

## EMNLP Tutorial 4

Philipp Koehn

This tutorial deals with statistical machine translation.

Given the following Centauri/Arcturan [Knight, 1997] parallel corpus:

- 1a. ok-voon ororok sprok .
- 1b. at-voon bichat dat .
- 2a. ok-drubel ok-voon anak plok sprok .
- 2b. at-drubel at-voon pippat rrat dat .
- 3a. erok sprok izok hihok ghrok .
- 3b. totat dat arrat vat hilat .
- 4a. ok-voon anak drok brok jok .
- 4b. at-voon krat pippat sat lat .
- 5a. wiwok farok izok stok .
- 5b. totat jjat quat cat .
- 6a. lalok sprok izok jok stok .
- 6b. wat dat krat quat cat .
- 7a. lalok farok ororok lalok sprok izok enemok .
- 7b. wat jjat bichat wat dat vat eneak .
- 8a. lalok brok anak plok nok .
- 8b. iat lat pippat rrat nmat .
- 9a. wiwok nok izok kantok ok-yurp .
- 9b. totat nmat quat olat at-yurp .
- 10a. lalok mok nok yorok ghrok klok .
- 10b. wat nmat gat mat bat hilat .
- 11a. lalok nok crrok hihok yorok zanzanok .
- 11b. wat nmat arrat mat zanzanat .
- 12a. lalok rarok nok izok hihok mok .
- 12b. wat nmat forat arrat vat gat .

Your assignment, translate this to Arcturan:  
farok crrok hihok yorok klok kantok ok-yurp .