

ODNOS MED DOLŽINO, OBVESTILNOSTJO IN POGOSTNOSTJO BESED

Članek preverja najprej statističnolingvistični zakon o odnosu med dolžino, obvestilnostjo in pogostnostjo besed na slovenskem gradivu. Da bi pokazal na možnost in koristnost aplikacij tega zakona na širšem jezikovnem področju, obravnava potem s tega stališča dva aktualna jezikovna pojava: (ne)kratkost izražanja ter univerbizacijo v terminologiji.

L'article vérifie d'abord la loi statistique-linguistique du rapport entre la longueur, l'information et la fréquence des mots en slovène. Pour montrer la possibilité et l'utilité des applications de cette loi dans un domaine linguistique plus large, il traite de ce point de vue deux phénomènes linguistiques actuels: l'abrégement/nonabrégement de l'expression et le figement dans la terminologie.

I.

Statistična leksikologija in makrolingvistika sta v preteklih desetletjih odkrili zakone, ki pogojujejo ustaljena razmerja med jezikovnimi sredstvi. V zadnjem desetletju je bila opravljena vrsta raziskav, ki so potrdile veljavnost nekaterih ključnih statističnolingvističnih zakonov in odkrile celo konstante ustreznih enačb. Najbolj načrtno in temeljito so to temo razvijali madžarski jezikoslovci in dosegli z njo več pomembnih uspehov.¹ Pri nas se s temi vprašanji nismo ukvarjali, deloma zaradi premajhne seznanjenosti s statističnim jezikoslovjem, deloma zaradi pomanjkanja temeljnih del, zlasti slovarja knjižnega jezika in frekvenčnega slovarja, pa seveda tehničnih pripomočkov. Osnovne podatke s prvimi statističnolingvističnimi poskusi in strokovno terminologijo sva objavila z M. Janežič v Jeziku in slovstvu leta 1962.²

Razmere se tudi v zadnjih letih niso bistveno spremenile, vendar začeta dela omogočajo analize nekaterih zakonov z našim gradivom in spodbujajo k razmišljanju o nadaljnjih poteh raziskav in o aplikaciji. Gre za nekatere, v glavnem štiri, temeljne zakonitosti, kakor jih je preprosto formuliral G. C. Lepschy po podatkih P. Guirauda: »Najpogost-

¹ F. Papp, Actes du X^e congrès international des linguistes IV, Bucarest 1970, 471—475.

² S. Suhadolnik in M. Janežič, Jezik in slovstvo VIII, 1962/63, št. 1/2, 45—49; št. 3, 75—79.

nejše besede so a) najkrajše, b) najstarejše, c) morfološko najbolj enostavne, č) semantično najširše.³

O veljavnosti citiranih statističnih zakonov v jezikoslovnih krogih menda ni več načelnih dvomov, je pa vsem jasno, da gole statistične ugotovitve ne bodo dale jezikoslovnih rezultatov. Torej bodo potrebne mnoge dodatne in ponovne analize z novim gradivom, ugotovitve bo treba ovrednotiti s povsem jezikoslovnega stališča, potem jih bo mogoče izkoristiti za boljše razumevanje delovanja jezika sploh in za analizo posameznih jezikovnih vprašanj.

Zato naj ta članek najprej s klasično metodo⁴ preveri tri temeljne statističnolingvistične zakonitosti na slovenskem gradivu, nato pa poskuša te ugotovitve aplicirati, in sicer na eno od splošnostilističnih vprašanj (kratkost in jedrnatost) ter eno od danes aktualnih vprašanj iz terminologije (univerbizacija).

II.

Preverjanje je bilo opravljeno na gradivu prvih treh črk Slovarja slovenskega knjižnega jezika /SSKJ/ I (7.500 besed) — kar zaradi hudo omejenega obsega sicer slabi verjetnostne račune, omogoča pa primerjavo z analizami, ki so bile s tem gradivom že narejene, npr. premike besedišča⁵ — razen tega na gradivu začasnega frekvenčnega slovarja (16.000 besed iz 500.000 izpisov) ter odzadnjega slovarja za prvo in drugo knjigo SSKJ (41.200 besed).⁶ Pri tem so bili zaradi poenostavitve homonimi šteti kot ista beseda, za pomenske enote pa vzete oznake iz SSKJ (pomeni, podpomeni, prenesena raba, terminološka raba; ne pa stalne zveze, fraze in pomenski odtenki). Iz istega vzroka se kot besede obravnavajo le sestavljene ali sestavljene enočlenske enote, ne pa dvo- ali veččlenske enote tipa *šivalni stroj*.

Ugotovitve so naslednje:

1. ZNAKI. Povprečna dolžina slovenske besede v slovarju je 8,7 znakov (črk), največ besed pa je v skupini z 8 znaki (16 %). Sorazmerno veliko besed obstoji iz 7, 9 in 10 znakov (40,5 %). Te štiri osrednje sku-

³ G. C. Lepschy: *La linguistica strutturale*, Torino 1966, 195. — Prim. podobno, a bolj specializirano formulacijo S. K. Šaumjana, *Leksikografičeskij sbornik V*, 1962, 49.

⁴ Pri seštevanju in računanju sta mi pomagali tov. Jana Hafner in Alenka Štular, za kar se jima lepo zahvaljujem.

⁵ S. Suhadolnik, VI. seminar slovenskega jezika, literature in kulture, Ljubljana 1970, 51—59.

⁶ Podroben opis teh del je v *Jeziku in slovstvu XI*, 1966, št. 6, 186—193.

pine zajemajo skupaj 56,3 % vseh slovenskih besed. Številne so še skupine s 5–6 znaki, manj številne so one z 11–12 znaki, redke oz. zelo redke so besede z 1–4 (4 %) oz. 13–26 znaki (7,8 %). Ali še drugače: slaba polovica besed ima po 1–8 znakov. Nesestavljene besede zasedajo skupine 1–16. Najdaljše med njimi so pridevniki in glagolniki, npr. *administracijski*, *birokratiziranje*, *centralizatorski*. Sestavljene besede zasedajo skupine 6–26 in popolnoma prevladujejo v zgornji četrtini lestvice. Med najdaljše spadajo: *buržoaznonacionalističen*, *dialektičnomaterialističen*, *krščanskosocialističen*. Podrobne številke so razvidne iz naslednje preglednice:

Znakov v besedi	Število gesel A–N	Odstotek	Znakov v besedi	Število gesel A–N	Odstotek
1	18	0,04	14	863	2,10
2	77	0,19	15	428	1,05
3	487	1,18	16	205	0,50
4	1063	2,60	17	111	0,27
5	2705	6,60	18	48	0,12
6	4138	10,03	19	26	0,04
7	5727	13,86	20	16	0,05
8	6587	15,97	21	6	0,01
9	5980	14,50	22	5	0,01
10	4937	11,96	23	1	
11	3683	8,97	24	1	
12	2575	6,25	25	—	
13	1535	3,72	26	1	
			41.221		100 %

V konkretni jezikovni rabi se povprečna dolžina besede skrajša na približno 6 znakov oz. niha nekako med 4 in 8 znaki glede na vrsto besedila. Vzrok za sorazmerno veliko skrajšanje besede v dejanski rabi je v dejstvu, da že prvih 4000 splošnih besed, ki so istočasno kratke besede, pokrije 97 % vseh uporabljenih besed,⁷ in se torej posebne, tj. dolge besede razmeroma redko uporabijo, razen v strokovnih oz. publicističnih tekstih; a tudi v teh pokrivajo splošne besede 64 % vseh uporabljenih besed.⁸ (O tem bo nadrobneje govor pri pogostnosti.)

2. POMEN in OBVESTILNOST. Slovenska beseda ima po podatkih v slovarju povprečno 1,5 pomena. Torej je pomensko precej obremenjena. Njeno ustrezno mesto v odnosu do mesta, ki ga ima najštevilnejša be-

⁷ G. Matoré (Histoire des dictionnaires français, Paris 1968, 194) meni, da obsega bazični slovar 7000–8000 besed.

⁸ P. Guiraud: Les caractères statistiques du vocabulaire, Paris 1954.

sedna družina glede na povprečno dolžino besede, je pomaknjeno na lestvici za eno stopnjo navzdol, tj. bliže dejanski jezikovni rabi, na mesto besed s 7 znaki. Pri tem je treba upoštevati, da so besede z 1—2 znakoma samó, besede s 3—4 znaki pa pretežno pomožne, skladišne; pomensko najbolj razvejane pa so tiste, ki obstojijo iz 5—6 znakov. Med temi so skoraj vse stare (domače) besede z največjimi pomenskimi obremenitvami (*baba, barva, bel, beseda, boben, bog, bogat, cel, celica, cena, cesta, cvesti, delati*), potem nove izposojenke, zlasti kulturne besede (*anketa, banka, blok, bomba, center, cirkus*) ter narečno-pogovorne besede, ki so prodrle v knjižni jezik, ker ta ni imel živih nevtralnih sinonimov zanje (*basati, fasati, fajn, fejest*). V skupini besed z nad 8 znaki povprečje pomenov razmeroma hitro pada in se v področju zloženih oz. sestavljenih besed približa značilnosti čistega enopomenskega leksema. Podrobnejši prikaz odnosov med dolžino besede in povprečjem njenih pomenov kaže naslednja preglednica:

Znakov v besedi	Število gesel A—C	Skupno število pomenov	Povprečje pomenov na eno geslo	Znakov v besedi	Število gesel A—C	Skupno število pomenov	Povprečje pomenov na eno geslo
1	4	46	11,50	13	254	273	1,16
2	17	53	1,94	14	128	144	1,12
3	98	319	3,25	15	63	69	1,09
4	208	493	2,37	16	40	45	1,11
5	493	989	2	17	15	15	1
6	805	1400	1,74	18	15	14	1,07
7	1038	1564	1,50	19	3	3	1
8	1145	1626	1,42	20	—	—	—
9	971	1272	1,31	21	1	1	1
10	792	1028	1,29	22	2	2	1
11	605	736	1,21	23	—	—	—
12	409	490	1,19	24	1	1	1
				7085 10.565 1,49 287 (homonimov in kazalk)			

Podatki v preglednici opozarjajo na določeno stalnost: postopno in skladno naraščanje oz. nato upadanje števila gesel glede na njihovo dolžino ter konstantno upadanje povprečja besednih pomenov glede na dolžino besede. Kljub temu analizirano gradivo ne potrjuje popolnoma Guiraud-Lepschyjeve⁹ formule $n \cdot s^2 = C$ o konstanti C odnosov med številom besed n in številom njihovih pomenov s , čeprav se zdi ta verjetnejša kot Pappova¹⁰ formula $n \cdot 2^s = C$, ker pogojuje prepočasen

⁹ G. C. Lepschy, *ibid.* 194.

¹⁰ F. Papp, *ibid.* 472.

padec. Morda je v formuli napaka, morda je vzrok v neenakosti gradiva ali pa samo v nezadostnosti našega gradiva. Pri drugih poskusih se je pokazalo, da bi se dala dognati tudi konstanta $/K/$ povprečja pomenov $/p/$ v odnosu do števila znakov $/z/$. Mogla bi se izraziti takole: $(p - 1) \cdot z^2 = K$. Toda formulo bi bilo treba preveriti na obsežnejšem gradivu. Zdaj konstanta niha med 20 in 30 oz. se vrtili okrog 25.

Poučen in deloma presenetljiv je račun o številu enopomenskih in večpomenskih besed. Po zelo shematičnem, v začetku članka omenjenem načinu štetja je stanje tako: po 1 pomen ima 74 % vseh besed, 2 pomena 17 %, 3 pomeni in več 9 %. Podrobna razdelitev kaže naslednje odnose:

Število pomenov	Število gesel A—C	Odstotek
1	5402	74,44
2	1206	16,62
3	551	4,84
4	139	1,91
5	70	0,96
6	28	0,38
7	19	0,26
8	12	0,17
9	6	0,09
10	4	0,06
11	4	0,06
12	5	0,07
13—18	10	0,14
kazalke	116	—
	7372	100 %

Če so te številke prave, pomeni, da ima slovenski jezik skoraj idealno pomensko razdelitev. Če pa pri računu upoštevamo, da imajo besede s posredno razlago v SSKJ navadno isto ali večje število pomenov in pomenskih usmeritev kakor njihove izhodiščne besede¹¹ in da pri štetju sploh nismo računali s pomenskimi odtenki (v slovarju so označeni s poševnimi črticami), s stalnimi zvezami in frazami, moramo po prvi korekturi izločiti iz skupine enopomenskih besed približno 10 %, po drugi pa nadaljnjih 15 % besed, kar pomeni, da je v slovenščini 64 oz. 49 % čistih enopomenskih leksemov. S tem se po eni strani spet vrnemo k trditvi o veliki pomenski obremenitvi besed, po drugi strani pa se s to številko približamo značilnosti, ki so jo ugotovili madžarski jezikoslovci na podlagi dejanskega stanja v svojem slovarju s 60.000 gesli in prej omenjene Pappove formule, čeprav se zdi trditev, češ da ima besedišče

¹¹ Slovar slovenskega knjižnega jezika I, Ljubljana 1970, XVI, § 58.

1/2 enopomenskih, 1/4 dvopomenskih, 1/8 tripomenskih itn. besed, zelo tvegana. Slovar se lahko širi le na področje enopomenskih besed, število večpomenskih bazičnih besed pa mora biti bolj ali manj stalno. Po prepričanju mnogih leksikografov je še ta odstotek enopomenskih besed visok, ker se tudi specialne besede uklonijo ekspresiji in metafori, kakor hitro se premaknejo iz čiste strokovne rabe,¹² in ker takih premikov slovarji ne zaznamujejo. Z drugimi besedami povedano: v slovarju je na vsake 2—3 besede ena večpomenska, v dejanski splošni rabi pa popolnoma prevladujejo večpomenske besede. Te besede imajo v povprečju po 5 pomena.

Vzporedno s pomenskostjo besed je treba obravnavati obvestilnost¹³ besed. Pojma se izključujeta. Informacija ali obvestilnost je obratno sorazmerna s številom pomenov in premo sorazmerna s številom znakov določene besede. Iz čim več enot je sestavljena besedna ali stavčna enota, tem večja je njena obvestilnost, tj. tem bolj se zmanjšuje nejasnost zaradi izločanja različnih možnosti, ki jih ustvarja npr. besedi večpomenskost. Osnovne, zlasti pa vodilne besede so slabo obvestilne.¹⁴ Torej je triznakovna beseda *gre* malo (nizko) obvestilna, ker implicira veliko pomenskih enot, večznakovne besede kot *hodi*, zlasti pa *odhaja* in *prihaja*, so zelo (visoko) obvestilne, ker so blizu enopomenskosti. Tudi oznaka *jug* je bolj obvestilna kakor *dol*, ker izključuje smeri *jugovzhod* in *jugozahod*. Prim. še: *vozi desno* = a) *vozi v desno stran*, b) *vozi po desni strani*, c) *vozi na (moji) desni strani*; *ta je bosa* = a) *ta ni obuta*, b) *ta ni podkovana*, c) *to ni res*. Obvestilnost se meri samo po tem, koliko določena enota zmanjša nejasnost, ko izloči nekatere, večino ali skoraj vse neustrezne možnosti.

5. POGOSTNOST. Gradivo potrjuje, da so najvišje na pogostnostni lestvici — to je, da so najpogosteje rabljene — pomožne besede in večpomenske besede. Te so tudi najkrajše in pomensko oz. funkcijsko najširše, kot smo prej ugotovili.¹⁵ V grobem vzeto, spadajo med vodilne in osnovne lekseme besede z 1 do približno 8 znaki oz. besede s povprečno 11,5 do 1,4 pomeni, med značilne lekseme pa besede z 9 do 26 znaki oz. besede z 1,3 do povprečno 1 pomenom. Vendar je v tej hierarhični razporeditvi precej izjem in neskladnosti, ki nastajajo ob naglem in ne-

¹² A. Vidovič-Muha, Naši razgledi 6. 3. 1970, 137—138.

¹³ Prim. R. Simeon: Enciklopedijski rječnik lingvističkih naziva I, Zagreb 1969, pod geslom *informacija*.

¹⁴ M. Zarebina, Język polski L, 1970, 45.

¹⁵ Prim. Zipfov zakon: Stevilo kontekstualnih pomenov kake besede ustreza kvadratnemu korenu njene pogostnosti (M. Ivić: Pravci u lingvistici, Ljubljana 1969, 209). Preizkus na našem gradivu z izračunano konstanto 1,3 je dal zadovoljive rezultate, zlasti na področju osnovnih leksemov.

kontroliranem zamenjavanju ali večanju besedišča v zadnjih desetletjih.¹⁶ Zato je prišlo med kratke, osnovne besede precej značilnih (nizkofrekvenčnih) besed, kakor *alk*, *boy*, *bide*, *city*. Obratna težnja, da bi se namreč kratke besede direktno uvrščale med visokofrekvenčne besede, ni opazna.

Besede, ki se zaradi katerega koli vzroka bistveno premaknejo na pogostnostni lestvici, težijo k spreminjanju in prilagajanju dolžine in oblike novemu položaju. Krajše besede se daljšajo oz. multiverbizirajo, daljše se krajšajo oz. univerbizirajo ter se formalno vključujejo v ustrezne besedne tipe. Pri teh spremembah odpadajo nebitveni elementi oz. se posebnosti zamenjujejo s splošnostmi. Seveda je ta potek navadno dolgotrajen in tudi ni enak za vse kategorije. Tipičen je za pogovorni jezik in žargone na eni ter za strokovni jezik na drugi strani. Besede, ki so formalno ali socialno blokirane, obdržijo izjémoma tudi v novem frekvenčnem okolju prvotno podobo.¹⁷ Besede pa, ki niso blokirane in jim jezik s popravki ne more uskladiti dolžine s pogostnostjo in obvestilnostjo, izpadajo ali izpadejo.¹⁸ Primeri za take premike so: a) *kup* → *kupčija*, *In* → *upanje*, *bil* → *bilka* (Pleteršniku je bila ta beseda še manjšalnica od *bil*), *pismonoša* → *poštar*, *ded* → *stari oče*, *did* → *otroški vrtec*, *drevi* → *danes zvečer*; b) *fiat 750* → *fičko*, *dinar* → *din*, *fakulteta* → *faks*, *tovarišica* (učiteljica) → *šica*, *ljudska milica* → *milica*, *radijski aparat* → *radioapararat* → *radio* ali *apararat*, *na svidenje* → *nasvidenje*; c) *drevored* → *drevored*, *šivalni stroj* → *šivalni stroj*, *farer* → *far* → *far* (slabš.) ali *duhovnik* (nevtr.), *zbogom* → *zbogom* (redko) ali *adijo* (pog.).

Podrobnejša analiza odnosov med pogostnostjo in dolžino besed je bila opravljena na prvih 2000 visokofrekvenčnih leksemih, in sicer tako, da obsega skupina A besede z rangom 1–100, skupina B 101–200, skupina C 201–300 in skupina Č 301–2000. (Skupina D v preglednici zajema besede z rangom nad 2000.) Povprečna dolžina besed v skupini A je 3,5. V njej je samo 12 besed z več kot 5 znaki, najdaljša je *življenje*. V tej skupini ni niti tuje niti sestavljene besede. Povprečna dolžina besed v skupini B je 4,8. V njej je 6 besed z več kot 7 znaki. V skupini C je povprečna dolžina besede 5,2. V njej je 10 besed z več kot 7 znaki. V tej zadnji skupini se pojavijo prve sestavljene besede, kakor *odgovoriti*, *prinesti*. Besede z rangom 1–500 so večinoma nepremakljive, mednje se vrivajo le posamezni najmočnejši značilni leksemi, ki odsevajo politično-socialne spremembe v družbi, npr. *socialističen*, *političen*, *slo-*

¹⁶ S. Suhadolnik, *ibid.* 52–53, 59.

¹⁷ A. Martinet: *Éléments de linguistique générale*, Paris 1967, 187–188, 193–194.

¹⁸ Prim. J. Vendryes: *Le langage*, Paris 1921, 250.

venski. Pri besedah z rangom 501—2000 je povprečna dolžina 6,4. Med temi 1700 leksemi je 140 besed z več kot 9 znaki, in sicer: 10 števnikov, 12 glagolnikov, 53 starih nesestavljenih besed, 58 tujk, specialnih besed in terminov, zlasti gospodarsko-političnih, ter 47 predpinskih glagolov in zloženih pridevnikov. Vse najdaljše besede so ali tujke ali termini: *narodnoosvobodilen*, *imperialističen*, *kapitalističen*, *materialističen*, *revolucionaren*. Povprečna dolžina prvih 2000 visokofrekvenčnih besed je 5 znakov, kar je malo pod povprečjem dolžine besed v konkretni jezikovni rabi. Posamezna povprečja od 3,5 do 6,4 v vseh štirih skupinah teh besed se tudi dobro ujemajo s povprečji dolžin tistih besed, ki se uporabljajo v posameznih zvrsteh knjižnega jezika, to je od 5,9 do 5,4. Prim. podatke v poglavju III/1 in opombo 27. Približne odnose med pogostnostjo, številom znakov in pomeni naj povzame naslednja preglednica:

Beseda	Znakov v besedi	Pomenov v besedi	Frekvenčno število izpisov	mesto skupine
a, vez.	1	5	255	B
da, vez.	2	20	4881	A
bi	2	5	2206	A
ah	2	5	20	C
bog	3	14	311	B
cel	3	15	254	B
ako	3	4	188	B
bič	3	5	25	Č
bolj	4	4	556	A
bati se	4	7	158	C
brat	4	10	110	Č
bled	4	5	56	Č
braniti	7	10	68	Č
bistven	7	2	19	Č
analiza	7	9	10	D
avtobus	7	1	7	D
bliskoma	8	1	10	D
bleščati se	8	5	7	D
ansambel	8	2	4	D
cigareta	8	1	5	D
bodočnost	9	2	17	Č
brezčuten	9	3	3	D
anomalija	9	2	2	D
bližnjica	9	1	2	D
centralizatorski	16	1	1	D
cerkvenoslovanski	17	1	—	D
avstromarksističen	18	1	6	D
antiimperialističen	19	1	2	D

III.

1. Po slovenskih slovnica in učbenikih se ponavlja le malo variirano napotilo: da sta *kratkost* in *jedrnatost* odliki dobrega stila¹⁹ in da vodi večanje števila besed v stavku do zmanjšane razumljivosti.²⁰ Ob takem stanju se je M. Kmecl po pravici začudil trditvi H. Arsena, češ da v nemški prozi »z naraščajočo dolžino stavka narašča tudi povprečna dolžina besede v takšnem stavku«, in ob tem menil, »da bi ustrezno opazovanje slovenskega jezika verjetno dalo nekoliko drugačne izide«. ²¹ O tem vprašanju lahko postavimo na osnovi prej obravnavanih zakonitosti tele trditve:

Kratka beseda oz. kratek stavek sta zaradi večpomenskosti in mnogih potencialnih možnosti razrešitve funkcionalna: a) v govornem jeziku (pogovorni jezik, žargon), kjer se besedno obvestilo (ali informacija) dopolnjuje še s tako imenovano sekundarno govorico (mimika, kretnja, intonacija, dialog, splošna situacija) in kjer je zaradi pričakovane reakcije obvestilo lahko minimalno, ter b) v lirski, zlasti sodobni pesniški izpovedi, kjer ne gre za pripovedovanje, poučevanje ali informiranje, ampak za neposreden, asociativen prenos čustvovanja, doživljanja, razpoloženja. Tak način izražanja se sicer lahko spremeni tudi v zelo obvestilno poročanje, zlasti če obstoji govorica po večini iz samostalnikov,²² vendar predstavlja ta tip praviloma komuniciranje na širokem spektru, pri čemer se informiraneec lahko vključuje na katerokoli od mnogih možnih realizacij, ki jih dopušča obvestilo,²³ ali pa vodi ta tip v popolno pomensko zaprtost govorice s končno stopnjo v čistem črkovno-glasovnem sporočilu.²⁴ Premočna obvestilnost v kratkem sporočilu pelje v nejasnost oz. premakne sfero dožemanja iz realnosti v čustvo ali celo v golo senzitivnost. To je leksikalna osnova pesniške (ne)komunikativnosti.

Vse druge oblike izpovedovanja od epskega opisa preko časopisnega poročila do znanstvene razprave težijo v želji po čim natančnejši formulaciji k visoko obvestilnim možnostim, k daljšim besedam in stavkom. Zaradi izbiranja daljših oz. specialnih besed in natančne sintaktične stave

¹⁹ A. Bajec: Rast slovenskega knjižnega jezika, Ljubljana 1951, 35—39. — J. Toporišič: Slovenski knjižni jezik II, Maribor 1966, 171—182.

²⁰ Tehnika in metodologija strokovnega pisanja, Ljubljana 1971, 146.

²¹ M. Kmecl, Jezik in slovstvo XVI, 1970/71, št. 1/2, 45—46.

²² Prim. S. Kosovel *Potopanje in Čirkus Kludsky, prostor št. 461*, Izbrane pesmi, Ljubljana 1931, 74, 156.

²³ Prim. J. Toporišiča interpretacijo Zajčeve pesmi *Veliki črni bik* in Salamunove pesmi *Gobice II* (Slovenski knjižni jezik IV, Maribor 1970, 45—46). — M. Grosman, Jezik in slovstvo XVII, 1971/72, št. 1/2, 2. — Z. Lešič: Jezik i književno djelo, Sarajevo 1971, 97.

²⁴ Prim. J. Pogačnika interpretacijo Zagoričnikovih pesmi v *Sodobnosti* XIX, 1971, št. 10, 950. — B. Vuletić, Umjetnost riječi XV, 1971, št. 2, 167—176.

pa je pri oblikovanju takega prenosa sporočila potreben večji napor, kot ga zahteva preprosto govorno sporočilo.

Kratek stavek torej avtomatično izključuje precizno obvestilo, dolg stavek in perioda pa avtomatično vključujeta daljše, nizkofrekvenčne in ozkopomenske besede. Zveza *intelektualno morje* zahteva nadaljnje pojasnilo, ker je za splošno razumevanje preveč posebna; *sinje oz. modro morje* zveni malo pleonastično, a spada po svoji naravi v kratki tip izražanja, saj ima pričakovani člen; *razburkano morje* je nevtralna, a v smer natančnega izražanja naravnana zveza.

Vsaka beseda, zlasti večpomenska, vsebuje celo vrsto potencialnih vezav in pomenskih usmeritev. Pri kratkih besedah je za ustvaritev jasnosti dvojna možnost: zamenjava te, malo obvestilne besede z zelo obvestilno ali pa pritegnitev novega besednega člena (*hiša* : *blagovnica* ali *blagovna hiša*). Pri visoko obvestilnih besedah oz. zvezah ostaja v glavnem samo druga možnost: daljšanje, širjenje z novimi visoko obvestilnimi členi. Zveza *otvoritev blagovne hiše* torej samo eksplicira to, kar je v predstavi besede *hiša* organsko implicitno. Sicer pa je realizacijo *otvoritev hiše* → *otvoritev blagovne hiše* mogoče predvideti že na osnovi Markovega zakona iz leta 1913.²⁵ Prim. tudi: *zelo pritiskati* (več pomenov) : *oršiti velik pritisk* (1 pomen); *razstaviti* (več pomenov) : *privediti razstavo* (1 pomen). Če pa novi element zveze ni dovolj točen oz. informativen, se ustvarja s tem nefunkcionalna zveza, ki jo jezik zavrača in sili v novo univerbizacijo. Prim.: *pogoj* → *prvi pogoj* → *predpogoj*; *jed* → *začetna jed* → *predjed*. Prepovedovanje univerbizacij v tem primeru je nepravilno in zato navadno tudi brezuspešno.²⁶

To so splošno jezikovne zakonitosti, in odstopanja so lahko le individualna. Zato Arsenovo poročilo o daljšanju besed v daljših nemških (pripovednih) stavkih ne prinaša dosti novega, zlasti ne nič tipično nemškega, če seveda odmislimo nemški besedotvorni način, pač pa prepričljivo potrjuje osnovni zakon o odnosu med dolžino in obvestilnostjo besed. Poskus na zelo omejenem slovenskem gradivu²⁷ kaže, da obstoji povprečen slovenski prosti stavek iz 6,5 besed, povprečna dolžina besede pa je 4,6 znakov; povprečen zložen stavek ima 15,9 besed, povprečje najdaljšega stavka pa znaša 37 besed oz. 47 besed, če se upošteva perioda kot enota. Dolžina in obvestilnost sta v tesni odvisnosti od dolžine stavka, načina pripovedi in stila, morda tudi od obdobja. Kakšni so odnosi glede na posamezne zvrsti jezika in avtorje, naj osvetli preglednica.

²⁵ V. Vinogradov, *Voprosy jazykoznanija* VII, 1958, št. 3, 14. — M. Ivić, *ibid.* 205, 208.

²⁶ Slovenski pravopis, Ljubljana 1962, pod geslom *predjed* in *predpogoj*.

²⁷ Analiza je zajela vsako prvo stran citiranega besedila oz. celo pesem.

Avtor in naslov dela	Vrsta besedila	Povprečno število			Število besed v najdaljšem stavku
		besed v prostem stavku	besed v zloženem stavku	znakov v besedi	
Prešeren, Kam Prešeren, Ukazi	osebna izpoved	4,1	12,4	3,9	23
		4	8,5	4,1	15
Prešeren, Janezu N. Hradeckemu	neosebna izpoved	6,6	19,3	4,6	51 oz. 89
Cankar, Otroci in starci Prežihov, Samorastniki	pripoved	7,5	19,2	4,6	28 oz. 62
		6,9	12,8	4,6	27
Cankar, Slovenska kultura Kermauner, Volja do življenja	strokovni opis	8,4	21,5	4,9	56
		8,5	17,3	5,4	61
povprečno		6,5	15,9	4,6	37 oz. 47

Iz tega sledi, da bi bilo treba presplošna stilistična opozorila »povprečno pismenim« o krajšanju sporočil oz. neprimernosti daljših stavkov, zlasti podredij, revidirati,²⁸ saj je bilo ponovno ugotovljeno, da je konstrukcija kratkih oz. samih glavnih stavkov z veliko obvestilnostjo negospodarna, neorganska in monotona,²⁹ ter opozoriti na različne možnosti v različnih jezikovnih rabah.

2. Težnja po krajšanju in univerbizaciji je v sodobnem besedišču zelo močna³⁰ in ji, kot smo videli, podlegajo praviloma besede, ki večajo svojo pogostnost. Pri tem gre zlasti za specialne izraze in termine, ki se začno uporabljati bolj na široko in prehajajo preko žargona ali pogovornega jezika v nevtralnno področje jezika. Nekaj časa imajo še ekspresivno noto, potem se popolnoma nevtralizirajo. Končno lahko vplivajo tudi v nasprotni smeri na strokovni izraz sam.³¹

V skladu z obravnavanimi zakonitostmi je, in tudi v skladu z zahtevo, naj bo glavno terminološko merilo pomenska jasnost,³² če se v stroki ohranjajo ali na novo določajo sistemsko najprimernejši izrazi, v širši poljudni rabi pa se dopuščajo pogovorno-žargonske oblike, npr. *trafo* : *transformator*, *faza* : *fazni vodnik*, *jod* : *jodova tinktura*, *eter* : *dietileter*, *jasmin* : *skobotovec* ali *nepravi jasmin*, *Titova 3* : *Titova cesta št. 3*. Da pa nazoni o funkciji terminoloških univerbizacij niso razčiščeni in se ne skladajo s statističnolingvističnimi zakonitostmi, kažejo nasprotujoče si

²⁸ J. Toporišič: Slovenski knjižni jezik II, 180, 189.

²⁹ E. Martinet, *ibid.* 192. — M. Kmecl, *ibid.* 46.

³⁰ A. Jedličká, *Slavica Pragensia* XI, 1969, 95–101.

³¹ Univerbizirane oblike prodirajo v knjižni jezik zlasti po časopisju. Prim. *prosvetnik*, *španec*, *počitnikar*, *velikanka* (ladja), *radialka* (Delo 19., 20., 21. XII. 1971).

³² Slovenski jezikovni priročnik za tehnike, Ljubljana 1969, 7.

težnje pri reševanju praktičnih terminoloških vprašanj pa tudi ustrezni stavki v strokovnih navodilih o tvorjenju terminov.

Vzemimo preprost primer iz železniške terminologije, ki je bila med prvimi strokovno obdelana in je dobila tudi svojo registracijo.³³ *Železniški vagoni so tovorni in osebni*. V podskupino prvih spadajo: *nizkostenski, visokostenski, odprti, zaprti vagon*. V podskupini drugih bi analogno morali biti: *službeni, poštni, jedilni, spalni, ležalni vagon*. V resnici se je namesto *spalnega vagona* vpeljal *spalnik*, namesto *ležalnega vagona* pa *kuše*. Delno se kot sinonima uporabljata tudi *ležalnik* in *jedilnik*, čeprav že imata pomene, ki so obravnavanima pojmom zelo blizu. Ker je terminologija v tem primeru prehitro sprejela pogovorne izraze oz. oblike, se je osnovna terminološka sistematika podrla, še preden je bila končana.

Sicer pa se ob *spalniku* vprašanje o oblikovanju terminov razširi v dvoje pomislekov: a) ali se res lahko pod pretvezo jasnosti uvaja eno samo obrazilo za določeno pomensko kategorijo, čeprav ima to obrazilo v splošni rabi precej drugačno pomensko usmerjenost ali stilno obarvanost,³⁴ ter b) ali ni zahteva po primarnosti obrazilnega tvorjenja terminov³⁵ sploh v neskladju z zakonom o obvestilnosti. Na nezadostno razmišljanje o teh vprašanjih kažejo prizadevanja nekaterih terminologov, da bi dali terminom čim prej univerbizirane, navadno kar žargonske oblike. Taki so predlogi, naj bi se namesto *Peltonova turbina* pisalo *peltonka*³⁶ ali namesto *dizelski motor* oz. *dizel motor* kar *dizel*. Terminološko tudi ni sprejemljiv tip *alu folija* namesto *aluminijasta* oz. *aluminijeva folija*.

Še bolj se čuti pomanjkanje sistemskosti pri reševanju terminoloških vprašanj na tistih področjih, kjer se oblikuje izrazje pravzaprav samo od sebe. Tam se mešajo umetno narejeni dvočlenski termini z narečnimi izposojenkami in pogovorno-žargonskimi elementi tako močno, da je zgrešen bistveni namen terminologije: jasnost in obvestilnost. Na to opozarja npr. pogostno sprejemanje ali tvorjenje terminov z obrazilom *-ec*: aer. *brezrepec* : *brezrepo letalo*, agr. *dalmatinec* : *dalmatinsko vino*, čeb. *lipovec* : *lipov med*, šport. *amerikanec* : *vrsta skoka v višino, internacionalec* : *državni reprezentant*, vet. *arabec* : *konj arabske pasme*. Te univerbizacije so tako agresivne, da terminologi dostikrat niti ne poskušajo

³³ H. Podkrajšek: Obrtna, trgovinska, tvorniška in železniška terminologija, Prevalje 1926. Slovar je jezikovno pregledal A. Breznik.

³⁴ Prim. podobne sporne terminološke univerbizacije: *končnik, preteklik; odjemalka, usmerjevalka; stojka*.

³⁵ Slovenski jezikovni priročnik za tehnike, 9, 38—39.

³⁶ Ibid. 11.

poiskati ali ustvariti nevtralnega terminološkega izraza, ampak se zadovoljujejo samo z njegovo razlago, npr. žarg. *internacionalec* = *tekmovallec*, ki nastopa na mednarodnih tekmovanjih kot reprezentant svoje države. In vendar je obrazilo *-ec* pomensko tako široko ter obvestilno šibko, da je komaj primerno za tvorbo novih terminov.³⁷

Prav tako pa terminologi premalo upoštevajo statističnolingvistične zakonitosti tudi pri terminih, ki so se že ali se še bodo uvrstili med visokofrekvenčne lekseme, in v odporu proti tujkam po nepotrebnem poudarjajo alternativo strokovni : nestrokovni jezik, domača : tuja beseda. Tako npr. odklanjajo besedo *bojler* in jo nadomeščajo z *vodnim grelnikom* ali *električnim grelnikom vode*, čeprav ne gre za čisto nov pojem.³⁸ Glede na to, da je pojem široko znan in beseda *bojler* zelo rabljena, da pojem ne spada v sistematiko tipa *spalni vagon* in da je *vodni grelnik* kot termin premalo informativen, *električni grelnik vode* pa predolg in pomensko preozek v primeri z *bojlerjem*, bi bilo treba tudi v terminologijo sprejeti izraz *bojler* in ga specializirati z *električni, plinski*. Prav tako je dvomljivo, če bo priporočani (*elektronski*) računalnik zares izločil sinonim *kompjuter*. Ta ima namreč zaradi velike obvestilnosti, večje kratkosti in paradigmatiskih možnosti, npr. *kompjutirati, kompjutiranje*, vse pogoje za obstoj in uveljavitev.

Pravi termin mora imeti dva jasna osnovna elementa: nosilec pomena in aktualizator. Nosilec pomena mora biti praviloma nizko obvestilen, aktualizator pa visoko obvestilen element, pa naj je to obrazilo, del sestavljenke oz. zloženke ali samostojna beseda. Ta, v končni stopnji multiverbizacijski sistem je najbolj izpeljan v botaniki in zoologiji, in sicer tako, da je splošna beseda sprejeta za oznako višje kategorije, s prilastki pa se delajo oznake za vrste, npr. *breza* : *puhasta breza*.³⁹ Ker pa ima ta način tvorjenja terminov tudi slabo lastnost, tj. preobremenjuje osnovne lekseme, seveda ne more veljati kot splošno načelo terminološke tvornosti. Zlasti zato ne, ker nastajajo terminološka poimenovanja istega pojma na več ravneh, na terminološki ravni navadno z določeno zamudo. Od tod nenehna sinonimno-stilna napetost v besedišču. Toda to je splošen jezikovni pojav,⁴⁰ ki ustvarja različne stilne možnosti v knjižnem jeziku ter spodbuja individualno besedno tvornost. Kljub temu lahko sprej-

³⁷ Beseda *amerikanec* ima po gradivu za Slovar slovenskega knjižnega jezika 11 pomenov, beseda *lipovec* pa 8 pomenov.

³⁸ Slovenski jezikovni priročnik za tehnike, 10.

³⁹ Vsi zgornji, posebej neoznačeni primeri so vzeti iz Slovarja slovenskega knjižnega jezika oz. iz gradiva zanj.

⁴⁰ Prim. *je ne sais pas* (oblika pisanega knjižnega jezika), *je sais pas* (oblika govornega knjižnega jezika), *SSepa* (substandardna oblika). Teorie informace a jazykověda, Praha 1964, 258.

memo okvirno trditev, da predstavlja tip *hokejska palica* idealni termin, iz aktualizatorja narejena *hokejka* njegov pogovorno-publicistični sinonim, iz nosilca pomena narejena *palica* pa žargonsko obliko pravega termina.⁴¹

Terminološka poimenovanja bi bilo treba torej oblikovati bolj v skladu z načeli, ki se kažejo iz odnosov med pogostnostjo, obvestilnostjo in dolžino jezikovnih elementov. Pa tudi zahteva po jasnosti terminov in živosti strokovnega pisanja⁴² bi zaslužila ovrednotenje s tega stališča.

RÉSUMÉ

Le but l'article est de vérifier la loi statistique-linguistique de l'accord entre la longueur du mot, sa signification, son information et sa fréquence.

La longueur moyenne du mot slovène est de 8,7 unités d'après le dictionnaire, et de 6 unités dans l'usage courant. La plupart des mots comprennent 8 unités. Le mot le plus long a 26 unités. Le mot slovène a 1,5 de signification en moyenne. Les formules de P. Guiraud $n \cdot s^2 = C$ et de F. Papp $n \cdot 2^s = C$ ne sont pas tout à fait confirmées en slovène. On pourrait établir la constante entre la moyenne dans le nombre des significations et celle de la longueur des mots d'après la formule $(p - 1) \cdot z^2 = K$. Selon un critère très rigoureux il y a 74 % de mots slovène avec une seule signification, et 64 % ou 49 % d'après un critère moins rigoureux. Le rapport entre la fréquence et les critères cités ci-dessus est vérifié.

Ensuite, on a analysé deux phénomènes linguistiques actuels à la base de cette loi. Le résultat est le suivant: a) Il faut modifier la loi stylistique de la nécessité de l'emploi de courtes propositions de la manière qu'elle puisse indiquer de différentes longueurs des mots et des propositions dans les diverses espèces de la langue. b) Il faudra coordonner la terminologie moderne avec l'information et la fréquence des mots et il faudra en même temps empêcher la pénétration des figements inutiles dans les langues spéciales.

⁴¹ J. Ružička, 1000 poučení zo spisovnej slovenčiny, Bratislava 1971, 81.

⁴² J. Toporišič: Slovenski knjižni jezik II, 190.