



UDK 811.163.6'342:316.774(497.4)

*Damjan Huber*

Filozofska fakulteta Univerze v Ljubljani

damjan.huber@ff.uni-lj.si

## RAZISKOVALNA IZHODIŠČA BESEDILNOFONETIČNE ANALIZE SLOVENSKEGA MEDIJSKEGA GOVORA (NA PRIMERU BESEDNOVRSTNE PREDVIDLJIVOSTI POUĐARKOV)

Besedilna fonetika je področje, ki je v slovenističnem jezikoslovju razmeroma slabo raziskano. V prispevku je predstavljen primer besedilnofonetičnega raziskovanja, tj. analiza poudarkov z vidika besednovrstne predvidljivosti, pri čemer je izpostavljena pomembnost posameznih raziskovalnih faz, kot so izbor govorcev, priprava, transkripcija in segmentacija gradiva ter dvostopenjsko potrjevanje raziskovalnih hipotez – slušna in računalniška analiza. Raziskava temelji na korpusu izbranih slovenskih pogovornih TV-oddaj, torej na avtentičnem in bolj ali manj spontanem medijskem govoru.

**Ključne besede:** jezikoslovna metodološka načela, besedilna fonetika, poudarek, besedne vrste, medijski govor

Prosody, or sentence phonetics, is a relatively poorly researched area in Slovenian linguistics. This article presents a piece of prosodic research: an analysis of sentence stress from the perspective of text genre predictability, emphasizing the importance of specific research phases, particularly the selection of speakers, preparation, transcription and segmentation of material, and a two-stage confirmation of the research hypotheses using audio and computer analysis. The research is based on a corpus of selected Slovenian television discussion programs—in other words, on authentic and relatively spontaneous media speech.

**Keywords:** methodological principles of linguistics, prosody, sentence stress, text genre, media speech

### 0 Uvod

V okviru *fonetike besedila* (SS 1976), *stavčne fonetike* (Toporišič 1992) oz. *besedilne fonetike* (SS 2000), ki jo Toporišič definira kot »s/lušne značilnosti povedi, eno- ali večstavčne« (Toporišič 1992: 308), raziskujemo premore, poudarke, intonacijo, register, hitrost govora in glasovno barvanje. Iz definicije je razvidno, da Toporišič v opredelitvi govorjenega jezika izhaja iz enote zapisanih besedil; kot raziskovalno enoto ima v mislih predvsem stavek: »Ker je najmanjše možno sporočilo poved s skladenjsko obliko stavka ali več stavkov, to fonetiko imenujemo tudi kar stavčno. V načelu bi besedilna fonetika morala zaobjeti slušne značilnosti kakega besedila v celoti; v sedanjem stanju vendar obravnava kvečjemu še slušne pojave med posameznimi povedmi, podrobneje pa komaj slušne značilnosti povedi, zlasti seveda stavka« (SS 2000: 533).

Če je še za drugo polovico 20. stoletja veljalo, da redke analize govornega jezika temeljijo predvsem na subjektivni oceni raziskovalcev in omejenem korpusu nesponatano tvorjenih besedil, tj. branem govoru zapisanih (literarnih) besedil (npr. Toporišič 1972), potem smo v zadnjem desetletju in pol priča kvantitativnemu in kvalitativnemu porastu (besedilnofonetičnih) raziskav govornega jezika (npr. Hribar 2001/2002, Valh Lopert 2005, Tivadar 2009, Volk 2011). Raziskave govora temeljijo na reprezentativnih besedilnih korpusih posnetega zvočnega gradiva, analiz pa se lotevamo z računalniškimi orodji, specializiranimi za transkripcijo in analizo govora (npr. Praat), s čimer lahko raziskujemo tudi realno (in spontano) govorno produkcijo, tj. najbolj avtentičen del govora.

## 1 Raziskovalna izhodišča

Pri analizi besedilnofonetičnih značilnosti si lahko pomagamo z različnimi metodološkimi pristopi. Tako lahko poudarke raziskujemo v razmerju do ločil v zapisu, glede na tipična mesta pojavljanja, besedni red, vloge, ki jih opravljajo v govoru, primerjalno z drugimi (besedilno)fonetičnimi značilnostmi, npr. intonacijo, preomor in formantnimi vrednostmi, glede na neverbalna sredstva, (ne)spontani govor, (ne)profesionalne govorce, spol in družbeni status govorcev, stilno (ne)zaznamovanost, (ne)znane informacije, (ne)dialoškost komunikacije idr. V prispevku je predstavljena raziskava, ki temelji na kombinirani slušno-eksperimentalni metodi preučevanja poudarkov z vidika besednovrstne predvidljivosti,<sup>1</sup> posebna pozornost pa je namenjena opisu poteka raziskave, tj. načrtovanju korpusa, zbiranju, pripravi, transkripciji in segmentaciji gradiva, analizi in interpretaciji podatkov ter sintezi rezultatov. Ker je besedilnofonetičnih raziskav, opravljenih na slovenskem gradivu, razmeroma malo (npr. Podbevšek 2006, Volk 2011), je eden izmed ciljev raziskave na slovenskem gradivu preveriti v tujini že uveljavljene pristope (npr. Cruttenden 1986, Horga 1996).

Poudarke<sup>2</sup> razumemo kot prozodično realizirano izpostavljenost dela govorne enote na površinski, tj. stavčni oz. besedilni ravni, ki pa je odraz informacijskega urejanja besedila na globinski ravni (Duběda 2005: 183), saj opravljajo različne vloge v govoru, od logičnega in subjektivnega izpostavljanja najpomembnejših delov govorne enote do organizacije besedila in emocionalnega izražanja govorcevih razpoložljivih stanj. Fonetično se uresničujejo s kombinacijo več akustičnih dejavnikov, npr. s

<sup>1</sup> Analiza poudarkov z vidika besednovrstne predvidljivosti je del širše raziskave govora TV-voditeljev in politikov v izbranih oddajah TV Slovenija, opravljene v okviru doktorskega dela *Poudarek in pavza v standardnem slovenskem govoru* (Huber 2013), v prispevku pa sta prilagojeno in dopolnjeno predstavljeni predvsem (pod)poglavji *Predstavitev gradiva in raziskave* ter *Besednovrstna predvidljivost poudarkov*.

<sup>2</sup> V slovenskem jezikoslovju sta se termina *naglas* in *poudarek* dolgo časa med sabo zamenjevala oz. je bilo govora predvsem o poudarku v smislu besednega poudarka, danes naglasa. Razlikovanje med poudarkom, ki je lastnost stavka, in naglasom, ki je lastnost besede, sta dokončno uvedla Rigler (1968) in Toporišič (SS 1976), o (stavčnem) poudarku pa je pisal že Breznik v razpravi *Besedni red v govoru*, kjer razpravlja o t. i. *prosti stavi* v okviru besednega reda ter o *govornem odstavku* kot raziskovalni enoti, kot merilo zakonitosti besednega reda v prosti stavi pa določi *stavčni poudarek* (Breznik 1908: 226).

spremembo trajanja posameznih glasov, frekvence osnovnega tona (F0), jakosti ali hitrosti govora (npr. Palková 1997: 156–57), ter ga uvrščamo med najpomembnejše akustične dogodke, ki spremljajo govor in hkrati odločilno vplivajo na razumevanje vsebine in pomena (kot okrepljen izgovor tistega dela govora, ki se govorniku zdi najpomembnejši za razumevanje sporočila). Med raziskovalci ni enotnega mnenja o tem, kateri od naštetih signalov je najpomembnejši za percepcijo poudarka, najpogosteje pa so za prepoznavanje izpostavljenih delov govora pomembne kombinacije vseh naštetih (akustičnih) parametrov: poudarjeni segment je lahko hkrati izražen z večjo izgovorno jakostjo, dlje traja, ima višji F0 in je jasneje artikuliran. V več raziskavah je bilo ugotovljeno, da je nemogoče analizirati vsak opazovani parameter posebej, saj so med sabo povezani in se kot skupek (prozodičnih) značilnosti izražajo v poudarku. Pri interpretaciji računalniško izmerjenih rezultatov si je priporočljivo pomagati tudi z drugimi znotrajjezikovnimi, npr. besedilnofonetičnimi (hitrost govora, premori, intonacija), ter zunajjezikovnimi značilnostmi, npr. s televizijsko sliko, ki pokaže tudi neverbalni del komunikacije (mimiko, gestiko idr.).

### 1.1 Hipoteza

Raziskave o besednovrstni predvidljivosti poudarjanja, ki bi temeljile na slovenskem gradivu, ne obstajajo, tuji raziskovalci pa navajajo različne rezultate analiz poudarkov z vidika besednih vrst. Enotni so si edino v tem, da so stavčnočlenske besede pogosteje poudarjene kot nestavčnočlenske (npr. Quirk idr. 1985, Đorđević 1984, Matulina Jerak 1989). Raziskovalno hipotezo, da med stavčnočlenskimi besednimi vrstami govorniki poudarjajo predvsem samostalnike, med nestavčnočlenskimi pa členke, smo postavili na podlagi že znanih tujih raziskav in subjektivnega opazovanja slovenskega medijskega (političnega) govora ter preverili na izbranem gradivu, tj. govoru politikov in voditeljev<sup>3</sup> TV-oddaj nacionalne televizije.

### 1.2 Gradivo

Zaradi sledljivosti in medsebojnih primerjav je zelo pomemben natančen opis gradiva in metodoloških načel raziskave. Jezik lahko opazujemo in opisujemo z različnimi pripomočki in z različnih vidikov, vendar je razumljivo, »da analiza tako širokega področja, kot sta jezik in komunikacija, ne more težiti po enciklopedični izčrpnosti in da morajo biti nekateri vidiki jezika potisnjeni v drugi plan ali kar povsem izpuščeni, če se analiza osredotoči na dovolj določen predmet« (Miščević 1983: 59).

Kot podlago za analizo in opis poudarkov v slovenskem medijskem govoru smo izbrali zbirko avtentičnega in bolj ali manj spontanega medijskega govora oz. govora

<sup>3</sup> Slovnica oblika za moški spol je v prispevku nevtralna ter velja tako za politike in voditelje kot tudi za političarke in voditeljice TV-oddaj. Kjer je to potrebno, so posebej zapisane tudi ženske oblike.

izbranih politikov in voditeljev TV-oddaj nacionalne televizije.<sup>4</sup> V veliki meri gre za dialoge in multilogе oz. monologe, ki prehajajo v dialoge, ter formalne in javne govorne položaje, ki se odvijajo v ljubljanskih studiih TV Slovenija. Specializirani korpus govora *Korp-2MeGo*<sup>5</sup> desetih oddaj oz. delov oddaj v skupnem trajanju dobrih dveh ur je nastal za namen širše raziskave, objavljene v Huber 2013, uporabili pa smo ga kot enega od virov za postavljanje in kot edini vir za preverjanje raziskovalnih hipotez.

### 1.2.1 Izbor gradiva

Govor politikov,<sup>6</sup> ki gostujejo v TV-oddajah, in govor novinarjev oz. TV-voditeljev<sup>7</sup> sta sestavini medijskega govora, ki ga gledalci pogosto povsem nezavedno dojemamo kot govorni model, vzor, ideal oz. standard: »Pod vplivom množičnih govornih medijev in šolstva ima danes večina ljudi neko predstavo o standardnem izgovoru. Ta predstava je bolj ali manj površna, bolj ali manj naivna, toda obstaja« (Škarić 1999: 205). Ker je medijski govor »praviloma dosledna aktualizacija predpisanega jezikovnega sistema, njegovi nosilci pa so govorniki, ki so kljub zavestnemu upoštevanju sistema hkrati podvrženi tudi vplivu spontanosti in njihovih sprememb v splošni govorni rabi« (Vitez 1999: 42), je idealna priložnost opazovanja, kako govor šolanih govorcev odstopa od nekega ideala oziroma jezikovne norme. Z razvojem govorjenih medijev v 20. stoletju je medijski govor postal dostopen množici poslušalcev in gledalcev ter s tem eden izmed načinov (posrednega) jezikovnega izobraževanja. Raziskava bo torej temeljila na analizi spontanega govora politikov ter spontanega in

<sup>4</sup> Slovenski medijski govor, kot ga razumemo v prispevku, oblikujejo govorniki, ki se pojavljajo v slovenskih govorjenih medijih, po Toporišičevi razvrstitvi funkcijskih zvrsti pa ga lahko uvrstimo med govorjena publicistična (in praktičnostrokovna) besedila (SS 2000: 28–31). Predvsem nacionalni radii in televizije imajo tudi vlogo nacionalnega povezovanja ljudi, zaradi česar so na RTV Slovenija razvili interno šolo (knjižnega) govora za svoje napovedovalce in voditelje oddaj, kar pomeni, da je »medijski govor načeloma »idealno govorno uresničevanje«, z normo torej najnatančnejše usklajena govorna realizacija« (Vitez 1999: 42).

<sup>5</sup> *Korpus 2-urnega medijskega govora.*

<sup>6</sup> N. Hribar (2001/2002: 315) govor politikov v TV-oddajah opredeljuje »kot vmesno stopnjo med zborna oblika knjižnega jezika in njegovo [...] pogovorno varianto. Govorniki namreč nastopajo v javnem govornem položaju [...] pred množičnim poslušalstvom [...], zaradi česar se predvideva, da bo njihov jezik zborni; ker pa gre za prosti govor [...], se govorniki [...] odmikajo od zborne oblike knjižnega jezika.«

<sup>7</sup> Izraza *novinar* in *voditelj* v raziskavi uporabljamo sinonimno za poimenovanje voditeljev analiziranih TV-oddaj, čeprav npr. T. Verovnik (2004: 166–70) razlikuje govor *voditeljev*, *posnete novinarske prispevke* ter *intervjuje v studiu in posnete pogovore/izjave*. Najmanj odstopov od pričakovane realizacije ugotavlja pri govoru voditeljev – še najpogostejša so odstopanja od pričakovanih stavčnofonetičnih potekov. Nekoliko več odstopov od pričakovane realizacije na stavčnofonetični ravni je zaznala pri posnetih novinarskih prispevkih, in sicer zelo malo pri voditeljih ter več pri govoru politikov in gospodarstvenikov.

delno nespontanega (branega) govora<sup>8</sup> voditeljev okroglih miz ter drugih pogovornih oddaj s politično vsebino na TV Slovenija. Pri izboru gradiva smo se v čim večji meri poskušali držati naslednjih kriterijev:

- V ospredju je analiza spontanega govora politikov v pogovornih oddajah s politično vsebino na nacionalni TV Slovenija z razmeroma dobro gledanostjo in so že razmeroma dolgo na sporedu.<sup>9</sup>
- V vsaki izbrani oddaji bo narejena analiza 10-minutnega govora 1 politika in 1 voditelja, s tem da bo večji poudarek na analizi (spontanega) govora politika.
- Posnetki oddaj morajo biti dovolj kakovostni za računalniško analizo.
- Zagotoviti je treba enakovredno zastopanost govork in govorcev.
- Izločiti je treba vse ekstreme. Pravorečna (ne)ustreznost, ki je pogosto moteča tudi za gledalce, je eden od pokazateljev govorčeve (ne)kredibilnosti. Izbrati je torej treba govorce, ki imajo čim bolj nevtralno izgovorjavo.
- Priporočljiva je regionalna oz. dialektalna raznovrstnost govorcev.

Izbrali smo 10 pogovornih TV-oddaj s politično vsebino, predstavljenih v tabeli 1, v skupni dolžini 122 minut in 35 sekund. Vse oddaje imajo podoben koncept: pogovor voditelja z gosti/politiki oz. gostom/politikom o aktualnem in polpreteklem političnem dogajanju; pogovor poteka v studiu brez občinstva. Vse analizirane oddaje so torej bile pogovorne, v katerih gosti (politiki) govorijo predvsem spontano, medtem ko voditelji oddaj govorijo tako spontano kot delno spontano oz. celo nespontano (berejo uvodne predstavitve gostov in obrazložitve tem oddaj). Oddaje so bile predvajane na TV Slovenija 1 ali 2 med 17. 5. 2006 in 1. 6. 2007 v večerno-nočnih terminih med 20.00 in 00.10. V osnovi gre za dva tipa pogovornih oddaj s politično vsebino:

- 1 voditelj se v studiu pogovarja z 1 politikom (oddaje *Intervju, Pogovor s predsednikom vlade, Vroči stol*);
- 1 voditelj se v studiu pogovarja z več (3–7) politikami oz. drugimi gosti (oddaje *Pod žarometom, Omizje, Studio City*).

<sup>8</sup> Glede na navezanost na pisno predlogo Tivadar (1999: 346) loči 5 načinov uresničitve knjižnega govora: 1. *branje*, 2. *polbranje*, 3. *govor na podlagi opornih točk*, 4. *prosti govor brez zapisa, a z miselno pripravo*, 5. *popolnoma prosti govor*. Pri govoru politikov lahko govorimo predvsem o prostem govoru, a z miselno pripravo (4), in o popolnoma prostem govoru (5), pri voditeljih pa o vseh petih stopnjah, najpogosteje pa verjetno o branju (1) in polbranju (2) pri napovedih oddaje in predstavitvah gostov ter o prostem govoru brez zapisa, a z miselno pripravo (4) pri moderiranju pogovora.

<sup>9</sup> Pri voditeljih na nacionalni televiziji lahko pričakujemo razmeroma skrbno govorico, saj RTV Slovenija kot edina medijska hiša pri nas že od začetka šestdesetih let 20. stoletja skrbi za sistematično izobraževanje svojih novinarjev.

Tabela 1: Analizirane oddaje.

Besedilo	Oddaja	Na sporedu	Število govorcev v oddaji	Govorec <sup>10</sup>	Čas trajanja govora za analizo	
					po govorcih	skupaj
Besedilo 1S	Intervju	7. 1. 2007, TV SLO 1, 21.30–22.25	2	GovPz1	6 min in 37 s	10 min in 45 s
				GovNm2	3 min in 23 s	
				hkratni govor	45 s	
Besedilo 2S	Pogovor s predsednikom vlade	19. 6. 2006, TV SLO 1, 20.00–20.55	2	GovPm3	8 min in 33 s	10 min in 46 s
				GovNm4	2 min in 13 s	
				hkratni govor	0 s	
Besedilo 3S	Vroči stol	1. 6. 2007, TV SLO 2, 20.50–21.40	2	GovPm5	6 min in 49 s	10 min in 35 s
				GovNm6	3 min in 42 s	
				hkratni govor	4 s	
Besedilo 4S	Pod žarometom	22. 8. 2006, TV SLO 1, 20.00–20.55	7	GovPz7	6 min in 52 s	11 min in 15 s
				GovNz8	3 min in 27 s	
				hkratni govor	56 s	
Besedilo 5S	Omizje	17. 5. 2006, TV SLO 1, 22.55–00.10	5	GovPz9	8 min in 49 s	22 min in 5 s
				GovPm10	7 min in 43 s	
				GovNm11	5 min in 12 s	
				hkratni govor	21 s	
Besedilo 6S	Vroči stol	13. 4. 2007, TV SLO 2, 20.50–21.40	2	GovPm12	7 min in 32 s	10 min in 22 s
				GovNm13	2 min in 43 s	
				hkratni govor	7 s	
Besedilo 7S	Omizje	6. 6. 2006, TV SLO 1, 22.55–00.10	6	GovPz14	5 min in 26 s	10 min in 18 s
				GovNz15	3 min in 56 s	
				hkratni govor	56 s	
Besedilo 8S	Pod žarometom	12. 9. 2006, TV SLO 1, 20.00–21.05	6	GovPm16	6 min in 43 s	10 min in 46 s
				GovNz17	3 min in 58 s	
				hkratni govor	5 s	
Besedilo 9S	Studio City	17. 7. 2006, TV SLO 2, 20.55–21.55	4	GovPz18	3 min in 0 s	6 min in 32 s
				GovNm19	3 min in 22 s	
				hkratni govor	10 s	
Besedilo 10S	Omizje	25. 10. 2006, TV SLO 1, 22.55–00.05	8	GovPz20	7 min in 7 s	19 min in 11 s
				GovPm21	7 min in 31 s	
				GovNz22	4 min in 2 s	
				hkratni govor	31 s	

<sup>10</sup> Šifre govorcev: Gov – govorec, P – politik, N – novinar oz. voditelj, m – moški, z – ženska, s številkami 1–22 so označeni posamezni govorcei.

V enem primeru nam ni uspelo zagotoviti 10 minut govora oddaje, v kateri smo analizirali govor 1 politika in 1 voditelja (*Besedilo 9S*), medtem ko smo v dveh oddajah (*Besedilo 5S* in *Besedilo 10S*) transkribirali po ok. 20 minut govora, saj smo v okviru teh oddaj analizirali govor 2 politikov in 1 voditelja. Vsi izbrani govorniki so znane medijske oz. politične osebnosti. Razen v *Besedilu 9S*, kjer smo analizirali za nekaj sekund več voditeljevega govora, v vseh drugih besedilih količinsko prevladujejo govori politikov; tudi v skupnem seštevku (tabela 2). Tako v kategoriji politikov kot pri voditeljih nekoliko več časa posvečamo analizi moških govorcev. Ugotovimo torej lahko, da korpus *Korp-2MeGo* bolj ali manj zadošča kriterijem, postavljenim pred začetkom raziskave.

**Tabela 2:** Trajanje govora glede na spol in vrsto govorcev.

Govorniki	Moški	Ženske	Moški + ženske
<b>Politiki</b>	44 min in 51 s	37 min in 51 s	82 min in 42 s
<b>Voditelji</b>	20 min in 35 s	15 min in 23 s	35 min in 58 s
<b>Hkratni govor</b>	/	/	3 min in 55 s
<b>Skupaj: 122 min in 35 s (2 uri, 2 minuti in 35 sekund)</b>			

### 1.2.2 Govorniki

V tabeli 3 lahko poleg šifer 15 govorcev (8 moških in 7 žensk) in trajanja analiziranega govora razberemo še naslednje demografske podatke: letnico rojstva in spol govornika, delo, ki ga opravlja (zaposlitev in izobrazba), kje je govornik preživel otroštvo oz. mladost in kje je obiskoval srednjo šolo, kje je študiral ter kje živi in dela. Gre za 3 politike in 3 političarke (vsak/a v 2 različnih oddajah) ter 5 voditeljev (4 so nastopili v 1 oddaji, 1 pa v 2 oddajah) in 4 voditeljice (vse samo v 1 oddaji). Vsi politiki se kot govorniki pojavijo v dveh oddajah, vsi voditelji (razen voditelja *GovNm6/GovNm13*) pa v eni oddaji. Prvi jezik vseh govorcev je slovenščina, vsi imajo univerzitetno izobrazbo družboslovne smeri, rojeni so med letoma 1941 in 1970, živijo in delajo predvsem v Ljubljani (nekateri so se vmes zaradi dela preselili v tujino), otroštvo in mladost so preživeli v različnih krajih po Sloveniji, nimajo govornih napak in v nevtralnem govornem položaju normalno poudarjajo in členijo govor; zaradi družbene (politiki) in medijske (voditelji) prepoznavnosti jih lahko uvrstimo med govorce z razmeroma velikim (tudi govornim) vplivom na gledalce oz. poslušalce. V analiziranih oddajah govorniki uporabljajo knjižni jezik, kot ga definira Toporišič (SS 2000: 15–19), in sicer voditelji TV-oddaj predvsem zbornega (branje oz. po spominu obnavljanje vprašanj in predstavitve gostov), politiki pa splošno- oz. knjižnopogovornega, ko v (delno) spontane odgovore na voditeljeva vprašanja vnašajo kratke nedoločnike, redukcije glasov idr.

**Tabela 3:** Podatki o govornicah.

Št.	Šifra govorca	Čas trajanja govora	Leto rojstva	Spol	Zaposlitev, izobrazba	Otroštvo, SŠ, VŠ, delo <sup>11</sup>
<i>Politiki oz. gosti oddaj</i>						
1.	GovPm3, GovPm12	16 min in 5 s	1958	M	politik, obramboslovec	Grosuplje, Stična, Ljubljana, Ljubljana
2.	GovPm5, GovPm16	13 min in 32 s	1955	M	politik, obramboslovec	Kranj, Kranj, Ljubljana, Ljubljana/Bruselj
3.	GovPm10, GovPm21	15 min in 14 s	1962	M	politik, družboslovec	Destričnik, Ptuj, Ljubljana, Maribor/ Ljubljana/Bruselj
4.	GovPz1, GovPz14	12 min in 3 s	1941	Ž	političarka, publicistka, filozofinja, sociologinja	Žiri, Škofja Loka, Ljubljana, Ljubljana
5.	GovPz7, GovPz20	13 min in 59 s	1941	Ž	političarka, pravnica	Ljubljana, Ljubljana, Ljubljana, Ljubljana
6.	GovPz9, GovPz18	11 min in 49 s	1957	Ž	političarka, pedagoginja	Bučkovci, Maribor, Ljubljana, Ljubljana
<i>Voditelji oddaj</i>						
7.	GovNm2	3 min in 23 s	1948	M	TV-voditelj, urednik, novinar	Stična na Dolenjskem, Stična, Ljubljana, Ljubljana
8.	GovNm4	2 min in 13 s	1970	M	TV-voditelj, urednik, novinar	Ljubljana, Ljubljana, Ljubljana, Ljubljana
9.	GovNm6, GovNm13	6 min in 25 s	1963	M	TV-voditelj, urednik, novinar	Štajerska, Štajerska, Ljubljana, Ljubljana
10.	GovNm11	5 min in 12 s	1967	M	TV-voditelj, urednik, novinar	Ljubljana, Ljubljana, Ljubljana, Ljubljana
11.	GovNm19	3 min in 22 s	1960	M	TV-voditelj, publicist, filozof	Logatec, Postojna, Ljubljana, Ljubljana
12.	GovNz8	3 min in 27 s	1965	Ž	TV-voditeljica, novi- narka	Podčetrtek, Celje, Ljubljana, Ljubljana
13.	GovNz15	3 min in 56 s	1965	Ž	TV-voditeljica, urednica, novinarka, zgodovinarica	Rače, Maribor, Ljubljana, Ljubljana
14.	GovNz17	3 min in 58 s	1966	Ž	TV-voditeljica, urednica, novinarka	Črnomelj, Črnomelj, Ljubljana, Ljubljana/ Zagreb
15.	GovNz22	4 min in 2 s	1969	Ž	TV-voditeljica, urednica, novinarka	Moravske Toplice, Ljutomer, Ljubljana, Ljubljana

<sup>11</sup> Kraji, v katerih je govorec preživel otroštvo oz. obiskoval osnovno šolo (*otročvo*), srednjo šolo (SŠ), kje je študiral (VŠ) in kje živi oz. je zaposlen po študiju (*delo*).



### 1.2.3 Transkripcija gradiva

V programu *Audacity*<sup>12</sup> smo ločili sliko in zvok, saj programi za analizo govora ne podpirajo slikovnih datotek, posnetke smo ustrezno razrezali in v digitalni obliki shranili v formatu .wav, stereo, frekvenca vzorčenja 22.050 Hz, resolucija 32 bitov.

Odvisno od namena raziskave se v jezikoslovju pogosto uporabljata predvsem dve transkripcijski orodji – *Transcriber*<sup>13</sup> v korpusnem jezikoslovju (npr. Verdonik in Zwitter Vitez 2011) in *Praat*<sup>14</sup> v fonetičnih raziskavah (npr. Volk 2011). *Transcriber* je bil že v osnovi zasnovan za transkribiranje TV-oddaj, saj omogoča zelo dobro pove-zavo med zvočnim posnetkom in transkripcijo ter vnosa metaoznak. Njegova prednost je nemoteno delovanje tudi pri daljših posnetkih (celo večurnih), vendar za potrebe fonetičnih raziskav ni najbolj uporaben, saj ne ponuja možnosti akustičnih analiz in zapisa hkratnega govora več kot dveh govorcev, s čimer je omejena možnost zapisovanja prekrivnega govora.

Po drugi strani *Praat* poleg transkribiranja ponuja tudi možnost akustične oz. fonetične analize govora (raziskovanje trajanja glasov, F0, jakosti, formantov, premorov idr.) ter zapisa hkratnega (spontanega) govora več kot dveh govorcev. Njegove prednosti so še preprosto predvajanje zvočnega posnetka in segmentiranje signala na manjše enote ter preprost prenos transkribiranega besedila v poljubni urejevalnik besedil, kjer lahko transkripcijo še naprej urejamo. Ima pa tudi nekaj pomanjkljivosti: zaradi številnih funkcij akustične analize govora ni najbolj uporabniško prijazen za transkribiranje, ne omogoča vnosa podatkov o govorcih in diskurzih ter vnosa meta-oznak, izhodna datoteka ni v XML-formatu, ampak v posebni *Praatovi* obliki, ki je sicer široko podprta v drugih orodjih (Verdonik in Zwitter Vitez 2011: 42–3).

Glede na predstavljene prednosti in pomanjkljivosti smo za transkripcijo in segmentacijo govora ter fonetično analizo poudarkov izbrali program *Praat*, ki ponuja možnost (delne) fonetične transkripcije in segmentacije spontanega in hkratnega go-vora več govorcev ter fonetično analizo govora. Pred transkripcijo smo posnetke iz sterea več govorcev lažje analize pretvorili v mono zvočni zapis, program pa pripravili za zapis govora po posameznih govorcih ter za označevanje poudarkov in drugih značilnosti posnetka (hkratnega govora, izrezanih delov oddaje ipd.).

#### 1.2.3.1 Obogatena ortografska transkripcija

Za različne raziskovalne namene so primerne različne vrste transkripcij. Glede na priporočila pobude EAGLES<sup>15</sup> je za *ortografsko* oz. *grafemsko transkripcijo* značilen standardni zapis govora z upoštevanjem nekaterih redukcij, dialektalnih oblik ter iz-

<sup>12</sup> Audacity – zvočni montažni program in snemalnik: <http://audacity.sourceforge.net>.

<sup>13</sup> Transcriber – program za prepisovanje govornega signala: <http://transag.sourceforge.net>.

<sup>14</sup> Praat – program za prepis in analizo govora: <http://www.praat.org>.

<sup>15</sup> EAGLES: <http://www.ilc.cnr.it/EAGLES96/home.html#ancora>.

pisovanjem števil in kratic. Za *citatno-fonemsko transkripcijo* je značilen fonemski zapis, ki izhaja iz zapisa besede v izolaciji (ne iz konteksta), za *širšo fonetično transkripcijo* je značilen fonemski zapis v kontekstu, medtem ko v *ožji fonetični transkripciji* označujemo tudi variante fonemov. V okviru *akustično-fonetične transkripcije* ločujemo vse segmente govora, ki so prepoznavni kot samostojni segmenti na sliki akustičnega valovanja ali na spektrogramu (Zemljarič Miklavčič 2008: 105–06).

Poleg naštetih osnovnih transkripcijskih možnosti obstaja več hibridnih možnosti zapisovanja govora, ko na podlagi raziskovalnih potreb oz. namena uporabe korpusa posamezne transkripcijske možnosti prilagodimo oz. vzamemo iz posameznih oblik transkripcije tiste možnosti, ki jih potrebujemo v posamezni raziskavi (Žganec Gros idr. 2003: 54). Tako smo se odločili za t. i. *obogateno oz. razširjeno ortografsko transkripcijo brez ločil*, kar pomeni, da govor zapisujemo v skladu z ustaljenim, standardnim zapisom – sicer brez ločil in velikih začetnic –, hkrati pa v transkripciji označujemo še mesta poudarkov ter nekatere druge (predvsem pogovorne) značilnosti govora, s čimer se delno prilagodimo govorjenemu jeziku (zapis, ki je čim bolj podoben dejanskemu izgovoru in hkrati čim bolj berljiv). V korpusu označujemo tudi identitete posameznih govorcev, hkrati govor in menjavanje govorcev, zapisujemo polleksikalne (*eee, mmm* idr.) in neleksikalne enote (smeh, kašljanje idr.), različna samopopravljanja ter (nerazumljive) fragmente. Pri transkripciji govora tako po eni strani poskušamo upoštevati načelo berljivosti (da bi bila transkripcija čim bolj berljiva, mora vsebovati čim manj metajezikovnih oznak), po drugi strani pa želimo označiti tudi raziskovani fizični pojav (poudarek) ter druge (meta)jezikovne značilnosti, ki vplivajo na zaznavanje raziskovane besedilnofonetične prvine (npr. podaljševanje glasov). V sami transkripciji rezultatov ne interpretiramo, ampak to naredimo pozneje v okviru analize označenih poudarkov. Posebnosti obogatene ortografske transkripcije, uporabljene v raziskavi, so:

- Brez uporabe velikih začetnic oz. z velikimi začetnicami zapisujemo le osebna imena in priimke ter druga lastna imena, npr. imena držav, prebivalcev, političnih strank in revij.
- Brez uporabe ločil, s čimer se izognemo subjektivni interpretaciji pri členitvi besedil oz. zapisu govora; nekatera ločila imajo druge, neskladenske vloge.
- Številke so zapisane z besedo.
- Kratice so zapisane, kot so izgovorjene, vendar skupaj, če gre za eno kratico, npr. *Esde* in *Sədə* za SD.
- Besede, ki so izgovorjene standardno, so zapisane v standardnem zapisu, npr. *delal* za izgovor /'de:laʉ/; izjeme so nekateri polglasniki (gl. naslednjo alinejo).
- Besede, ki so izgovorjene pogovorno, so zapisane tako, kot jih slišimo, vendar v okviru veljavnega črkopisa: a) polglasnik namesto drugega glasu v standardnem izgovoru ali v primerih, ko ni del standardnega izgovora, npr. sevada (seveda); b) končni *-u* namesto *-il/-el* v deležnikih, npr. želu (želel); c) končni *-i* namesto *-aj*: tuki (tukaj); č) končni *-ou* namesto *-el*, npr. prišou (prišel); d) različni izpadi samoglasnikov (redkeje soglasnikov), npr. lejte (glejte), tolko/tolk (toliko); e) raba *-e/-ə* namesto *-a-*, npr. zdej/zdəj (zdaj); f) izpadi in hkrati vrivanja drugih glasov,

npr. jes/jəs/jəst (jaz).

- Podaljšane glasove zapisujemo z *oo*, *eee*, *aaaa*, *sssss* idr., število črk pa je odvisno od dolžine trajanja glasov.

#### 1.2.4 Segmentacija gradiva

Osnovno enoto segmentacije smo poimenovali *segment*,<sup>16</sup> ki ga razumemo kot govorni tok oz. jezikovni niz, na obeh straneh zamejen s premorom, ki ga ne glede na njegovo dolžino izgovori isti govorec. Osnovni in edini pogoj za določanje segmentov je obstoj premora, kar pomeni, da lahko segment tvori tako osamljena beseda (npr. »iiiiinn«, Besedilo 2S, GovPm3) kot tudi več stavkov, ki so sicer deli različnih povedi (npr. »**povPREčnega evropskega** razvoja zdaj smo na osemdesetih odstotkih se pravi **zmanjšal** smo tudi«, Besedilo 2S, GovPm3). Tako razumljen segment je v raziskavi osnova za raziskovanje poudarkov. Znotraj enega segmenta:

- se lahko zgodi, da *ni označen noben slušno zaznan poudarek*, je pa v tem primeru poudarek največkrat slušno zaznan na besedi, ki je del drugega segmenta v okviru skupine segmentov, ki ji pripada raziskovani segment (npr. »SKLOpu pravil igre //<sup>17</sup> v kaTERih tudi mi tekmuje«, Besedilo 2S, GovPm3). Kljub vsemu sta lahko v segmentu brez slušno zaznanega poudarka označena po en jakostno najbolj izpostavljen zlog in/ali en zlog, ki smo mu v okviru segmenta ali skupine segmentov izmerili najvišji F0 (npr. »SKLOpu pravil igre«, 82,22 dB, Besedilo 2S, GovPm3);
- je lahko označen *en slušno zaznan poudarek*. V tem primeru sta znotraj tega segmenta označena po en jakostno najbolj izpostavljen zlog in en zlog, ki smo mu v okviru segmenta izmerili najvišji F0, ni pa nujno, da so vse tri kategorije označene na istem zlogu oz. celo znotraj iste besede (npr. »z **močno** gospoDARsko delegacijo v Rusiji srečanje«, 81,43 dB, 123,33 Hz, Besedilo 2S, GovNm4);
- je lahko označenih *več slušno zaznanih poudarkov*, in sicer tako v primeru, ko segment ni razdeljen na dele segmenta (npr. »tako DA v tem trenutku v poslanskem klubu ni več ne zadnjega predsednika stranke in vlade«, Besedilo 3S, GovNm6), kot tudi v primeru, ko je segment sestavljen iz več delov segmenta (npr. »**dober veČER** / z **JELkom** Kacinom se bomo **pogovarjali** / **zakaj** je razpadlo nekoč mogočno politično podjetje Liberalna **demokraCIja**«, Besedilo 3S, GovNm6). V drugem primeru je lahko znotraj tega segmenta označenih po več jakostno najbolj izpostavljenih zlogov in več zlogov, ki smo jim v okviru posameznih delov segmenta izmerili najvišji F0.

Segment je lahko sestavljen iz več *delov segmenta*, če znotraj segmenta obsta-

<sup>16</sup> V jezikoslovju se uporabljajo različna poimenovanja za osnovne enote segmentacije govora, npr. *vloga*, še manjši del pa *izjava*. Izjava je osnovna enota govora, ki približno ustreza pojmu povedi v pisnem jeziku in je prozodično, semantično ter skladenjsko prepoznavna enota. Vloga pomeni govor enega govorca, dokler ga ne prekine in prevzame vlogo govornika drug govorec, sestavljena pa je lahko iz ene ali več izjav (npr. Kranjc 1999: 63). Namesto izjave se pogosto uporablja tudi *izrek*, ki je opredeljen kot poved s komunikacijsko funkcijo (npr. Volk 2011). Za *segment* smo se odločili zaradi specifične segmentacije, ki je opisana v nadaljevanju.

<sup>17</sup> // – razmejuje segmente, / – razmejuje dele segmentov.

ja več pomensko in/ali strukturno povezanih delov segmenta, npr. stavkov, ki imajo vsak svojo glagolsko obliko, vrinjenih (lahko tudi brezglagolskih) razlag, (pre)dolгих segmentov idr. (npr. »ne more *imet* na ČElu **predsednikaa** / v katerega se bo **KDORKoli** bo imel **potrebo zaletaval** / in **stranki** KARkoli *očital* za nazaj«, Besedilo 3S, GovNm5). Del segmenta je torej lahko osnovna enota za raziskovanje poudarkov, če je govor med dvema premoroma (segment) sestavljen iz več pomensko in/ali strukturno povezanih delov segmenta in če je v tem delu segmenta bil slušno zaznan poudarek. Več segmentov lahko tvori *skupino segmentov*, če jih družijo skupen slušni poudarek in/ali pomenska in/ali strukturna povezanost, premori med segmenti pa niso predolgi, da bi razbili pomensko in/ali strukturno povezanost govornega toka (npr. »favorita // in // po **DOLgem** oklevanju«, Besedilo 6S, GovPm12).

Transkripcijo v Praatu smo začeli s prepisovanjem daljših, tj. 60-sekundnih delov govora. Sledilo je označevanje poudarkov – besede s slušno (avditivno) zaznamimi poudarki so v transkripciji označene **krepkeje**. Označevanje poudarkov v korpusu je potekalo ročno in s pomočjo nekaterih avtomatskih pripomočkov v Praatu. Sledilo je razmejevanje (segmentacija) na krajše segmente s pomočjo premorov. Premor razumemo kot vsakršno prekinitev govornega toka; lahko so *tih* (*prazni*) ali *zapolnjeni* z različnimi vrstami zvočnega signala (podaljševanje glasov, hkratni govor, mašila, neverbalni dogodki, npr. smeh). V transkripciji so označeni premori, ki so dolgi 50 ms ali več. Segmentaciji je sledilo ponovno poslušanje krajših segmentov ter preverjanje in po potrebi korigiranje že označenih poudarkov.

V tako pripravljene gradivu smo izvedli meritve jakostno in višinsko izpostavljenih delov posameznih segmentov, delov segmentov oz. skupin segmentov. Zlogi, ki so glede na meritve v Praatu izgovorjeni z največjo jakostjo (intenziteto) v posameznem segmentu, so v transkripciji označeni z VELIKIMI TISKANIMI ČRKAMI, rezultati meritev pa so navedeni v decibelih (dB). Zlogi, ki imajo glede na meritve v Praatu izmerjeni najvišji F0 v določenem segmentu, so v transkripciji zapisani *ležeče* (*in podčrtano*), rezultati meritev pa so navedeni v hercih (Hz), npr.: »**dober veČER** spoštovani **gledalcii** mojaa nocojšnjaa« (82,17 dB, **150,12 Hz**; Besedilo 1S, GovNm2).

Če je bilo na podlagi slušne (avditivne) analize v govoru označenih več poudarkov znotraj enega segmenta oz. znotraj več segmentov, sta v končni računalniški analizi najpogosteje upoštevana le po en jakostni in en frekvenčni vrh. Slušni poudarek, največja jakost in najvišji F0 so lahko označeni v istih zlogih iste besede (npr. »**PRED** in po zadnjih volitvah ne«, Besedilo 6S, GovPm12), v različnih zlogih iste besede (npr. »prestopniki z Liberalne **demokraciJE**«, Besedilo 6S, GovPm12) ali v različnih besedah istega segmenta (npr. »se *tega* SOočjenja ne **bojimo**«, Besedilo 6S, GovPm12) ali različnih segmentov (npr. »gre ZA // predlog **sporazuma**«, Besedilo 6S, GovPm12).

Opisan potek transkripcije in segmentacije 60 sekund posnetka je v povprečju

trajal ok. 2 uri, kar ob nekaj več kot 2-urnem gradivu pomeni ok. 250 ur transkripcije in segmentiranja govora v Praatu.<sup>18</sup>

## 2 Besednovrstna predvidljivost poudarkov

Mnenja raziskovalcev o besednovrstni predvidljivosti poudarjanja so deljena, saj nekateri zagovarjajo tezo, da je normalno in nezaznamovano poudarjanje sestavni del slovnice določenega jezika, drugi pa so prepričani, da je poudarjanje povsem odvisno od konteksta in situacije izrekanja ter zato povsem nepredvidljivo. Quirk idr. (1985) menijo, da so najpogosteje poudarjeni samostalniki, šele nato glagoli, prislovi, pridevniki in zaimki, Đorđević (1984) ugotavlja, da so največkrat poudarjeni števnik, sledijo prislovi, pridevniki, glagoli, samostalniki in zaimki, medtem ko je Ž. Matulina Jerak (1989) prepričana, da so pojmi, ki opredeljujejo, bolj poudarjeni od pojmov, ki se opredeljujejo, kar pomeni, da so poudarjeni predvsem pridevniki, zaimki in glagoli. V splošnem naj bi za vse jezike veljalo, da je poudarjanje močno odvisno od besedne vrste, ki ji beseda pripada, saj niso vse besedne vrste enako verjetni nosilci poudarka. Tako Cruttenden (1986) v svoji raziskavi ugotavlja, da je za stavčnočlenske besede (besede odprtega razreda) značilno, da so bolj pričakovano v jedru kot slovnične (besede zaprtega razreda).

Ivas (1993: 305–06) je v raziskavi hrvaškega govora ugotovil, da so bili v TV-odajah, v katerih so nastopali politiki, največkrat poudarjeni samostalniki (26,6 %), sledijo pridevniki (23,9 %), glagoli (15,6 %), prislovi (13,7 %), zaimki (12,4 %), števnik (2,7 %), vezniki (2,5 %) in predlogi (2,3 %). Dobljene odstotke je primerjal s skupnimi odstotki zastopanosti določenih besednih vrst v vsem analiziranem korpusu (samostalnikov je bilo 25,4 %, glagolov 16,9 %, zaimkov 14,5 %, pridevnikov 12,5 %, veznikov 11,2 %, predlogov 9 %, prislovov 8,4 % in števnikov 1,7 %) in zaključil, da so glede na skupno število pojavitev v korpusu najmanjkrat poudarjeni vezniki in predlogi, največkrat pa pridevniki. Šuštaršič (1995: 164–78) v svoji primerjalni analizi slovenščine in angleščine ugotavlja, da je v slovenščini poudarek pogosto na prislovih, nedoločnih in osebnih zaimkih ter na glagolih. Toporišič (SS 2000: 448, 541–43) v slovnici omenja, da je v slovenščini težišče pogosto na členkih, saj stojijo pred tistim delom stavka, ki ga posebej poudarjajo.

Ker so mnenja glede besednovrstne predvidljivosti poudarjanja različna, smo preverjali, v kolikšni meri so glede na besedno vrsto predvidljivi poudarki v analiziranem korpusu. V raziskavi je vsaka poudarjena beseda, ki je bila slušno zaznana, obravnavana kot samostojna enota, kar pomeni, da so v primeru poudarjenosti besed, kot so *bom*, *je* ali *se*, tudi te besede samostojno uvrščene med glagole ipd. Besede smo razvrstili v dve osnovni skupini, in sicer govorimo o *stavčnočlenskih*

<sup>18</sup> Za primerjavo: ortografska transkripcija (torej brez označevanja fonetično-fonoloških značilnosti govora) ene minute posnetka, vključenega v *Korpus govornje slovenščine GOS*, je trajala ok. 30 minut (Verdonik in Zwitter Vitez 2011: 44).

ter *nestavčnočlenskih* oz. *slovnično-modifikacijskih besedah*<sup>19</sup> (tabela 4). Preverjamo hipotezo, da med stavčnočlenskimi besednimi vrstami vsi analizirani govorniki poudarjajo predvsem samostalnike, med nestavčnočlenskimi pa členke.

**Tabela 4:** Besednovrstna predvidljivost poudarkov – simboli.

Besedne vrste in simboli		
stavčnočlenske besede (BS)	samostalniki (BSS)	
	pridevniki (BSPd)	
	prislovi (BSPr)	
	glagoli (BSG)	
	zaimki (BSZ)	
	števniki (BSŠ)	
nestavčnočlenske besede (BN)	slovnične besede (BNS)	predlogi (BNSP)
		vezniki (BNSV)
	modifikacijske besede (BNM)	členki (BNMČ)
		medmeti (BNMM)

V korpusu *Korp-2MeGo* je poudarjena 4401 beseda od 16.629 izgovorjenih, kar pomeni, da je poudarjenih 26,5 % vseh izgovorjenih besed v korpusu: 31,6 % (3856) vseh stavčnočlenskih in 12,3 % (545) vseh nestavčnočlenskih besed. Izmed posameznih besednih vrst dosegajo najvišje odstotke poudarjenosti prislovi (poudarjenih je 40,8 % vseh prislovov v korpusu oz. 524 od 1284), samostalniki (38,6 % oz. 1401 od 3633), pridevniki (37,1 % oz. 488 od 1316), števniki (33,0 % oz. 93 od 282), zaimki (31,2 % oz. 539 od 1729) in glagoli (20,5 % oz. 811 od 3948), torej vse stavčnočlenske besedne vrste. Precej manj pogosto so poudarjene nestavčnočlenske besede, še najpogosteje členki, tj. modifikacijska besedna vrsta (20,1 % oz. 265 od 1321). Preostale nestavčnočlenske besedne vrste si sledijo takole: vezniki (12,3 %), medmeti (11,5 %) in predlogi (5,4 %). Glede na predstavljene rezultate smo samo delno potrdili hipotezo, saj med nestavčnočlenskimi besednimi vrstami govorniki res najpogosteje poudarjajo členke, med stavčnočlenskimi besednimi vrstami pa so samostalniki z 38,6 % takoj za prislovi (40,8 %) na drugem mestu (tabela 5).

<sup>19</sup> Delitev besednih vrst glede na skladijsko vlogo je prilagojeno povzeta po A. Vidovič Muha (1984: 144, 2013: 36); zaradi primerjav s tujimi raziskavami npr. ne govorimo o samostalniških in pridevniških besedah, ampak o samostalnikih in pridevnikih, posebej so izločeni zaimki in števniki, izpuščeni pa so povedkovniki.

Tabela 5: Besednovrstna predvidljivost poudarkov – rezultati.

	<b>BSS</b>	<b>BSPd</b>	<b>BSPr</b>	<b>BSG</b>	<b>BSZ</b>	<b>BSŠ</b>	<b>BNSP</b>	<b>BNSV</b>	<b>BNMČ</b>	<b>BNMM</b>	<b>BS</b>	<b>BN</b>	<b>BS+BN</b>
	%	%	%	%	%	%	%	%	%	%	%	%	št.
<b>Pm</b>	39,6	33,4	35,9	21,2	30,0	29,0	6,6	11,3	19,8	9,1	31,1	11,8	1550
<b>Pz</b>	37,9	48,3	43,8	19,4	30,6	42,7	4,5	11,5	22,1	22,2	31,7	12,9	1443
<b>Pm+Pz</b>	38,9	39,3	39,6	20,3	30,3	35,4	5,6	11,4	21,1	15,0	31,4	12,4	2993
<b>Nm</b>	33,5	31,9	39,5	20,0	30,5	32,9	6,3	13,4	12,3	0,0	29,7	10,3	739
<b>Nz</b>	44,0	32,7	50,0	22,4	37,0	21,6	3,0	16,9	25,3	0,0	35,2	14,5	669
<b>Nm+Nz</b>	38,0	32,2	43,8	21,1	33,3	29,0	4,9	14,9	17,9	0,0	32,1	12,1	1408
<b>Pm+Nm</b>	37,4	32,9	37,1	20,8	30,1	30,7	6,5	11,9	17,1	7,1	30,7	11,3	2289
<b>Pz+Nz</b>	40,1	43,5	45,4	20,2	32,3	36,1	4,0	12,8	23,0	16,7	32,8	13,3	2112
<b>P+N</b>	38,6	37,1	40,8	20,5	31,2	33,0	5,4	12,3	20,1	11,5	31,6	12,3	4401

Med govorkami in govorci ni zaslediti večjih razlik, saj so govorci poudarili 25,6 % (2289 od 8932) vseh izgovorjenih besed, govorka pa 27,4 % (2112 od 7697). Nekoliko večje razlike se kažejo znotraj posameznih besednih vrst, in sicer predvsem pri pridevniki (govorka so poudarile 43,5 % vseh izgovorjenih pridevnikov, govorci le 32,9 %), členkih (govorka 23,0 %, govorci 17,1 %), prislovi (govorka 45,4 %, govorci 37,1 %), števnikih (govorka 36,1 %, govorci 30,7 %) ter medmetih (govorka 16,7 %, govorci 7,1 %). V celotnem korpusu nekoliko pogosteje poudarjajo voditelj/ice (27,0 %) v primerjavi s politiki/političarkami (26,2 %), in sicer predvsem voditeljice (30,0 %), manj politiki (26,1 %) in političarke (26,4 %) ter voditelji (24,7 %). Pri poudarjanju stavčnoočlenskih in nestavčnoočlenskih besed med politiki/političarkami in voditelji/cami ni večjih razlik, še največje so pri poudarjanju voditeljev (stavčnoočlenske 29,7 %, nestavčnoočlenske 10,3 %) in voditeljic (stavčnoočlenske 35,2 %, nestavčnoočlenske 14,5 %). V okviru posameznih besednih vrst velja pri kategoriji voditelj/ice izpostaviti predvsem prislove (voditeljice 50,0 %, voditelji 39,5 %), samostalnike (voditeljice 44,0 %, voditelji 33,5 %), zaimke (voditeljice 37,0 %, voditelji 30,5), členke (voditeljice 25,3 %, voditelji 12,3 %) in števnik (voditelji 32,9 %, voditeljice 21,6 %), pri kategoriji politiki/političarke pa pridevnike (političarke 48,3 %, politiki 33,4 %), števnik (političarke 42,7 %, politiki 29,0 %) in prislove (političarke 43,8 %, politiki 35,9 %). V povprečju velja, da voditelji raje poudarijo kakšno besedo manj, voditeljice pa kakšno več, medtem ko pri politikih in političarkah ni večjih razlik, saj so vsi poudarili od 23,9 % do 27,7 % vseh izgovorjenih besed v korpusu. Podobne ugotovitve veljajo tudi ločeno po stavčnoočlenskih in nestavčnoočlenskih besedah.

Ivas (1993) svojo raziskavo hrvaškega medijskega govora sklene z ugotovitvijo, da so glede na skupno število pojavitev v korpusu najmanjkrat poudarjeni vezniki in predlogi, največkrat pa pridevniki in negacije. Glede na rezultate, pridobljene v predstavljeni raziskavi, se lahko povsem strinjamo z Ivasom glede predlogov, ki so tudi v predstavljeni raziskavi poudarjeni najmanjkrat (5,4 %), prav tako so vezniki med najmanj pogosto poudarjenimi (12,3 %), nekoliko drugače pa je s pridevniki, ki so v hrvaščini najpogosteje poudarjeni, v naši raziskavi pa so s 37,1 % na tretjem mestu, in sicer z majhnim zaostankom za prislovi (40,8 %) in samostalniki (38,6 %).

O pogosto poudarjenih prislovi v slovenščini poroča tudi Šuštaršič (1995), ki je do tega podatka prišel na podlagi nekoliko drugačnih metodoloških izhodišč (primerjalna analiza slovenščine in angleščine), kljub vsemu pa se v obeh raziskavah nakazuje tendenca o zelo pogostem poudarjanju prislovov v slovenščini. S Toporišičem (SS 2000) se lahko strinjamo, da so v slovenščini razmeroma pogosto poudarjeni tudi členki, saj so glede na opravljeno raziskavo najpogosteje poudarjana nestavčnoočlenska besedna vrsta v korpusu *Korp-2MeGo*.

### 3 Za naprej ...

Besedilnofonetične značilnosti so v slovenskem jezikoslovju razmeroma slabo raziskane. Z raziskavo, predstavljeno v prispevku in v Huber 2013, ki temelji na



opisanem metodološkem pristopu, smo sicer prišli do marsikatero nove ugotovitve v slovenski besedilni fonetiki, še vedno pa to področje ponuja veliko možnosti za nadaljnja raziskovanja. Tako bi bilo pri poudarkih smiselno še temeljiteje raziskati njihove vloge, ki jih opravljajo v govoru, povezave z besednim redom, intonacijo, hitrostjo govora, inherentnimi vrednostmi glasov, ne nazadnje bi bilo zanimivo raziskati tudi povezavo med poudarki in neverbalno komunikacijo, pri dodatnem potrjevanju rezultatov pa nam lahko pomagajo tudi perceptivni testi, izvedeni na večjem številu oseb. Raziskava bi lahko skupaj z nadaljnjimi analizami poudarkov in drugih besedilnofonetičnih značilnosti pripomogla k (pre)oblikovanju teoretičnih besedilnofonetičnih izhodišč, ki bi veljala za različne vrste govornih položajev.

#### VIRI IN LITERATURA

- Anton BREZNIK, 1908: Besedni red v govoru. *Dom in svet* 21. 222–30, 258–67.
- Alan CRUTTENDEN, 1986: *Intonation*. Cambridge: University Press.
- Tomáš DUBĚDA, 2005: *Jazyky a jejich zvuky: Univerzálie a typologie ve fonetice a fonologii*. Praga: Karlova univerza.
- Branivoj ĐORĐEVIĆ, 1984: *Gramatika srpskohrvatske dikcije sa praktikumom*. Beograd: Univerzitet umetnosti.
- Damir HORGA, 1996: *Obrada fonetskih obavijesti*. Zagreb: Znanstvena biblioteka hrvatskoga filološkog društva.
- Nataša HRIBAR, 2001/2002: Vprašanje skladenjske zapletenosti povedi v govornem jeziku politikov. *Jezik in slovstvo* 47/7–8. 315–29.
- Damjan HUBER, 2013: *Poudarek in pavza v standardnem slovenskem govoru: Doktorska disertacija*. Ljubljana: FF.
- Ivan IVAS, 1993: *Izričajna jezgra u hrvatskom jeziku: Doktorski rad*. Zagreb: FF.
- Simona KRANJC, 1999: *Razvoj govora predšolskih otrok*. Ljubljana: FF.
- Željka MATULINA JERAK, 1989: O »logičkom« akcentuiranju u njemačkoj rečenici. *Fonološki i fonetski aspekti govorenog jezika: Zbornik radova*. Ur. D. Horga. Zagreb: Društvo za primijenjenu lingvistiku Hrvatske. 57–67.
- Nenad MIŠČEVIĆ, 1983: *John Langshaw Austin. Jezik kot dejavnost*. Prev. E. Bahovec, M. Rotar in J. Šumić - Riha. Ljubljana: Univerzum (Analecta).
- Zdena PALKOVÁ, 1997: *Fonetika a fonologie češtiny*. Praga: Karlova univerza.
- Katarina PODBEVŠEK, 2006: *Govorna interpretacija literarnih besedil v pedagoški in umetniški praksi*. Ljubljana: Slavistično društvo Slovenije.
- Randolph QUIRK, Sidney GREENBAUM, Geoffrey LEECH, Jan SVARTVIK, 1985: *A Comprehensive Grammar of English Language*. London, New York: Longman.
- Jakob RIGLER, 1968: Problematika naglaševanja v slovenskem knjižnem jeziku. *Jezik in slovstvo* 13/6. 192–99.
- Ivo ŠKARIĆ, 1999: *V iskanju izgubljenega govora*. Prev. S. Stergaršek. Ljubljana: Šola retorike.
- Rastislav ŠUŠTARŠIČ, 1995: Naglas in poudarek v angleščini in slovenščini. *Slavistična revija* 43/2. 157–82.

- Hotimir TIVADAR, 1999: Fonem /v/ v slovenskem govornem knjižnem jeziku. *Slavistična revija* 47/3. 341–61.
- , 2009: Fonetika in realno gradivo – izbira(nje), pridobivanje, urejanje in analiza govornih medijskih besedil. *Infrastruktura slovenščine in slovenistike*. Ur. M. Stabej. Ljubljana: FF. 365–69.
- Jože TOPORIŠIČ, 1972: Členitev besedila s premori. *Slavistična revija* 20/1. 149–58.
- , <sup>1</sup>1976, <sup>2</sup>2000: *Slovenska slovnica*. Maribor: Obzorja.
- , 1992: *Enciklopedija slovenskega jezika*. Ljubljana: CZ.
- Alenka VALH LOPERT, 2005: *Kultura govora na Radiu Maribor*. Maribor: Slavistično društvo Maribor.
- Darinka VERDONIK, Ana ZWITTER VITEZ, 2011: *Slovenski govorni korpus Gos*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Tina VEROVNIK, 2004: Govorjeni knjižni jezik v televizijskih dnevnoinformativnih oddajah: Študija primera. *Poti slovenskega novinarstva – danes in jutri*. Ur. M. Poler Kovačič in M. Kalin Golob. Ljubljana: FDV. 157–73.
- Ada VIDOVIČ MUHA, 1984: Nova slovenska skladnja J. Toporišiča. *Slavistična revija* 32/2. 142–55.
- , <sup>2</sup>2013: *Slovensko leksikalno pomenoslovje*. Ljubljana: FF (Razprave).
- Primož VITEZ, 1999: Od idealnih jezikovnih struktur k strategiji realnega govora. *Slavistična revija* 47/1. 23–48.
- Jana VOLK, 2011: *Intonacija v spontanem neformalnem govoru slovenskih govorcev na dvojezičnem področju Slovenske Istre: Doktorska disertacija*. Ljubljana: FF.
- Jana ZEMLJARIČ MIKLAČIČ, 2008: *Govorni korpusi*. Ljubljana: FF.
- Jerneja ŽGANEC GROS, France MIHELČ, Simon DOBRIŠEK, 2003: Govorne tehnologije: Pridobivanje in pregled govornih zbirk za slovenski jezik. *Jezik in slovstvo* 48/3–4. 47–59.

## SUMMARY

The *Praat* program for speech analysis is used to analyze sentence stress from the point of view of text genre predictability. Sentence stress is understood as: a) the acoustic, semantic, syntactic, and wider prosodic emphasis of parts of units of speech at the sentence or text level, b) an expression of the informational organization of the text at a deep level, and c) an important marker and organizer of speech production. At the acoustic level, sentence stress can be determined by the intensity of speech, raised F0, extended duration of sounds, changed tempo of speech, and other factors.

For the purposes of this research, a corpus was compiled of ten discussion programs with political content on TV Slovenia: in other words, the authentic and relatively spontaneous media speech of selected politicians and TV presenters (three male and three female politicians, and five male and four female presenters), totaling 122 minutes and 35 seconds. In addition to the results of the analysis of sentence stress from the point of view of genre predictability, the article presents the methodological foundations for studying prosodic characteristics, such as corpus planning, collecting material, putting forward hypotheses, the transcription and segmentation of material, audio and computer analysis, and interpretation of the results.

In the corpus analyzed, 31.6% of lexical words and 12.3% of grammatical words are stressed. The most emphasized items are all lexical words: adverbs (40.8% of those in the corpus), nouns (38.6%), adjectives (37.1%), numerals (33.0%), pronouns (31.2%), and verbs (20.5%). Grammatical words are stressed far less often—most frequently particles (20.1%), and less often conjunctions (12.3%), interjections (11.5%), and prepositions (5.4%).