

# GRADNJA IN ANALIZA PETJEZIČNEGA KORPUSA PODNASLOVOV GOVOROV TED

MIHA HELBL, ŽIGA DOMEVŠČEK

---

# POTEK PREDSTAVITVE

- Uvod
  - Kratka predstavitev TED
  - Metodologija – gradnja korpusa
  - Metodologija – analiza podnaslovov
  - Hipotezi
  - Analiza korpusa
  - Analiza podnaslovov
  - Zaključek
-

# UVOD

- Raziskava prevodov prostovoljnih prevajalcev za nepridobitno organizacijo TED
  - Prikaz razhajanja med izvirnimi angleškimi transkripcijami in slovenskimi podnaslovi
  - Gradnja petjezičnega korpusa:
    - Slovenščina,
    - angleščina,
    - nemščina,
    - italijanščina,
    - francoščina.
  - Analiza podnaslovov
  - Prikaz najpogostejših besed v slovenskem in angleškem delu korpusa
-

# KRATKA PREDSTAVITEV TED

- Nepridobitna organizacija TED ustanovljena leta 1984
    - Prvotno širjenje idej na področjih tehnologije, zabave in oblikovanja (**T**echnology, **E**ntertainment, **D**esign)
    - Danes vsa področja od znanosti do gospodarstva in do globalnih vprašanj
  - Možen brezplačen ogled govorov na spletu
    - Podnaslovi v 110 jezikih
    - 106 prevajalcev za slovenščino z več kot 300 govori
  - Projekt odprtega prevajanja
    - Na željo obiskovalcev spletišča
-

# SMERNICE IN MNOŽIČENO PREVAJANJE

- Uskladitev smernic za podnaslove
  - Najmanj dva prostovoljca pri določenem projektu (lektor in prevajalec)
  - Prostorske in časovne omejitve
  - Interes, ki združuje člane prevajalski projektov.
  - Lektoriranje, prevajanje, zbiranje idej, virov, gradnja pomnilnikov prevodov.
-

# METODOLOGIJA – GRADNJA KORPUSA

TED270						
Jezik	Naslov	Govorec	Prevajalec	Lektor	Opis	Čas snemanja
ANG	We're all hiding something. Let's find the courage to open up	Ash Beckham	/	/	In this touching talk, Ash Beckham offers a fresh approach to empathy and openness. It starts with understanding that everyone, at some point in their life, has experienced hardship. The only way out, says Beckham, is to open the door and step out of your closet.	2013 septembe TEDxBoulder
SLO	Vsi nekaj skrivamo. Najdimo pogum, da se odpremo.	Ash Beckham	Klavdija Cernilogar	Nika Kotnik	V tem ganljivem govoru Ash Beckham ponudi svež pristop k empatiji in odprtosti. Začne s premiso, da vsak človek v toku svojega življenja doživi težke trenutke. Edini izhod, pravi Beckhamova, je, da odpremo vrata in izstopimo iz svoje omare.	2013 septembe TEDxBoulder
NEM	Alle haben Geheimnisse. Finden wir den Mut, uns zu öffnen.	Ash Beckham	Ludmila Schäfer	Sabrina Hettrich	In diesem berührenden Vortrag bietet Ash Beckham einen erfrischenden Denkansatz zu Empathie und Offenheit. Es beginnt mit dem Verständnis, dass jeder in seinem Leben schwere Zeiten durchlebt. Der einzige Weg hinaus, sagt Beckham, ist sich zu öffnen und das Geheimnis preiszugeben.	2013 septembe TEDxBoulder
ITA	Tutti noi nascondiamo qualcosa. Cerchiamo il coraggio per uscire allo scoperto. Ash Beckham	Ash Beckham	Azzurra Selvaggia Magrin	Paul Meighan	In questo emozionante discorso, Ash Beckham offre un approccio fresco all'empatia e alla franchezza. Inizia spiegando che tutti, ad un certo punto della loro vita, hanno fatto provato delle difficoltà. L'unica via d'uscita, afferma Beckham, è uscire allo scoperto.	2013 septembe TEDxBoulder
FRA	On cache tous quelque chose. Trouvons le courage de nous ouvrir	Ash Beckham	Diane Wolf	Beddiyuh Med Ali	Dans cette conférence touchante, Ash Beckham nous offre une nouvelle approche de l'empathie et de l'ouverture. Ça commence par la compréhension que chacun d'entre nous, à un moment de notre vie, a connu des difficultés. La seule solution, explique Beckham, est d'ouvrir la porte et sortir de son placard.	2013 septembe TEDxBoulder

- Zbiranje metapodatkov

# METODOLOGIJA – ANALIZA PREVODOV

- Omejitev na besedila v slovenščini in angleščini
- Analiza 20 govorov kjer je bila razlika v številu segmentov največja (20 %)
  - Segment – en podnaslov, ki je lahko tako enovrstičen kot dvovrstičen
- Opazovala sva uporabljene prevajalske strategije
  - Izpuste – diskurzivni označevalci, vsebinski izpusti, izpuste ponavljanja
  - Konkretizacijo
  - Posploševanje
  - Zamenjava samostalnika z zaimkom

# HIPOTEZA

- Zaradi prostovoljne narave prevajalskega projekta TED bo prihajalo do velikega odstopanja pri količini slovenskih podnaslovov
- V primerjavi z angleškim izvirnikom bo v slovenskih podnaslovih več izpustov oziroma zgostitev besedila kot dopolnjevanja – kognitivno lažje izpuščati kot dodajati



# ANALIZA KORPUSA

- Petjezičen korpus s 131 govori (skupno 968 559 besed)
- Analizirala sva število pojavnic in različnic za slovenski in angleški jezik
  - Skupno slovenski jezik: 181.536 pojavnic in 222.276 različnic
  - Skupno angleški jezik: 226.935 pojavnic in 270.741 različnic
- Razvidno, da je bilo v slovenskih podnaslovih veliko izpuščanja
- Seznam najpogostejših besed
  - Angleški jezik (pojavnice): people, say, know, think, make, year, time, world, good, laughter
  - Slovenski jezik (različnice): ljudje, stvari, smeh, leta, način, svet, aplavz, hvala, primer, videti
- And **people** turned around and were like, oh my God, I didn't even notice her sitting there. (**laughter**)
- To je tudi storila, **ljudje** so se ozrli in imeli reakcijo, kot da so jo prvič opazili, da sedi tam. (**smeh**)
- Five **years** in grad school, a few **years**, you **know**, I'm at Northwestern ...
- Pet let sem prebila na podiplomskem študiju, potem sem prebila nekaj let na univerzi Northwestern ...

# ANALIZA PREVODOV

- Analizo strategij krajšanja prevodov sva razdelila na 6 večjih sklopov
- Najpogostejši so bili izpusti diskurzivnih označevalcev v slovenskih podnaslovih (46%)
  - Angleški podnapisi so transkripcije
- Druga najpogosteje uporabljena strategija so bili vsebinski izpusti (20 %)
- Posploševanje (14 %)
- Zamenjava samostalnika z zaimkom (11 %)
- Izpust ponavljanja (5 %)
- Najredkejša pa je bila konkretizacija (4 %)

# PRIMERI IZPUŠČANJA DISKURZIVNIH OZNAČEVALCEV

<b>Well, it's very simple.</b>	<b>To je enostavno.</b>
<b>And then finally -- you know, /.../</b>	<b>Na koncu --najboljša možnost-- /.../</b>
<b>Now, when psychologists show you bars, /.../</b>	<b>Ko vam psihologi kažejo stolpce, /.../</b>
<b>So thank you very much.</b>	<b>Najlepša hvala.</b>
<b>»Oh, I created this great product, /.../«</b>	<b>»Ustvaril sem odličen produkt, /.../«</b>

# PRIMERI VSEBINSKEGA IZPUŠČANJA

**/.../ a pair of safety glasses, /.../**

**/.../ ker lahko s parom očal, /.../**

**/.../ extra-virgin olive oils /.../**

**/.../ vrst ekstra deviškega olja /.../**

**My goal is to better understand these violent actors /.../**

**Ta nasilna gibanja želim bolje razumeti /.../**

**/.../that they can actually have experiences in their heads/.../**

**/.../ da lahko dejansko doživlja dogodke**

**/.../ from your life.**

**/**

# PRIMERI POSPLOŠEVANJA

**You failed the pop quiz, and you're hardly five minutes into the lecture.**

**To je enostavno. Po petih minutah predstavitve ste že padli na preizkusu znanja.**

**(Laughter) Because she was the main reason we were leaving the country.**

**Na koncu --najboljša možnost-- ker je bila ona glavni razlog za naš odhod.**

**and it's not because they whipped some up, tried it and went, "Yuck."**

**Ko vam psihologi kažejo stolpce, Ne zato, ker bi ga kdo naredil, poizkusil in se zgrozil.**

**so that they can feel better about the worlds in which they find themselves.**

**zato, da se lahko bolje počuti v danih okoliščinah.**

**Like Sir Thomas, you have this machine.**

**Kot g. Thomas imate tudi v to sposobnost.**

# PRIMERI ZAMENJAVE SAMOSTALNIKA Z ZAIMKOM

<b>There's something curious about professors in my experience -- /.../</b>	<b>Nekaj zelo zanimivega je na njih po mojih izkušnjah -- /.../</b>
<b>But if I do this with amnesiac patients, /.../</b>	<b>Na koncu --najboljša možnost-- Če pa to naredim pri naših bolnikih, /.../</b>
<b>He has been working on the topic for 20 years.</b>	<b>Ko vam psihologi kažejo stolpce, Na tem je delal že 20 let.</b>
<b>We need to know what makes These organisations tick.</b>	<b>Vedeti moramo, kaj jih vodi.</b>
<b>On the morning of the hemorrhage, I could not walk, talk, read, write or recall any of my life.</b>	<b>To jutro nisem mogla hoditi, govoriti, brati, pisati ali se spomniti česarkoli v svojem življenju.</b>

# PRIMER IZPUŠČANJA PONAVLJANJ

<b>A system of cognitive processes, largely non-conscious cognitive processes, that help them change</b>	<b>Sistem miselnih procesov, večinoma nezavednih, /.../</b>
<b>Third: make somebody else really, really rich.</b>	<b>Tretje: Nekoga drugega naredite zelo bogatega.</b>
<b>/.../ where a tiny, tiny fraction of the world /.../</b>	<b>/.../ kjer se lahko le zelo majhen del sveta /.../</b>
<b>WK: In Malawi, Kasungu. In Kasungu.</b>	<b>WK: V Malaviju v mestu Kasung.</b>
<b>You see the thing is, the thing is /.../</b>	<b>Vidite, stvar je v tem /.../</b>

## PRIMERI KONKRETIZACIJE

Here's two different futures that I invite you to contemplate.

To je enostavno. Razmislite o dveh različnih življenjskih poteh, /.../

from the one they like the most, to the one they like the least.

Na koncu --najboljša možnost-- po vrsti, glede na to, koliko so mu všeč.

/.../ from the one you like the most to the one you like the least."

Ko vam psihologi kažejo stolpce, slike glede na to, koliko so vam všeč."

So he recognizes that I need help and he gets me help.

Najlepša hvala. Kolega dojame, da potrebujem pomoč in mi jo pošlje.



# ZAKLJUČEK

- Namen projekta je bil preučiti odstopanja pri podnaslovnih prevodih govorov TED
  - Želela sva prikazati kakšni pristopi krajšanja izvirnega besedila so najpogostejši s strani prostovoljnih prevajalcev za TED
  - Ugotovila sva, da se odstopanja niso pojavljala tako pogosto kot sva pričakovala
  - Razlog za to vidiva v natančnih smernicah s strani TED in doslednemu pregledu in lekturi prevajalske ekipe
  - Dopolnitev korpusa s preostalimi govori
-

**HVALA ZA POZORNOST**

---