

USPOREDBA REZULTATA PRISTUPNIKA IZ RAZLIČITIH DIJALEKTALNIH REGIJA NA ZADACIMA VEZANIM UZ TEKSTOVE PISANE ČAKAVSKIM NARJEČJEM NA ISPITU DRŽAVNE MATURE IZ HRVATSKOG JEZIKA

Josip Šabić

Nacionalni centar za vanjsko vrednovanje obrazovanja
Petračićeva 4, 10000 Zagreb
josip.sabic@ncvvo.hr

Sažetak

Jedan od suvremenih izazova u području konstrukcije pismenih provjera znanja je razumijevanje razlika u uspjehu različitih skupina ispitanika na pojedinim zadacima. Razumijevanje ovih razlika je od posebne važnosti na ispitima visokog rizika kao što su ispiti državne mature. Cilj ovog rada bio je ispitati razlike u uspješnosti rješavanja zadataka državne mature iz Hrvatskoga jezika vezanih uz polazne tekstove pisane čakavskim narječjem među pristupnicima iz različitih dijalektalnih regija (čakavske, kajkavske i štokavske). Usporedbom prosječne riješenosti zadataka vezanih uz čakavske tekstove te analizama diferencijalnoga funkcioniranja zadataka (engl. *differential item functioning*) i diferencijalnoga funkcioniranja čvora zadataka (engl. *differential bundle functioning*) pronađene su vrlo male razlike u rezultatu u korist pristupnika iz čakavske regije u odnosu na pristupnike iz kajkavske i štokavske regije. Na temelju ovih rezultata, u stručnoj raspravi donesena je preporuka o nastavku uporabe izvornih tekstova na čakavskom, kajkavskom i štokavskom narječju na ispitima državne mature iz Hrvatskog jezika, pri čemu trebaju biti podjednako zastupljena sva narječja kako pristupnici ne bi bili u neravnopravnom položaju zbog pripadnosti svojega razgovornog jezika jednomu narječju.

Ključne riječi: državna matura, hrvatski jezik, čakavsko narječje, diferencijalno funkcioniranje zadataka, diferencijalno funkcioniranje čvora zadataka

UVOD

Razumijevanje razlika u uspjehu različitih skupina ispitanika na pojedinim zadacima je jedan od suvremenih izazova u konstrukciji pismenih provjera znanja (npr. Wyse i Mapuranga, 2009). Neki zadatak može funkcionirati različito za dvije skupine ispitanika iz različitih razloga. Ponekad se skupine razlikuju po svojstvu koje je relevantno za konstrukt koji se pismenom provjerom znanja pokušava za-

hvatiti. Npr. može se razlikovati nastavni plan i program po kojem su dvije skupine učenika pohađale nastavu nekog školskog predmeta, a pismena provjera ispituje gradivo tog predmeta po određenom nastavnom planu i programu. Može se očekivati kako će na nekim zadacima razlika u uratku tih dviju skupina ispitanika biti veća, a proizlaziti će iz toga što određeni zadaci ispituju dio gradiva kojemu jedna skupina učenika nije poučavana u školi. Zadatak također može osim konstrukta koji je željeni predmet mjerenja ili interesa istraživača mjeriti i neko drugo svojstvo koje je irelevantno (Zumbo, 1999). Primjerice, ako je neki problemski zadatak u pismenoj provjeri znanja iz matematike sastavljen vrlo kompleksnim rječnikom, tada taj zadatak ne mjeri samo matematičko znanje nego i znanje rječnika. Takav zadatak zbog toga može biti pristran prema određenim skupinama ispitanika, npr. prema ispitanicima kojima hrvatski jezik nije materinji ili prema ispitanicima koji ostvaruju niži rezultat na testu rječnika. Na uspješnost u rješavanju matematičkih problemskih zadataka mogu utjecati i drugi čimbenici poput roda i sociokulturalnih čimbenika (Hyde i Mertz, 2009). Ako se utvrdi da je neki zadatak pristran u odnosu na neku od skupina, tj. da diskriminira neku skupinu ispitanika prema svojstvu irelevantnom za predmet mjerenja, moguće je izostaviti rezultat ispitanika na tom zadatku pri formiranju ukupnoga rezultata i ne upotrebljavati taj zadatak u budućnosti. Također, moguće je iskoristiti tu informaciju u revidiranju specifikacija pismene provjere znanja za buduće primjene.

Ispiti državne mature iz Hrvatskoga jezika sastavljeni su prema ispitnom katalogu u kojem je opisana struktura ispita te područja i obrazovni ishodi koji se njima ispituju (Nacionalni centar za vanjsko vrednovanje obrazovanja, 2011). Viša razina ispita iz Hrvatskoga jezika usklađena je s nastavnim planom i programom za gimnazije (Ministarstvo kulture i prosvjete, 1994). U ispitu državne mature iz Hrvatskoga jezika na višoj razini školske godine 2011./2012., prvih deset zadataka je bilo vezano uz dva polazna teksta pisana čakavskim narječjem. Ako bi se našle razlike u uspješnosti rješavanja zadataka vezanih uz čakavske tekstove između pristupnika iz regija u kojima su dominantna različita narječja (čakavsko, kajkavsko i štokavsko), takav nalaz bi mogao potaknuti predmetne stručnjake na raspravu o mogućim promjenama ispitnog kataloga i strukture ispita, kako bi se u budućnosti ujednačio broj zadataka u ispitu koji su vezani uz tekstove pisane različitim narječjima. Naime, zadaci svakog ispita državne mature iz Hrvatskoga jezika za višu razinu predstavljaju uzorak iz mogućeg skupa svih zadataka koji ispituju različite obrazovne ishode predviđene ispitnim katalogom te nastavnim planom i programom. Za svaki se ispit zadatke bira na način da u određenom omjeru budu zastupljena pojedina područja ispitivanja. Ako se pokaže potreba, zadatke se u ispit može birati i na temelju nekih dodatnih kriterija, primjerice na način da u svakom ispitu bude određeni omjer zadataka vezanih uz čakavske, kajkavske i štokavske tekstove.

Različite skupine ispitanika mogu se razlikovati po prosječnoj riješenosti pojedinih zadataka te se ova razlika za svaki zadatak može izraziti veličinom učinka. Međutim, ako se uradak različitih skupina ispitanika na pojedinim zadacima

uspoređuje na temelju postotka riješenosti bez kontroliranja uratka ispitanika na čitavom ispitu, istraživač se izvrgava opasnosti od Simpsonova paradoksa (1951). Ovaj paradoks se događa kada skupina A ima veći postotak riješenosti nekog zadatka nego skupina B, dok pripadnici skupine B s istim ukupnim rezultatom na ispitu kao pripadnici skupine A ipak imaju u prosjeku veću riješenost tog zadatka. Drugim riječima, na temelju postotka riješenosti zadatka činilo bi se da je skupina A u povoljnijem položaju, a taj zadatak zapravo diskriminira u korist skupine B.

Diferencijalno funkcioniranje zadatka (DFZ) je pojava kada ispitanici jednakih razina razvijenosti konstrukta koji se ispitom ispituje, a koji pripadaju različitim skupinama, imaju različitu vjerojatnost točnoga odgovaranja na zadatak (Zumbo, 1999). Pri analiziranju ove pojave, ispitanici na nekome ispitu se obično podijele u dvije skupine, npr. prema spolu, dobi ili prema regionalnoj pripadnosti. Zatim se uspoređuje uradak na pojedinim zadacima onih ispitanika iz dvije skupine koji imaju jednaku razinu razvijenosti konstrukta koji se mjeri, odnosno jednak ukupan rezultat na ispitu ili primarnoj dimenziji koju ispit mjeri. Uobičajena statistička metoda kojom se testira DFZ je Mantel-Haenszelov postupak (npr. Roussos, Schnipke i Pashley, 2000). Analize DFZ su uobičajene nakon provedbe velikih međunarodnih ili nacionalnih ispitivanja kao što su PISA i državna matura (npr. Le, 2009; Jokić, Ristić Dedić i Šabić, 2011).

Međutim, statistički pristup analizama pojedinačnih zadataka kako bi se utvrdilo njihovo diferencijalno funkcioniranje pati od nedostatka statističke snage, a također i od pogreške tipa I (Roussos i Stout, 1996). Drugim riječima, ako diferencijalno funkcioniranje u populaciji stvarno postoji, ponekad ga nećemo uspjeti potvrditi statističkim postupkom, a često se događa i obratno - da statističke analize pokazuju na postojanje DFZ i kada u populaciji razlike nema. Jedan je od razloga ovakvim nalazima to što svaki zadatak nema samo jednu nego mnoge karakteristike (tip zadatka, sadržaj, kognitivnu razinu, složenost leksika itd.) koje mogu utjecati na uspješnost rješavanja zadatka od strane različitih skupina. Kako se očekuje da će obrasci u kojima se javlja DFZ biti vidljiviji u skupovima zadataka nego na pojedinim zadacima, istraživači su domislili metode za analiziranje uratka ispitanika na skupovima zadataka koji imaju neku karakteristiku koja je potencijalno važna za objašnjenje DFZ. Gomilanjem zadataka koji imaju određenu karakteristiku povećava se vjerojatnost utvrđivanja je li ta karakteristika povezana s DFZ.

Metodologija koju su predložili Shealy i Stout (1993) sastoji se od konfirmatornog pristupa statističkom testiranju hipotetičkih izvora DFZ. Prema ovom pristupu, neki zadatak diferencijalno funkcionira zato što ne mjeri samo primarnu dimenziju (θ) koju bi ispit trebao mjeriti (npr. znanje iz predmeta Hrvatski jezik; tj. onu dimenziju koja se obično nakon primjene ispita boduje) nego i neku dodatnu dimenziju (η) koju se ispitom zapravo ne želi mjeriti. Kako zadatak ne mjeri samo jednu dimenziju, kaže se da su takvi podaci multidimenzionalni. DFZ se javlja kada se dvije skupine ispitanika razlikuju po distribucijama na dimenziji η uz kontrolu vrijednosti na dimenziji θ . Zadatak će diferencijalno funkcionirati za dvije skupine ako te dvije

skupine nakon izjednačavanja po primarnoj dimenziji imaju različite distribucije na sekundarnoj dimenziji. Usto, karakteristična krivulja zadatka (tj. krivulja koja prikazuje odnos razine θ kod ispitanika i vjerojatnosti točnog odgovaranja na zadatak) mora biti osjetljiva na postojanje sekundarne dimenzije. Oba navedena uvjeta moraju biti zadovoljena da bi došlo do pojave DFZ.

Kako bi se ispitalo jesu li skupovi zadataka iz nekog ispita povezani s nekom sekundarnom dimenzijom, Douglas, Roussos i Stout (1996) predlažu uporabu čvorova zadataka kao jedinica analize. Čvorovi zadataka (engl. *item bundles*) su skupovi zadataka koji imaju neku zajedničku karakteristiku za koju se smatra kako bi mogla predstavljati sekundarnu dimenziju. Diferencijalno funkcioniranje čvorova zadataka (DFČZ) tako predstavlja ekstenziju DFZ. Ovaj statistički postupak ima veću statističku snagu od analiza DFZ.

Shealy i Stout (1993) su autori metodologije i računalnog programa SIBTEST (engl. *The Simultaneous Item Bias Test*; Simultano testiranje pristranosti zadataka). U SIBTEST paradigmi koja je razvijena u okviru teorije odgovora na zadatak, čvor zadataka predstavlja dio ispita za koji se pretpostavlja kako mjeri primarnu i još neku sekundarnu dimenziju, a ostatak ispita je dio ispita koji mjeri samo primarnu dimenziju (Gierl, Bisanz, Bisanz, Boughton i Khaliq, 2001; Stout, Bolt, Froelich, Habing, Hartz i Roussos, 2003). Dvije skupine ispitanika se izjednačavaju prema uratku na ostatku ispita te se uspoređuje njihov uradak na odabranom čvoru zadataka. SIBTEST metodologijom se dobivaju stabilni i pouzdani nalazi čak i na relativno malim uzorcima (250 ispitanika u svakoj od skupina). Za usporedbu, Mantel-Haenszelov postupak se obično ne provodi ako u nekoj od skupina ima manje od 500 sudionika (npr. Mazor, Clauser i Hambleton, 1992). Preduvjet za provođenje SIBTEST procedure je jednodimenzionalnost ispita koja se može provjeriti, primjerice, metodom glavnih komponenti (npr. Deng, Wells i Hambleton, 2008). SIBTEST procedurom se izračunava ponderirana srednja razlika između dvije grupe ispitanika te se ta razlika statistički testira. Pomoću regresijske procedure za korekciju, aritmetičke sredine u ovoj proceduri se korigiraju zbog razlika u distribucijama primarne dimenzije među skupinama (Gierl, Gotzmann i Boughton, 2004).

Analize SIBTEST metodologijom mogu se provesti i na razini pojedinačnog zadatka. U SIBTEST proceduri, parametar diferencijalnog funkcioniranja β za pojedini zadatak se interpretira kao prosječna razlika u vjerojatnosti točnoga odgovora na zadatak između dva po slučaju odabrana ispitanika od kojih je jedan iz prve, a jedan iz druge skupine, pod uvjetom da su te dvije skupine izjednačene po rezultatu na primarnoj dimenziji, tj. konstrukt koji ispit mjeri. Primjerice, vrijednost parametra β od 0,1 znači da je procijenjena prosječna vjerojatnost točnoga odgovora na zadatak (uprosječena za sve bodovne kategorije ukupnoga rezultata na primarnoj dimenziji) za 0,1 veća za slučajno odabranog ispitanika iz prve skupine nego za slučajno odabranog ispitanika iz druge skupine, uz uvjet da ta dva ispitanika imaju jednak rezultat na primarnoj dimenziji (Stout i sur., 2003). Ako u ispitu od 50 zadataka postoji čvor od 6 zadataka, a vrijednost parametra β tog čvora iznosi 2,0,

to znači da je utjecaj tog čvora na ukupan rezultat u ispitu takav da će slučajno izabrani ispitanik iz prve skupine imati u prosjeku za 2 boda veći ukupan rezultat na ispitu nego slučajno odabrani ispitanik iz druge skupine (Stout i sur., 2003). Veličina diferencijalnog funkcioniranja čvora se može interpretirati u terminima praktičnih implikacija koje rezultat na ispitu ima za ispitanike. Vrijednosti parametra β imaju normalnu distribuciju čija aritmetička sredina predstavlja pravu vrijednost β u populaciji. Procjenjuje se i standardna pogreška parametra β , koja se označava kao $SE(\beta)$. U oko 2/3 slučajeva prava vrijednost parametra β se nalazi u intervalu $\beta \pm 1 SE(\beta)$.

Cilj ovoga rada je ispitati postoje li razlike u uspjehu na zadacima vezanima uz čakavske polazne tekstove između pristupnika iz regija u kojima je dominantno narječje čakavsko, kajkavsko i štokavsko. Analize će se provesti bez kontrole primarne dimenzije koju ispit mjeri, a zatim i uz kontrolu primarne dimenzije, tj. prije i nakon izjednačavanja ispitanika iz različitih regija prema ukupnom rezultatu na ostatku ispita. U tu svrhu analizirat će se razlike u rezultatu pristupnika iz različitih regija analizama varijance, te analizama diferencijalnog funkcioniranja zadataka i diferencijalnog funkcioniranja čvora zadataka.

METODA

Ispit državne mature iz Hrvatskoga jezika na višoj razini

Ispit ljetnoga roka državne mature 2011./2012. iz Hrvatskoga jezika na višoj razini sastojao se od dviju ispitnih cjelina¹. Prva ispitna cjelina je *Književnost i jezik* i čine je skupine zadataka zatvorenoga tipa. Druga je ispitna cjelina *Školski esej*. U ovom istraživanju analize su provedene na rezultatima prve ispitne cjeline kako bi se osigurala jednodimenzionalnost rezultata koja je potrebna za analize DFZ i DFČZ. Prvom ispitnom cjelinom ispitivalo se razumijevanje književnoga teksta, te književnoteorijska, književnopovijesna i jezična znanja. Bila su ukupno 72 zatvorena zadatka. Od toga je 70 zadataka bilo višestrukog izbora s 4 ponuđena odgovora te je na tim zadacima bilo moguće ostvariti 0 ili 1 bod. Dva su zadatka bila politomna te je na tim zadacima bilo moguće ostvariti maksimalno 5 bodova (riječ je o zadacima u kojima se od pristupnika tražilo povezivanje različitih sadržaja jedne liste sa sadržajima druge liste po nekom principu). Ukupan mogući broj bodova je bio 80. Prvih 40 zadataka je bilo vezano uz 8 polaznih tekstova (po 5 zadataka uz svaki tekst). Ostali zadaci nisu bili vezani uz polazne tekstove. Prvih deset zadataka u ispitu bilo je vezano uz dva pjesnička teksta pisana čakavskim narječjem: Ujevićeva je antologijska lirska pjesma *Oproštaj* objavljena 1914. u almanahu *Hrvatska*

1 Ispiti državne mature 2011./2012. objavljeni su na mrežnim stranicama Nacionalnoga centra za vanjsko vrednovanje obrazovanja (www.ncvvo.hr)

mlada lirika, i napisana je jezikom sličnim Marulićevu, kojega Ujević u pjesmi i izrijekom priziva, a drugi je tekst bio ulomak Hektorovićeve narativnog, putopisno koncipiranog renesansnog spjeva *Ribanje i ribarsko prigovaranje* (1568.). Svih deset zadataka vezanih uz tekstove pisane čakavskim narječjem bili su dihotomni, tj. na njima je bilo moguće ostvariti 0 ili 1 bod. Za tih deset zadatka provedena je analiza razlika uratka pristupnika iz regija u kojima je dominantno narječje čakavsko, kajkavsko i štokavsko.

Sudionici

Istraživanje je provedeno na rezultatima pristupnika iz gimnazija i strukovnih škola ispitu državne mature iz Hrvatskoga jezika na višoj razini 2011./2012. Ispitu su također pristupili i kandidati koji su prethodnih godina završili srednju školu i pristupnici iz inozemstva, no njihovi rezultati nisu korišteni u analizama razlika uratka pristupnika iz različitih regija.

Analize su provedene na sekundarnim bazama u kojima je postojala varijabla županije. Na temelju dijalektoloških karata i literature (Vukušić, 1993; Lončarić, 1996; Lisac, 2003) u prvu su skupinu odabrani pristupnici iz dviju županija koje su dominantno čakavske (Splitsko-dalmatinska i Istarska županija; 3236 pristupnika), u drugu su skupinu odabrani pristupnici iz dviju županija koje su dominantno kajkavske (Krapinsko-zagorska i Varaždinska županija; 1237 pristupnika), a u treću su skupinu odabrani pristupnici iz dviju županija koje su dominantno štokavske (Osječko-baranjska i Brodsko-posavska županija; 2149 pristupnika).

Prije provođenja analiza, pristupnici iz različitih regija su podijeljeni u skupinu pristupnika koji su pohađali gimnazijske programe te u skupinu pristupnika koji su pohađali strukovne programe. Ovakva homogenizacija pristupnika prije provođenja daljnjih analiza je potrebna jer bi razlike u rezultatu pristupnika iz različitih gimnazijskih i strukovnih programa mogle sakriti razlike između rezultata pristupnika iz različitih regija ili čak navesti na zaključak kako je stvarni smjer tih razlika obratan.

Postupak

Usporedba postotaka točnih odgovora na zadacima između pristupnika iz različitih regija u svrhu zaključivanja o tome jesu li pristupnici iz neke od regija bili favorizirani tim zadacima nije potpuno opravdana. Naime, prije usporedbe uratka pristupnika iz različitih regija na zadacima, pristupnike je potrebno izjednačiti prema uratku na ostatku ispita, tj. prema uratku na onim zadacima koji nisu vezani uz čakavske polazne tekstove. Na taj se način u ovim analizama uspoređuju oni pristupnici koji su na ispitu iz Hrvatskoga jezika za višu razinu imali podjednako znanje iz predmeta Hrvatski jezik, tj. podjednak ukupan rezultat na ispitu (nakon izostavljanja zadataka koji su vezani uz čakavske tekstove). Ovakve analize se provode u okviru paradigme DFZ i DFČZ (npr. Stout i sur., 2003.). Stoga su osim usporedbe

uratka različitih skupina sudionika na temelju prosječnih rezultata analizama varijance u ovom istraživanju provedene analize DFZ te DFČZ.

REZULTATI I RASPRAVA

Usporedba prosječne riješenosti zadataka

Ispitu ljetnog roka državne mature 2011./2012. iz Hrvatskoga jezika na višoj razini je pristupilo ukupno 21707 pristupnika ($M = 49,01$, $SD = 11,464$, asimetričnost = 0,011, stan. pogreška asimetričnosti = 0,017, spljoštenost = -0,519, stan. pogreška spljoštenosti = 0,033). Ovakvi rezultati pokazuju platikurtičnost distribucije rezultata. Statistički postupci koji se temelje na usporedbi aritmetičkih sredina između različitih skupina ispitanika prilično su robusni i otporni na platikurtičnost distribucije (DeCarlo, 1997).

U Tablici 1 nalazi se deskriptivna statistika ukupnog rezultata na ispitu za pristupnike iz gimnazija i strukovnih škola iz čakavske, štokavske i kajkavske regije. Također, u istoj tablici se nalazi deskriptivna statistika rezultata za pristupnike iz tih istih skupina na 10 zadataka vezanih uz tekstove pisane čakavskim narječjem.

Kao mjere veličine učinka u radu su korišteni η^2 i Pearsonov r (Field, 2005). Prvom dvosmjernom analizom varijance utvrđena je velika vrijednost veličine učinka vrste škole ($F(1, 6616) = 2695,92$, $\eta^2 = 0,29$) na ukupan rezultat na ispitu, dok učinak regije nije postojao ($F(2, 6616) = 56,43$, $\eta^2 = 0,01$). Učinak interakcije također nije postojao ($F(2, 6616) = 1,84$, $\eta^2 = 0,00$). Drugom dvosmjernom analizom varijance također je utvrđena velika vrijednost veličine učinka vrste škole ($F(1, 6622) = 1243,90$, $\eta^2 = 0,16$) na rezultat na 10 zadataka vezanih uz tekstove pisane čakavskim

Tablica 1. Deskriptivna statistika rezultata na ispitu za pristupnike iz gimnazija i strukovnih škola iz različitih dijalektalnih regija

		Svi zadaci			Zadaci vezani uz čakavske tekstove	
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Gimnazije	Čakavska	1934	55,67	9,824	7,58	1,629
	Kajkavska	649	55,26	10,059	7,40	1,726
	Štokavska	1210	52,83	9,765	7,09	1,651
	Ukupno	3793	54,69	9,926	7,39	1,666
Strukovne škole	Čakavska	1302	42,27	8,239	6,02	1,657
	Kajkavska	588	43,02	8,633	5,84	1,578
	Štokavska	939	39,99	8,433	5,53	1,713
	Ukupno	2829	41,67	8,471	5,82	1,673

Tablica 2. Rezultati usporedbi pristupnika iz različitih regija iskazani veličinom učinka (Pearsonov r)

Usporedbe	Gimnazije	Strukovne škole
čakavska regija – kajkavska regija	0,02 (0,05)	-0,04 (0,06)
čakavska regija – štokavska regija	0,14 (0,15)	0,14 (0,14)
kajkavska regija – štokavska regija	0,12 (0,09)	0,17 (0,05)

Prvi broj predstavlja vrijednost veličine učinka na ukupan rezultat. Unutar zagrada nalazi se vrijednost veličine učinka na rezultat na 10 zadataka vezanih uz tekstove pisane čakavskim narječjem

narječjem, dok učinak regije nije postojao ($F(2, 6622) = 54,09, \eta^2 = 0,01$). Učinak interakcije također nije postojao ($F(2, 6622) = 0,00, \eta^2 = 0,00$).

U Tablici 2 su prikazani rezultati usporedbi pristupnika iz različitih regija iskazani veličinom učinka (Pearsonovim r).

Vrijednosti veličina učinka iz Tablice 2 na ukupan rezultat na ispitu variraju od zanemarivih do malih vrijednosti (Cohen, 1988). I kod pristupnika iz gimnazija i kod pristupnika iz strukovnih škola, vrijednost veličine učinka pri uspoređivanju rezultata pristupnika iz čakavske i kajkavske regije je zanemariva. Vrijednosti veličina učinka pri uspoređivanju rezultata pristupnika iz čakavske i štokavske regije te kajkavske i štokavske regije su vrlo male.

Vrijednosti veličina učinka iz Tablice 2 za rezultat na 10 zadataka vezanih uz tekstove pisane čakavskim narječjem također variraju od zanemarivih do malih vrijednosti. I kod pristupnika iz gimnazija i kod pristupnika iz strukovnih škola, vrijednost veličine učinka pri uspoređivanju rezultata pristupnika iz čakavske i kajkavske regije te rezultata pristupnika iz gimnazija iz kajkavske i štokavske regije je zanemariva. Vrijednosti veličine učinka pri uspoređivanju rezultata pristupnika iz čakavske i štokavske regije su vrlo male.

Provjera jednodimenzionalnosti ispita

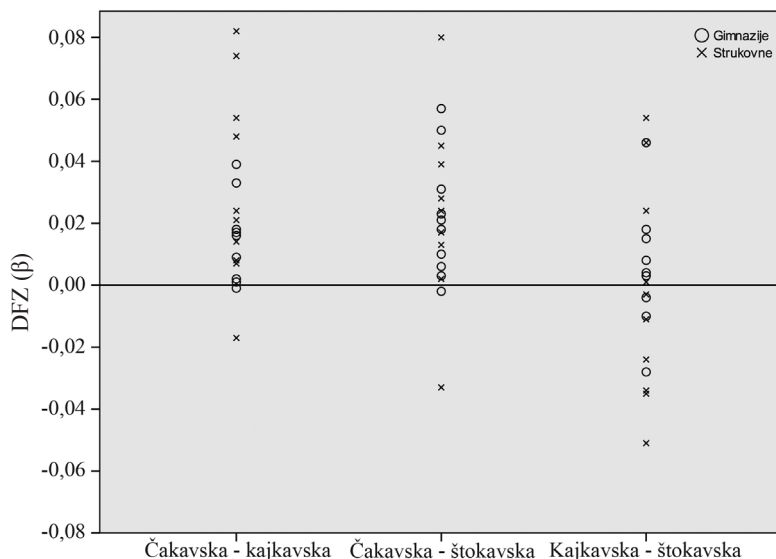
U računalnom programu SPSS 17.0 for Windows provedena je analiza glavnih komponenti na rezultatima svih pristupnika na zadacima u ispitu. Pri tome su zadaci vezani uz pojedini polazni tekst tretirani kao jedan politomni zadatak. Prva glavna komponenta objašnjava 17,4% varijance rezultata, a druga 3,4%. Vrijednost karakterističnog korijena uz prvu glavnu komponentu je preko pet puta veća od vrijednosti karakterističnog korijena uz drugu glavnu komponentu. Niska vrijednost varijance objašnjene prvom komponentom u ispitu iz Hrvatskog jezika na višoj razini je vjerojatno uzrokovana različitim profilima pristupnika koji izlaze na ispit te heterogenim ispitnim materijalom kojim se ispituju različita područja gradiva.

Analiza glavnih komponenata je zatim provedena samo za zadatke vezane uz polazne tekstove (40 zadataka koji su vezani uz 8 polaznih tekstova), pri čemu su zadaci vezani uz isti tekst tretirani kao jedan politomni zadatak. Ovom analizom je dobiveno jednokomponentno rješenje, pri čemu je prva komponenta objašnjavala 36,7% varijance, a karakteristični korijeni svih ostalih komponenti su bili manji od 1.

Sve analize DFZ i DFČZ su provedene uz kontroliranje rezultata pristupnika na ostatku ispita (62 zadatka). Analize su također ponovljene i uz kontroliranje rezultata pristupnika na zadacima koji su vezani uz polazne tekstove (30 zadataka), pri čemu su 32 zadatka koja nisu vezana uz polazne tekstove izostavljena iz analize. Uz obje strategije kontroliranja rezultata pristupnika dobiveni su vrlo slični nalazi, pa su u tekstu koji slijedi prikazani samo rezultati dobiveni uz kontroliranje rezultata pristupnika na ostatku ispita (tj. na 62 ispitna zadatka).

Diferencijalno funkcioniranje zadataka

Rezultati analiza diferencijalnoga funkcioniranja pojedinih zadataka vezanih uz čakavske tekstove provedenih pomoću računalnoga programa SIBTEST prikazani su na Slici 1. Parametar β (os ordinata) upućuje na veličinu diferencijalnoga funkcioniranja zadataka između dvije skupine. Pozitivna vrijednost parametra β upućuje na razliku u korist prve skupine iz para, a negativna vrijednost na razliku u korist druge skupine iz para.



Slika 1. Diferencijalno funkcioniranje zadataka vezanih uz tekstove pisane čakavskim narječjem za pristupnike iz različitih regija

Iz Slike 1 je vidljivo kako je 9 zadataka diferencijalno funkcioniralo u korist pristupnika iz gimnazija iz čakavske regije u odnosu na pristupnike iz gimnazija iz kajkavske regije, tj. u ovoj usporedbi 9 je zadataka imalo pozitivnu vrijednost parametra β . Treba napomenuti kako je riječ o vrlo malim vrijednostima parametra β . Jedan je zadatak imao negativan parametar β , koji upućuje na diferencijalno funkcioniranje zadatka u korist pristupnika iz gimnazija iz kajkavske regije.

Slični rezultati dobiveni su i u usporedbi rezultata pristupnika iz gimnazija iz čakavske i štokavske regije. I u ovoj je usporedbi 9 zadataka diferencijalno funkcioniralo u korist pristupnika iz gimnazija iz čakavske regije, a jedan zadatak u korist pristupnika iz gimnazija iz štokavske regije. Usporedbe rezultata pristupnika iz gimnazija iz kajkavske i štokavske regije pokazuju kako je 7 zadataka diferencijalno funkcioniralo u korist pristupnika iz gimnazija iz kajkavske regije, a 3 zadatka u korist pristupnika iz gimnazija iz štokavske regije. Analizama rezultata pristupnika iz strukovnih škola pronađeni su slični nalazi kao i analizama rezultata pristupnika iz gimnazija.

Iz Slike 1 je razvidno da se u usporedbi rezultata kajkavske i štokavske skupine vrijednosti parametara β grupiraju oko nule, tj. da se u većem broju zadataka poništava slučajno izmjereno diferencijalno funkcioniranje za zadatke s karakteristikom za koju nije očekivano da uzrokuje DFZ između tih skupina. Također, u usporedbama rezultata čakavske skupine s druge dvije skupine pristupnika, moguće je da neke druge karakteristike uzrokuju diferencijalno funkcioniranje zadataka u neočekivanom smjeru, kao što je to slučaj za po jedan zadatak u obje usporedbe učenika iz gimnazija, te za po jedan zadatak u obje usporedbe učenika iz strukovnih škola. Dobiveni rezultati pokazuju kako analize DFZ pomažu u prepoznavanju obrasca kako pristupnici iz različitih skupina rješavaju zadatke koji posjeduju određenu karakteristiku te koji su zadaci primjereniji za pojedine skupine.

Diferencijalno funkcioniranje čvora zadataka

Rezultati analiza diferencijalnoga funkcioniranja čvora zadataka provedenih pomoću računalnoga programa SIBTEST prikazani su u Tablici 3. Parametar β upućuje na veličinu diferencijalnoga funkcioniranja čvora zadataka između dvije skupine. Pozitivna vrijednost parametra β upućuje na razliku u korist prve skupine iz para, a negativna vrijednost na razliku u korist druge skupine iz para.

Primjerice, parametar β za prvi par iz Tablice 3 (pristupnici iz gimnazija iz čakavske i kajkavske regije) iznosi 0,150. To znači da će se rezultati dva po slučaju odabrana pristupnika iz ove dvije skupine (pod uvjetom da su skupine izjednačene prema uratku na ostatku ispita) u prosjeku razlikovati za 0,150 bodova nakon rješavanja čvora zadataka koji su vezani uz čakavske tekstove. Slične su razlike pronađene i između pristupnika iz čakavske i štokavske regije. Pristupnici iz kajkavske i štokavske regije nisu se razlikovali (vrijednost parametra β je zanemariva i iznosi

Tablica 3. Diferencijalno funkcioniranje čvora zadataka za pristupnike iz različitih regija

	Gimnazije		Strukovne škole	
	β	Standardna pogreška	β	Standardna pogreška
čakavska regija – kajkavska regija	0,150	0,067	0,313	0,074
čakavska regija – štokavska regija	0,215	0,052	0,241	0,068
kajkavska regija – štokavska regija	0,070	0,074	-0,038	0,080

0,070). Regionalne razlike u rezultatu na čvoru zadataka vezanih uz tekstove na čakavskom narječju pronađene su i za pristupnike iz strukovnih škola. Ove razlike su u istom smjeru kao i razlike pristupnika iz gimnazija, te su nešto veće.

Ovakve analize su posebno važne u kontekstu ispita vanjskog vrednovanja i ispita visokog rizika kao što su ispiti državne mature. Kako bi se osiguralo pravedno ispitivanje, nužno je sastaviti takve ispite koji neće diskriminirati pristupnike iz različitih grupa (prema spolu, prema regionalnoj pripadnosti i sl.) na temelju nekih svojstava koja nisu bitna za ono što se zapravo želi mjeriti. Ovakve analize služe boljem razumijevanju varijance ukupnoga rezultata na ispitu i povećanju valjanosti i pouzdanosti mjerenja. Još uvijek u edukacijskim i drugim mjerenjima stručnjaci uglavnom nastoje konstruirati jednodimenzionalne ispite koji mjere jedan predmet mjerenja (Deng i sur., 2008).

S obzirom na populacijsku prirodu podataka, nije uputno razmatrati statističku značajnost utvrđenih razlika među pristupnicima iz različitih regija. Naime, pristupnici iz pojedine regije koji su pohađali gimnazijske ili strukovne programe zapravo predstavljaju populaciju pristupnika iz tih regija i programa koji su na ljetnom roku državne mature 2011./2012. izašli na ispit iz Hrvatskoga jezika na višoj razini. Osim toga, u ovako velikim skupinama pristupnika postoji mogućnost da se i vrlo male razlike pokažu kao statistički značajne. Smislenije je stoga razmatrati rezultate u terminima praktičnih implikacija koje imaju za pristupnike.

Dakle, prema rezultatima analiza DFČZ, pristupnici iz regije u kojoj je dominantno narječje čakavsko će, u prosjeku, u ukupnom rezultatu iz ispita Hrvatski jezik na višoj razini imati malu bodovnu prednost u odnosu na pristupnike iz regija u kojima su dominantna narječja kajkavsko ili štokavsko, ako je riječ o pristupnicima koji inače imaju podjednaku razinu znanja iz predmeta Hrvatski jezik. Rezultati koji ukazuju na postojanje vrlo malih razlika dobiveni su i usporedbom prosječnih rezultata pristupnika iz različitih regija na zadacima vezanim uz čakavske tekstove bez kontroliranja njihova uratka na ostatku ispita. Veličine učinka tih razlika su ili vrlo male ili ne postoje. Veličine učinka razlika na zadacima vezanim uz čakavske tekstove su podjednako velike kao i veličine učinka razlika na ukupnom rezultatu na ispitu.

Rezultati analiza prikazanih u ovom radu prezentirani su predmetnim stručnjacima, profesorima hrvatskog jezika. U stručnoj raspravi je donesen zaključak o nastavku korištenja izvornih tekstova na čakavskom, kajkavskom i štokavskom narječju na ispitima državne mature iz Hrvatskog jezika. Na ispitima trebaju biti podjednako zastupljena sva narječja kako pristupnici ne bi bili u neravnopravnom položaju zbog pripadnosti svojega razgovornog jezika jednomu narječju. U budućnosti je stoga moguće pri određivanju strukture ispita državne mature iz Hrvatskoga jezika uzeti u obzir i varijablu koliko je polaznih tekstova unutar ispita pisano kojim narječjem te na taj način kontrolirati utjecaj narječja kojim su tekstovi pisani na ukupan rezultat pristupnika iz različitih regija. Može se očekivati kako bi se izrazito niske vrijednosti veličina učinaka pri usporedbi rezultata pristupnika iz različitih regija kao i vrlo male vrijednosti DFZ i DFČZ vezanih uz čakavske tekstove dobile i na ispitima iz Hrvatskoga jezika na osnovnoj razini. Metodologija korištena u ovom istraživanju trebala bi se u budućnosti koristiti u analizama međugrupnih razlika (primjerice rodnih) na svim ispitima državne mature.

LITERATURA

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- DeCarlo, L.T. (1997). On the meaning and use of kurtosis. *Psychological methods*, 2, 292-307.
- Deng, N., Wells, C., Hambleton, R. (2008). A Confirmatory Factor Analytic Study Examining the Dimensionality of Educational Achievement Tests. U: H. Marx, T. Levine (ur.), *NERA Conference Proceedings, Rocky Hill, Connecticut, 22-24 October 2008*, rad pod brojem 31. Rocky Hill, CT: University of Connecticut.
- Douglas, J., Roussos, L.A., Stout, W.F. (1996). Item bundle DIF hypothesis testing: Identifying suspect bundles and assessing their DIF. *Journal of Educational Measurement*, 33, 465-485.
- Field, A.P. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- Gierl, M.J., Bisanz, J., Bisanz, G., Boughton, K., Khaliq, S. (2001). Illustrating the utility of differential bundle functioning analyses to identify and interpret group differences on achievement tests. *Educational Measurement: Issues and Practice*, 20, 26-36.
- Gierl, M.J., Gotzmann, A., Boughton, K.A. (2004). Performance of SIBTEST when the percentage of DIF items is large. *Applied Measurement in Education*, 17, 241-264.
- Hyde J.S., Mertz J.E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences*, 106, 8801-8807.
- Jokić, B., Ristić Dedić, Z., Šabić, J. (2011). *Analiza sadržaja i rezultata ispita državne mature iz Fizike*. Zagreb: Nacionalni centar za vanjsko vrednovanje obrazovanja i Institut za društvena istraživanja.
- Le, L.T. (2009). Investigating Gender Differential Item Functioning across Countries and Test Languages for PISA Science items. *International Journal of Testing*, 9, 122-133.

- Lisac, J. (2003). *Hrvatska dijalektologija 1, Hrvatski dijalekti i govori štokavskog narječja i hrvatski govori torlačkog narječja*. Zagreb: Golden marketing – Tehnička knjiga.
- Lončarić, M. (1996). *Kajkavsko narječje*. Zagreb: Školska knjiga.
- Mazor, K.M., Clauser, B.E., Hambleton, R.K. (1992). The effect of sample size on the functioning of the Mantel-Haenszel statistic. *Educational and Psychological Measurement*, 52, 443-451.
- Ministarstvo kulture i prosvjete (1994). *Nastavni program za gimnazije*. Zagreb: Glasnik Ministarstva kulture i prosvjete Republike Hrvatske.
- Nacionalni centar za vanjsko vrednovanje obrazovanja (2011). *Ispitni katalog iz Hrvatskoga jezika za državnu maturu u školskoj godini 2011./2012*. Zagreb: Nacionalni centar za vanjsko vrednovanje obrazovanja.
- Roussos, L.A., Schnipke, D.L., Pashley, P.J. (2000). *A formulation of the Mantel-Haenszel differential item functioning parameter with practical implications. (LSAC Statistical Report 96-03)*. Newtown, PA: Law School Admission Council
- Roussos, L., Stout, W. (1996). A Multidimensionality-Based DIF Analysis Paradigm. *Applied Psychological Measurement*, 20, 355–371.
- Shealy, R., Stout, W.F. (1993). An item response theory model for test bias. U: P.W. Holland, H. Wainer (ur.), *Differential item functioning*, 197-239. Hillsdale NJ: Erlbaum.
- Simpson, E.H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13, 238-241.
- Stout, W., Bolt, D., Froelich, A.G., Habing, B., Hartz, S., Roussos, L. (2003). *Development of a SIBTEST Bundle Methodology for Improving Test Equity with Applications for GRE Test Development. Graduate Record Exam Board Report No. 98-15P, Educational Testing Service Report 03-06*. Princeton, NJ: Educational Testing Service.
- Vukušić, S. (1993). Jezična sudbina Hrvata u Istri (jedan vidik ili skica za studiju). *Društvena istraživanja*, 4-5, 699-706.
- Wyse, A.E., Mapuranga, R. (2009). Differential Item Functioning Analysis Using Rasch Item Information Functions. *International Journal of Testing*, 9, 333–357.
- Zumbo, B.D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

COMPARISON OF RESULTS OF CANDIDATES FROM DIFFERENT
DIALECTAL REGIONS ON ITEMS RELATED TO THE TEXTS
WRITTEN IN ČAKAVIAN DIALECT IN CROATIAN LANGUAGE
STATE FINAL EXAMINATION

Summary

One of the modern challenges in the field of written examination construction is the understanding of the difference in the success of different groups of examinees on individual items. Understanding these differences is of particular importance in high-stakes examinations such as the state final examinations. The aim of this study was to examine the differences in the results between examinees from different dialectal regions (Čakavian, Kajkavian and Štokavian) on the Croatian language state final examination items which were related to the texts written in the Čakavian dialect. Based on a comparison of the average results on items related to Čakavian texts and an analysis of differential item functioning and differential bundle functioning, very small differences were found in favor of examinees from the Čakavian region in comparison to examinees from the Kajkavian and Štokavian region. Based on these findings, the professional discussion on the Croatian language examination resulted in a recommendation to continue using authentic texts in Čakavian, Kajkavian and Štokavian dialects in Croatian language state final examinations in which all dialects should be equally represented to ensure equal treatment of all candidates regardless of their native dialect.

Key words: state final examination, Croatian language, Čakavian dialect, differential item functioning, differential bundle functioning

Primljeno: 27. 01. 2014.