

Evaluating and improving syntactic lexica by plugging them within a parser

Elsa Tolone, Benoît Sagot, Éric de La Clergerie

LIGM, Université Paris-Est, France & FaMAF, Universidad Nacional de Córdoba, Argentine

INRIA Paris-Rocquencourt

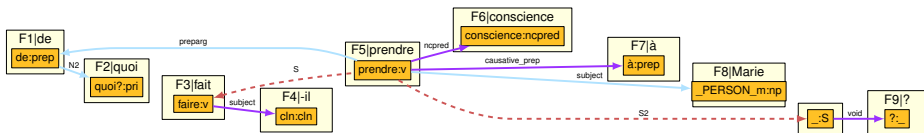
<http://alpage.inria.fr>



LREC'12
Istanbul
May 23-25 2012

Deep parsing with FRMG (and LEFFF)

FRMG is a deep wide-coverage TAG parser for French, co-developed with LEFFF lexicon



de quoi fait-il prendre conscience à Marie ?
What does he make Mary become aware of ?

To be tried at <http://alpage.inria.fr/parserdemo>

Evaluating lexica ?

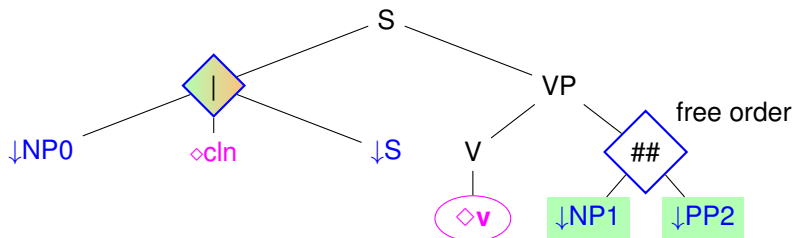
FRMG depends on LEFF syntactic lexicon (Sagot), and largely co-developed with it, but

- can we use FRMG with other lexica ?
- is it easy to plug a new lexicon ?
- can we then run fine-grained evaluations of syntactic lexica ?
- can we use feedback information to improve a lexicon ?

We tried to answer these questions, with 3 new lexica: LGLEX, DICOVALENCE, and NEW LEFF

FRMG: a French meta-grammar

- A large-coverage Meta-grammar for French abstract descriptive layer, constraint-based, modularity, inheritance
- generation of a TAG/TIG grammar extended domain of locality, capture of subcategorization frames
- with **factorized** trees
 - ▶ current version: 323 trees (and only 38 verbal trees)
 - ▶ one tree \equiv many ordinary TAG trees
 - ▶ \Rightarrow one verbal tree stands for many subcat frames, arg positions, realizations, ...



FRMG

hypertag #286

arg0	<code>arg0</code>	<code>extracted -</code> <code>fun <code>fun0</code></code> <code>kind <code>kind0</code> subj nosubj</code> <code>pcas -</code> <code>real <code>real0</code> - CS N2 PP S cln prel pri</code>
arg1	<code>arg1</code>	<code>extracted -</code> <code>fun <code>fun1</code></code> <code>kind <code>kind1</code> - acomp obj prepacomp prepobj</code> <code>pcas <code>pcas1</code> + - apres à avec de par ...</code> <code>real <code>real1</code> - CS N N2 PP S V adj cla ...</code>
arg2	<code>arg2</code>	<code>extracted -</code> <code>fun <code>fun2</code></code> <code>kind <code>kind2</code> - prepacomp prepobj prepscomp</code> <code> prepvcomp scomp vcomp wh-</code> <code> comp</code> <code>pcas <code>pcas2</code> - + apres à ...</code> <code>real <code>real2</code> - CS N N2 PP S ...</code>
cat	<code>v</code>	
diathesis	<code>active</code>	
refl	<code>refl</code>	
ctrsubj	<code>ctr</code>	
imp	<code>imp</code>	

Coupling FRMG with a lexicon: Hypertags

FRMG

hypertag #286

arg0	arg0	[extracted - fun fun0 kind kind0 subj nosubj pcas - real real0 - CS N2 PP S cln prel pri]
arg1	arg1	[extracted - fun fun1 kind kind1 - acompl obj prepacomp prepobj pcas pcas1 + - apres à avec de par ... real real1 - CS N N2 PP S V adj cla ...]
arg2	arg2	[extracted - fun fun2 kind kind2 - prepacomp prepobj prepscomp prepvcomp scomp vcomp wh-comp pcas pcas2 - + apres à ... real real2 - CS N N2 PP S ...]
cat	v	
diathesis	active	
refl	refl	
ctrsubj	ctr	
imp	imp	

LEFF

hypertag «**promettre**»

arg0	[fun subj kind subj - pcas -]
arg1	[fun obj kind obj scomp pcas -]
arg2	[fun objà kind prepobj - pcas à -]
refl	-
ctrsubj	subj
imp	-

Coupling FRMG with a lexicon: Hypertags

FRMG

hypertag #286

arg0	[arg0]	extracted - fun [fun0] subj kind [kind0] subj nosubj pcas real [real0] - CS N2 PP S cln prel pri
arg1	[arg1]	extracted - fun [fun1] obj kind [kind1] acomp obj prepacomp prepobj pcas [pcas1] + apres à avec de par ... real [real1] - CS N N2 PP S V adj cla ...
arg2	[arg2]	extracted - fun [fun2] objà kind [kind2] prepacomp prepobj prep- scomp prepvcomp scomp vcomp whcomp pcas [pcas2] + apres à ... real [real2] - CS N N2 PP S ...
cat	v	
diathesis	active	
refl	[refl]	
ctrsubj	[ctr] subj	
imp	[imp]	

LEFF

hypertag «promettre»

arg0	[fun subj - kind subj - pcas -
arg1	[fun obj kind obj scomp pcas -
arg2	[fun objà kind prepobj - pcas à -
refl	-
ctrsubj	subj
imp	-

ALEXINA is a lexical formalism

- with an *intensional* level for lemma
- and the generation of an *extensional* level for forms
- the descriptions use a set of primitive features, and macros

LEFFF is a wide-coverage morphosyntactic and syntactic lexicon for French, covering all categories

LEFFF is partially factorized:

one entry may cover several meanings and several subcat frames

⇒ 5,736 entries for 5,450 distinct ones (intensional level)

Freely available at <http://gforge.inria.fr/projects/alexina/>

- Intensional level

```
promettre v55 100;Lemma;v;  
<Suj:cln | scompl | sinf | sn,  
  Obj:( cla | de-sinf | scompl | sn) ,  
  Objà:( cld | à-sn) >;  
@CtrlSujObj , cat=v;  
%actif ,%passif,%ppp_employé_comme_adj,%passif_impersonnel
```

- Extensional level

```
promet 100 v  
[pred="promettre_____1<Suj:cln | scompl | sinf | sn,  
_____  
_____Obj:( cla | de-sinf | scompl | sn) ,  
_____  
_____Objà:( cld | à-sn)>" ,  
@CtrlSujObj , @pers , cat=v , @P3s]
```

- A macro

```
@CtrlSujObl = [ctrsubj = suj];
```

Resulting from the conversion of LADL tables (Gross)

- 1 first, into Lglex format (Constant and Tolone)
- 2 then, into alexina format (Sagot and Tolone)

- Wide-coverage lexicon

kind	#tables	#entries	#lemma
verbs	67	13,867	5,738
pred. nouns	78	12,696	8,531

- Fine-grained: many entries for some verbs
53 entries for *tenir* (LEFFF: 6 entries)
- Some features can't be represented in ALEXINA and/or exploited in FRMG
for instance: added determiner for predicative nouns in FRMG
missing: semantic restrictions

Freely available at <http://infolingua.univ-mlv.fr>

Developed by **Mertens** and **van den Eynde**

- based on pronominal approach (**Benveniste**)
- fine grained: one entry for one meaning
- small coverage
3,738 verbs for 8,313 entries

Evolution of LEFFF towards a more semantic lexicon

- finer-grained: one meaning per entry
- automatic merge of LEFFF and DICOVALENCE, plus manual validation of 505 verbs (986 entries):
 - ▶ the 100 most frequent lemmas
 - ▶ the *dubious lemmas*: more output entries than the sum of corresponding LEFFF and DICOVALENCE entries
- still a wide-coverage lexicon
7,933 verbs, for 12,613 entries

The lexica at a glance

Verbs from each lexicon:

Lexica	#Entries	#Lemmas	Ratio
LEFFF	7,108	6,827	1.04
LGLEX	13,867	5,738	2.41
DICOVALENCE	8,313	3,738	2.22
NEW LEFFF	12,613	7,933	1.58

Other categories imported from LEFFF, shared by all lexica

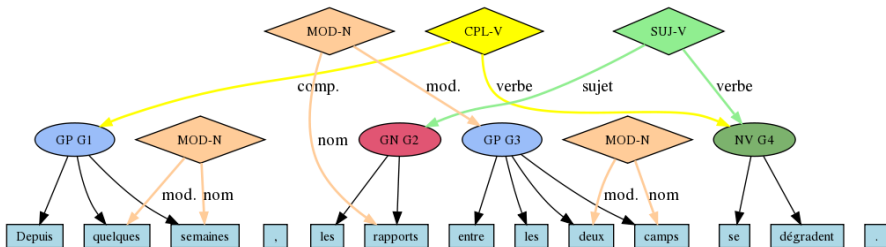
Category	#Int. Entries	#Lemmas	#Ext. Entries
nouns	41,816	41,592	86,675
adjectives	10,556	10,517	34,359
adverbs	4,111	3,676	4,155
prepositions	260	259	728
proper nouns	52,499	52,202	52,571
other	1,007	854	1,589

EASy/Passage evaluation campaign

French parsing evaluation campaigns organized within EASy and Passage actions.

We use the EASy reference corpus as benchmark

- around 4K sentences, manually annotated (but with errors !)
- various styles: journalistic, literacy, medical, mail, oral, questions
- constituency and dependency based format (shallow level)
 - ▶ 6 kinds of *chunks*: GN, NV, GA, GR, GP, PV
 - ▶ 14 kinds of *relations*: SUJ-V, AUX-V, COD-V, ATB-SO , CPL-V, MOD-V, MOD-N, MOD-A, MOD-R, MOD-P, COMP, COORD, APPOS, JUXT



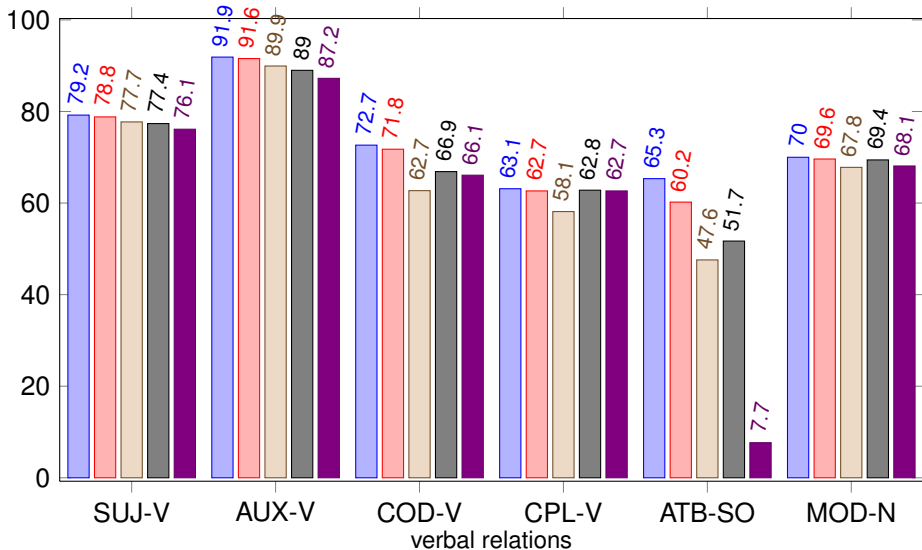
Overall results

Setting:

- each sentence segmented with **SXPIPE**, no prior tagging, use of a lexicon
- **FRMG** returns either
 - ▶ full parses (possibly by relaxing some agreement constraints)
 - ▶ sequences of partial parses, covering the sentence
 - ▶ nothing in case of timeout (100s)
- whenever possible, **FRMG** returns a shared dependency forest of all possibilities
- then heuristic-based disambiguation and conversion to Passage format

Lexicon	Cover. (%)	Chunks (%)	Rels (%)	Time (s)
LEFF	83.45	89.03	66.76	0.35
NEW LEFF	82.19	88.74	66.09	0.55
LGLEX	80.61	87.89	63.19	1.10
DICOVALENCE	71.44	88.08	64.49	0.38
Old DICOVALENCE	65.69	87.06	62.72	0.42

Analysis per verbal relation (F-measure)



LEFF NEW LEFF LGLEX DICOVALENCE Old DICOVALENCE

Experiments on French TreeBank

Evaluation on CONLL dependency version of FTB (journalistic style)
richer set of verbal dependencies, but still shallow level

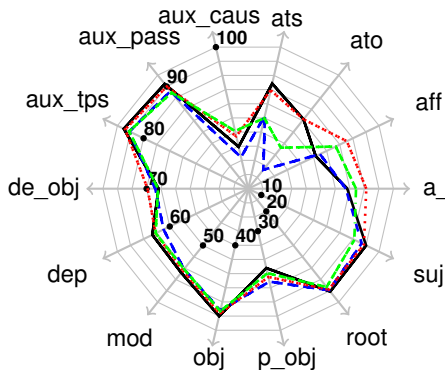
1	de	de	P	P	5	de_obj
2	quoi	quoi?	PRO	PROWH	1	obj
3	fait	faire	V	V	5	aux_caus
4	-il	-il	CL	CLS	5	subj
5	prendre	prendre	V	V	0	root
6	conscience	conscience	N	NC	5	obj
7	à	à	P	P	5	mod
8	Marie	marie	N	NPP	7	obj
9	?	?	PONCT	PONCT	5	ponct

On FTB test part (1200 sentences)

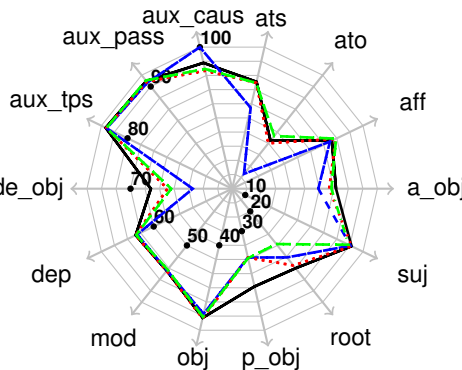
Lexicon	Cover. (%)	LAS (%)	Time (s)	New LAS (%)	δ (%)
LEFF	89.53	82.21	0.61	85.82	4.4
NEW LEFF	88.76	81.36	0.94	83.83	3.0
LGLEX	86.73	78.75	1.95	81.82	3.9
DICOVALENCE	75.28	79.38	0.69	-	-
MST	-	88.20	-	-	-

Analysis per CONLL verbal relation

Recall



Precision



— Leff - - - LGLex
..... NewLeff - · - · Dicovale

— Leff - - - LGLex
..... NewLeff - · - · Dicovale

Original motivation: find lexical entries that are incorrect or incomplete through full parse failures:

a form is suspect if occurring more often than expected in failed sentences, in co-occurrence with non-suspect forms

~> fix-point iterative algorithm, close to EM (Expectation-Maximization) and return the best sentences where a form is the main suspect
=> WEB-based interface to browse suspects, lexical info, and sentences

May be used for any lexica, but can also be adapted for contrasting lexica

a verb is suspect for lexicon L if occurring more often than expected in failed sentences that succeed for LEFFF, in co-occurrence with non-suspect verbs.

Tried on a 100Ksent. toy corpus (wikipedia, wikisource, europarl, AFP news) but could be tried on CPC (100Mwords) or even bigger (700Mwords)

Some suspects (for LGLEX)

A first typology of errors on the first 15th suspects for LGLEX:

- missing entries in the expected LADL table
 - ▶ **réaffirmer** (to reaffirm 28), **réélire** (to reelect, 10), **mixer** (to mix, 7), **zapper** (to omit, 4)...
la mémoire ... qui zappe les détails ...
the memory ... which omits the details ...
- existing entries, but missing coding information
 - ▶ **susciter** (to spark off, 41; 36DT & 38R), **recruter** (to recruit, 14; 38R), ...
- mandatory args (for LGLEX), but missing ones in the sentences
kidnapper (to kidnap, 12; 36DT N0 V N1 Prep N2) : *Les deux Italiens ont été kidnappés le 18 décembre*
- misc. situations

For NEW LEFFF: most significant error is *estimer* (to consider) : missing clausal argument (to consider that S)

Conclusion

- Relatively easy to plug new (alexina-based) lexica into **FRMG**
- Rather good results for all lexica, even if lower than with **LEFFF** (better than **FRMG+LEFFF** in 2007 Passage campaign)
- Room for needed and normal co-adaptation **FRMG**-lexicon

Future:

- specific training per lexica
- completing lexica and/or merging information (error mining, evaluation)
- better factorization of lexical entries in **LGLEX**, delaying use of more semantic entries at disambiguation time
- assign probabilities to entries and/or frames
- enrich **FRMG** with some new features, to take into account richer lexical information

Thank you