

Ohranjanje jezikovne zahtevnosti besedil pri prevajanju testov PISA

ŠPELA ARHAR HOLDT, IZTOK KOSEM

ZAVOD ZA UPORABNO SLOVENISTIKO TROJINA & FILOZOFSKA FAKULTETA UNIVERZE V
LJUBLJANI

PISA (The Programme for International Student Assessment)

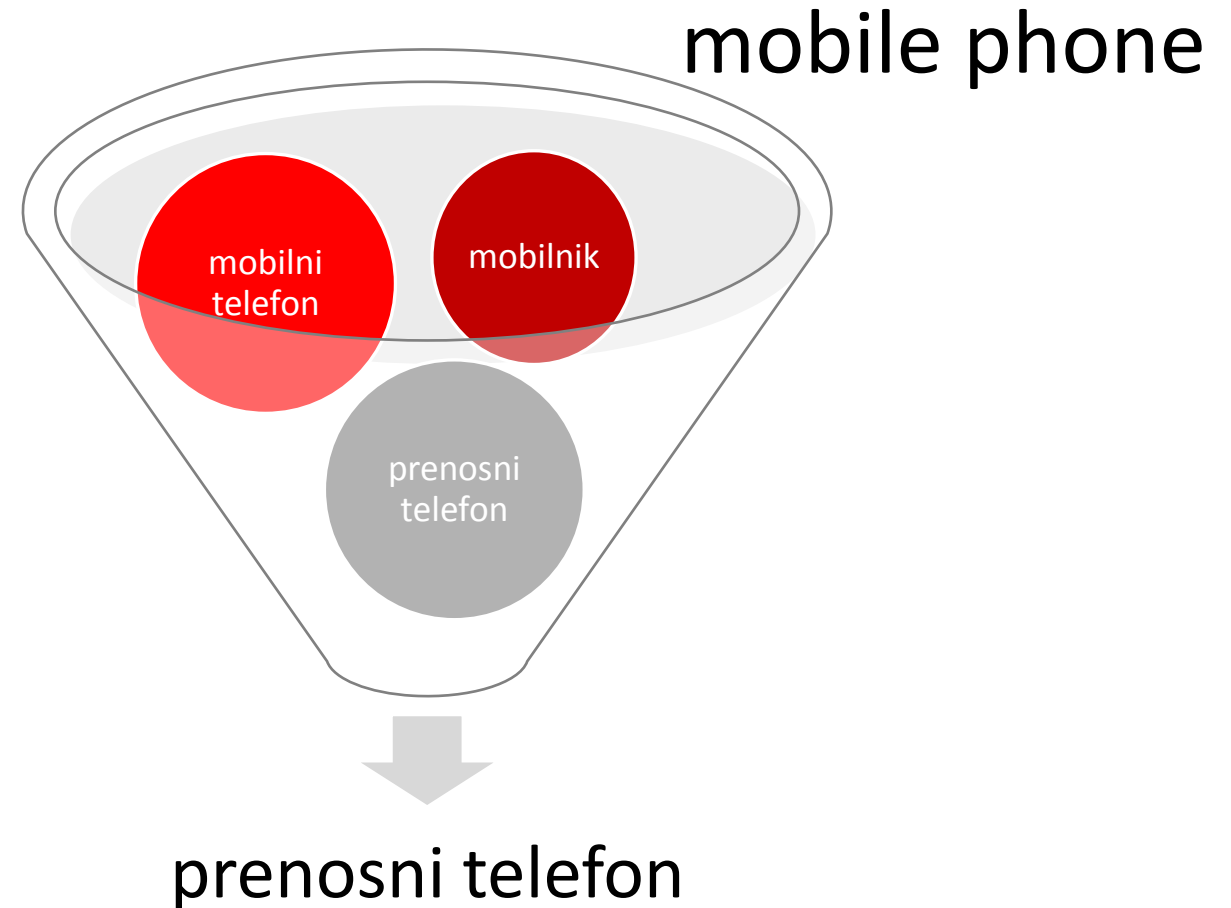
- Mednarodna raziskava, ki primerja uspešnost 15-letnih učencev na področju matematike, naravoslovja in bralne pismenosti.
- Raziskavo je leta 2000 lansirala OECD, v Sloveniji prvič izvedena 2006, od takrat na vsaka tri leta.
- Med dejavniki vpliva se omenjajo tudi prevodi testov v nacionalne jezike (npr. Grisay et al., 2007; Arffman, 2012; Solano-Flores et al., 2013).
- Jezikovna zahtevnost besedil: predpisana, vendar ni opredeljena metodologija, ki bi omogočila preverjanje zahtevanega stanja.

Raziskava

- Po naročilu Pedagoškega inštituta.
- Na testih bralne pismenosti za leti 2009 in 2012.
- Statistična primerjava jezikovnih pojavov, kakor se kažejo v referenčnih korpusih jezika izvirnika in prevoda.
- Ali se v slovenskih prevodih testov PISA pojavlja sprememba jezikovne zahtevnosti?
- Je mogoče razviti postopek, ki bi težavna mesta prevoda sistematično detektiral?

Jezikovna zahtevnost

- Povezujemo jo s tipičnostjo jezikovnih pojavov v rabi.
- Če med jezikovnimi možnostmi izberemo tisto, ki je v primerjavi z izvirnikom v rabi manj tipična/običajna/pogosta, zvišamo jezikovno zahtevnost testa.
- Ali obratno.



Primerjava korpusnih podatkov

1,2-milijardni referenčni
korpus Gigafida
za slovenski jezik

(Polnopomenske)
besede in besedne
zveze, ki se
pojavljajo v
besedilih
analiziranih testov.

2,1-milijardni referenčni
Oxford English Corpus
za angleški jezik

Podatki o širših
kolokacijah za
kombinacije besed,
ki se ne pojavljajo
neposredno
skupaj.

Test *Varnost prenosnih telefonov*

slovenska beseda	bes. vrsta	pog. Gigafida	pog. na milijon	angleška beseda	bes. vrsta	pog. OEC	pog. na milijon
mladi	sam.	1430	1,2	(the) young	sam.	78121	37,7
možgani	sam.	54190	45,7	brain	sam.	145120	70,0
nakazati	gl.	28049	23,6	suggest	gl.	562670	271,4
namen	sam.	241383	203,4	purpose	sam.	253924	122,5
napačen	prid.	56106	47,3	wrong	prid.	323375	156,0
napravica	sam.	4734	4,0	gadget	sam.	12190	5,9

Test Varnost prenosnih telefonov

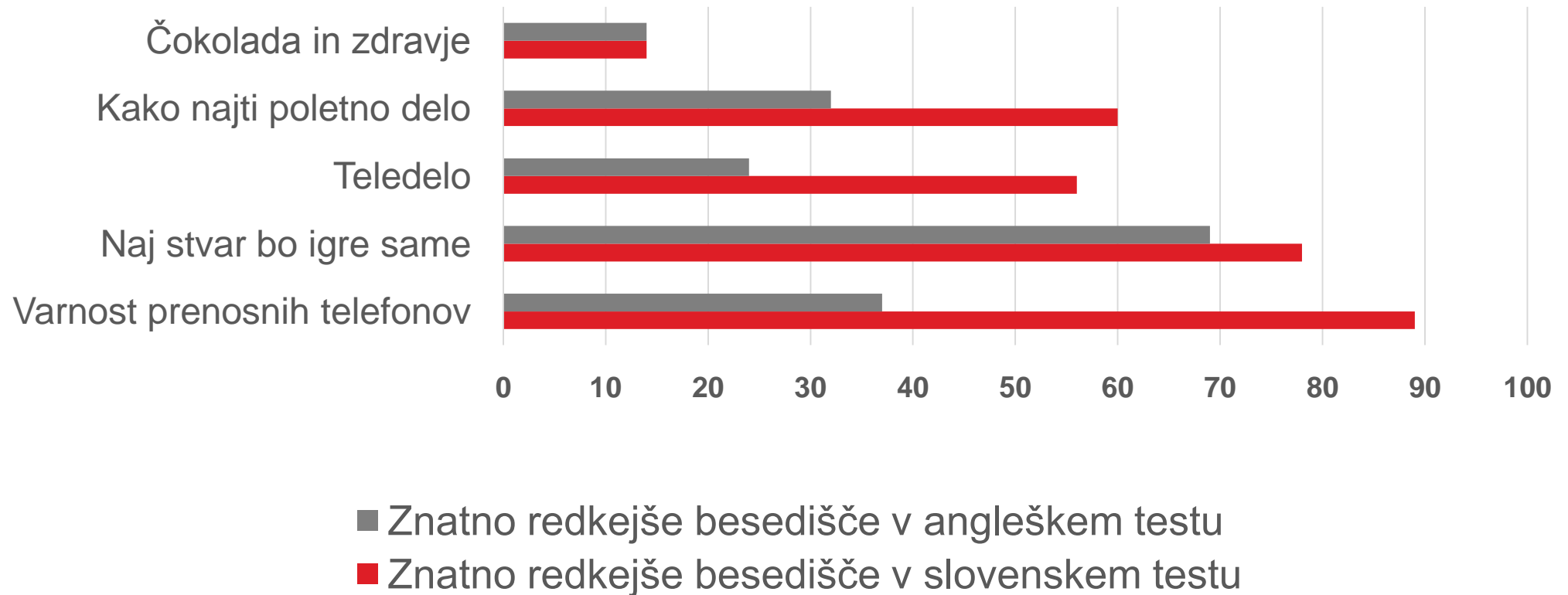
slovenska beseda	bes. vrsta	pog. Gigafida	slovnice razlike med jezikoma (dovršnost-nedovršnost)	bes. vrsta	pog. OEC	pog. na milijon	
med	sam.	1430	ng	sam	78121	37,7	
ogani	sam.	54190	45,7	brain	sam	145120	70,0
nakazati	g			suggest			
namen	sam.		4	purpose	sam	253924	122,5
napačen	prid.	56106	47,3	wrong	prid.	323375	156,0
napravica	sam.	4734	4,0	gadget	sam	12190	5,9

označenost podatkov

razlike v pomenski členjenosti

izhodišče za kvalitativno analizo

Redkejše besedišče



Analiza redkega besedišča

- Težave predvsem na kolokacijski ravni, kjer najdemo več primerov, pri katerih je slovenska kolokacija statistično šibkejša kot angleška, kot obratno.
- Rezultat so besedila, ki na prvi pogled delujejo jezikovno neproblematična, vendar so zaradi vrste atipičnih kombinacij manj tekoča in »naravna«, kot to velja za angleško edicijo testa.
- Zveze, ki jih v korpusu Gigafida ni moč najti, se pojavljajo tudi v delu testa z vprašanji oziroma nalogami za učence.

Telecommuting vs. teledelo

TELECOMMUTING

The way of the future

Just imagine how wonderful it would be to “telecommute” to work on the electronic highway, with all your work done on a computer or by phone! No longer would you have to jam your body into crowded buses or trains or waste hours and hours travelling to and from work. You could work wherever you want to – just think of all the job opportunities this would open up!

Molly

TELEDELO

Delo prihodnosti

Predstavljajte si, kako krasno bi bilo imeti “teledelo” in potovati po elektronski avtocesti, pri čemer bi vse delo opravili na računalniku ali prek telefona! Ne bi se vam bilo treba več gnesti na natrpanih avtobusih ali vlakih ali zapravljati dolgih ur za vožnjo v službo in domov. Lahko bi delali, kjer bi hoteli - samo pomislite, koliko priložnosti za delo bi to odprlo!

Maja

Telecommuting vs. teledelo

TELECOMMUTING

The way of the future

Just **imagine** how wonderful it would be to “telecommute” to **work** on the electronic highway, with all your work done on a computer or by phone! No longer would you have to jam your body into crowded buses or trains or waste **hours** and hours travelling to and from work. You could work wherever you want to – just think of all the job opportunities this would **open up!**

Molly

TELEDELO

Delo prihodnosti

Predstavljajte si, **kako krasno** bi bilo imeti “teledelo” in potovati po elektronski avtocesti, pri čemer bi vse delo opravili na računalniku ali prek telefona! Ne bi se vam bilo treba več **gnesti na natrpanih avtobusih** ali vlakih ali **zapravljati dolgih ur** za vožnjo v **službo** in domov. Lahko bi delali, kjer bi **hoteli** - samo **pomislite**, koliko **priložnosti za delo** bi to odprlo!

Maja

Telecommuting vs. teledelo

TELECOMMUTING

The way of the future

Just imagine how wonderful it would be to “telecommute” to work on the electronic highway, with all your work done on a computer or by phone! No longer would you have to jam your body into crowded buses or trains or waste hours and hours travelling to and from work. You could work wherever you want to – just think of all the job opportunities this would open up!

Molly

DELO NA DALJAVO

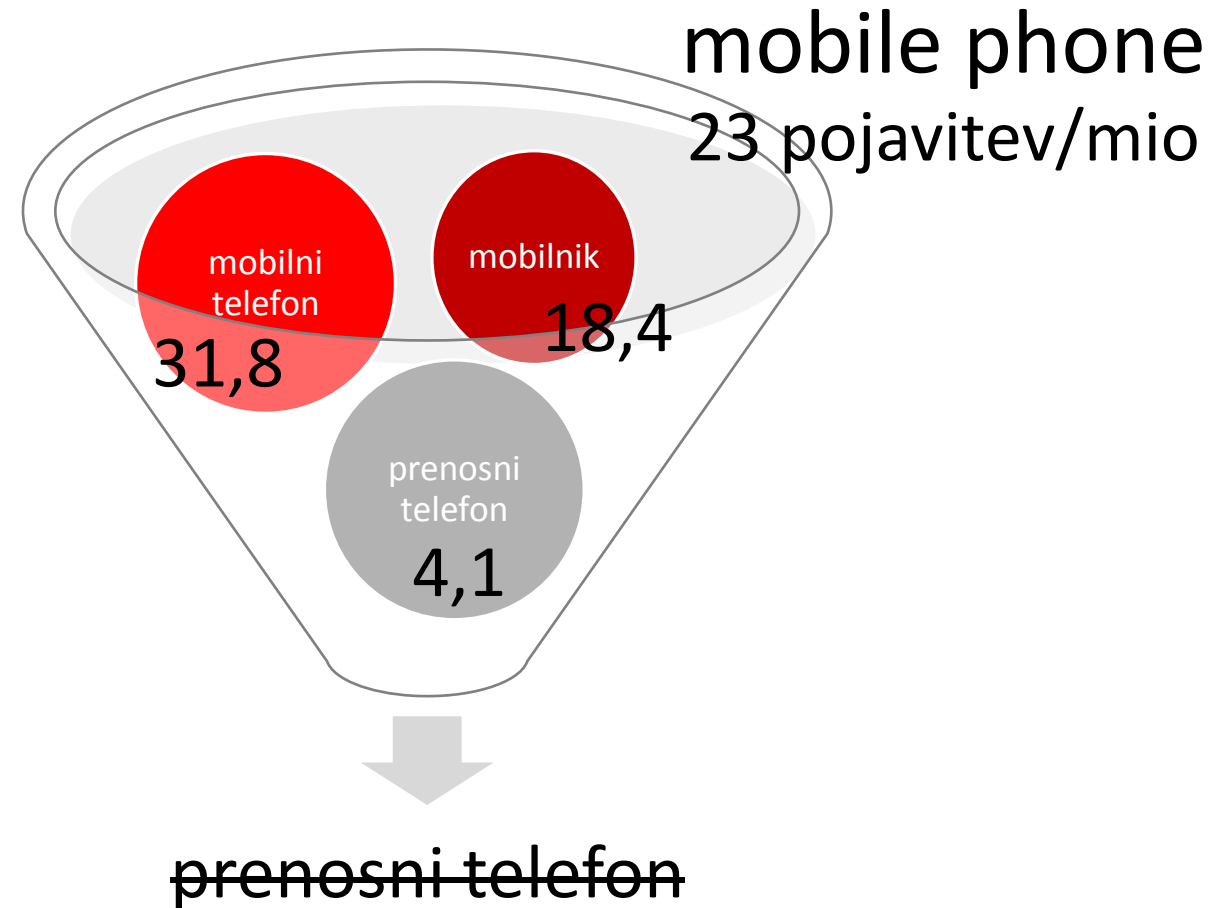
Delo prihodnosti

Predstavljajte si, kako čudovito bi bilo delati na daljavo in se peljati v službo po elektronski avtocesti, pri čemer bi vse delo opravili na računalniku ali prek telefona! Ne bi se vam bilo treba več drenjati na prepolnih avtobusih ali vlakih ali zapravljati časa za vožnjo v službo in domov. Lahko bi delali, kjer bi želeli - samo pomislite, koliko možnosti za zaposlitev bi to odprlo!

Maja

Podpora prevajanju

- Predlagani postopek je lahko dobrodošla povratna informacija za prevajalca.
- Ključno pa je zagotoviti, da ima slednji že v prvem koraku na voljo jezikovne podatke, ki jih potrebuje.
- Pri slovarski predstavitvi kolokacij in sinonimije je nujno vključiti podatke o pogostnosti.



Sklep

- S predlaganim postopkom je mogoče identificirati potencialno problematična mesta prevoda in omogočiti prevajalcu širši, sintetični pogled na zahtevnost izvirnika in besedila.
- Postopek omogoča izboljšanje stanja z relativno nizkim finančno-časovnim vložkom.
- Ne vemo pa še, v kolikšni meri lahko izboljšave prevoda dejansko vplivajo na rezultate testiranj bralne pismenosti PISA v Sloveniji.
- Diskusija o predlagani metodologiji v prispevku.

pozornost
in dober
tek!



Literatura

- Inga Arffman.** 2012. Unwanted Literal Translation: An Underdiscussed Problem in International Achievement Studies. *Education Research International*, 2016 (ID 503824): 1–13.
- Špela Arhar Holdt, Iztok Kosem in Polona Gantar.** 2016. Corpus-Based Resources for L1 Teaching: The Case of Slovene. V: A. Marcus-Quinn in T. Hourigan, ur. *Handbook on Digital Learning for K-12 Schools*. Springer, v tisku.
- Averil Coxhead.** 2000. A New Academic Word List. *TESOL Quarterly* 34(2): 213–238.
- Vojko Gorjanc, Polona Gantar, Iztok Kosem in Simon Krek,** ur. 2015. *Slovar sodobne slovenščine: problemi in rešitve*. Znanstvena založba Filozofske fakultete UL.
- Aletta Grisay, John H.A.L. de Jong, Eveline Gebhardt, Alla Berezner in Beatrice Halleux-Monseur.** 2007. Translation equivalence across PISA countries. *Journal of Applied Measurement*, 8(3): 249–266.
- Adam Kilgarriff, Pavel Rychlý, Pavel Smrz in David Tugwell.** 2004: The Sketch Engine. V.: G. Williams in S. Vessier, ur. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004 Lorient, France July 6–10, 2004*, str. 105–116, Lorient. Universite de Bretagne - sud.
- Iztok Kosem, Mojca Stritar Kučuk, Sara Može, Ana Zwitter Vitez, Špela Arhar Holdt in Tadeja Rozman.** 2012. *Analiza jezikovnih težav učencev: korpusni pristop*. Trojina, zavod za uporabno slovenistiko.
- Simon Krek, Polona Gantar, Špela Arhar Holdt in Vojko Gorjanc.** 2016. Nadgradnja korpusov Gigafida, Kres, ccGigafida in ccKres. *Konferenca Jezikovne tehnologije in digitalna humanistika 2016*, v pripravi.
- Nataša Logar Berginc, Miha Grčar, Marko Brakus, Tomaž Erjavec, Špela Arhar Holdt in Simon Krek.** 2012. *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, uporaba*. Ljubljana, Trojina, zavod za uporabno slovenistiko; Fakulteta za družbene vede.
- Paul Nation, Robert Waring.** 1997. Vocabulary size, text coverage and word lists. V N. Schmitt, M. McCarthy (ur.) *Vocabulary: Description, Acquisition and Pedagogy*, str. 6–19. Cambridge: Cambridge University Press.
- OECD.** 2010. *Translation and adaption guidelines for PISA 2012*. Dostop 5. 3. 2016: <http://www.oecd.org/pisa/pisaproducts/49273486.pdf>
- Tadeja Rozman, Irena Krapš Vodopivec, Mojca Stritar Kučuk in Iztok Kosem.** 2012. *Empirični pogled na pouk slovenskega jezika*. Trojina, zavod za uporabno slovenistiko.
- Guillermo Solano-Flores, Luis Ángel Contreras-Niño in Eduardo Backhoff.** 2013. The measurement of translation error in PISA-2006 items: An application of the theory of test translation error. V: M. Prenzel, M. Kobarg, K. Schöps in S. Rönnebeck, ur. *Research on PISA*, str. 71–85. Springer Netherlands.