

Dr Dušanka Vuletić

Fakultet za defektologiju Sveučilišta u Zagrebu

O UČESTALOSTI NEKIH ELEMENATA GOVORA

Svaki jezični sustav ima svoje karakteristike s obzirom na semantiku, morfologiju, sintaksu i fonetiku. Neki su im elementi zajednički u svim spomenutim kategorijama čak i kad je riječ o udaljenim jezičnim skupinama. Tako ćemo vokale u velikoj većini slučajeva moći razlučiti od konsonanata i u japanskom, kineskom i drugim dalekim jezicima, ma da nećemo razumjeti ni riječi, niti moći ma i približno odrediti ili samo naslutiti kakva je sintaktička konstrukcija iskaza. Moći ćemo možda prepoznati globalno melodiju i znati je li riječ o pitanju, uskliku, neutralnoj afirmaciji.¹ Moći ćemo prepoznati akcentuirane slogove, grupe slogova pod jednim akcentom. Dakle, riječ je o zvukovnoj realizaciji koju često prevodimo na poznato, na ono što smo navikli slušati u materinskom jeziku. Pri tome zanemarujemo razlike, osobito one sitnije, slično svodimo na poznato.

Naš govorni sustav ima svoje specifičnosti u odnosu prema ostalim sustavima, razlikuje se od njih manje ili više u svim gramatičkim kategorijama. U ovom bismo se radu osvrnuli samo na neke karakteristike hrvatskoga književnog jezika iz područja fonetike. Željeli smo doznati koja je učestalost pojedinih glasova u našem jeziku, odnose između skupina glasova, prosječni broj glasova

u slogu, te prosječni broj slogova i glasova u leksičkim i fonetičkim riječima.

Osobito nas je zanimao prvi dio, to jest glasovna analiza. Naime, takva su ispitivanja već vršena u nas, no na manjem materijalu (Maretić), ili su pak brojeni znakovi, a ne glasovi (Matković, Sinković). Mi smo iskoristili mogućnost koje pruža rad više ljudi pa smo skupili i više materijala na kojem smo s većom sigurnosti mogli donositi zaključke. Kvantificiranje glasova i kategorija glasova može nam, s logopedskog stajališta, pomoći pri dijagnosticiranju stupnja oštećenosti govora kod dislalije.

Osim toga primijetili smo da pri čitanju tekstova ne samo na televiziji i radiju već i u logopedskim ambulancama, veoma često dolazi do zabune između slova i glasova. U logopedskim ambulancama, zbog »hiperkorektnosti« govora, ta se zabuna proteže i na govor bez predodčenog teksta. Tako logoped koji inače posve normalno govori 'radiću', 'bese stre' itd, najednom u uvjetima govorne rehabilitacije počinje lomiti jezik trudeći se da izgovori 't' u 'radit ću', 'z' u 'bez sestre' zamjeravajući pacijentu čak i asimilaciju 'z' u posljednjem primjeru, što je u tom položaju posve normalno.

¹ Branko Vuletić, Ljudskost govora, Zbornik Zagrebačke slavističke škole, Međunarodni slavistički centar SR Hrvatske i Filozofski fakultet u Zagrebu, Zagreb 1973, str. 234.

Krivimo zbog toga prošireno uvjerenje prema kome pišemo kako govorimo, što dovodi do mnogih zabluđa. Naime, govor prevodimo u pismo, u slova prema pravopisnim normama, a zbog navedena uvjerenja pri čitanju — jer naš pravopis nije fonetski — vraćamo pismo u prvobitni govor, već u neku njegovu izmijenjenu, artikulacijski često prilično tešku formu.

Primijetili smo također da se u logopedskim ambulantom često, tobože u funkciji korekcije, akcentuiraju slogovi i riječi koji ne mogu nositi akcenat. I dok nekome, tko je već dobro usvojio govor, takva odstupanja neće nanijeti štete, djeci koja tek usvajaju govor, pružaju se pogrešni modeli, stvara se pogrešan sustav, pa i pogrešni pojmovi.

Poučeni tim iskustvom željeli smo slične pojave spriječiti kod budućih logopeda, za što nam je također pružio priliku ovaj rad pomoću analiziranja leksičkih i fonetičkih riječi.

U toku šk. god. 1973/74. sa studentima Logopedije IV godine Fakulteta za defektologiju izvršili smo djelomičnu analizu hrvatskoga književnog jezika s obzirom na glasove, slogove, leksičke i fonetičke riječi. Ograničili smo se na te kategorije jer smo analizu izvršili na pisanom tekstu, a kako se on po dužini i konstrukciji rečenice razlikuje od govornog izraza, nismo dublje ulazili u daljnju analizu.

Svaki je student odabrao tekst iz beletristike, udžbenika, časopisa ili novina. Osim originalnih tekstova na našem jeziku, uključeni su i prijevodi. Uvjet je bio da tekst bude napisan ijekavicom. Analiza je izvršena na 28 tekstova. U toku rada broježani su podaci prikazivani posebno za svaku stranicu, odnosno za ulomke na

stranicama s više teksta, kako bi kontrola bila lakša. U toku rada izvršene su tri kontrole svih ispitivanih elemenata u svim tekstovima, a u nekima i više. Zbog netočnosti morali smo odbaciti 10 analiziranih tekstova, pa smo za ovaj rad uzeli u obzir samo 18, i to 12 djela iz područja beletristike, (među njima 4 prijevoda), 2 iz udžbenika, 2 novinska teksta, te dva iz časopisa esejističkog sadržaja. Posve je sigurno da su se pogreške u brojenju potkrale i ovdje, no kako je riječ o opsežnom materijalu, vjerujemo da se one kompenziraju.

U svakom djelu slučajno je odabran ulomak i definiran brojem od 5000 slogova s manjim odstupanjima kako se ne bi lomile fonetičke riječi.

Kao jedinica za slog uzimao se jedan sonorni impuls, naravno u zvučnoj realizaciji teksta pa, da ne spominjemo jasne slučajeve, u riječima kao 'lijepo', 'vrijeme', 'riječ', skupinu 'ije' brojili smo kao jedan slog.²

U okviru ulomka definiranog s 5000 slogova izbrojeni su glasovi — ne slova. Pri tome smo uzimali u obzir glasovne promjene u riječi i u susretu riječi koje pravopis ne bilježi, a za što smo naknadno dobili potvrdu u Silićevu i Rosandićevu udžbeniku »Osnove fonetike i fonologije hrvatskog književnog jezika«.³

Tako smo, na primjer, 'n' u 'jedanput' brojili kao 'm', riječ 'protestni' čitali smo 'protesni' analogijom prema 'mjesni'. Skupinu 'ts' kao u 'bratski' bilježili smo kao 'c', a tako i 'ds' u riječi 'odsjek' koju smo čitali 'ocjek'. Riječ 'predškolski' čitali smo 'prečkolški', 'najjednostavniji' kao 'najednostavniji'.

Osim asimilacija i ispadanja glasova uzeli smo u obzir i glasove koji se ne bilježe, pa smo participe kao 'bio',

^{2,3} Ova je analiza izvršena prije negoli je izišao iz štampe udžbenik Silić—Rosandić, Osnove fonetike i fonologije hrvatskog književnog jezika, Školska knjiga, Zagreb 1974, u kojoj smo našli potvrdu za takvu izgovornu normu, a to vrijedi i za ostala načela koja smo uzeli pri brojenju glasova. Možda smo bili nešto radikalniji od autora udžbenika, pa smo i susret glasova T + S brojili kao njihovu asimilaciju C. Na pr. 'predškolski' kao prečkolški, str. 46, te 29, 30 i 31.

'radio' čitali kao 'bijo', 'radijo', a prema tome i druge vrste riječi kao 'čio', 'mio'. Bilježili smo i glas 'dz' — zvučni par afrikate 'c', kao u riječi 'predznak'.

U susretu riječi vodili smo, također, brigu o promjenama koje se zbivaju, a ne bilježe se u pismu. Tako smo, na primjer, 'još danas' čitali kao 'joždanas', bez sestre' kao 'besestre', 'iz svijeta' kao 'isvijeta', 'pjevat ću' kao 'pjevaću', 'našeg grada' kao 'našegrada' itd.

Leksičke su riječi brojene u skladu s razmacima u tekstu.

Za fonetske su riječi uzete skupine pod jednim akcentom, bez obzira na broj slogova. Tako, na primjer, rečenica 'Iskrali su se iz sobe' ima dvije fonetičke riječi, a pet leksičkih.

Kad smo izbrojili u okviru 5000 slogova pojedine glasove, leksičke i fonetičke riječi, rangirali smo glasove prema njihovoj učestalosti, potražili odnose između pojedinih kategorija glasova, te odnose između slogova, leksičkih i fonetičkih riječi.

Analizom su obuhvaćena ukupno 90102 sloga.

Ti su slogovi sastavljeni od 196634 glasa.

Ti se slogovi i glasovi nalaze u 41820 leksičkih riječi.

Fonetičkih riječi našli smo ukupno 29064.

Glasove smo rangirali po učestalosti. Usporedbe radi prikazujemo uz naš rang glasova i rangove koje su našli drugi autori. Poredak se, uz manja odstupanja, uglavnom slaže.

Uz naš rang prikazujemo Maretićev, jer se njegov način brojenja gotovo podudara s našim. Maretić kaže: »Ni u jeziku nijesu svi glasovi jednako obični, već se jedni nalaze češće, dru-

gi rjeđe, pa tako je i s našim glasovima. Hoteći da saznam koliko je koji glas u našem običan, radio sam ovako: iz Vukova prijevoda Novoga zavjeta izabrah deset različnih mjesta, pa u njima bez ikakvoga prekida brojah i bilježah sve glasove koliko se puta koji nalazi. Na svakom od tih mjesta uzeh 1.000 glasova, pa kad je gdje taj broj navršen, tamo već dalje ne brojah, pa makar brojeći stao baš usred koje riječi. Tako za ovaj račun izbrojih 10.000 glasova, a to čini oko 2.000 riječi«, i dalje »Samo se po sebi razumije da u ovakome poslu glasove treba brojiti onako kako se izgovaraju, a ne kako su napisani; dakle, 's njim' treba uzimati kao š njim, — bratski, gradski kao bracki, gracki, — mislio, radio kao mislijo, radijo.«⁴ Za razliku od Maretića, mi nismo vršili asimilaciju po mjestu tvorbe u prefiksima i prijedlozima, jer se ona u današnjem standardnom izgovoru ne vrši. Tako u našem brojenju 's njim' ostaje 'snjim' prema pravilu: »Ni s ni z ne mijenja se u š i ž: a) kada se nađe u složenoj riječi kao završni suglasnik prefiksa (predmetka): sljubiti, razljutiti, raznjihati; b) kada se nađe pred suglasnicima lj i nj u kojima je j dobiveno prema starom jat: posljednji, sljebenik, sljeme, sljepoća, snjegovi, ozljeda.«⁵ Maretić je u svom brojenju razlikovao 'r' kao konsonantski i kao vokalski, što smo mi propustili.

Sličnim su se poslom bavili Matković i Sinković, koji kažu: »Niti u jednom jeziku nije ispunjen uslov vjerojatnosti pojavljivanja i nezavisnosti znakova. Svaki jezik karakteriziran je potpuno određenom statističkom strukturom.«⁶ Kao što i sami kažu, autori su brojili znakove, a ne glasove, i to tako da je slovo 'lj' brojeno kao 2 znaka, to jest 'l' i 'j', a to

4, 9, 10 Tomo Maretić, Gramatika hrvatskoga ili srpskoga književnog jezika, Matica hrvatska, Zagreb 1963, str. 22.

⁵ Težak, Babić, Pregled gramatike hrvatskosrpskog jezika, Školska knjiga, Zagreb 1966, str. 48.

⁶ Matković, Sinković, Teorija informacije, I dio, Sveučilište u Zagrebu, Zagreb 1970, str. 103.

vrijedi i za 'nj' i 'dž', pa ih prema tome u rangu i nema, a ukupan broj glasova 'l', 'n', 'j' promijenjen je (zanimarujemo 'd' i 'ž' jer se veoma rijetko javlja glas 'dž').

Donosimo i ljestvicu koju Škarić preuzima iz prijašnjeg Matkovićeve rangiranja znakova. Uspoređujući prosjek distribucije riječi prema Guiraudu s distribucijom fonema Škarić kaže: »Slični neekviprobabilni odnosi se javljaju u distribuciji glasova (fonema)«,⁷ to jest (slobodno interpreti-

ramo), neki se glasovi javljaju češće, a neki veoma rijetko. I taj je rang načinjen na temelju brojenja znakova, pa kao ni kod Matkovića i Sinkovića nema slova 'lj', 'nj', 'dž'.

Vrijednosti koje donose Matković i Sinković, te Škarić, pretvorili smo u postotke kako bi usporedba bila lakša. Postoci nisu izraženi s jednakim brojem decimala jer smo morali ostati vjerni brojkama koje smo našli u citiranim djelima. Kako tu nije bilo ujednačenosti, ostavili smo i na-

UČESTALOST

Glas (znak)	GLASOVA		ZNAKOVA	
	Naš rang %	Maretić %	Matković %	Škarić %
A	12,0330	10,79	9,60	9,0
I	10,1874	10,35	7,42	7,8
O	9,5943	9,60	7,54	6,9
E	8,9831	10,99	7,70	7,3
N	5,4375	4,57	4,64	4,3
S	4,7001	4,64	4,20	3,9
T	4,6772	4,00	3,67	3,5
R	4,5531	3,88	3,82	3,6
J	4,2627	5,03	4,35	3,9
U	4,0883	4,29	3,64	3,4
M	3,8076	3,67	3,13	2,9
D	3,4312	4,28	3,19	2,8
K	3,3326	3,18	2,98	2,6
V	3,3138	4,02	3,06	2,9
L	2,9227	1,94	3,06	2,9
P	2,7879	2,26	2,04	1,8
Z	1,7423	1,47	1,44	1,3
G	1,7296	1,92	1,66	1,6
B	1,5084	1,73	1,55	1,4
Š	1,1905	1,50	0,86	0,8
Č	1,1097	1,20	0,84	0,8
C	0,8620	0,56	0,67	0,6
Ć	0,7018	0,91	0,49	0,5
Lj	0,6865	0,90	—	—
H	0,6642	0,73	0,65	0,6
Nj	0,6484	0,63	—	—
Ž	0,6301	0,64	0,52	0,5
F	0,2039	—	0,11	0,1
Đ	0,1977	0,31	—	0,2
Dž	0,0107	—	—	—
Dz	0,0027	—	—	—

U Maretićeve rangu iznijeli smo zbroj konsonantskog »r« (3,43% i vokalskog »r« (0,45%).

⁷ 10, ¹² Ivo Škarić, *Kibernetika i jezik, Savremena lingvistika 7—8, Zagreb 1973, str. 23.*

še postotke sa 4 decimale. Treba također istaknuti da su znakovi brojeni pomoću stroja, te su prema tome ušli u konačan zbroj bijeli prostori među riječima i interpunkcija, pa tako u Matkovićevu i Sinkovićevu rangu na bijele prostore otpada 18,17%, a u rangu koji donosi Škarić prema Matkoviću 22,9%, što ne utječe na rang znakova, ali zato utječe na postotak svakoga pojedinog glasa umanjujući ga.

Usporedimo li rangove, primijetit ćemo da nekih velikih razlika nema. Možda ponajviše udara u oči mjesto glasa — znaka 'e' koji se kod Maretića nalazi na prvom mjestu, a u ostale tri ljestvice na prvom mjestu nalazi se 'a' (kao glas i znak). Kod Maretića 'e' je za 0,20% učestaliji od 'a'! U nas se 'e' nalazi na četvrtom mjestu, a 'a' je od njega učestaliji oko 3%. U ostala dva ranga 'e' se nalazi na drugom mjestu.⁸ Ostale su razlike još mnogo manje i, da smo zaokruživali decimale, velik bi ih dio nestao.

Gledajući sve rangove možemo reći da se vokali nalaze na vrhu ljestvice (osim vokala 'u' koji se nalazi među učestalijim konsonantima, a za polovicu je rjeđi od najmanje zastupljenog vokala u navedenoj skupini). Veoma su rijetki 'f', 'd', 'dž' i 'dz'. Maretić čak i ne bilježi 'f' i 'dž', tumačeći: »Glasovi su f, dž veoma rijetki, njihov je decimalni broj manji od 0,01; oba se nalaze ponajviše u tuđim riječima, sasvim rijetko u pravima narodnim (kako su npr. ufati se, fijuhati, otadžbina, srdžba).«⁹ Kako u našem jeziku danas ima mnogo tuđica, povećao se i rang tih glasova u našoj ljestvici. U Matkovićevima možemo registrirati samo povećanje znaka 'f'.

Između te dvije ekstremne kategorije — najčešćih i najrjeđih glasova — opažamo nagli pad postotka između skupine vokala 'a', 'i', 'o', 'e' i konsonanta koji zatim postepeno opa-

daju do dna ljestvice. Ti prvi, najčešći konsonanti jesu (s malim varijacijama u svim ljestvicama) 'n', 's', 't', 'r', slijedi sonant 'j' i vokal 'u', a za njima ostali glasovi.

S logopedskog stajališta zanimljiv nam je posebno rang onih glasova koji su najpodložniji odstupanjima u izgovoru. Bila bi to skupina 's', 'z', 'c', a često uz nju i 'š', 'ž', 'č', 'ć', 'dž', 'đ' — dakle poremećaj glasova koji nazivamo sigmatizmom. Na temelju ranga glasova možemo zaključiti da je kod osobe koja neispravno izgovara 's', 'z', 'c' govor oštećen 7,5%; da dodatna skupina donosi još 4,5%, a nije rijedak slučaj oštećenja obje skupine kod iste osobe, što iznosi 12%. Naravno, ti su postoci samo gruba procjena, jer nije svejedno na koji su način ti glasovi oštećeni.

Rang glasa 'r' govori nam da je kod rotacizma izgovor oštećen 4,5%. Rang je manji negoli kod sigmatizma, no bez obzira na oblik oštećenja, više ometa razumljivost (ili više dekoncentrira slušača), jer ga okolina lakše uočava. Dakle, rangiranje glasova pomaže nam pri ustanovljivanju nekih proporcija, no nije i ne može biti jedini kriterij za određivanje odstupanja.

U daljnjoj analizi potražili smo odnos između konsonanta i vokala; mada u našem jeziku ima samo 5 vokala, našli smo da oni čine 44,8854% govora. Maretić je našao da su vokali zastupljeni u našem jeziku sa 46,47%. Ako iskoristimo njegov podatak o vokalskome 'r', naš bi se postotak povećao na 45,33%, naravno na štetu konsonantskoga 'r'.

Odnos konsonanata prema vokalima jest 1,2278 : 1.

Sve smo zvučne glasove suprotstavili bezvučnima i našli da oni čine 79,77% svih glasova, ili da se na 4 zvučna glasa javlja tek jedan bezvu-

⁸ Uspoređujemo ovdje rang glasova i znakova jer se kod vokala ne zbivaju promjene u zvukovnoj realizaciji koje bi utjecale na njihov rang.

čan. Zvučnost smo definirali po vibriranju glasiljki.

Nazali su zastupljeni s 9,8935%, što je opet zanimljivo za patologiju govora. Ako u govoru nema nazalnosti, znamo da je oštećenje od 10%, dok je pri otvorenoj rinolaliji nazalnost prisutna u 90% govora. Naravno, i tu važe ograničenja i modifikacije koje smo spomenuli uz poremećaje izgovora, pa i u većoj mjeri, jer veći broj činilaca djeluje na sam poremećaj.

Okluzivi su zastupljeni sa 17,4668%, sonanti s 15,7388%, frikativi s 9,1311% i afrikate s 2,8846%.

Dakle, redosljed bi bio: vokali, okluzivi, sonanti, nazali, frikativi, afrikate.

Promatrajući ljestvicu učestalosti glasova i u njoj pojedine skupine, primijetili smo da su samo vokali i dijelom sonanti, kao pojedini glasovi, relativno ujednačeno zastupljeni u svojim skupinama. Ostale su skupine raspršenije. Tako, na primjer u skupini frikativa glas 's' ima veoma visoko mjesto na ljestvici (4,7%), dok je njegov zvučni par 'z' zastupljen sa samo 1,7423%, a najbliži mu je na ljestvici iz skupine frikativa.

Tražeci dalje odnose između glasova i drugih ispitanih kategorija, našli smo da su naši slogovi u prosjeku sastavljeni od 2,1823 glasa. Naravno, raspon je velik; tako Škarić navodi da »Slogovi u hrvatskom ili srpskom jeziku mogu biti sastavljeni od vokala s jednim do pet konsonanata.«¹¹

Leksičke su riječi sastavljene u prosjeku od 4,7011 glasova, što se slaže s Maretićevom opaskom da u 10000 glasova ima otprilike 2000 riječi.

Fonetička je riječ sastavljena od 6,7655 glasova u prosjeku.

Leksička je riječ sastavljena prosječno od 2,1545 sloga.

Nas je više zanimao broj slogova u fonetičkoj riječi, tj. u skupini slogova pod jednim akcentom, koja može biti sastavljena od jedne ili više leksičkih riječi. Našli smo da je u prosjeku leksička riječ trosložna — sastavljena je od 3,1001 sloga. Škarić je klasificirao učestalost fonetičkih riječi (akcenatskih jedinica) prema broju slogova i našao je da su najčešće upravo trosložne, da slijede dvosložne, pa četverosložne, peterosložne, jednosložne, a zatim one veće od pet slogova, pa zaključuje da je najfrekventnija ritmička forma u našem jeziku daktil.¹² Naš prosjek tako dobiva potvrdu o svome realnom postojanju.

Osim rezultata koje smo izložili, pomoću skupljanja podataka postigli smo i to da su studenti naučili pretvarati pisani tekst u govor, prepoznati fonetske promjene, posebno one koje se u pismu ne bilježe. Na taj nas je posao uputio govor što ga slušamo putem sredstava javnih komunikacija i u logopedskim ambulancama. A to je govor profesionalaca od kojih uče drugi. Čini se da ljudi postaju sve osjetljiviji prema tom problemu, jer ne bi inače izišao Silićev i Rosandićev udžbenik u kome su sadržana i pravila čitanja. Da bismo ih aplicirali, potrebna su uvjeravanja, insistiranje, jer ih pismeni prihvaćaju s nevjericom, kao novost ili čak novotariju. Podsjetimo se samo da je i Maretić takav način čitanja — govorni — smatrao toliko normalnim da nije smatrao potrebnim da iznese sva aplicirana pravila, već samo neka, kao ilustraciju. Jer logično je da je prvotno nastao govor, a potom je nastalo pismo. No pismo se mora vratiti u prvobitni govor, a ne u neku treću tvorevinu.

ON THE FREQUENCY OF SOME SPEECH ELEMENTS

Summary

The purpose of this work was to find out to what extent the individual sounds are represented in the Croatian literary language, the range of their occurrence, the relation between the sound groups, the average number of sounds in the syllables, lexical and phonetic words, as well as the relation of the syllables to the lexical and phonetic words. The other one, but not less important aim was to teach the students of logopaedy to transform a written text into the speech, paying regard to the changes that are not recorded in writing. There had been started the analysis of 28 works, but 10 were later discarded due to incorrectness, so that the data were used from the passages from 18 works altogether, among which 12 having been taken from the literature field, 2 from the textbooks, 2 from the newspapers and 2 from the journals. The number of the syllables was limited to about 5.000 in each passage. Within that range the sounds were counted paying regard to the sound changes taking place in the word and in linking the words but which are not registered by the orthography, being as well registered the sounds not recorded in our writing. By the analysis have been comprised 90.102 syllables altogether, in which are contained 196.634 sounds, 41.820 lexical words and 29.064 phonetic words. Vowels figure out about 45 per cent of the sound material, there following occlusives, sonants, nasals, fricatives and affricates. In the studied groups are comparatively equally frequent vowels and sonants, while the frequency of the individual sounds in the other groups is very uneven. There are about 80 per cent of sonorous sounds in our speech and about 10 per cent of nasal sounds.

A lexical word on an average consists of 2 syllables, a phonetic one of 3 syllables.

The results of this study could contribute to appraisal of the manifestation of some speech disorders. The work itself has been of use to the students of logopaedy at the Faculty of Defectology in Zagreb, as a preparation for logopaedic work in the sense of consideration for speech, which nowadays more and more becomes a slave of writing. But speech had been originated first and writing appeared later. Writing must be restored to the original speech, not to some third product.