

Evaluating WebLicht's annotations of learner English

Julia Müller & Christian Mair (University of Freiburg)



SPONSORED BY THE



Federal Ministry of Education and Research

(1) "The economy demands ***for** people who are very flexible."

The syntax tree shows the following structure:

- ROOT: demands (NNS) [Incorrectly tagged]
- NMOD: The (DT), economy (NN)
- PRD: demands (NNS) -> for (IN), people (NNS), who (WP), are (VBP), very (RB), flexible (JJ)
- AMOD: very (RB), flexible (JJ)
- SBAR: who (WP), are (VBP)
- PMOD: for (IN), people (NNS)

- 'demands' correctly identified as root
- but incorrectly assigned a noun tag

token	morphmajorword...
The	determiner
economy	noun
demands	verb ✓
for	adv_conj_pcl_pr... [Incorrectly tagged]
people	noun

- 'demands' correctly tagged as verb
- tagger struggles to decide on a tag for 'for'

token	pos
The	DT
economy	NN
demands	NNS ✗
for	IN
people	NNS

- 'demands' incorrectly tagged as a plural noun,
- 'for' as preposition or conjunction

(2) "They had the role *, to represent the Empire."

The left tree shows a correct parse of the sentence with a comma. The right tree shows an alternative parse structure.

- parser performs well despite the overuse of the comma

(3) "They ***conect** with bigots."

token	pos	lemma
They	PRP	they
conect	VBP	conect
with	IN	with
bigots	NNS	bigot
.	.	.

- tagger correctly identifies 'conect' as a verb despite the spelling mistake
- due to the misspelling, it can't provide the correct lemma