From deep syntax to semantic frames: feedback from the French FrameNet project

Marie Candito,
Paris Diderot University

joint work with
Marianne Djemaa, Philippe Muller, Laure Vieu,
G. de Chalendar, B. Sagot, P. Amsili (for the French FrameNet)
C. Ribeyre, D. Seddah, G. Perrier, B. Guillaume (Deep syntax project)
Olivier Michalon, Alexis Nasr (Semantic parsing)

Institut für Sprache und Information, Düsseldorf
December, 21st 2017

Plan

1. French FrameNet
    - Motivation
    - Methodology
    - Evaluation and stats
    - Feedback: What was difficult?
2. "Deep" syntax: reducing syntactic variation
3. Deep syntax for FrameNet parsing

Motivation

Treebanking / parsing background at Alpage / LLF

No large scale French data for semantic analysis

⟶ let's build a resource:

- relying on corpus evidence
  - respecting "natural" distribution of phenomena
  - for quantitative analysis and data-driven methods
- including generalizations
  - over lexical items (to limit lexical variation)
  - over predicate-to-argument relations (semantic roles)
- focusing on syntactico-semantic linking patterns
- capitalizing on previous work for other languages

⟶ let's use FrameNet (Baker et al. 1998; Fillmore 2007)

FrameNet

Large scale "Instanciation" of the "Frame Semantics" (Fillmore, 82) for English
developed at Berkeley (Baker et al. 98, Fillmore et al. 03 ...)

- frames = schematic representations of speakers' knowledge of the situations underlying the meanings of LUs (Fillmore, 2007)
- a frame
  - can be evoked by lexical units ($\longrightarrow$ triggers)
    - e.g.: Commitment frame evoked by *promise.v*, *promise.n*, *oath.n* ...
  - whose semantic valency is describing using frame-specific semantic role names
    - Commitment; Speaker, Addressee, Message

FrameNet

- Frame annotation = occurrence of a trigger evoking a frame, plus annotation of
  - role fillers: which portions of texts fill which semantic roles
  - syntactic information on role fillers

John ***PROMISED*** *Bill that the rights would be respected*

$\longrightarrow$ allowing to extract syntactico-semantic valency patterns

= linking patterns

FrameNet: characteristics

Structure: frames et relations entre frames

- $\approx$ 1200 frames
- linked by frame-to-frame relations (inheritance, perspective...)
- each frame-to-frame relation instance specifies role-to-role mappings
    - Complaining inherits Statement
    - Speaker $\leftrightarrow$ Complainer, Message $\leftrightarrow$ Complaint, Addressee $\leftrightarrow$ Addressee, ...

English lexicon

- $\approx$ 13600 senses : lemma+pos / frame pairs

FrameNet: characteristics

Annotations

- lexicographic : examples from BNC, chosen in order to
  - capture variety of syntactic realization of semantic roles
  - and derive linking generalizations
  - 174000 annotated frame instances
- full-text : complete annotation of running text
  - 28000 annotated frame instances only
  - but much better training data (Das et al., 10))
  - because of natural distribution of senses and linking patterns

FrameNet's key characteristics: Variable granularity of roles

- Well-known difficulty of defining a fixed and limited set of roles
- FrameNet's answer = Frame-specific semantic roles

  - $\oplus$ But coarser granularity can be derived from role-to-role mappings in frame-to-frame relation instances

  - $\ominus$ though in FrameNet's present form
    - lack of systematicity in relation instances (Ovchnnikova et al. 10; Osswald et VanValin, 14)
    - some relations are fuzzy ("Using")

FrameNet's key characteristics: "semantics first" philosophy

Empirically grounded study of syntax/semantic interface is one of FrameNet's core objectives

But FrameNet's documentation (Rupenhoffer et al. 05) uses primarily semantic criteria for defining

- the perimeter of a frame (which generalizes over several lexemes)
- the set of roles of a frame (≠Levin's classes, VerbNet...)
  - e.g. "Core" roles are those "necessary to the central meaning of the frame" (Fillmore, 2007)

FrameNet's key characteristics: "semantics first" philosophy

- $\ominus$ lack of formal (syntactic) criteria to define the perimeter of frames
- $\ominus$ semantic granularity of the frames is partially arbitrary
  - certain frames contain synonyms only,
  - others allow more semantic variation

FrameNet's key characteristics: "semantics first" philosophy

- $\ominus$ lack of formal (syntactic) criteria to define the perimeter of frames
- $\ominus$ semantic granularity of the frames is partially arbitrary
  - certain frames contain synonyms only,
  - others allow more semantic variation

- $\oplus$ semantic treatment of syntax/semantic mismatches ($\neq$ VerbNet)
- $\oplus$ a frame can generalize over LUs with different POS
  - e.g. the Causation frame can be evoked by
    - *because.c*, *for.c*,
    - *cause.v*, *result.v*, ...
    - *cause.n*, *consequence.n*,
    - *because of.prep*, ...

FrameNet's key characteristics: "semantics first" philosophy

- ⊕ roles less dependent on syntax, cf. converse verbs example
  - 2 frames Commerce_buy Commerce_sell
  - linked by "perspective" relation, and mapped roles
    - sell.v, sale.n, auction.n,...
    - buy.v, purchase.v, purchase.n ...
- FrameNet / Propbank(Nombank) / VerbNet roles:

I *Buyer(Arg0/Agent)* **bought** a Sovereign acoustic guitar *Goods(Arg1/Theme)* for 20 pounds *Money(Arg3/Asset)* from an absolute prat *Seller(Arg2/Recipient)*

China canceled its *Seller(Arg0/Agent)* **sale** of a uranium conversion facility *Goods(Arg1/Theme)* to Iran *Buyer(Arg2/Recipient)*

ASFALDA French FrameNet project

ANR funding, Oct 2012 to March 2016
Partners: Alpage, CEA-List, LIF, LLF, MELODI (IRIT)

The objective was to build

- a French FrameNet
- and a FrameNet semantic parser

Methodology: coverage issues

Full coverage (of course) not reachable...

Which frames to work on? Which lemmas? Which corpus to annotate?

Methodology: coverage issues

Full coverage (of course) not reachable...

Which frames to work on? Which lemmas? Which corpus to annotate?

Preliminary full-text annotation experiments, using English frames
Feedback: extremely difficult to understand the perimeter of a frame

$\longrightarrow$ Decision to focus on a few notional domains, but fully described

Methodology: coverage issues

Full coverage (of course) not reachable...

Which frames to work on? Which lemmas? Which corpus to annotate?

Preliminary full-text annotation experiments, using English frames
Feedback: extremely difficult to understand the perimeter of a frame

$\longrightarrow$ Decision to focus on a few notional domains, but fully described

- Commercial transactions
  - very much discussed in FrameNet litterature
- Cognitive stances
  - Position of a Cognizer concerning the truth value of a proposition
  - stative: who knows/believes what
  - inchoative (discover.v) / causative (convince.v)
- Verbal communication
- Causality

Methodology: overview

- Frame structure and lexicon built in parallel
  - 10 people participated
  - initially for 7 domains, but only 4 later annotated on corpus
  - addition of syntactic criteria to define frames
- Corpus annotation
  - Pilot annotation to build the annotation guidelines
  - 4 domain experts + 6 annotators
  - 2 independent annotation + adjudication
  - $\longrightarrow$ again some further modifications of frames and lexicon!
- Extraction of data-driven lexicon

Frames for French: using syntactic criteria to define frames and roles

- FrameNet's definition of roles is a bit fuzzy
- Typology of roles (Core / Peripheral / Extra-thematic) is documented too briefly

Fillmore 2007: precise account of semantic / syntactic mismatches

but lack for precise criteria of what is a syntactic valent and a semantic valent

Frames for French: using syntactic criteria to define frames and roles

$\longrightarrow$ after many attempts, we have simply come up with (Djemaa, 16):

Semantic valents of a LU are necessarily

- subcategorized (syntactic valents in canonical use of the LU)
  - (criteria for French from Bonami, 99)
- semantically selected by the LU
  - $\neq$ subjects of raising verbs
- two types of roles only :
  - Core role: semantically mandatory (criteria from Bonami, 99)
  - Peripheral role: subcategorized but not semantically mandatory

Frames for French: using syntactic criteria to define frames and roles

$\longrightarrow$ after many attempts, we have simply come up with (Djemaa, 16):

Semantic valents of a LU are necessarily

- subcategorized (syntactic valents in canonical use of the LU)
  - (criteria for French from Bonami, 99)
- semantically selected by the LU
  - $\neq$ subjects of raising verbs
- two types of roles only :
  - Core role: semantically mandatory (criteria from Bonami, 99)
  - Peripheral role: subcategorized but not semantically mandatory
- for LUs canonically used as modifiers, the syntactic governor is a semantic valent too
  - A due to B
  - A because B
  - A responsible for B

Frames for French: using syntactic criteria to define frames and roles

Criteria to delimitate frames:

1. All LUs in the same frame have the same set of Core roles
   - To avoid frame multiplication, more precisely roles organized into
     - Core sets (at least one is semantically mandatory)
     - or mutually exclusive sets (at most one is semantically mandatory)

   Example: frame FR_Attributing_cause

- Cognizer ATTRIBUE Effect to Cause_eventuality / Responsible_entity
- Cognizer EXPLIQUE Effect by Cause_eventuality / *Responsible_entity

NB: in corpus, non-local realizations of roles are annotated too

2. Decision to split frames in two must not create artificial polysemy

Frames for French

- 105 frames for French built from English frames (FN 1.5)
  - 47 non modified English frames
  - 38 adapted to French / new frame delimitation criteria
  - 20 new frames created in order to complete a domain

Domain	Nb frames	NB LUs (with at least one annotation)
Commercial Transactions	19	99
Cognitive stances	44	442
Verbal Communication	47	411
Causality	11	285
All	105	873

- Note: some frames belong to several domains (e.g. FR_Attempt_suasion)

Lexicon

Previous coverage strategies:

- Exhaustive frame (senses) coverage for a given lemma
  - SALSA German FrameNet (Burchardt et al., 2006)
  - produces WSD-oriented data
- versus Exhaustive lexical coverage for a given frame
  - Berkeley FrameNet
  - information extraction oriented
  - $\longrightarrow$ chosen in French FN, for the 4 selected domains

Lexicon

Lexical coverage:

- define full French lexicon for each covered frame
  - $\longrightarrow$ e.g. 19 triggers for Commerce_buy
  - $\longrightarrow$ e.g. 110 triggers for Causation

  - manual validation of French lexicons obtained by automatic projection from English (Padò, 2007; Mouton et al., 2010)
  - extension using monolingual French lexicons (tables LADL, LVF, LexConn (Roze et al., 12), Casoar (Benamara et al., 11), French-TimeBank (Bittar, 10), le Dictionnaire Electronique des Synonymes)
  - use of kwics on target corpora

Corpus annotation

Possible coverage strategies:

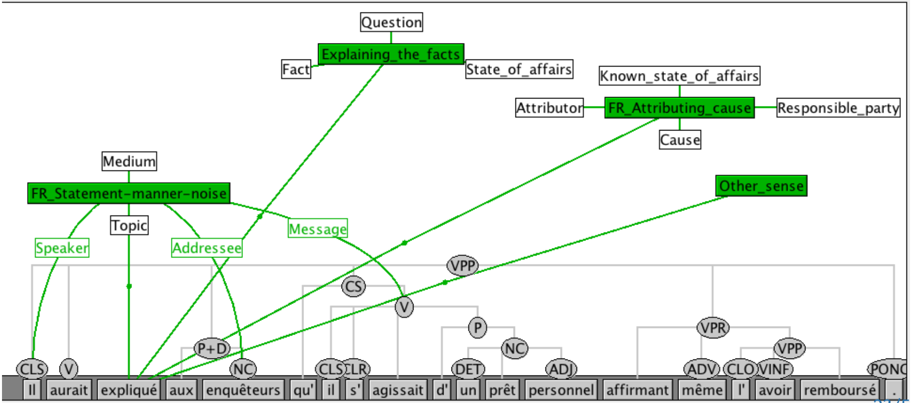
- Either select examples to maximize diversity of grammatical realizations
  - lexicographic approach
  - lexicographic examples of Berkeley FrameNet
- or annotate running text: all lemmas of a text / first *xxx* occurrences of given lemma
  - preserves natural distribution
    - of senses
    - of grammatical characteristics of role fillers
  - used in full-text annotations of Berkeley FrameNet
  - SALSA German FrameNet
  - $\longrightarrow$ French FN, first 100 occurrences in target corpus

Corpus annotation

- Annotation on syntactic (dependency) trees, in order to:
  - speed up annotation (but bias)
  - extract syntactico-semantic patterns
- 2 preexisting treebanks:
  - French Treebank (Abeillé et al. 03) and Sequoia (Candito & Seddah, 12)
  - $\approx$ 21500 sentences
  - Mainly news, plus medical, Europarl, Fr-wikipedia narrative texts
- Annotation at most first 100 occurrences of covered lemma+pos
- "Out_of_domain" dummy frame when sense is not covered

Corpus annotation: Graphical tool

- Pilot annotation to develop annotation guide
- 2 independent annotations + adjudication (for 75% of data)
- 1 single annotation (by expert) for remaining 25%
  - Automatic pre-annotation of possible frames of the lemma
  - Salto tool (Burchardt et al., 2006)



Evaluation: inter-annotator agreement

Between 2 independent annotations (75% of all annotations):

- for a trigger occurrence: Fscore for the frame choice
- for a frame chosen by both annotators: Fscore for role fillers

	Nb trigger occurrences	% of Nouns	% of Verbs	Fscore inter-annotateur		
				Frame	Exact Role	Partial Role
	17667	36	50	85.9	77.2	81.9
Break-down by domain						
Commercial	3307	60	40	92.0	73.4	80.4
Causality	7691	30	48	79.2	74.2	80.4
Cognitive stances	7886	28	62	90.6	81.1	86.0
Communication	2221	23	76	89.6	82.3	87.5
Break-down by trigger POS						
V	8834	-	-	87.6	82.8	87.1
N	6234	-	-	86.8	68.3	72.5
other	2509	-	-	77.7	74.6	82.1

Stats (1.2 release)

http://asfalda.linguist.univ-paris-diderot.fr/frameIndex.xml

- $\longrightarrow \approx$ 16200 annotated frames (plus 8750 occurrences "Other_sense")

	Nb distinct frames	Nb distinct lemma+POS	Nb senses	Nb annotated frames ($\neq$ Other_sense)
ALL	105	873	1109	16167
Fully covered lemma+POS		490		7213
Commercial	19	90	99	2930
Causality	11	243	285	3895
Cognitive	44	372	442	5426
Communication	47	347	411	5233
N	-	296	346	5282
V	-	446	594	9165
PREP	-	35	43	674
ADV	-	26	42	407
CONJ	-	22	28	301
ADJ	-	43	48	234

Semantic roles

- 421 frame-specific roles (for the 105 frames)
- grouped into 40 “macro-roles” defined at the domain level

Frame structure and lexicon: feedback

- Division into frames is clearly the most difficult part
  - When starting from English frames: Difficulty to understand differences between frames
  - Attempt to specify distinctive characteristics of frames
- Automatic induction with or without manual intervention could be an option
- do we really want discrete classes of LUs (frames) ?

Corpus FrameNet annotation: typical problems

Berkeley FrameNet: chosen examples $\longrightarrow$ Annotation difficulties not much documented

Very wide range of phenomena to handle

- divergences between litteral expression and interpretation
  - multi-word expressions (for triggers)
  - metaphors, ellipsis...
- lexical semantics
  - polysemy
  - nouns referring to a participant
- morpho-syntax / semantics interface
  - syntax/semantic mismatches (documented in Fillmore, 2007)
  - elliptic coordinations: arg cluster, head gapping...
  - ambiguous status of reflexive “se” clitic
    - diathesis alternation marker
    - frozen
    - true reflexive ...

Focus: predicative noun referring to a participant

(English examples for ease of reading)

Relational nouns:

*The precise **CAUSES** of multiple sclerosis are not yet known.*

SYNTACTICALLY:

- *causes* is monovalent: one PP

Focus: predicative noun referring to a participant

(English examples for ease of reading)

Relational nouns:

*The precise **CAUSES** of multiple sclerosis are not yet known.*

SYNTACTICALLY:

- *causes* is monovalent: one PP

SEMANTICALLY: predicate could be

- monovalent: entity_being_the_cause_of(Effect)

Focus: predicative noun referring to a participant

(English examples for ease of reading)

Relational nouns:

*The precise **CAUSES** of multiple sclerosis are not yet known.*

SYNTACTICALLY:

- *causes* is monovalent: one PP

SEMANTICALLY: predicate could be

- monovalent: entity_being_the_cause_of(Effect)
- $\longrightarrow$ bivalent: causality_relation(Cause, Effect)

Focus: predicative noun referring to a participant

(English examples for ease of reading)

Relational nouns:

*The precise **CAUSES** of multiple sclerosis are not yet known.*

SYNTACTICALLY:

- *causes* is monovalent: one PP

SEMANTICALLY: predicate could be

- monovalent: entity_being_the_cause_of(Effect)
- $\longrightarrow$ bivalent: causality_relation(Cause, Effect)

(frame Causation; Cause, Effect)

*The precise **CAUSES** of multiple sclerosis are not yet known*

Focus: predicative noun referring to a participant

In annotation guidelines, explicit distinction of:

- referential uses (most occurrences)
- predicative uses
  - $\longrightarrow$ in which we consider the trigger does not refer to the participant

Distinction pertaining for any noun:

- *Have you seen the unicorn?*
- *This animal is a unicorn.*

Focus: predicative noun referring to a participant

Referential uses

(frame Causation; Cause, Effect)

*The exact **CAUSES** of multiple sclerosis are not yet known*

*Liberia is still suffering from the **CONSEQUENCES** of 14 years of war*

Focus: predicative noun referring to a participant

Typical cases of predicative uses:

(frame Causation; Cause, Effect)

- **Apposition**:
  *First **CAUSE** of unemployment , short-term contract terminations are rising.*
- **Copulative sentence**:
  *Chronic venous insufficiency is the **CONSEQUENCE** of various abnormalities.*
- **Inverted copulative sentence**:
  *The **RESULT** of these measures has been to partially fill the deficit.*

Event/Participant ambiguity

Typical case of a predicate that can refer to a role:
event/participant ambiguous nominalisations (or event/result).

(Commerce_buy; Buyer, Goods)

- *(...) to multiply its* ***PURCHASES*** *of car parts (...)*
  - event, buying act
- *How many times have you put down your* ***PURCHASES*** *willy-nilly, and trusted the cashier to do his job right?*
  - *purchases* refers to bought goods

Normalizing syntactic representations: “Deep” syntax

(Candito et al., 14; Perrier et al. 14)

see also Enhanced dependencies in Universal Dependencies (Schuster and Manning, 16; Candito et al. 2017)

Bottom-up approach from syntactic parses

- how to better take advantage of syntactic parses?
- (almost) without resorting to semantic disambiguation
- use of formal criteria mostly
- to complete and normalize syntactic valency of verbs, adjectives

Normalizing syntactic representations: “Deep” syntax

(Candito et al., 14; Perrier et al. 14)

see also Enhanced dependencies in Universal Dependencies (Schuster and Manning, 16; Candito et al. 2017)

Bottom-up approach from syntactic parses

- how to better take advantage of syntactic parses?
- (almost) without resorting to semantic disambiguation
- use of formal criteria mostly
- to complete and normalize syntactic valency of verbs, adjectives

One step towards semantic representations
$\neq$ orthogonal to approaches that jointly build syntactic/semantic representation

- cf. categorial grammar, frame-augmented LTAGs (Kallmeyer and Osswald, 13)...

Normalizing syntactic representations: “Deep” syntax

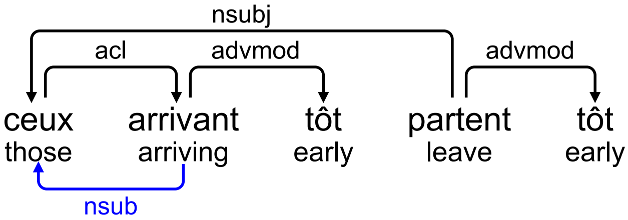
Three main enhancements, concerning very well known phenomena:

- recovering non-local arguments
- neutralizing syntactic alternations (diathesis alternation)
- by-passing of syntactic markers

“Deep” syntax: recovering non-local arguments

“subjects” of non finite verbs: **cases fully determined by syntax**

- raising/control verbs: *Paul seems/wants to sleep*
- control nouns, adjectives ...
- noun-modifying participles: *people liking sports / born in 45*



$\longrightarrow$ dependency graphs

"Deep" syntax: recovering non-local arguments

"subjects" of non finite verbs: **cases not fully determined by syntax**

- e.g. infinitive adverbial clauses

Il mangera avant de jouer
He will-eat before to play
« He will eat before playing »

When main verb is active, with non expl subject
subject of infinitive = subject of main verb
in most cases (83% on Sequoia corpus)

Counter-example:
D'autres photos ont subi des retouches pour **accentuer** *le drame*
Other photos have undergone modifications to **accentuate** *the drama*

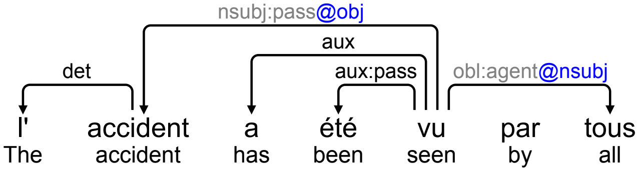
“Deep” syntax: recovering non-local arguments

Arguments shared by coordinated predicates

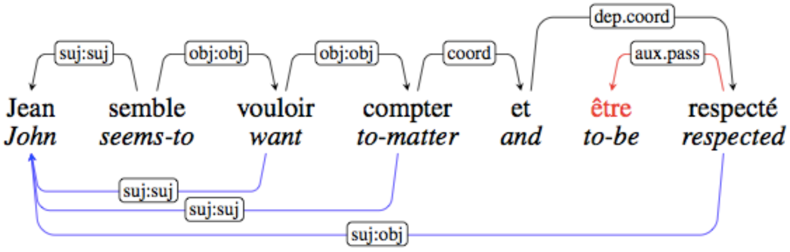
- **Paul** *is* **starving** *and* **wants** *to* **eat**
- *Paul is cooking and selling pancakes*
- *Paul is sleeping and selling pancakes*

"Deep" syntax: Neutralizing syntactic alternations

- recover canonical grammatical functions
  - ≈ the function you would get in active personal voice
- cheap way to limit linking diversity
- for French:
  - massive for passive
  - but also: impersonal, causative, mediopassive



Deep syntax: interaction of phenomena



Frequent in corpus

Attested example:
Les **députés** *du Bundestag ne peuvent pas être* **appelés** *à* **témoigner** *ou être* **arrêtés** *pour une infraction*

The deputees of-the Bundestag (not) can not be called to testify nor be arrested for an infraction

Deep syntax: obtaining deep syntactic graphs

Gold data: manually validated annotation on Sequoia (3099 sentences)

- iterative work annotation / graph-rewriting rules development
- Grew system (Guillaume et al. 2012) / OGRE system (Ribeyre et al. 2012)

Deep syntax: obtaining deep syntactic graphs

Gold data: manually validated annotation on Sequoia (3099 sentences)

- iterative work annotation / graph-rewriting rules development
- Grew system (Guillaume et al. 2012) / OGRE system (Ribeyre et al. 2012)

Quantitative assessment of enhancements:

- 200 sentences set
- $\approx$ 1000 edges correspond to arguments of verbs
- set A: 18.9% are not present in surface dependency tree
- set N: 13.9% have a "neutralized" label
- $A \cup N$ represent 27.7% of the 1000 argumental edges

Deep syntax: obtaining deep syntactic graphs

Pseudo-gold data: deterministic graph-rewriting rules applied to French Treebank (Ribeyre et al, 2014)

- Evaluation on 200 sentences shows quality is quite good (Fscore=97.7)

Deep syntax: obtaining deep syntactic graphs

Pseudo-gold data: deterministic graph-rewriting rules applied to French Treebank (Ribeyre et al, 2014)

- Evaluation on 200 sentences shows quality is quite good (Fscore=97.7)

Deep syntax parsing:

- pipeline surface parsing + deterministic rules
- or direct learning of graph parser (Ribeyre et al., 2015)

FrameNet parsing

- WSD task: frame selection for an ambiguous trigger
- SRL task: role identification

Challenges

- Generalization over data
  - WordNet (e.g. Johansson et Nugues, 2007)
  - Distributed representations (e.g. Hermann et al. ACL 2014)
- Joint models (e.g. Yang and Mitchell 2017)

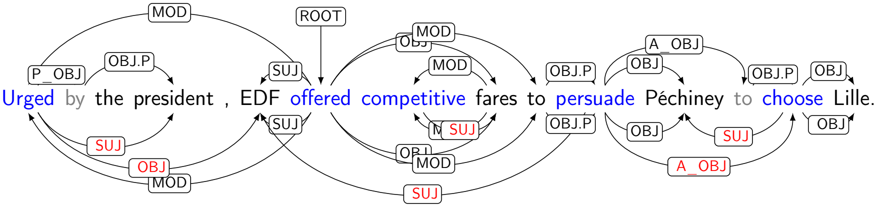
Deep syntax for FrameNet parsing

Joint work **Olivier Michalon**, Corentin Ribeyre, Alexis Nasr (Michalon et al. Coling 2016)

- syntactic features known to be quite useful for SRL
  - since Gildea et Jurafsky, 2002
  - still true with neural networks approach (Hermann et al. 14; Yang and Mitchell 17)
- is it worth using deep syntax ?

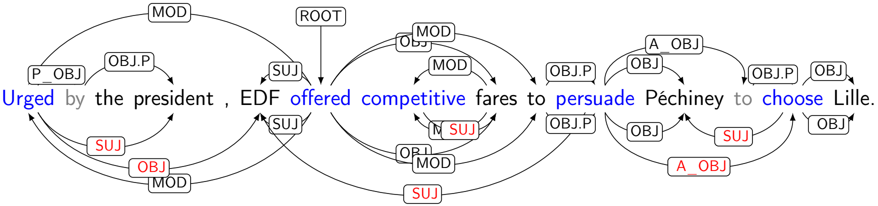
Deep syntax for FrameNet parsing

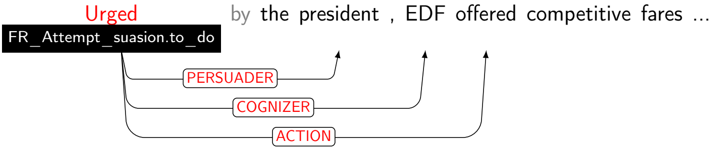
(arcs for determiners and punctuations not shown)



Deep syntax for FrameNet parsing

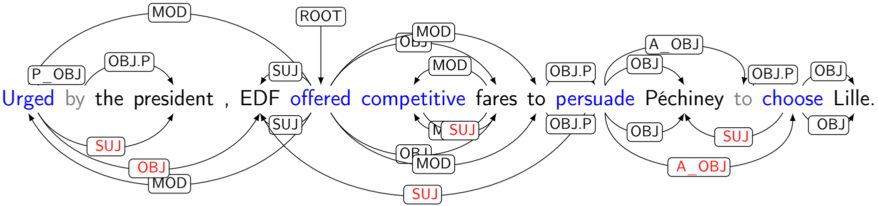
(arcs for determiners and punctuations not shown)

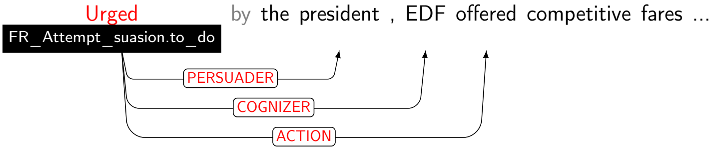




Deep syntax for FrameNet parsing

(arcs for determiners and punctuations not shown)





Syntactic path between "Urged" et "EDF" :

- surface: -mod,+suj
- deep: +obj

Measuring the normalizing effect

Syntactic path between

- a predicate
- (the syntactic head) of a role filler

For a given role, deep syntactic paths are more regular:

The distributions P(path to role filler | frame-specific role) are less scattered when using deep syntax.

cf. average entropy over all roles decreases:

- 1.65 with "classical" syntactic paths
- 1.32 with "deep" syntactic paths

Measuring the normalizing effect

5 most frequent paths,
for the role fillers of verbal triggers

surface syntax		deep syntax	
(+suj)	25.0%	(+suj)	33.1%
(+obj)	17.0%	(+obj)	32.8%
(-mod)	8.0%	(+a_obj)	4.7%
(+obj,+obj.cpl)	4.4%	(-mod)	3.2%
(+a_obj,+obj.p)	4.1%	(+mod,+obj.p)	2.5%
Total	58.6 %	Total	76.2 %

Impact for FrameNet parsing

Very basic system (pipeline WSD + SRL, supervised linear classification)

- WSD : one classifier per ambiguous lemma
- SRL : one classifier per frame

Positive impact for FrameNet SRL, in particular for verbal triggers

	Prec.		Recall		F-measure	
Input syntax	surf	deep	surf	deep	surf	deep
WSD (gold frame $\neq$ Other_sense)	80.1	80.7	80.1	80.7	80.1	80.7
SRL (for gold role filler heads)	81.4	86.4	59.1	66.1	**68.5**	**74.9**

Prec.		Recall		F-measure	
surf	deep	surf	deep	surf	deep
80	80.5	80.8	80.9	80.4	80.7
75.7	80.3	51.6	59.0	**61.3**	**68.0**

Table: FastSem results for **verbs**, using **gold** (top) and **predicted** (bottom) surf and deep syntax.

Perspectives

- Deep syntax:
  - Currently checking whether improvement carries over using a neural architecture

Perspectives

- Deep syntax:
  - Currently checking whether improvement carries over using a neural architecture
- FrameNet construction:
  - very difficult and time consuming
  - coverage will never be sufficient
  - $\longrightarrow$ Frame induction
    - WSD + deep syntactic subcats + distributed representation of fillers...

Perspectives

- Deep syntax:
  - Currently checking whether improvement carries over using a neural architecture
- FrameNet construction:
  - very difficult and time consuming
  - coverage will never be sufficient
  - $\longrightarrow$ Frame induction
    - WSD + deep syntactic subcats + distributed representation of fillers...
- Saturation of annotated frames
  - a lot of roles are not instantiated
  - DNI flag for definite null instanciation
  - recover role fillers in preceding sentences

Perspectives

- Deep syntax:
  - Currently checking whether improvement carries over using a neural architecture
- FrameNet construction:
  - very difficult and time consuming
  - coverage will never be sufficient
  - $\longrightarrow$ Frame induction
    - WSD + deep syntactic subcats + distributed representation of fillers...
- Saturation of annotated frames
  - a lot of roles are not instantiated
  - DNI flag for definite null instanciation
  - recover role fillers in preceding sentences
- In the long term, still a long way for NLU
  - in particular: factivity of eventualities

Thank you!