Integrating compound recognition in constituency parsing with an hybrid approach

Matthieu Constant

Université Paris-Est, LIGM, CNRS

PARSEME Meeting
17 September 2013

Integration of realistic MWE recognition in statistical parsing

Motivations

- Non-compositionality
- Improve statistical parsing accuracy [Nivre and Nilsson 2004; Cafferkey 07; etc.]

Limitations of the talk

- MWE = compounds (non compositional token sequences)
- Constituency parsing

Integration of realistic MWE recognition in statistical parsing (Cont.)

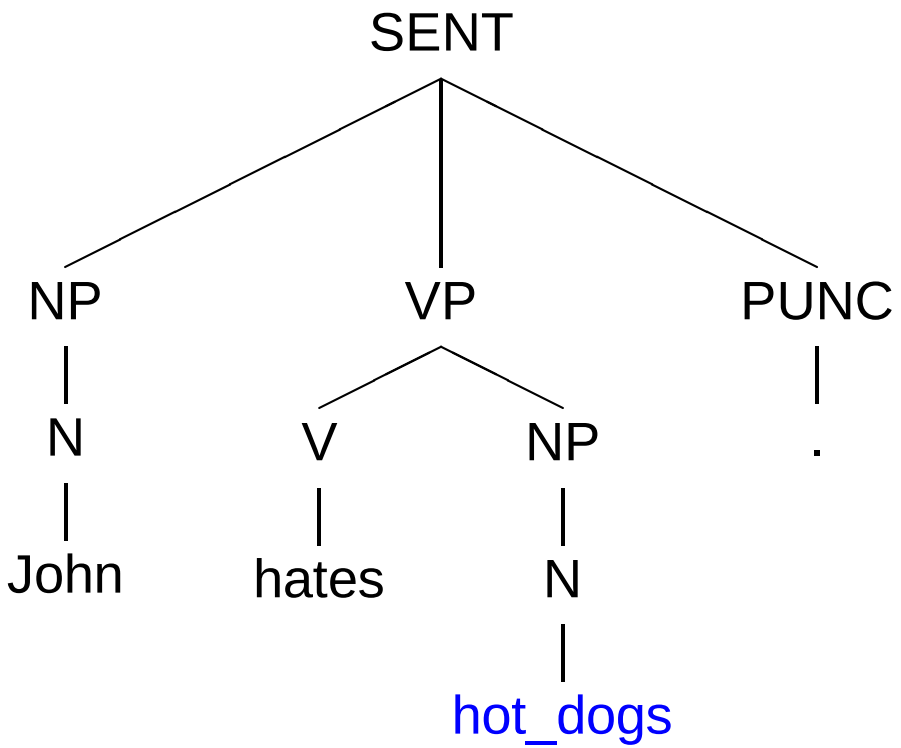
Content of the talk

- Brief overview of existing approaches
- Our experiments (in collaboration with A. Sigogne, J. Le Roux and P. Watrin)
- Perspectives

Experiments on French

- French Treebank (FTB) with compound annotations (15% of tokens are part of a compound)
- MWE resources available
- Some reference works on MWE+Parsing (Arun and Keller 2005, Green et al. 2011,2013)

Example



Compound recognition

Traditional Cues

- Strong lexical association $\rightarrow$ lexical resources, statistical criteria
- Syntactic information $\rightarrow$ local patterns, parsing

Supervised compound recognition

- Segmentation task with discriminative models: e.g. CRF
- Combination of different resources: annotated corpus, lexica, POS taggers, NE Recognizers, etc. [Vincze et al. 2011; Constant et al. 2012]
- Joint compound recognition and linguistic analysis [Green et al. 2011; Constant and Sigogne 2011]

Constituency Parsing and Compound recognition

State-of-the-art parsers

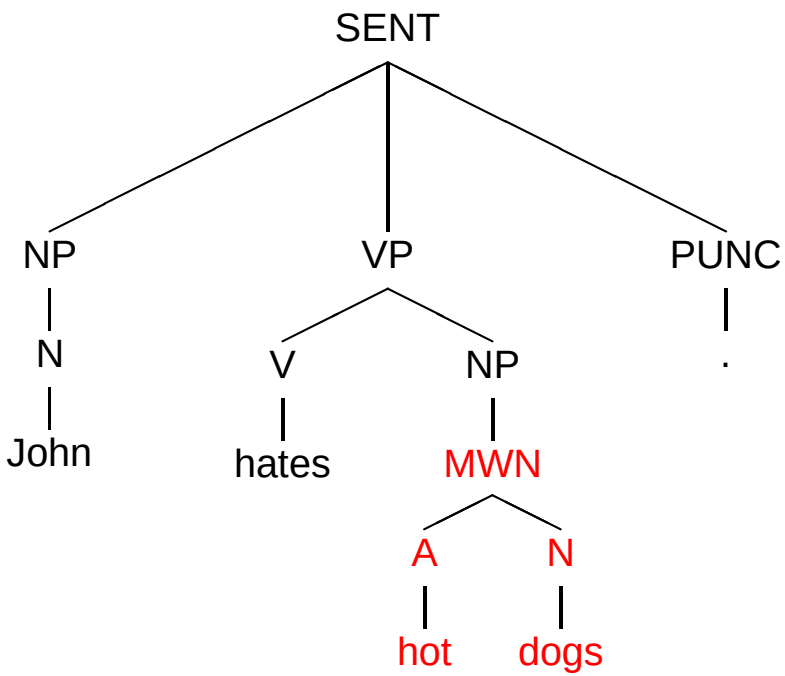
- Nonlexicalized strategies: Probabilistic Context-Free Grammars with Latent Annotations (PCFG-LA) [Matsuzaki 2005, Petrov 2006]
- Reranking with discriminative models [Charniak et Johnson 2005]
- State-of-the-art for French [Seddah et al. 2009; Green et al. 2011; Le Roux et al. 2011]

Where to integrate compound recognition?

- Before parsing
- During parsing (joint approach)
- After parsing
- Combinations?

Joint compound recognition and parsing (baseline)

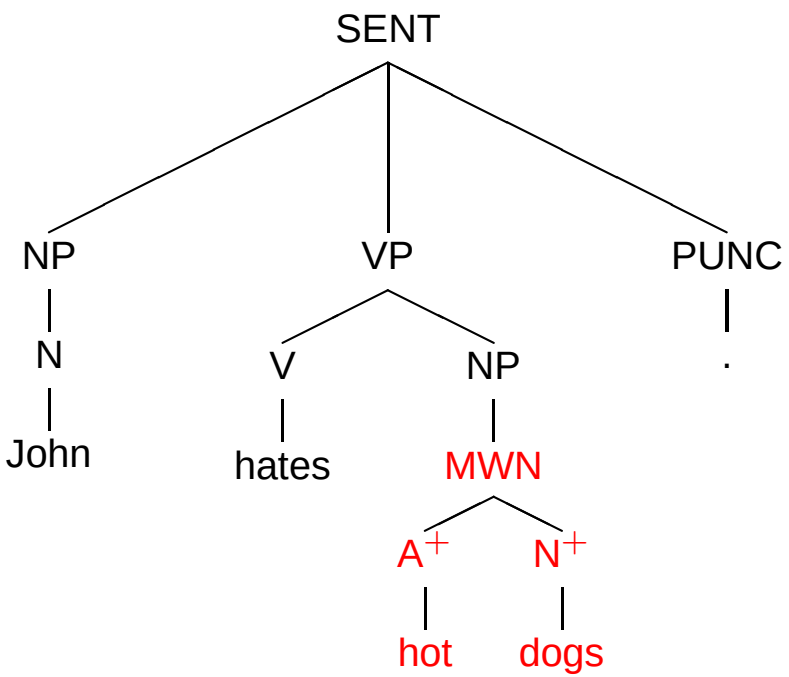
- Compound recognition integrated in the grammar [Arun and Keller 2005, Green et al. 2011]
- Compounds are annotated with a specific nonterminal node



Joint compound recognition and parsing (cont.)

Our experiments

- Test different specific POS tagsets for compound items
- Small but significant improvement in MWE recognition accuracy
- No differences in general parsing accuracy as compared with baseline



Prerecognition

Pregrouping strategy [Nivre and Nilsson 2004; Arun and Keller 2005]

- Compound prerecognition (John hates hot dogs)
- Grouping compound as a single token (hot dogs $\rightarrow$ hot_dogs)
- Most of experiments with gold compound annotation

Realistic Experiments

- Improving parsing accuracy
  - Shallow parsing [Korkontzelos and Manandhar 2010]
  - "Deep" constituency parsing [Cafferkey 2007]
- Example from [Korkontzelos and Manandhar 2010]
  - Without MWEs: He threw (the fire) wheel up
  - With MWEs: He threw (the fire_wheel) up

Prerecognition (Cont.)

Our experiments

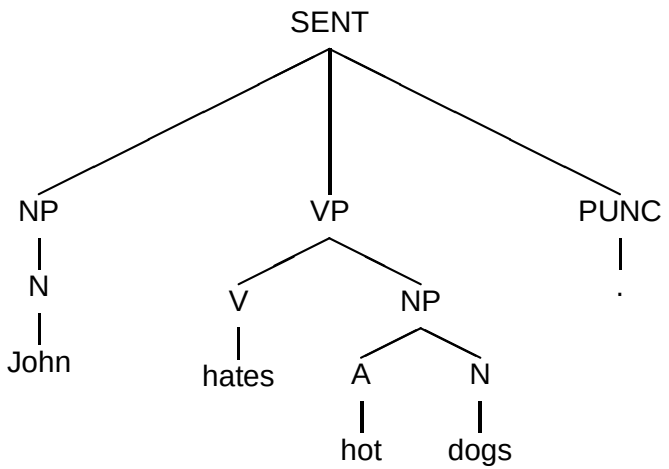
- Use of a CRF-based prerecognizer
- CRF features: POS and word ngrams, lexicon-based, etc.
- Grammar training on treebank annotated with gold compounds
- Evaluation: undoing compounds as for the baseline

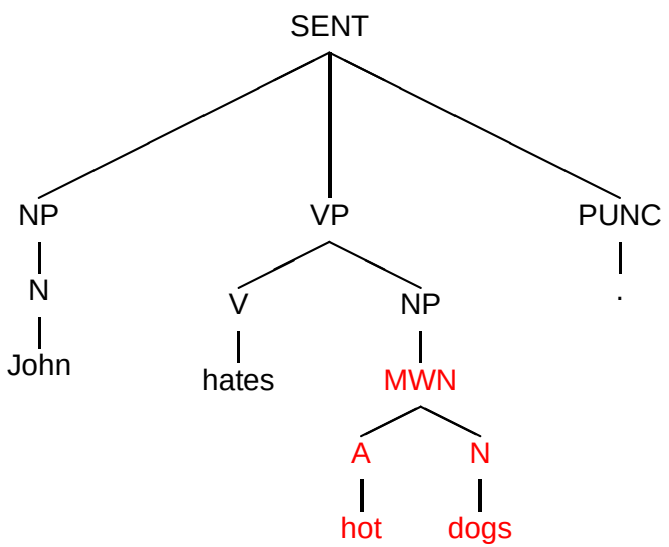
Conclusions

- CRF-based recognizer = state-of-the-art for French as compared with [Green et al. 2011]
- Precognition may greatly improve parsing accuracy (if good tuning)
- Preliminary experiments showed that false compounds may cause side effects on parsing

Reranking strategy

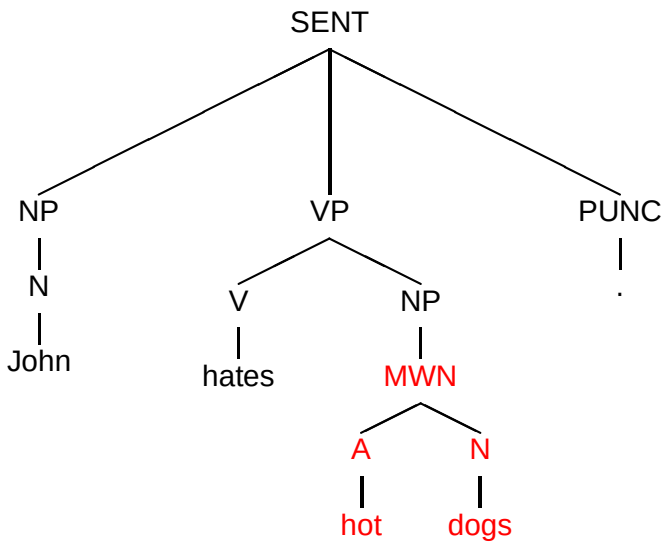
n-best parses

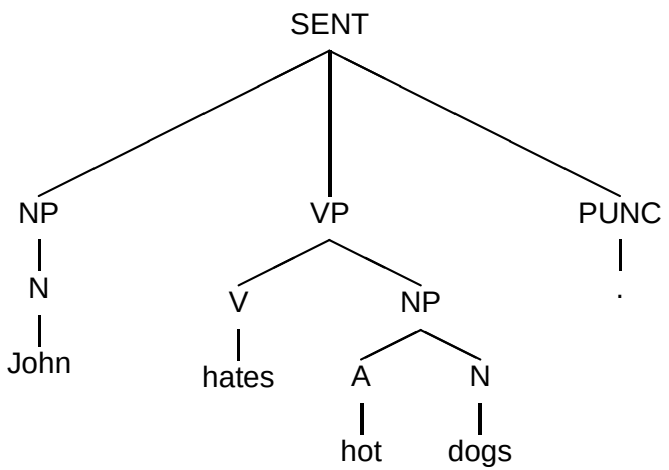
1. 

2. 

Reranking strategy (Cont.)

Result

1. 

2. 

Reranking strategy (Cont.)

Our experiments

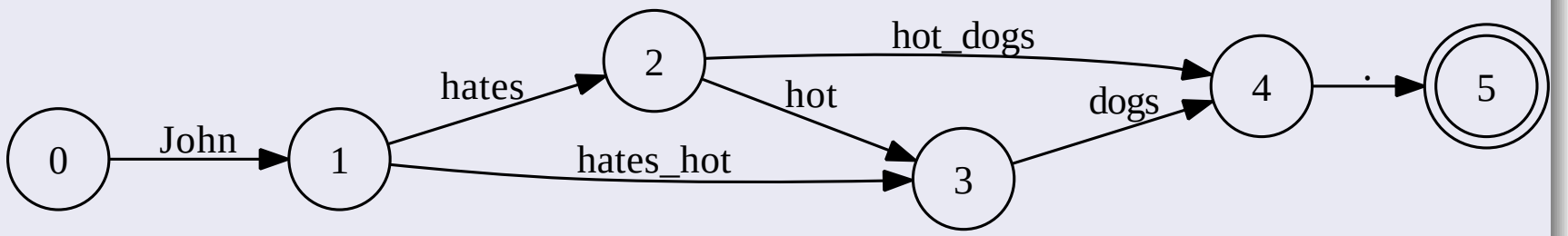
- Use of n-best joint MWE+parser with reranker
- Use of MWE-dedicated features (e.g. based on lexicon).

Conclusions

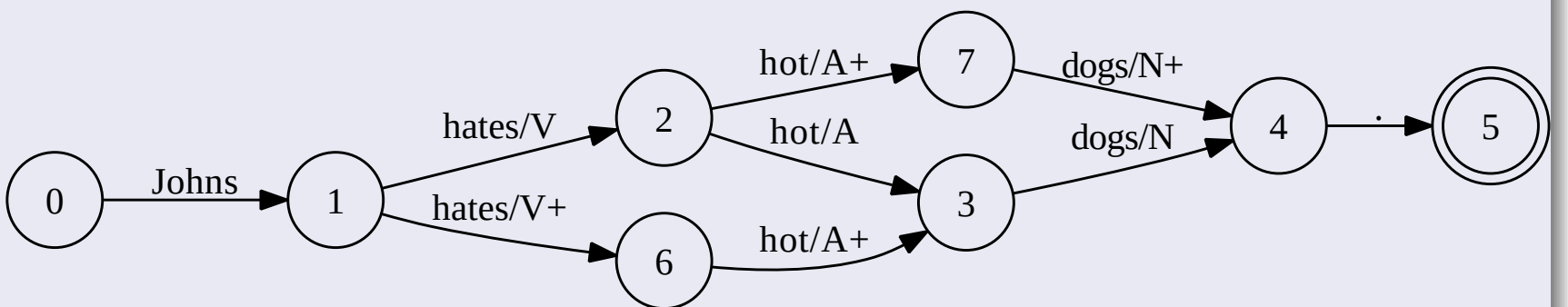
- Small but significant improvement in all metrics
- BUT... the MWE-dedicated features are useless when added to standard non local features (Charniak and Johnson 2005)

Ambiguous prerecognition strategy

m best compound segmentation



m best compound-aware POS tagging



Ambiguous prerecognition strategy (cont.)

Our experiments

- The compound prerecognizer outputs its *m* best analyses
- The parser is in charge of selecting the best analysis as well as the best parse.
- related with [Goldberg and Tsarfaty 2008]

Conclusions

- Results not as good as expected
- Oracle results are promising

Perspectives

Improving Combinations

- Combining everything
- Better selection of the *n*-best parses
- Sequence of rerankers

Extending to dependency parsing

- Most of previous works on golden MWE segmentation [Nivre and Nilsson 2004; Eryigit et al., 2011]
- Preliminary experiments on SPMRL shared task on parsing task: first rank on French MWE+Parsing track (with M. Candito and D. Seddah)

Perspectives (Cont.)

Extensions to...

- Other MWEs $\rightarrow$ annotated corpus ?
- Other PARSEME languages

Better evaluation

- Current evaluation: binary (0 or 1)
- Why not weigh compounds with respect to their non compositionality degree?
- How? Linguistic criteria, statistical criteria, cognitive criteria?

THANKS!

Questions ?