

Razdvoumljanje besednega  
pomena pri strojnih prevajalnikih  
Amebis Presis, Google Translate  
in MT@EC

- Ana je jabolko.  $\neq$  Ana is an apple.
- Ana je jabolko. = Ana eats an apple.
  
- Troje povesti  $\neq$  Troy to take the lead
- Troje povesti = Three stories

- Why does Humpty Dumpty love autumn?  
Because he had a great **fall**.
- Vsi čakajo soboto, samo Franci **Petek**.

Izvirnik:

“**Enoglasna pritrditev** načrtu, pa smo šli.”

Google Translate:

“**Monophonic mounting** plan, we went.”

# Večpomenskost



# Vendar ...

Dvoumnost ni problematična, kadar imamo v ciljnem jeziku samo en prevod:

ribja glava = fish head

glava družine = head of the family

# Eksperiment

Cilj je bil izmeriti natančnost razdvoumljanja za:

- tri prevajalnike (Google Translate, Amebis Presis, MT@EC)
- tri besedilne žanre (književnost, publicistika, spletni dnevniki)
- smer prevajanja slovenščina → angleščina

# Metoda

- Skupni obseg testnih besedil okoli 900 besed
- Književnost, publicistika in spletni dnevniki
- Orodja za predobdelavo podatkov: Excell, Besana, SSKJ, [thefreedictionary.com](http://thefreedictionary.com) in [sloWNet](http://sloWNet).



# Predstavljena verzija pregibnika Amebis Besana 4.12

Za preizkus kakovosti pregibanja v polje spodaj vpišite besedo ali besedno zvezo. Če nimate nameščene slovenske tipkovnice, pišite *kožo* ali *ko^so* in ne *kaso*, *kasha* ali *kos^o*.

 Pregibaj

Oblika za pregibanje

## povesti

glagol

palnopomenski glagol

nedoločnik




povesti

namenilnik

povest

povednik

sedanjik







|              |  ednina |  dvojina |  množina |
|--------------|--|---|---|
| prva oseba   | povedem  | povedeva  | povedemo  |
| druga oseba  | povedeš  | povedeta  | povedete  |
| tretja oseba | povede   | povedeta  | povedejo  |

velelnik

sedanjik

|             |  ednina |  dvojina |  množina |
|-------------|--|---|---|
| prva oseba  |  | povediva  | povedimo  |
| druga oseba | povedi   | povedita  | povedite  |

deležnik na -l

|  |  ednina |  dvojina |  množina |
|--|--|---|---|
|  moški spol   | povedel  | povedla   | povedli   |
|  ženski spol  | povedla  | povedli   | povedle   |
|  srednji spol | povedlo  | povedli   | povedla   |

## povest

samostalnik


občno ime

 ženski spol

 ednina  dvojina  množina

## Zadetki iskanja


**povést** -i ž (ê)

SSKJ 

- lit. *krajše, vsebinsko, izrazno manj zahtevno pripovedno delo, navadno v prozi*: napisati, prebrati povest; zgodovinska povest; povest v verzih; povesti in romani / okvirna povest / povest se dogaja v 18. stoletju
- daljša zgodba z izmišljenimi sestavinami*: ded je rad pripovedoval povesti; njegove lovske povesti sem že vse slišal / povedati komu svojo življenjsko povest *zgodbo*
- zastar. *zgodovina*: zanimati se za povest jezika Slovencev
  - ekspr. o kupovanju čevljev je povedal celo povest *zelo obširno je pripovedoval o njem*; eksp. to je pa dolga povest *o tem bi se dalo veliko povedati*; eksp. kaj bi še pravil: stara povest *tako se navadno zgodi, tako stvari navadno potekajo*

## Zadetki iskanja

**povèsti** -vêdem dov., povêdel in povêdel povêdla, stil. povêl povêla (é)

SSKJ 

- knjiž. *kot spremljevalec, vodnik narediti, da kdo kam gre; popeljati*: spoštljivo ga je povedel v hišo; povedel jo je v drugo sobo / sam ga je povedel po stopnicah / povedel jih je v boj
  - vznes. povedel jih je zmagi naproti *pod njegovim vodstvom so zmagali*; knjiž. povsod so šli za njim, kamorkoli jih je povedel v *vsem so ga podpirali*
- publ. *pri tekmovanju na prvo mesto, pred druge*: na ovinku je povedel kolesar Partizana // *doseči prednost v igri*: v začetku je povedlo domače moštvo / povesti z ena proti nič

# sloWTool

kot|

Slovenian

Number of hits: 6

[Link](#)

POS: Noun 10: eng-30-08544410-n POS:IN: factorum CLUSTER:0: recodingz APPROVED:

SYNONYM (SLV): **kot**

SYNONYM (ENG): **corner**

SYNONYM (FR): **coin**

DEFINITION: *a place off to the side of an area*

USAGE: *he tripled to the rightfield corner*

USAGE: *the southeastern corner of the Mediterranean*

→ [HYPERNYM]: *območje, področje, country, aire, région, région, aire, area, campagne, pays, zone*

POS: Noun 10: eng-30-09257011-n POS:2 POS:IN: factorum CLUSTER:0: recodingz APPROVED:

SYNONYM (SLV): **kot, vogal, vogel**

SYNONYM (ENG): **corner**

SYNONYM (FR): **coin**

DEFINITION: *a projecting part where two sides or edges meet*

USAGE: *he knocked off the corners*

→ [HYPERNYM]: *del, kos, piece, part, élément, partie, rôle, morceau, part, part, partie, cadre, pièce*

STVAR: K.ričina 2012-07-23 22:47:17

POS: Noun 10: eng-30-13887509-n POS:2 POS:IN: factorum CLUSTER:0: Anglez APPROVED:

SYNONYM (SLV): **kot**

SYNONYM (ENG): **angle**

SYNONYM (FR): **angle**

DEFINITION: *the space between two lines or planes that intersect; the inclination of one line to another; measured in degrees or radians*

→ [ENG\_DERIVATIVE]: *nagniti se, nagibati se, zanašati se, nagniti, nagibati, nagniti (se), nagibati (se), slant, lean, angle, tilt, tip, pencher*

→ [ENG\_DERIVATIVE]: *oster, robot, oglas, angulaire, angular, angulate, anguleux*

→ [ENG\_DERIVATIVE]: *oglat, angulaire, anguleux, angular*

→ [HYPERNYM]: *prostor, local, espace, space*

→ [ENG\_DERIVATIVE]: *angulate*

→ [ENG\_DERIVATIVE]: *angle*

← [HYPERNYM]: *pravi kot, right angle, angle droit*

← [HYPERNYM]: *magnetna deklinacija, magnetna variacija, variation, magnetic variation, magnetic declination, déclinaison magnétique terrestre, déclinaison magnétique*

← [HYPERNYM]: *vile, entrejambe, fork, crotch*

← [HYPERNYM]: *pazduha, axil*

← [HYPERNYM]: *helix angle*

← [HYPERNYM]: *complementary angles*

← [HYPERNYM]: *kotna razdalja, angular distance*

← [HYPERNYM]: *plane angle*

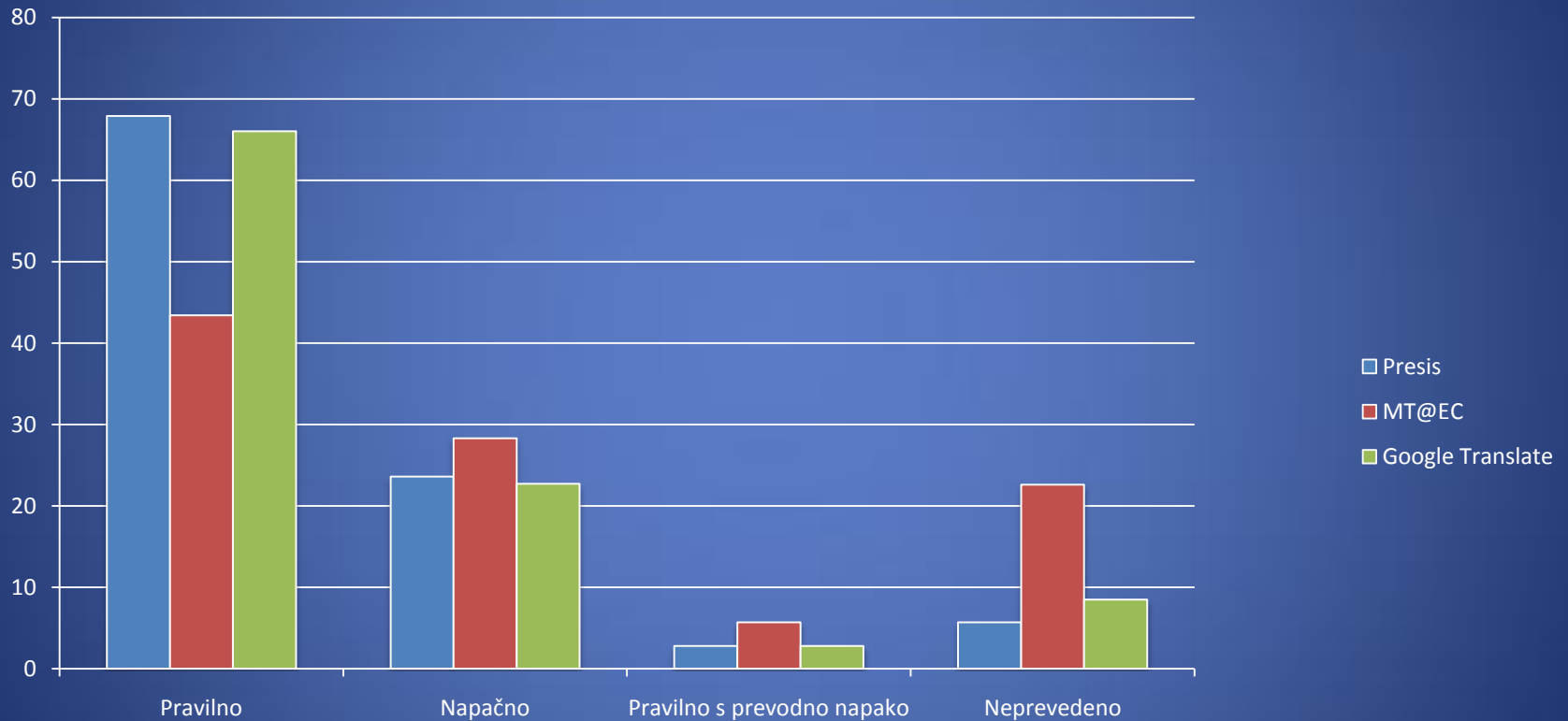
← [HYPERNYM]: *spherical angle*

← [HYPERNYM]: *angle solide, solid angle*

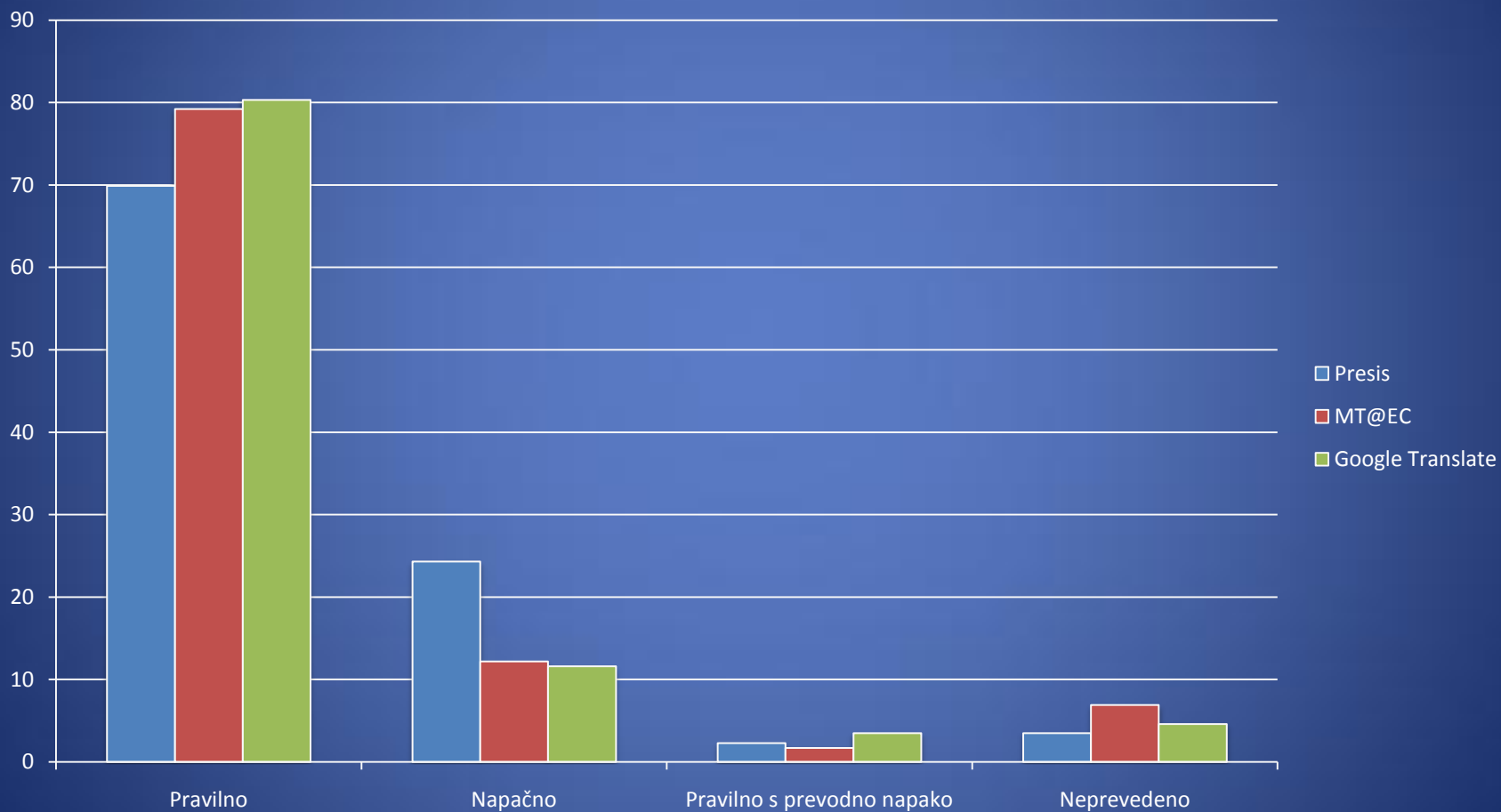
← [HYPERNYM]: *naklon, inclinaison, inclination of an orbit, inclination*

| Presis                     | KNJIŽNA BESEDILA | PUBLICISTIČNA BESEDILA | SPLETNA BESEDILA | SKUPAJ |
|----------------------------|------------------|------------------------|------------------|--------|
| Pravilno                   | 67,9 %           | 69,9 %                 | 73,3 %           | 70,4 % |
| Napačno                    | 23,6 %           | 24,3 %                 | 19,8 %           | 22,8 % |
| Pravilno s prevodno napako | 2,8 %            | 2,3 %                  | 2,6 %            | 2,5 %  |
| Neprevedeno                | 5,7 %            | 3,5 %                  | 4,3 %            | 4,3 %  |
| Google Translate           | KNJIŽNA BESEDILA | PUBLICISTIČNA BESEDILA | SPLETNA BESEDILA | SKUPAJ |
| Pravilno                   | 66,0 %           | 80,3 %                 | 75,9 %           | 75,2 % |
| Napačno                    | 22,7 %           | 11,6 %                 | 14,6 %           | 15,4 % |
| Pravilno s prevodno napako | 2,8 %            | 3,5 %                  | 1,7 %            | 2,8 %  |
| Neprevedeno                | 8,5 %            | 4,6 %                  | 7,8 %            | 6,6 %  |
| MT@EC                      | KNJIŽNA BESEDILA | PUBLICISTIČNA BESEDILA | SPLETNA BESEDILA | SKUPAJ |
| Pravilno                   | 43,4 %           | 79,2 %                 | 61,2 %           | 64,3 % |
| Napačno                    | 28,3 %           | 12,2 %                 | 21,6 %           | 19,2 % |
| Pravilno s prevodno napako | 5,7 %            | 1,7 %                  | 3,4 %            | 3,3 %  |
| Neprevedeno                | 22,6 %           | 6,9 %                  | 13,8 %           | 13,2 % |

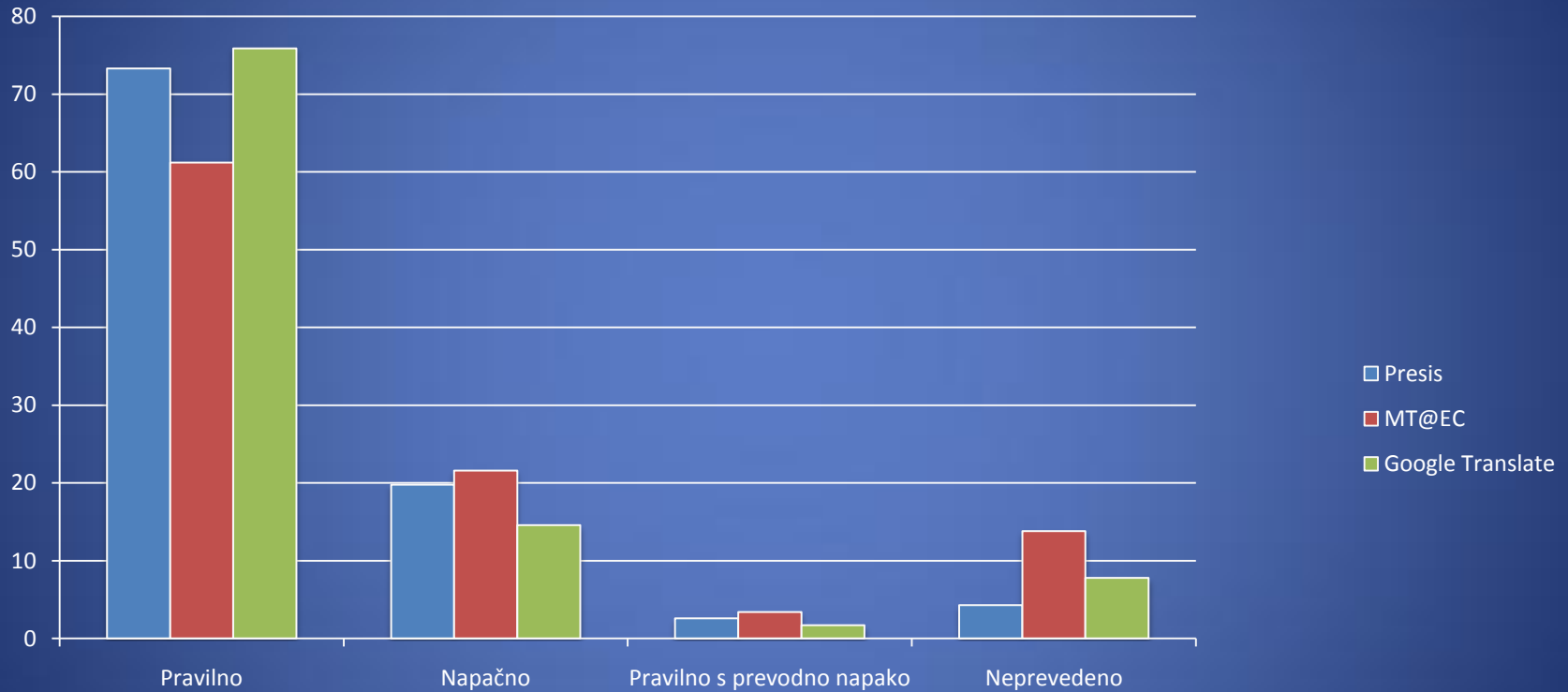
# Natančnost RBP v književnih besedilih



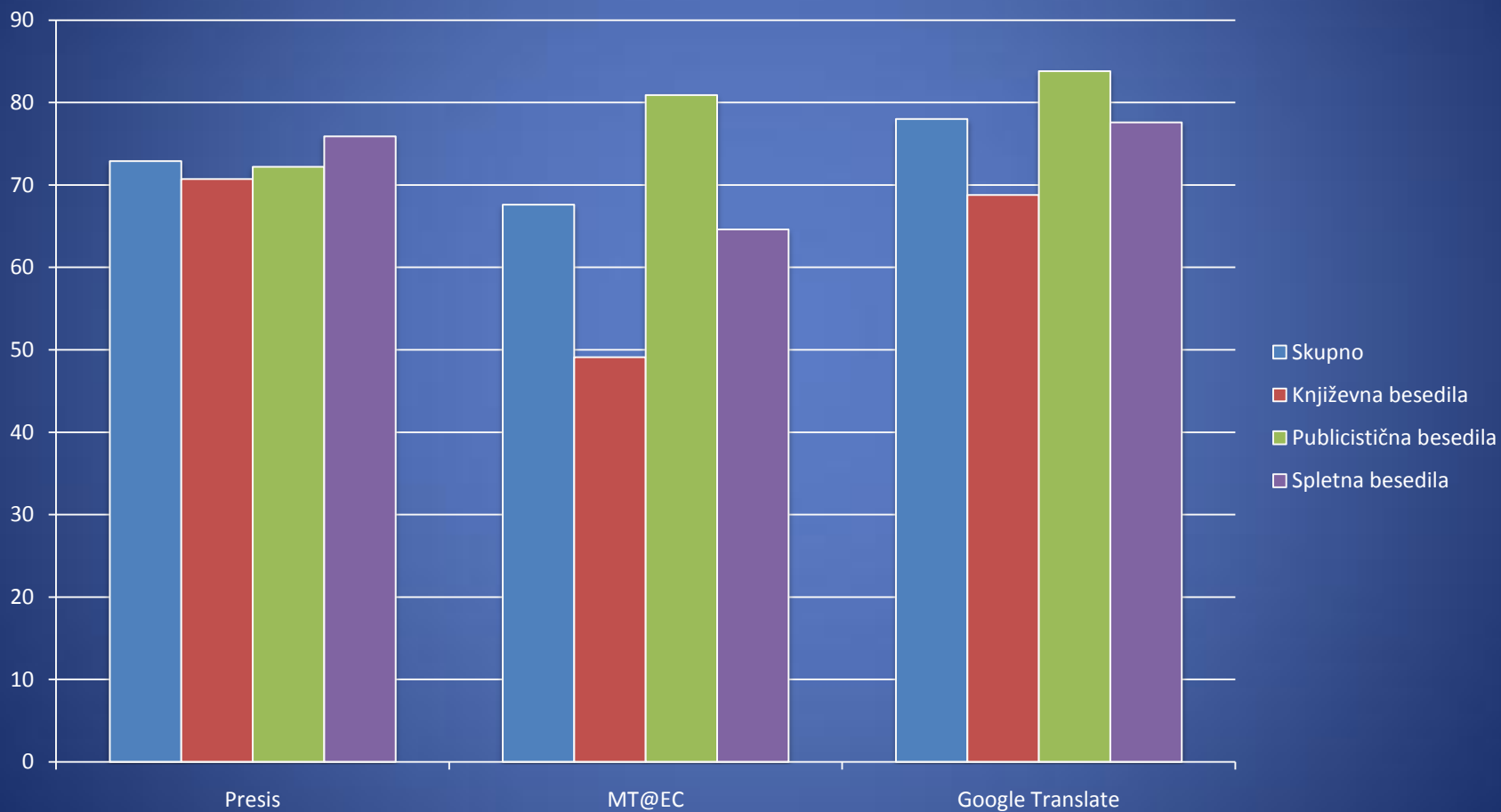
# Natančnost RBP v publicističnih besedilih



# Natančnost RBP v spletnih dnevnikih



# Primerjava po žanrskih kategorijah





-“How much data is needed?”

-“More.”

-Franz Josef Och

Hvala za pozornost!! 😊