

Multilingual and Parallel Corpora Translation Studies

Amir Zeldes

amir.zeldes@georgetown.edu

Reminder: the PTB tag set (NB: variations)

CC	Coordinating conjunction
CD	Cardinal number
DT	Determiner
EX	Existential there
FW	Foreign word
IN	Preposition or conjunction
JJ	Adjective
JJR	Adjective, comparative
JJS	Adjective, superlative
LS	List item marker
MD	Modal
NN	Noun, singular or mass
NNS	Noun, plural
NNP	Proper noun, singular
NNPS	Proper noun, plural
PDT	Predeterminer
POS	Possessive ending
PRP	Personal pronoun

PRP\$	Possessive pronoun
RB	Adverb
RBR	Adverb, comparative
RBS	Adverb, superlative
RP	Particle
SYM	Symbol
TO	to
UH	Interjection
VB	Verb, base form
VBD	Verb, past tense
VBG	Verb, gerund or present participle
VCN	Verb, past participle
VBP	Verb, non-3rd person sg. present
VBZ	Verb, 3rd person sg. present
WDT	Wh-determiner
WP	Wh-pronoun
WP\$	Possessive wh-pronoun
WRB	Wh-adverb

Searching for POS tags

- To combine POS and token searches, use:

English	<input [pos='\"RP\"]"/' type="text" value="[pos=\" vvz\"]=""/>
German	<input type="text"/>

- Similarly across languages:

English	<input &="" rp\"="" type="text" value="[pos=\" word='\"out\"]"/'/>
German	<input &="" ptkvz\"="" type="text" value="[pos=\" word='\"heraus\"]"/'/>

midwife , with lilac ribbons in her cap , came **out** of Anna 's boudoir . She approached
Karenin , and 🔍

Im Salon war niemand ; aus Annas Wohnzimmer kam auf das Geräusch seiner Schritte
die Hebamme in einer Haube mit lila Bändern **heraus** . 🔍

Operators

- `a+` → `a, aa, aaa ...`
- `.` → `b, q, 3, - ...`
- `yea*h` → `yeh, yeah, yeaah, yeaaaah`
...
- `anyways?` → `anyway, anyways`
- `(girl|boy)s?` → `girl, girls, boy, boys`

`[word=".*s" & lemma!= ".*s" & pos="NN?P?S"]`

Operators apply to tokens too!

- CQP allows you to search for patterns of tokens/annotations too!

- Say we want to find uses of “a lot of”
- But people could say “a whole lot of” ...
- Other things?

`[word="a"] [pos="JJ"]* [word="lot"] [word="of"]`

- How do you make sure there's at least one adjective?

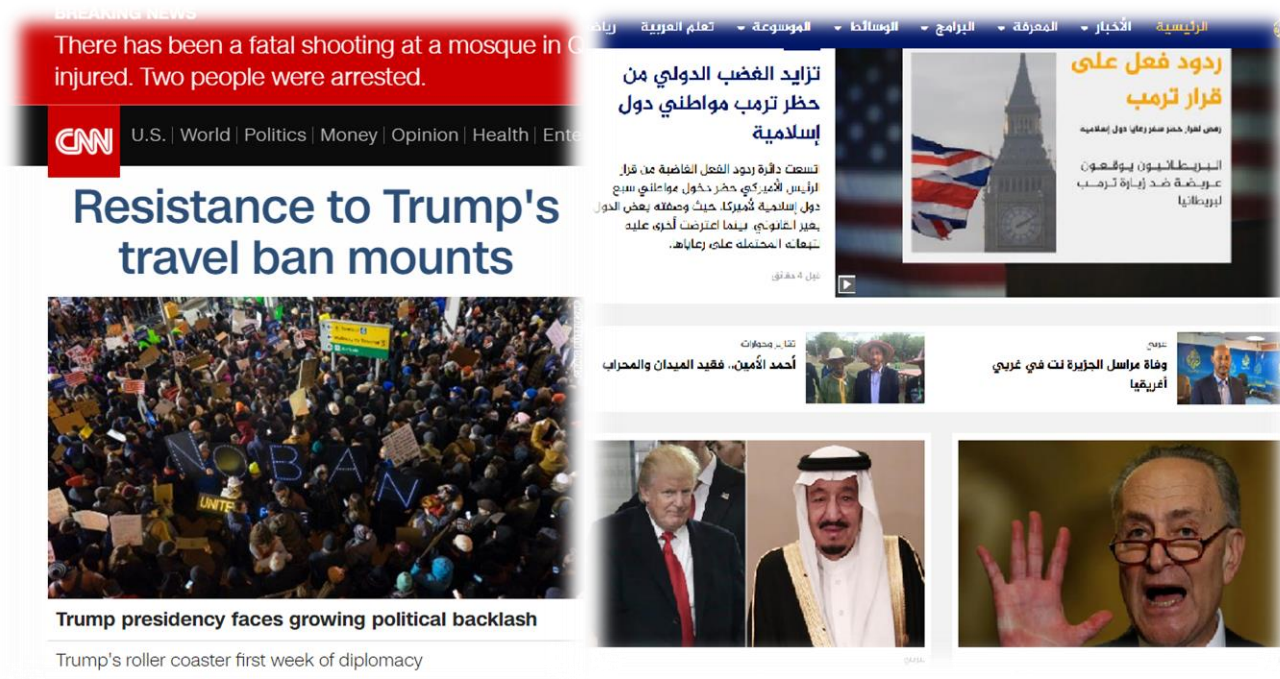
What's 'a good lot'?

Anna Karenina:

- You must have a thousand limes here , and each one of them would yield **a good lot of bast**
- Ведь их тысяча лип , из каждой **два хороших лубка** выйдет

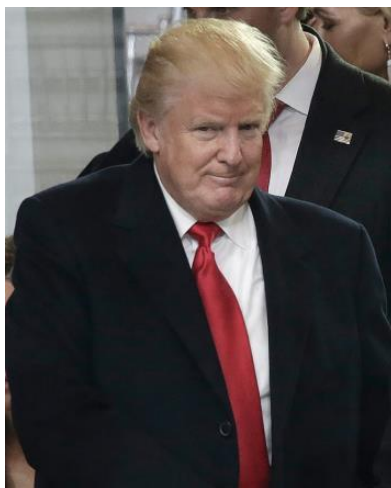
The same text in other languages?

- What are other languages saying today?
- What is it like to read the world in another language?



What's noteworthy if you speak Arabic?

- What did king Salman bin Abdulaziz discuss with US President Donald Trump?



- What is the Iranian response like?

A Gift for Extremists

- From his end, Iranian foreign affairs minister Mohammad Javad Zarif responded in a tweet that Trump's order 'will go down in history as a great gift for extremists and their defenders.'



- However he also said: "by contrast to the United States, our decision is not retroactive, whoever has a valid travel visa to Iran will be welcome."

Assignment – translation log

- We'll be talking more about translation this week
- Pick a short source text in your L2 and translate it into English (about 100 words)
 - Make a record of every sentence
 - How did you begin translating it?
 - What dilemmas did you run into?
 - Is this easy, average or difficult to translate?
 - Go back to your log and pick one of the hardest sentences
 - Describe in detail what makes it hard
 - Were there things you felt 'can't be translated'?

Translation log

Sent	Dif	Process	Issues
1	easy	headline	للمتطرفين - The extremists or extremists? foreign secretary title in Iran?
2	med	SVO sentence, start at subject	Name transliteration?
		Figure out Twitter construction	in a tweet/a tweet using twitter...
		'Go down in history part' readily translatable	Order/decree for marsūm - should I use Eng. Executive Order? Is that faithful?
3	med	Translate embedding verb	على عكس unlike/by contrast?
		Direct speech, unlike clause, then 'whoever clause'.	سيكون محل ترحيب will be welcome or welcomed? Will be in a place of greeting...
		Framework first, then figure out type of visa being mentioned	

Assignment – translation log

- Submitting:
 - Send me a single Word document or PDF with:
 - Source text
 - Your translation, with numbered sentences
 - Your sentence-by-sentence log (brief comments on each sentence)
 - Detailed analysis of a hard sentence (1-2 paragraphs)

Translation



- **translation**

noun 1. a written communication in a second language having the **same meaning** as the written communication in a first language

[WordNet]

Translation

- Munday's overview:
 - Why have there been so few 'fundamental or new' ideas about translation?

Evaluation criteria IoL (see also: ATA, ITI)

- Accuracy
 - Information
 - Comprehension
- Appropriateness
 - Register
 - Vocabulary
- Cohesion
 - Rhetorical devices
 - Pronominalization, topic and discourse markers
- Technical accuracy

CHARTERED INSTITUTE OF LINGUISTS
Supporting and developing language professionals worldwide

Verifying accuracy?

- What are some options to 'prove' accuracy
 - Reverse translatability
 - Selection of correct original
 - Truth value equivalence



Exercise

- Join classmates with an overlapping language
 - One should translate sentence 1 of the IoL test text
 - Another should translate sentence 2

Exercise

1. A large study on the future of work in the UK predicts the rise of the "mobile worker" moving - laptop and mobile in tow - between office, home, hotel, airport lounge or motorway service station as the needs of a job demand.
2. Today, more than 5 million people, almost a fifth of employees, already spend some time working at home or on the move, according to the report published by the Economic and Social Research Council and the Tomorrow Project, a charity studying future trends.

Exercise

- Now switch sentences:
 - Translate you partner's sentence back

Next

- For Wednesday, please prepare your reverse translation of your partner's translated sentence
 - Do not consult the English original!
- The translation log assignment is due Monday next week by e-mail