

Multi-Word Sketches Released

Vojtěch Kovář

Lexical Computing Ltd.
`vojtech.kovar@sketchengine.co.uk`

SKEW 3
March 21, 2012

Outline

- 1 Motivation
- 2 What we can do
- 3 How can we use it?
- 4 Conclusions

Motivation

- Word sketches are very useful
 - but for isolated words only
- So far no collocations for multiwords
 - New York
 - hot water
 - take place
 - get off

What we can do

■ Filtering collocations

- according to various criteria
- subcorpus, word forms, ...
- also particular collocations

■ Collocates for New York

- are actually collocates for York-n filtered by “New”

■ Collocates for get off

- collocates for get-v filtered by “off”

■ → Headword and “filterwords”

How can we get it?

- Two options
- Successive filtering of the collocates
 - → start with “York”, then click on a link at “New”
 - “Show links to multiword sketches” option
 - “MW links” switch on the word sketch page
- Entering a multiword directly as lemma
 - system will try to guess the headword
 - according to frequency and PoS
 - description is given with the result
- Demo

Conclusions

- Word Sketch for multiword expressions
 - straightforward usage
- Limitations
 - often not enough data for multiwords >2
 - sketch grammar may need tweaking
 - only filtering collocates of one headword
 - \rightarrow we cannot do “very young man”
- Future directions
 - include collocations of all components of the multiword
 - your feedback