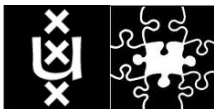


How Synchronous is Adjunction in Translation Data?

Sophie Arnout Khalil Sima'an



Institute for Logic, Language and Computation
University of Amsterdam

SSST8, Oct. 25, 2014

Can adjunction explain translation data?

Adjunction

Expands syntactically complete phrases with optional material

here: *adjuncts* ~ *modifiers*

overall , *on the basis of this financial perspective* , *new financial commitments that can be entered into next year* could rise by 2.9% *compared with this year* .

Linguistically motivated description of synchronous data for modelling / evaluation

Cross-linguistic validity?

- one **Canadian** job in three || **au Canada**, un emploi sur trois
- this was a **European** disaster || cette catastrophe était *Européenne*

To what degree is adjunction synchronous in translation data?

Empirical measures of adjunct equivalence

Robust and interpretable measures

- adjunct translation-equivalence types: equivalent, divergent, uncertain
- competitive adjunct linking strategy based on word alignment coverage

Measurements on empirical data

- French-English data
- effect of labelling, word alignments, data set, sentence length
- adjuncts identified through heuristics on dependency labels

Results

Measures

Effect on adjunct equivalence of

- labelling, word alignments, data, length
 - data: 10% more equivalent adjuncts in EP, News than Canadian Hansards
- adjunct equivalence an indicator for more literal translations?

Conclusion

How synchronous is adjunction in Fr-En translation data?

- 40 – 50% adjunct equivalence with heuristics on automatic labels
 - ▶ need better identification means
- 70% with gold labels
 - ▶ factors behind adjunct divergence