Google Translate
&
TectoMT

Martin Majliš

Plan

- What is Google Translate?

- What is TectoMT?

- Why combine?

- Results

- TODO

Google Translate

- Google service
- 34 languages
- AJAX API
- http://translate.google.com/

Google Translate API

- http://ajax.googleapis.com/ajax/services/language/translate?v=1.0&q=it's lunchtime!&langpair=en%7Ccs

- {"responseData": {"translatedText":"přestávka na oběd, je to!"}, "responseDetails": null, "responseStatus": 200}

TectoMT

- a highly modular extendable NLP software system

- composed of numerous (mostly previously existing) NLP tools integrated into a uniform infrastructure

- aimed at (not limited to) developing MT system

- http://ufal.mff.cuni.cz/tectomt/

Google Translate + TectoMT

Why combine?

Google Translate + TectoMT

Why combine? (cont.)

- Google
  - Great amount of text
  - Computing power - 450k servers (2006)
- TectoMT
  - Linguistic approach
  - Hypothesis: translation choices are conditioned rather by governing/dependent words than by linear predecessors/followers

Translation
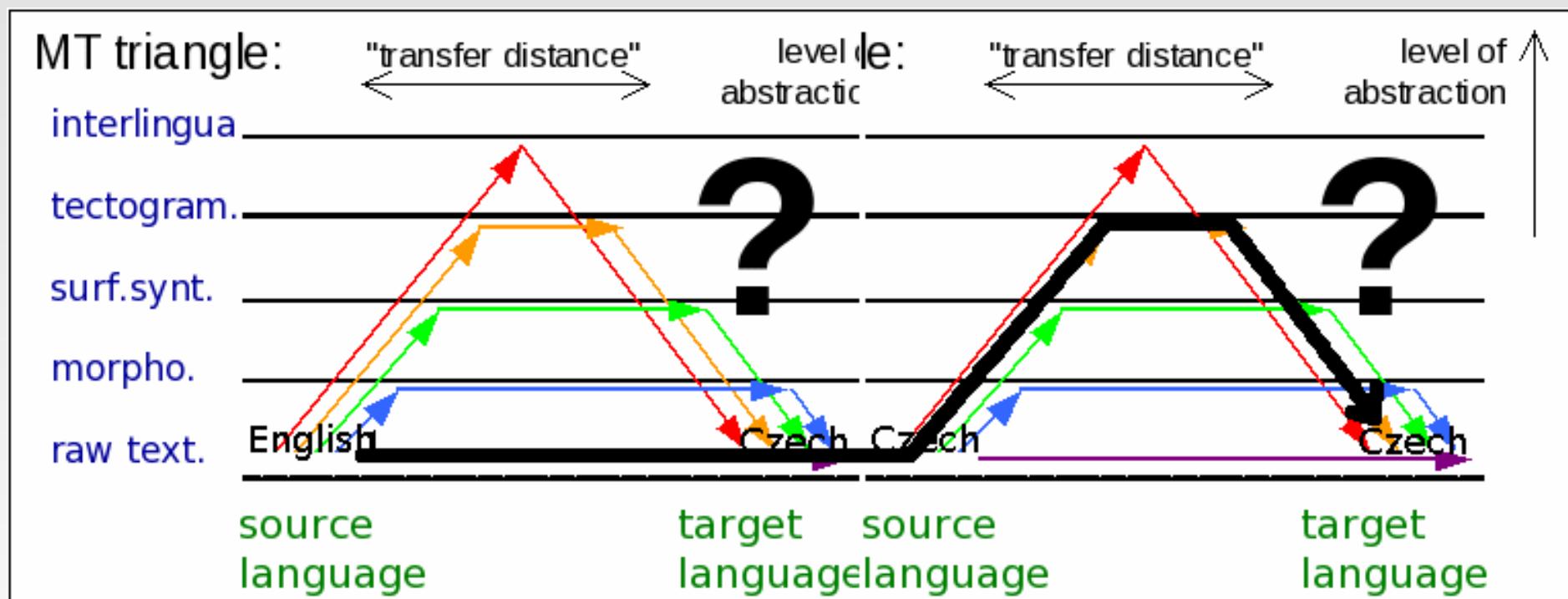
Examples

- This time the fall in stocks on Wall Street is responsible for the drop.

  - GT: Tato doba je pokles zásob na Wall Street je zodpovědný za pokles.

  - GT+TMT: Tato doba je pokles zásob na wall street, je zodpovědný za pokles.

Examples (cont.)

- ## Stocks fall in Asia

  - ### GT: Zásoby pád v Asii

  - ### GT+TMT: Zásoba pád v asie.

Examples (cont.)

- The stock exchange in Taiwan dropped by 3.6 percent according to the local index.

  - GT: Burze, na Tchaj-wanu klesla o 3,6 procent v závislosti na místní rejstřík.

  - Burze tchaj - wan klesl o 3, 6 procent na místní rejstřík.

Examples (cont.)

- Despite the fact that it is a part of China, Hong Kong determines its currency policy separately, that is, without being dependent on the Chinese Central Bank.

    - GT: Navzdory skutečnosti, že je součástí Číny, Hong Kong určuje její měny politice samostatně, tedy bez závislosti na čínské centrální bance.

    - GT+TMT: Navzdory skutečnosti, že je součástí číny, hong kong určuje její měny politice samostatně, tedy bez na závislosti čínské centrální bance.

Google Translate + TectoMT

Errors

- ## Name entity
  - Kevin Rudd => Kevin Rudd => kevin rudd

Numbers!
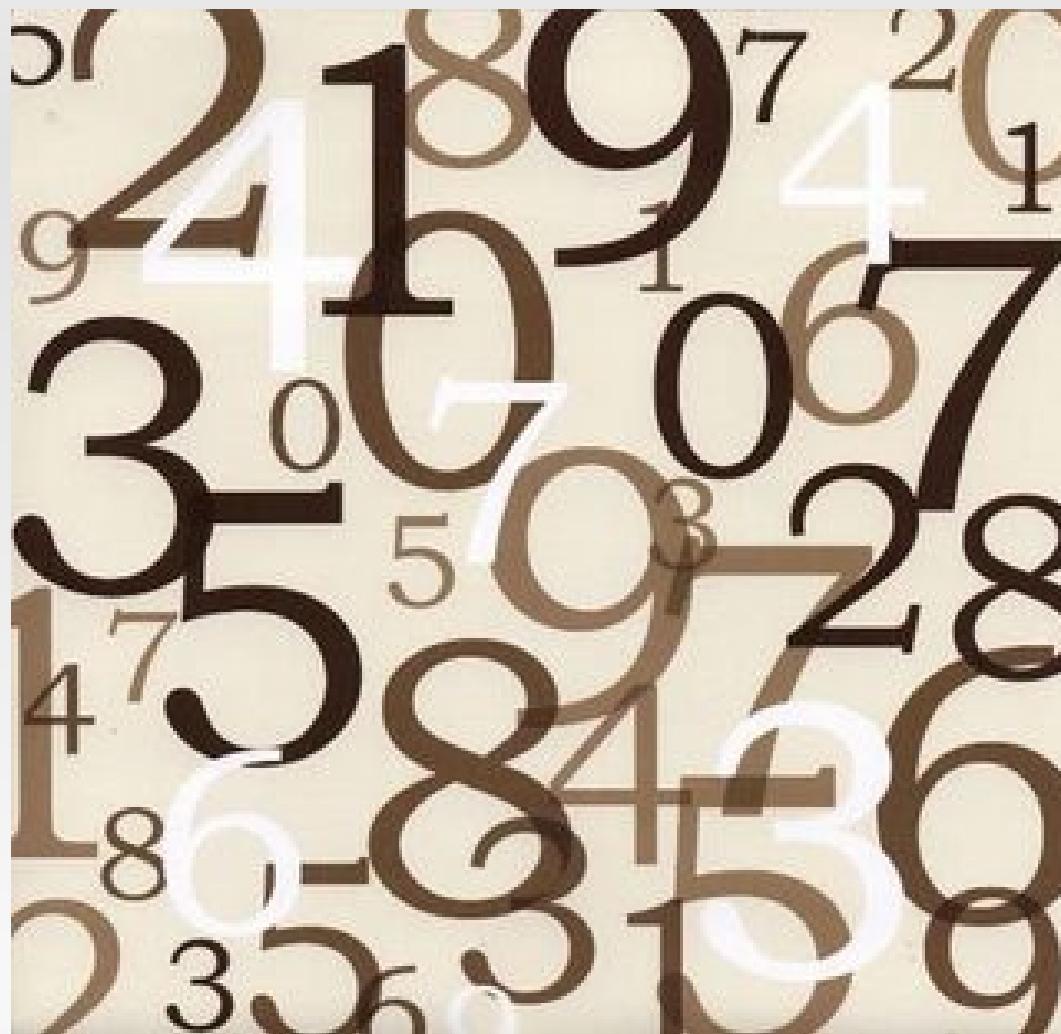
Results

	BLEU	NIST
Google Translate	5.2542	0.1543
GT + Parahrasing	4.4241	0.0953

TODO

- ## Somehow improve :)

  - Analyse english sentence

  - Translate subtrees

  - Merge subtrees

  - Analyze + Parahrase czech sentence

Questions?

Google Translate + TectoMT

Thank you!