

# Bi-LSTM TAG Supertagging for French and German

Tatiana Bladier

21 February 2018, Düsseldorf

TreeGraSP Workshop  
WP3 Semantic Parsing

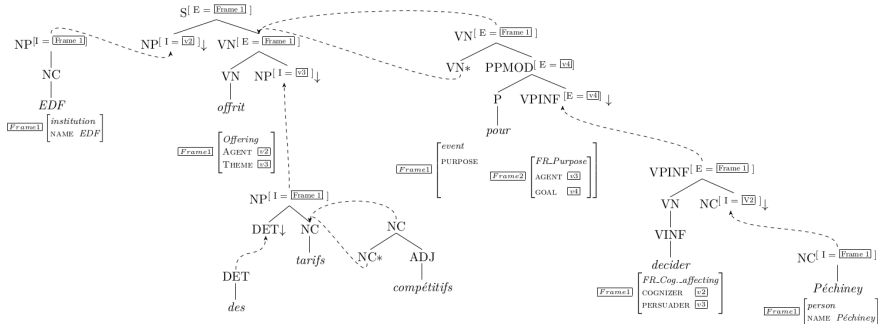
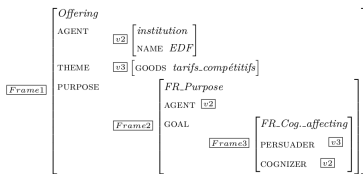
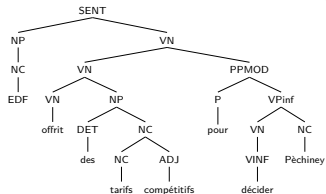


# Agenda

- ▶ TAG Semantic Parsing: General Idea
- ▶ From French TreeBank (FTB) to TAG
- ▶ Bi-LSTM TAG Supertagging
- ▶ Working Package 3: Next steps

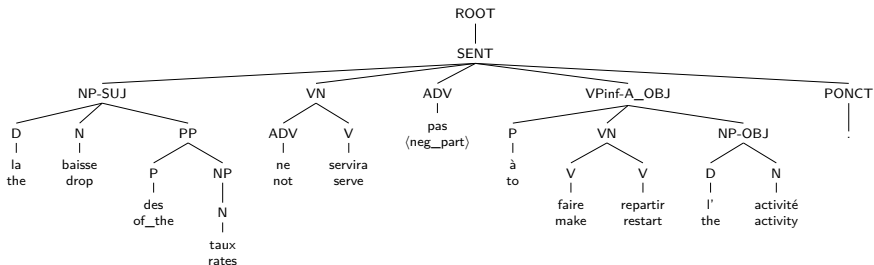
# TAG Semantic Parsing: General Idea

adapted from Kallmeyer & Osswald (2014)



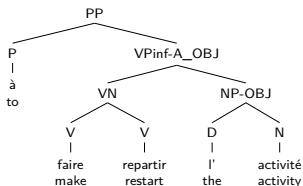
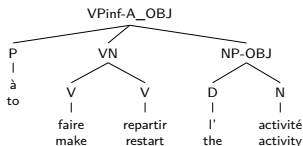
# Resources: French TreeBank and French FrameNet

- ▶ FTB: Latest version version 1.0 from 2016
- ▶ FTB: 21.550 sentences (approximately 664.500 tokens) from the newspaper *Le Monde*, years 1990 to 1993 (Abeillé 2003)
- ▶ FTB: trees have a relatively flat structure compared to TiGer and Penn Treebank
- ▶ French FrameNet: approximately 20.000 annotated sentences, based on French TreeBank and Sequoia Treebank (Djemaa et al. 2016)

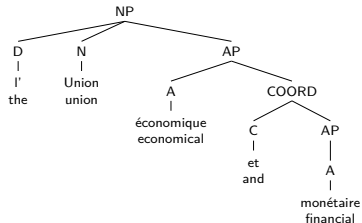
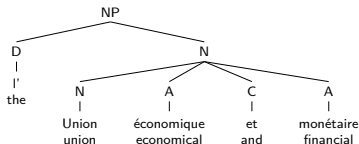


# FTB Preprocessing: Complement Raising and Undoing compounds

- ▶ FTB-Preprocessing steps as described in Candito et al. (2010), powered by the discodop parser (van Cranenburgh et al. 2016)
- ▶ Complement raising

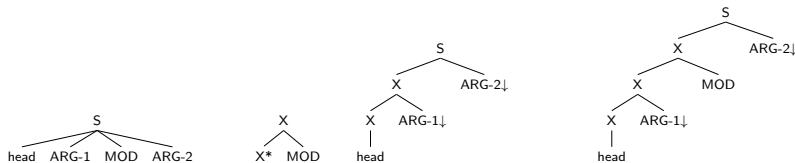


- ▶ Undoing some compounds with regular patterns (3.072 from 6.125 (Candito 2010))



# Sister-Adjunction

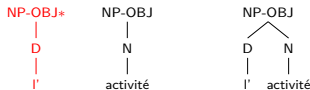
- ▶ Trees in FTB are relatively flat and allow any ordering of arguments and modifiers



- ▶ Only a complete binarization of FTB trees would solve such problems
- ▶ Such binarization should be linguistically motivated  
⇒ possible inconsistencies and incorrect linguistic structures
- ▶ Might not be necessary if we use supertags
- ▶ In order to preserve flat tree structures we decided to use the sister-adjunction operation
- ▶ We do not apply the traditional TAG adjunction

# Sister adjunction

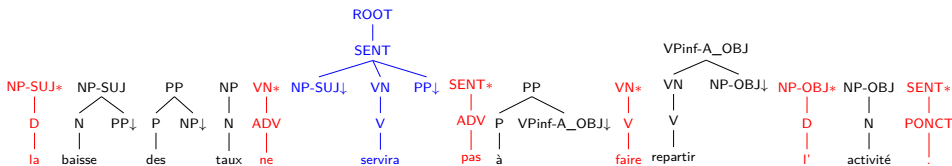
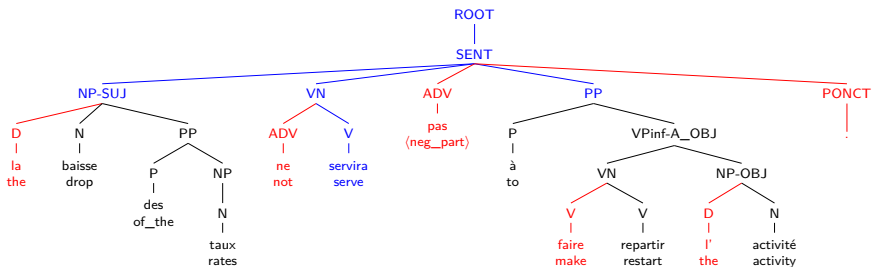
- ▶ Notion of sister-adjunction (adapted from Kaeshammer (2012))
- ▶ Sister-adjointing trees have the root level
- ▶ Left sister-adjunction (relative to the position of the head)



- ▶ Right sister-adjunction (relative to the position of the head)



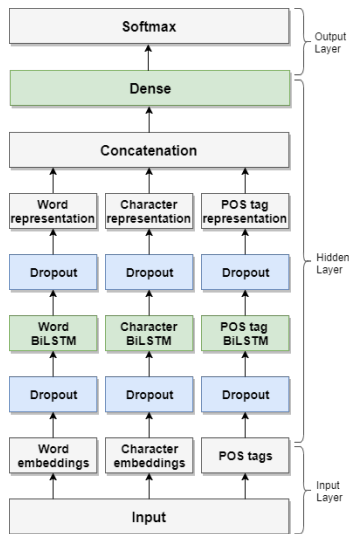
# Extraction of Supertags





# Bi-LSTM TAG Supertagging: Architecture

- ▶ Supertagging  $\Rightarrow$  a task of assigning a supertag to each word in a sentence (Chen 2010)
- ▶ Supertagging is “almost parsing” (Bangalore and Joshi, 1999), since it performs much of a parsing disambiguation
- ▶ Supertagging can increase the TAG parsing efficiency by reducing syntactic ambiguity and sentence complexity (Sarkar 2007)
- ▶ Bi-LSTM Supertagging Architecture (adapted from Samih (2017))



# Supertagging Experiments: Challenges

- ▶ Big number of supertags (compared to for example CCG) (Kasai, 2017)
- ▶ Almost a half of the supertags appear only once
- ▶ Longer sentences (and smaller number of sentences) in French corpus
- ▶ No comparable results for French TAG supertagging

Parameters	French	German Kaeshammer (2012)
Supertags	4488	3701
Supertags appear once	1722	1706
POS tags	39	57
Words	675316	885468
Sentences	21550	50000
Words per sentence (average)	31.3	17.7

# Supertagging Experiments: First Results

- First results for German (French is on the way):

System	Accuracy
Bäcker & Harbusch (2002) (German, HMM-based)	0.783 (test set)
Westburg, A. (2016) (German, TiGer)	0.8292 (dev set)
Kasai et al. (2017) (English, Penn Treebank, Bi-LSTM)	<b>0.8932</b> (test set)
Word model + character model (Bi-LSTM)	0.8693 (dev set)
Word model + character model + POS model (Bi-LSTM)	<b>0.8784</b> (dev set)
Word model + character model + POS model (GRU)	0.8774 (dev set)

# Experiment Results: Some Errors

- ▶ Some proper names are not correctly recognized
  - ★ *Mitbegründer Jobs*:  
supertag (NP (NN Jobs)), gold supertag (NP (NE Jobs))
- ▶ Predicted elementary tree instead of an auxiliary tree (and vice versa)
  - ★ *dessen Softwarehouses*:  
supertag (NP (DP ) (NN Softwarehouses))  
gold supertag (NP\* (NN Softwarehouses))
- ▶ Erroneously recognized attributes of the supertag
  - ★ *in die Gruppe zurückgeholt*:  
supertag (S (NP )(VP (NP )(VVPP zurückgeholt)))  
gold supertag (S (NP )(VP (VVPP zurückgeholt)))
- ▶ Constructions with conjunctions
  - ★ *verkaufen und verbuchen*:  
supertag (S (NP )(VP (NP )(VVINF verbuchen)))  
gold supertag (VP (NP )(VVINF verbuchen))

# Working Package 3: Next steps

- ▶ Experiments with French supertags
- ▶ Parsing with extracted supertags (Sarkar 2007)
- ▶ Supertagging Combined with Semantic Labeling (Chen 2010)

# References (I/II)

Abeillé, A., L. Clément, and F. Toussnel (2003). "Building a treebank for French", in A. Abeillé (ed) Treebanks, Kluwer, Dordrecht. (p.165-187)

Bäcker, J., & Harbusch, K. (2002). Hidden Markov model-based supertagging in a user-initiative dialogue system. In Proceedings of the Sixth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+ 6) (pp. 269-278).

Bangalore, S., & Joshi, A. K. (1999). Supertagging: An approach to almost parsing. Computational linguistics, 25(2), 237-265.

Candito, M., Crabbé, B., & Denis, P. (2010). Statistical French dependency parsing: treebank conversion and first results. In Seventh International Conference on Language Resources and Evaluation-LREC 2010 (pp. 1840-1847). European Language Resources Association (ELRA).

Chen, J. (2010). Semantic labeling and parsing via tree-adjoining grammars. Supertagging: Using Complex Lexical Descriptions in Natural Language Processing, 431-448.

Djemaa, M., Candito, M., Muller, P., & Vieu, L. (2016). Corpus annotation within the french framenet: a domain-by-domain methodology. In Tenth International Conference on Language Resources and Evaluation (LREC 2016).

Kaeshammer, M. (2012). A German treebank and lexicon for tree-adjoining grammars (Master's thesis, Universität des Saarlandes, Saarbrücken, Germany).

# References (II/II)

Kallmeyer, L., & Osswald, R. (2014). Syntax-driven semantic frame composition in Lexicalized Tree Adjoining Grammars. *Journal of Language Modelling*, 1(2), 267-330.

Kasai, J., Frank, B., McCoy, T., Rambow, O., & Nasr, A. (2017). Tag parsing with neural networks and vector representations of supertags. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 1712-1722).

Samih, Y. (2017) *Dialectal Arabic processing Using Deep Learning* (Doctoral thesis, Heinrich Heine Universität Düsseldorf, Düsseldorf, Germany).

Sarkar, A. (2007). Combining supertagging and lexicalized tree-adjoining grammar parsing. *Complexity of Lexical Descriptions and its Relevance to Natural Language Processing: A Supertagging Approach*, 113.

van Cranenburgh, A., Scha, R., & Bod, R. (2016). Data-oriented parsing with discontinuous constituents and function tags. *Journal of Language Modelling*, 4(1), 57-111.

Westburg, A. (2016). *Supertagging for German - an implementation based on Tree Adjoining Grammars* (Master's thesis, Heinrich Heine Universität Düsseldorf, Düsseldorf, Germany).