na 9. mjestu.

Nastavljajući pregledavanje rang-lista do 17. mjesta, na kojem se u KEL nalazi prva leksička jedinica, zapažamo da dolazi do manjih odstupanja. Naime, do 17. mjesta u KEL nalazimo i jedinice *are* i *an* kojih nema do tog broja u WFLSE, u kojem do spomenutog broja nalazimo još *it* i *was* kojih nema do tog broja u KEL. Međutim, odstupanja nisu velika, *it* se npr., u KEL nalazi već na 18. mjestu /u WFLSE je na 14. mjestu/, a *are* se u WFLSE nalazi na 17. mjestu. /u KEL je na 11. mjestu/.

Prva leksička jedinica u KEL jest *fig*. Nimalo ne začuđuje što ova skraćenica ima ovako visoku frekvenciju (296) i visok položaj na rang-listi (rang 17), iz razloga što su u knjigama iz oblasti elektronike veoma česte referencije na crteže kojima, zbog prirode ove znanosti, obiluju spomenute knjige.

Iza prve leksičke jedinice od rednog broja 18 do 50 nalazi se 19 gramatičkih jedinica i 14 leksičkih jedinica. Prikazat ćemo ih u sljedećoj tabeli.

| R  | Različnica   | R  | Različnica  |
|----|--------------|----|-------------|
| 18 | <i>it</i>    | 34 | <i>may</i>  |
| 19 | <i>which</i> | 39 | <i>we</i>   |
| 20 | <i>at</i>    | 40 | <i>will</i> |
| 21 | <i>or</i>    | 42 | <i>has</i>  |
| 23 | <i>from</i>  | 45 | <i>two</i>  |
| 25 | <i>on</i>    | 46 | <i>have</i> |
| 28 | <i>can</i>   | 48 | <i>one</i>  |
| 29 | <i>if</i>    | 49 | <i>all</i>  |
| 30 | <i>when</i>  | 50 | <i>than</i> |
| 31 | <i>not</i>   |    |             |

| R  | Različnica | R  | Različnica |
|----|------------|----|------------|
| 22 | frequency  | 36 | used       |
| 24 | signal     | 37 | circuit    |
| 26 | system     | 38 | data       |
| 27 | shown      | 41 | voltage    |
| 32 | output     | 43 | current    |
| 33 | input      | 44 | systems    |
| 35 | t          | 45 | amplifier  |

Dakle, među prvih 50 različenica nalazi se čak 15 leksičkih jedinica. Ove leksičke jedinice pripadaju isključivo stručnom vokabularu, pa je ovo još jedan od dokaza o reprezentativnosti korpusa.

Od rednog broja 51 do 100 broj leksičkih jedinica brže raste na račun gramatičkih tako da leksičkih jedinica ima 29, a gramatičkih 21. I ove gramatičke i leksičke jedinice prikazat ćemo u odvojenim tablicama.

| R  | Različnica | R   | Različnica |
|----|------------|-----|------------|
| 52 | such       | 73  | thus       |
| 53 | these      | 76  | into       |
| 54 | but        | 77  | other      |
| 56 | also       | 79  | since      |
| 62 | its        | 82  | must       |
| 63 | same       | 87  | because    |
| 64 | each       | 91  | any        |
| 65 | more       | 93  | there      |
| 68 | then       | 97  | first      |
| 70 | between    | 100 | only       |
| 73 | thus       |     |            |

| R  | Različnica   | R  | Različnica |
|----|--------------|----|------------|
| 51 | power        | 81 | x          |
| 55 | high         | 83 | use        |
| 56 | time         | 84 | bandwidth  |
| 58 | noise        | 85 | collector  |
| 59 | line         | 86 | order      |
| 60 | transistor   | 88 | response   |
| 61 | carrier      | 89 | control    |
| 66 | transmission | 90 | type       |
| 67 | gain         | 92 | effect     |
| 69 | value        | 94 | amplitude  |
| 71 | function     | 95 | feedback   |
| 74 | phase        | 96 | called     |
| 75 | modulation   | 98 | impedance  |
| 78 | point        | 99 | number     |
| 80 | information  |    |            |

Iz tablica je evidentno da od rednog broja 51 do 100 broj leksičkih jedinica raste u odnosu na gramatičke jedinice. Dok je od rednog broja 1 do 50 bilo 35 gramatičkih jedinica i 15 leksičkih jedinica, dotle u sljedećih 50 riječi ima 21 gramatička jedinica i 29 leksičkih jedinica. Dakle, u prvih 100 riječi ima 56 gramatičkih jedinica i 44 leksičke jedinice. Dakle, među prvih 100 riječi u KEL ima znatno više leksičkih jedinica nego na istom broju riječi u BC i WFLSE. Već smo istakli da se prva leksička jedinica u KEL nalazi na 17. mjestu dok su u WFLSE javlja znatno kasnije, tj. na 65. mjestu. Uzrok ove nepodudarnosti treba također tražiti u duljini teksta, tj. što je tekst dulji neleksičke jedinice se dulje "vuku".

Bilo je interesantno utvrditi kakav je dalje odnos gramatičkih i leksičkih jedinica od rednog broja 101 do 200. Rezultati analize iznose se u sljedećim tablicama.

## GRAMATIČKE JEDINICE R 101-200

| R   | Različnica | R   | Različnica |
|-----|------------|-----|------------|
| 101 | zero       | 154 | would      |
| 106 | some       | 155 | been       |
| 109 | however    | 157 | often      |
| 118 | usually    | 158 | should     |
| 120 | no         | 172 | therefore  |
| 126 | they       | 178 | up         |
| 131 | over       | 179 | where      |
| 137 | through    | 184 | many       |
| 140 | both       | 192 | was        |
| 149 | most       | 193 | while      |
| 152 | their      | 195 | does       |

## LEKSIČKE JEDINICE R 101 - 200

| R   | Različnica  | R   | Različnica      |
|-----|-------------|-----|-----------------|
| 102 | case        | 151 | s               |
| 103 | e           | 153 | v               |
| 104 | frequencies | 156 | change          |
| 105 | low         | 159 | energy          |
| 107 | required    | 160 | amplifiers      |
| 108 | wave        | 161 | diagram         |
| 110 | operation   | 162 | p               |
| 111 | signals     | 163 | small           |
| 112 | using       | 164 | spectrum        |
| 113 | valves      | 165 | bit             |
| 114 | block       | 166 | characteristics |
| 115 | method      | 167 | device          |
| 116 | constant    | 168 | diode           |
| 117 | n           | 169 | figure          |
| 119 | emitter     | 170 | fm              |
| 121 | maximum     | 171 | form            |
| 122 | resistance  | 173 | analysis        |

|     |          |     |                |
|-----|----------|-----|----------------|
| 123 | source   | 174 | characteristic |
| 124 | digital  | 175 | common         |
| 125 | example  | 176 | eq.            |
| 127 | base     | 177 | instruction    |
| 128 | desired  | 180 | c              |
| 129 | equation | 181 | elements       |
| 130 | message  | 182 | f              |
| 132 | process  | 183 | k              |
| 133 | side     | 185 | plot           |
| 134 | large    | 186 | receiver       |
| 135 | obtained | 187 | temperature    |
| 136 | r        | 188 | communication  |
| 138 | applied  | 189 | design         |
| 139 | b        | 190 | necessary      |
| 141 | load     | 191 | pulse          |
| 142 | shown    | 194 | different      |
| 143 | v        | 196 | level          |
| 144 | circuits | 197 | m              |
| 145 | error    | 198 | made           |
| 146 | field    | 199 | program        |
| 147 | file     | 200 | results        |
| 148 | memory   |     |                |
| 150 | possible |     |                |

Iz gornjih tablica vidimo da broj leksičkih jedinica i dalje raste tako da ih od rednog broja 101 do 200 ima 78 dok gramatičkih jedinica ima 22.

*Bibliografija*

1. Istočnica B a b i ć, "Struktura minimalnog stručnog vokabulara građevinske struke za potrebe nastave na građevinskim fakultetima", neobjavljeni magistarski rad, Zagreb, 1984.
2. A.B. C a r l s o n, *Communication Systems. An Introduction to Signals and Noise in Electrical Communication*, Rensselaer Polytechnic Institute, 1986.
3. Milica G a č i ć, "Frekvencijsko-leksička analiza engleskog kao jezika struke krivičnih disciplina (u visokoškolskoj nastavi)", doktorska disertacija, Zagreb, 1982.
4. John L. H i l b u r n, Paul. M. J u l i c h, *Microcomputers, Microprocessors: Hardware and Software Applications*, Tokyo McGraw-Hill, 1983.
5. Henry K u č e r a, W.N. F r a n c i s, *Computational Analysis of Present-Day American English*, Brown University Press, 1967.
6. Svetlana L e k o v i ć, "Frekvencijska struktura engleskog stručnog vokabulara (prostornog planiranja)", neobjavljeni magistarski rad, Dubrovnik, 1986.
7. G.M. M i l l e r, *Modern Electronic Communication*, Monroe Community College, 1978.
8. J. M i l l m a n, H a l k i a s, *Electronic Devices and Circuits*, New York, McGraw - Hill, 1967.
9. J. M i l l m a n, H. T a u b, *Pulse, Digital and Switching Waveforms*, New York, McGraw - Hill, 1965.
10. F.H. R a v e n, *Automatic Control Engineering*, University of Notre Dame, 1968.
11. S h i l l i n g, B e l o v e, *Electronic Circuits: Discrete and Integrated*, New York, McGraw - Hill, 1968.
12. G. W i e d e r h o l d, *Database Design*, Tokyo, McGraw - Hill, 1983.
13. Arne Z e t t e r s e n, *A Word-Frequency List of Scientific English*, Indiana University Press.

*Sanja Čurković Kalebić*: A WORD-FREQUENCY ANALYSIS OF ENGLISH  
IN ELECTRONICS

S u m m a r y

Having mentioned the basic principles of linguistic statistics and the well-known word-frequency analyses this paper goes on describing the results of the word-frequency analysis of English in electronics. It pays a special attention to the data obtained from the frequency list concerning the distribution of the frequencies within the corpus, the rank and cumulative frequencies and the relation between grammatical and lexical words. Those data have been compared to the ones provided by similar analyses of the so-called "Brown Corpus" and "A Word-Frequency List of Scientific English". Some regularities of linguistic statistics have been re-examined as well.