

ika bi, sa svoje strane, morala biti spremna da prihvati racionalno utemeljen dijalog
anjima jezika, ne popuštajući pred pritiscima svakodnevne (i nerijetko politikantske)
matike, i znajući da je izgradnja jezične politike uvijek dugotrajan i veoma
pleksan posao čiji se rezultati ne mogu očekivati u kratkom roku. To ujedno znači da
politika pružala podršku i teorijskim i primijenjenim i istraživačkim aspektima nauke o
u, jer se bez znanstvenih spoznaja ne može voditi uopće nikakva jezična politika."

itics, for its part, needs to welcome a rationally-based dialogue about language
stions, not yielding to the pressures of day-to-day (and sometimes demagogic)
edieny, and realizing that building up a language policy is always a long and
ly complex task that cannot be expected to give results in a short time. This also
ns that politics should give its support to theoretical and applied research in the
nce of language, for without scientific findings no language policy whatever can
arried out."

Dubravko Škiljan

atsko društvo za primijenjenu lingvistiku – HDPL
atian Applied Linguistics Society – CALS

www.hdpl.hr

ISBN 978-953-7494-02-5



9 789537 494025



Jezična politika i jezična stvarnost

Branko Tošović, Arno Wonisch

Gralis-Korpus

Zagreb: Hrvatsko društvo za primijenjenu lingvistiku, 2009. S. 117–125.

Language Policy and Language Reality



Urednica / Editor
Jagoda Granić



HDPL

1. ELEKTRONIČKI KORPUSI

U cilju prikupljanja građe i sistemskog proučavanja jezika razvijaju se u posljednje vrijeme dosta intenzivno raznorodni *on-line* korpusi za automatsko traženje i nalaženje informacije o jezičnim oblicima, značenjima i kategorijama. Postoje dvije osnovne vrste elektroničkih korpusa: jednojezični i višejezični. Prvi služe za proučavanje samo jednog jezika. Takvi korpusi obično se realiziraju kao nacionalni projekti, pa stoga neki i u naslovu imaju ekspliciranu takvu orijentaciju. Drugi su namijenjeni za proučavanje dvaju ili više jezika te se stoga obično nazivaju paralelnim korpusima. Aktivnost na izradi jednojezičnih korpusa u posljednje je vrijeme dala vrijedne rezultate tako da sada postoje dva elektronička korpusa za hrvatski jezik (Hrvatski nacionalni korpus, Hrvatska mrežna riznica), jedan za srpski (Korpus savremenog srpskog jezika na Matematičkom fakultetu Univerziteta u Beogradu) i jedan za bosanski/bošnjački (Korpus bosanskih tekstova na Sveučilištu u Oslu). Problem je njihova korištenja u tome što (a) imaju različitu koncepciju, (b) neizbalansirano odražavaju funkcionalnostilsku raslojenost, (c) posjeduju nejednak fond, (d) jedan od njih uopće ne funkcionira (bosanski/bošnjački). Za razliku od jednojezičnih korpusa, višejezičnih ima mnogo manje. Razlog je činjenica da je u svakoj nacionalnoj korpusnoj lingvistici materinski jezik primaran i da je teže i složenije razviti višejezični korpus jer se u njemu udvostručuje, utrostručuje sve ono što postoji u jednojezičnom korpusu. Nama za sada nije poznat neki značajniji paralelni korpus za bosanski/bošnjački, hrvatski i srpski jezik (dalje B, K, S), točnije, bilo je pokušaja, ali sa skromnim ili nikakvim rezultatima. Između ostalog, iz tih razloga, a u cilju svestranog, kompleksnog i objektivnog istraživanja više slavenskih jezika, zasnovan je *Gralis-Korpus*. Tako je, recimo, za utvrđivanje identičnosti, sličnosti i različitosti B, K, S neophodno imati dobar, reprezentativan korpus. Stoga smo inicirali izradu paralelnog *Gralis-Korpusa* koji bi istraživačima omogućio da na svim razinama i u svim funkcionalnim stilovima analiziraju podudarnosti, sličnosti i razlike u cjelovitim tekstovima, njihovim dijelovima i elementima. U tu svrhu započeo je na Sveučilištu u Grazu 2007. godine s radom

GRALIS Korpus
Institut für Slawistik | Universität Graz |

2. GRALIS-KORPUS – VIŠEJEZIČNI PARALELNI KORPUS



Riječ je o višjezičnom paralelnom korpusu koji obuhvaća tri južnoslavenska jezika (B, K, S), jedan istočnoslavenski jezik (ruski), jedan zapadnoslavenski jezik (poljski), s orijentacijom i na druge jezike, prije svega južnoslavenske (slovenski, bugarski i makedonski). U okviru *Korpusa*

razvijaju se dva podkorpusa – *Speech-Korpus* i *Text-Korpus*, pri čemu je *Speech-Korpus* podijeljen na *Wort-Korpus* (izgovor izoliranih riječi), *Fix-Korpus* (izgovor zadanih tekstova) i *Frei-Korpus* (snimanje spontanog govora u svim područjima života i rada). Odnos između *Text-Korpusa* i *Speech-Korpusa* trebao bi biti 90% : 10%.

Gralis-Korpus je namijenjen stručnjacima za B, K, S i jezikoslovcima šireg profila (osobito s područja opće i sistemske lingvistike te sociolingvistike), kao i onima koje zanimaju intralingvistički, interlingvistički i ekstralingvistički suodnosi između B, K, S. Taj se korpus može široko i svrsishodno primjenjivati u nastavi, posebice na visokoškolskim ustanovama. *Gralis-Korpus* će biti od koristi svima onima koji se u praksi susreću s teškoćama vezanim za B, K, S (studentima, lektorima, filmskim radnicima, političarima i dr.). Na osnovi takvog korpusa može se prići izradi programa za automatsko mjerenje funkcionalne bliskosti i udaljenosti jezika.

Najveći dio *Gralis-Korpusa* čini paralelni informacijsko-istraživački sustav za bosanski/bošnjački, hrvatski i srpski jezik sastavljen od tekstova koji imaju bar dvije verzije (bosansku/bošnjačku i hrvatsku, bosansku/bošnjačku i srpsku, hrvatsku i srpsku). Cilj je korpusa omogućiti što dublje, šire i svestranije proučavanje podudarnosti, sličnosti i razlika između hrvatskog, srpskog i bosanskog/bošnjačkog jezika. Budući da se radi o bliskim jezicima i budući da njihov odnos daje povoda za raznorazne spekulacije, izradom takvog korpusa dobit će se reprezentativna i raznovrsna građa za objektivnu ocjenu identičnosti, sličnosti i razlika između njih. *Gralis-Korpus* treba pokazati kako funkcioniraju B/K/S-jedinice svih razina (fonetsko-fonološke, ortoepske, gramatičke i stilističke) na opsežnoj konkretnoj građi, u prirodnom okruženju. *Gralis-Korpus* će biti maksimalno reprezentativan (kako bi omogućio relevantne i pouzdane informacije) i uravnotežen (kako bi adekvatno odražavao globalnu diferencijaciju jezika, posebno funkcionalnostilsku). Kao teoretska osnova za tipološku obradu i prezentaciju tekstova služi knjiga *Funkcionalni stilovi* (Tošović, 2002). Prema koncepciji izloženoj u njoj, *Gralis-Korpus* je podijeljen na pet funkcionalnih stilova

(književnoumjetnički, publicistički, znanstveni, administrativni, razgovorni). Cilj je projekta napraviti takav korpus (1) koji neće ovisiti o ikakvim vanjskim okolnostima, (2) koji će moći pratiti brzinu i kvalitetu promjena informacijskih tehnologija i (3) koji će se moći stalno razvijati, usavršavati i dorađivati. Budući da se kvaliteta svakog korpusa određuje (a) dubinom i širinom obilježavanja, (b) mogućnostima traženja i nalaženja informacije, (c) reprezentativnošću, proporcionalnošću i uravnoteženošću te (d) dostupnošću, u izradi korpusa će se o tome voditi najviše računa. Izrada *Gralis-Korpusa* bit će maksimalno prilagođena programskim standardima koji se koriste u svijetu (TEI, XML i dr.), kako bi se postigla kompatibilnost i usporedivost s drugim sličnim korpusima te omogućila šira primjena. Rad na *Gralis-Korpusu* i konačan je (stavljanjem u "on-line pogon") i beskonačan (stalnim popunjavanjem, ispravljanjem i produbljivanjem). *Gralis-Korpus* će se razvijati: 1. kvantitativnim proširivanjem novim tekstovima i sadržajima, 2. kvalitativnim poboljšavanjem (širim i dubljim obilježavanjem), 3. formalnim usavršavanjem (osvježavanjem *www-dizajna*), 4. funkcionalnim ubrzavanjem (boljim sustavom pretraživanja i nalaženja informacija) te 5. programskim obogaćivanjem primjenom novih softverskih paketa (Kofler i Wonisch, 2007; Thomann, 2007).

2.1. Gralis Speech-Korpus: Wort-Korpus, Fix-Korpus i Frei-Korpus

Gralis Speech-Korpus prvi je on-line govorni korpus za bosanski/bošnjački, hrvatski i srpski jezik. Kao što je već rečeno, podijeljen je na nekoliko dijelova: *Wort-Korpus*, *Fix-Korpus* i *Frei-Korpus*. *Wort-Korpus* obuhvaća audiomaterijal nastao izgovaranjem izoliranih riječi. *Fix-Korpus* sadrži snimke kraćih tekstova (otprilike dvadesetak rečenica). Dosad je najviše materijala snimljeno na osnovi teksta *Jutro*, sastavljenog od 18 rečenica, u kojem nema fonetsko-fonoloških, leksičkih i gramatičkih razlika između B, K i S. U okviru *Fix-Korpusa* izrađen je *Akzentarium*, *Transkriptarium*, *Suprasemgentarium* i *Valorisarium* (Forić, 2007; Lehner, 2007; Midžić, 2007; Tošović, 2007a; 2007b).

U *Akzentariumu* se daje naglasak izgovorenih riječi. Također pruža mogućnost obilježavanja akcenta i dobivanja informacija o njemu na osnovi unesenih riječi. U *Transkriptariumu* se ukazuje na izgovor riječi u rečenici. *Suprasemgentarium* prezentira intonacijsku strukturu rečenice. *Valorisarium* služi da stručnjaci za

	Treffer: 1	Thema: Jutro
1		Jutros su me vrlo rano ptice probudile.
2		Sunce je tek bilo izašlo iza brda.
3		U daljini se prostirala velika šuma.
4		U njoj su rasli borovi, jele, breze i omorike.
5		Na livadama oko šume pasle su ovce i krave.
6		Na travi je i dalje ležala rosa.
7		U obližnjem parku ptice su cvrkutale.
8		Grad se postepeno budio.
9		Ulice su postajale sve bučnije.
10		U dvorištu su se pojavila dva psa.
11		Oni su se veselo igrali, povremeno lajući.
12		Iza dvorišta su dopirali zvuci tramvaja.
13		Đaci su se spremali za odlazak u školu.
14		Neki od njih imali su na sebi džempere.
15		Trgovine su počele s radom.
16		Poštari su raznosili pisma.
17		Jutro je sve više prelazilo u dan.
18		Ja se spremam za rad.

Akzentarium

Wort (z.B. ja) suchen

á è ì ó ù f

Sprache auswählen: hrvatski

Gralis-Akzentor:
hrvatski

0 Treffer

Wörterbücher:
hrvatski

0 Treffer

administrativnom, publicističkom, znanstvenom i razgovornom), pri čemu je trenutačno težište na snimanju i obradi TV dnevnikâ koji se istovremeno emitiraju u Zagrebu, Beogradu, Sarajevu i Podgorici (Tošović, 2007). U *Speech-Korpusu* tekstualnu informaciju prati audiosnimka i(li) videosnimka. Korisniku se nudi više mogućnosti pretrage – u obilježenim rečenicama, u sklopovima koji imaju transkripciju, akcentuaciju i(li) suprasegmentaciju itd. Za sav snimljeni materijal koji se uključuje u

Gralis-Korpus traži se pismena suglasnost, počevši od pojedinaca koji nešto govore, pa do radijskih, televizijskih i izdavačkih kuća čiji se prilozi segmentiraju i unose u korpus (Forić, 2007).

Za obradu podataka o govornicima i pojedinim audiosnimkama primjenjuju se posebne procedure, koje

TV dnevnici
7. svibnja (ponedjeljak) 2007. godine

1. Bosanski TV dnevnik	19.00 30 min.
2. Crnogorski TV dnevnik	19.30 33 min.
3. Hrvatski TV dnevnik	19.30 24 min.
4. Srpski TV dnevnik	19.30 38 min.
	125 min.

Jednostavna pretraga

Gralis Fix-Korpus

Rečenica ☐ samo obilježene/obeležene rečenice

(S)pol Maternji/materinski jezik M(j)esto rođenja M(j)esto stanovanja Nacionalnost

Pretraga

Rezultat pretrage:

Dobijeni rezultat: 2 (0% des Korpus)

Dobijeni rezultat: 1 - 2 od 2

	m/ž	Nacionalnost	Godina rođenja	M(j)esto rođenja M(j)esto Država	M(j)esto stanovanja	Maternji/materinski jezik
1	f	hrvatska	1964	Split	Hrvatska Graz	hrvatski
2	m	hrvatska	1964	Split	Hrvatska Graz	hrvatski

fonetiku, fonologiju i prozodiju provjere jesu li transkripcija, akcentuacija i intonacija korektno napravljene, odnosno obilježene (Forić, 2007; Tošović, 2007b). *Frei-Korpus* predstavlja zbirku slobodno izgovorenih tekstova koji pripadaju svim funkcionalnim stilovima (književnoumjetničkom,

počinju popunjavanjem *on-line* evidencije (Midžić, 2007; Wonisch i Just, 2007). Ona ima tri dijela. To su: Ispitanik, Snimka i Korpus. U prvom se unose osnovne informacije o govorniku (spol, nacionalnost, religija, godina i mjesto rođenja, mjesto stanovanja, zanimanje, obrazovanje, mjesto školovanja, materinski jezik, važnija mjesta boravka, strani jezik itd.), u drugi dolaze podaci o temi, mjestu, datumu, okolnostima snimanja, uređaju, dužini i formatu snimke, operatoru, a u treći dolaze podaci o tome za koji će se tip istraživanja dobiveni materijal koristiti (za slušnu, spektralnu ili neku drugu analizu). Pri dnu ispitanik pismeno daje suglasnost za objavljivanje izgovornog teksta u *Gralisu* pod određenim kodom. Na početku se za svakog ispitanika bira zaporka, koja se sastoji od 1) naziva grada, gdje je snimka napravljena, 2) troznamenastog broja i 3) slova **b**, **k** ili **s** za oznaku bosanskog/bošnjačkog, hrvatskog ili srpskog jezika (npr.: split_001k). Nakon unosa zaporka otvara se detaljan formular s podacima o govorniku i snimci. Ako treba dodati, recimo, mjesto ili jezik koji još nisu u bazi podataka, ulazi se u rubriku "Novo m(j)esto ili šifru/zaporku un(ij)eti" (Midžić, 2007).

2.2. Gralis Text-Korpus

Gralis Text-Korpus čine tekstovi koji imaju najmanje dvije jezične verzije (**b** i **k**, **b** i **s**, **k** i **s**). Za korištenje sučelja nudi se sljedeći izbor: b/k/s, engleski, njemački. U pretrazi se na početku bira "primarni jezik" (kao oznake jezika uzete su internacionalne kratice **bh**, **hr**, **sr**), što podrazumijeva jezik

koji je u rezultatima pretrage na prvom mjestu. Program nudi dvije vrste pretrage: jednostavnu i složenu (uz pomoć tzv. CQP-sintakse). Dobiveni pojmovi (odnosno riječi) obilježavaju se žutom ili plavom bojom. Iznad tekstova nalaze se naslovi na trima jezicima, koji strelicom upućuju na original (ako se radi o internetskoj adresi, ukazuje se na poveznicu). Osim pretrage po tekstovima korpus nudi i mogućnost dobivanja čestotne liste riječi (Wonisch, 2007).

Pretraga

Primarni jezik ☒ Bosanski/Bošnjački ☐ Hrvatski ☐ Srpski

Izbor materijala ☒ Korpusni tekstovi ☐ Spisak/Lista riječi

Način pretrage ☒ Jednostavna pretraga ☐ Pretraga pomoću CQP-sintakse

Pretraga

PUTNA MREŽA U BOSNI I HERCEGOVINI

HR	Dobre prometne veze EU i susjednih zemalja, nesmetan i siguran promet ljudi i roba, su osnova za bolju trgovinu, održivi ekonomski razvoj, konkurenciju i dr. To je posebno važno jer se očekuje do 2020. godine povećanje prometa za više od 100 %.
SR	Dobre transportne veze EU i susjednih zemalja, nesmetan i siguran transport ljudi i roba, su osnova za bolju trgovinu, održivi ekonomski razvoj, konkurenciju i dr. To je posebno važno, jer se očekuje do 2020. godine povećanje transporta više od 100 %.
BS	Dobre prometne veze EU i susjednih zemalja, nesmetan i siguran promet ljudi i roba, su osnova za bolju trgovinu, održivi ekonomski razvoj, konkurenciju i dr. To je posebno važno jer se očekuje da će se do 2020. godine povećati promet za više od 100 %.

3. PRIPREMA I IZRADA KORPUSA

Izrada Korpusa ima dva smjera. U prvom se prikupljaju tekstovi i obavlja njihova gruba obrada s ciljem stvaranja nelematiziranog korpusa (sirovog) u zatvorenom internet-forumu ("BKS-Forum"), dostupnom samo suradnicima na izradi korpusa. Na osnovi nekoliko tekstova iz sirovog korpusa stvara se probni korpus (*Wartekorpus*, doslovno 'Korpus na čekanju'), na kojem se provjerava koncepcija i njezina praktična primjena. U drugom se smjeru priprema lematizirani korpus na taj način što se iz sirovog korpusa biraju reprezentativni uzorci, računalno obrađuju i stavljaju u *Gralis*. Dakle, prvo se sastavlja sirovi korpus (od originalnih tekstova), iz kojeg se biraju tekstovi za obilježavanje i lematizaciju.

Budući da *Gralis-Korpus* predstavlja paralelni korpus ponajprije suvremenog standardnog bosanskog/bošnjačkog, standardnog hrvatskog i standardnog srpskog jezika, u prvom razdoblju rada unose se i obrađuju tekstovi napisani od 1991. do danas. U drugoj fazi radit će se na tekstovima od 1981. do 1990, u trećoj od 1961. do 1980, a u četvrtoj od 1941. do 1960.

U *Korpus* se proporcionalno i uravnoteženo uključuju originalni i prevedeni tekstovi (B↔K↔S, drugi jezik/jezici → B, K, S). Tekstovi se prikupljaju u suradnji s nakladničkim kućama, uredništvima listova i časopisa, proizvođačima i distributerima filmova, autorima, prevoditeljima. Jedan dio preuzima se s interneta.

Izrada *Korpusa* realizira se u sklopu istraživačkog projekta (2006-2009) koji financira Austrijski fond FWF za potporu znanstvenih ispitivanja (*Fonds zur Förderung der wissenschaftlichen Forschung*) i međunarodne suradnje s najpoznatijim stručnjacima za slavensku korpusnu lingvistiku, osobito s onima koji dolaze s bosanskog/bošnjačkog, hrvatskog i srpskog govornog područja.

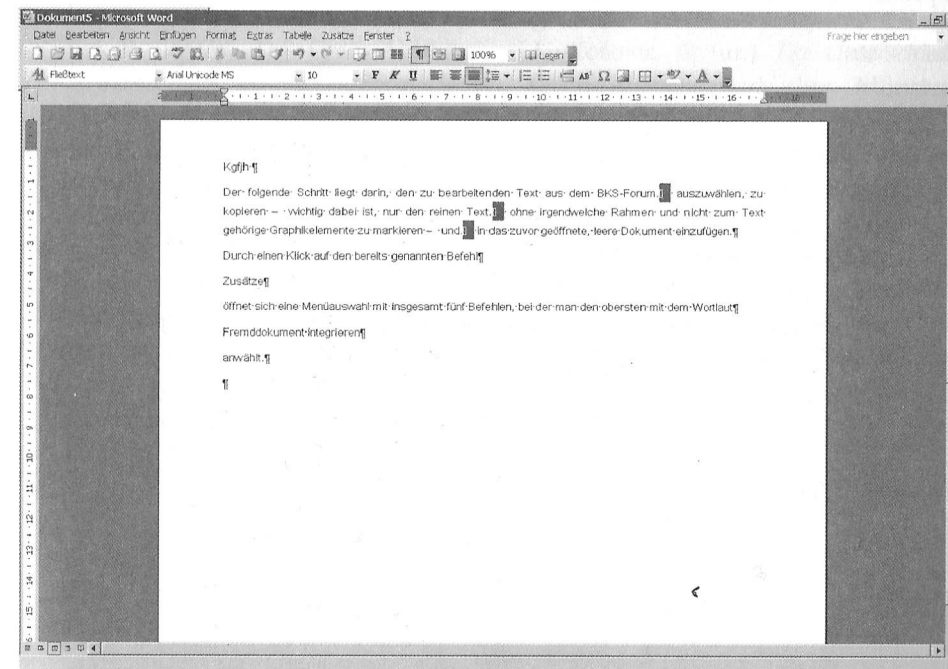
O autorskim pravima strogo se vodi računa i za unošenje materijala u *Korpus* traži se pismena suglasnost nakladnika, uredništava listova i časopisa, proizvođača i distributera filmova, autora tiskanih i elektroničkih verzija tekstova, prevoditelja i (ako nisu živi) njihovih pravnih nasljednika.

U pripremi *Korpusa* obavljene su intenzivne konzultacije s vodećim stručnjacima za korpusnu lingvistiku. U tom cilju organiziran je u ljetnom semestru 2006. godine seminar "Slavenska korpusna lingvistika". Gostovala su poznata imena (navodimo ih kronološkim redom): Dmitrij Dobrovoljski iz Moskve (govorio o ruskim paralelnim korpusima), Marko Tadić iz Zagreba (predstavio Hrvatski nacionalni korpus), Duško Vitas iz Beograda (prezentirao Korpus savremenog srpskog jezika na Matematičkom fakultetu u Beogradu), Cvetana Krstev iz Beograda (održala predavanje o elektronskom korpusu i rječniku), Tomaž Erjavec iz Ljubljane (prikazao Korpus slovenskog jezika FIDA plus i *Multext-East*), Bernhard Kettemann iz Graza (izlagao o temi *Korpus von*

Intelligent Design Texten), Danko Šipka iz Chandlera, SAD (govorio o tekstualnim korpusima u primijenjenoj slavistici), Damir Čavar (predstavio Hrvatsku jezičnu mrežnu riznicu), Stefan Schneider iz Graza (demonstrirao *Online-Korpus BADIP – Banca dati dell'italiano parlato*), Miloš Utvić iz Beograda (govorio o obradi teksta, obilježavanju, paralelizaciji i vertikalizaciji u izradi korpusa), Jana Zemljarič-Miklavčič (objasnila *Korpus govorjene slovenščine*). Tijekom semestra vođeni su intenzivni razgovori o koncepciji *Gralis-Korpusa*, posebno s Tomažom Erjavcem, Markom Tadićem i Duškom Vitasom. S nekim od njih održava se stalna konzultacija.

4. PROGRAMSKA OSNOVA

Kao što je *Gralis-Korpus* podijeljen na dva kompleksa (*Speech-Korpus* i *Text-Korpus*), tako je i programska osnova razdvojena. *Speech-Korpus* se razvija kao MySQL baza podataka koju je izradila Olga Lehner iz Graza (Lehner, 2007). To je vrlo fleksibilan sustav koji omogućuje proširivanje, poboljšavanje i mijenjanje govornog korpusa, a također da se u njemu jednostavno nalazi raznorodna informacija. U izradi *Text-Korpusa* koristi se više programa. Upravljanje *Text-Korpusom* vrši se pomoću *IMS Corpus Workbench* (CWB), čiji su autori Oliver Christ i Bruno Maximilian Schulze iz Instituta za strojnu obradu jezične informacije Sveučilišta u Stuttgartu (*Institut für Maschinelle Sprachverarbeitung der Universität in Stuttgart, IMS*). U tom paketu nalazi



se tzv. CQP (*Corpus Query Processor*), koji daje mogućnost široke pretrage korištenjem proširenih regularnih izraza i strukturnih oznaka u obilježenom korpusu. Oba sustava (*IMS Corpus Workbench* i *CQP*) rade pod *Linuxom*. Licenca za njihovo korištenje dobivena je u lipnju 2006. godine. Sučelje je za korištenje i pretraživanje korpusa u Institutu u Grazu (*Institut für Informationsverarbeitung in den Geisteswissenschaften, INIG*) na osnovi modela *Asset management* za upravljanje metapodacima digitalnih resursa (AM). Pomoću njega vrši se jednostavno integriranje u *web*-kontekst (što je posebno pogodno za edukativne forume, *Frameworks* itd.).

U AM-u se primjenjuju standardizirane procedure za obradu sadržaja. On omogućuje definiranje prava na pristup (npr. transformacije u XSLT-u, procedure vertikalizacije i analize i sl.), suradnju u obradi i upravljanju *Aseta* (koji služi kao kontejner u kojem se čuvaju podaci: tekstovi, audio/video/datoteke itd.), nudi mogućnost za uređenje zaporki. Hubert Stigler (autor *Aseta*), napravio je, prema CQP-u, *web-interface* i program za paralelizaciju i vertikalizaciju tekstova na trima jezicima, za dizajn je zaslužna Martina Semlak (iz INIG-a), a za administriranje servera na kojem se nalazi *Gralis-Korpus* zadužen je Dieter Schicker (iz istog Instituta). U generiranju *Gralis Text-Korpusa* primjenjuju se sljedeće procedure: pretvaranje dokumenta pomoću makroa u TEI-format, validiranje tročlanog teksta na osnovi standarda XML, generiranje vertikaliziranih, ujednačenih datoteka. Markiranje odlomaka vrši se primjenom makroa i regularnih izraza (ako je neophodno, granice segmenata tijekom obrade interaktivno se, premještanjem ili brisanjem, korigiraju). Korpusni tekstovi, kao i vertikalizirane datoteke, validiraju se u *Bulk-Modusu*, a onda se prelazi na dalju obradu pomoću *IMS Corpus Workbench*a.

Izvori

<http://www.gewi.uni-graz.at/gralis/>
<http://www.gewi.kfunigraz.ac.at/gralis/>
<http://riznica.ihj.hr/>
<http://www.hnk.ffzg.hr/>
<http://www.korpus.matf.bg.ac.yu/prezentacija/korpus.html>

Referencije

- FORIĆ, S. (2007). *Das Gralis Speech-Korpus*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- KOFER, S. i WONISCH, A. (2007). *Das Gralis-Rezensarium*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- LEHNER, O. (2007). *Die technische Entwicklung des Gralis Speech-Korpus*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- MIDŽIĆ, M. (2007). *Die Aufnahmeevidenz des Gralis Speech-Korpus*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- THOMANN, R. (2007). *Das Gralis-Anketarium*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- TOŠOVIĆ, B. (2002). *Funkcionalni stilovi. Funktionale Stile*. Graz: Institut für Slawistik der Universität Graz.
- TOŠOVIĆ, B. (2007). *Die Gralis Audio-VideoTools*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- TOŠOVIĆ, B. (2007a). *Das Gralis-Akzentarium*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- TOŠOVIĆ, B. (2007b). *Das Gralis-Korpus*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- WONISCH, A. (2007). *Das Gralis Text-Korpus*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.
- WONISCH, A. i Just, A. (2007). *Aufnahme und Bearbeitung von Audiodateien für das Gralis Fix-Korpus*. U: Tošović, B. (ur.). *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*. Münster-Hamburg-Berlin-Wien-London: LIT Verlag.