

Extending the adverbial coverage of a NLP oriented resource for French

Elsa Tolone & Stavroula Voyatzi

Laboratoire d'Informatique Gaspard-Monge
Université Paris-Est, France
{elsa.tolone;voyatzi}@univ-paris-est.fr

SUMMARY

We are presenting work on extending the adverbial entries of *LGLex*, a NLP oriented syntactic resource for French. Adverbs were extracted from the Lexicon-Grammar tables of both simple adverbs ending in *-ment* 'ly' (Molinier and Levrier, 2000) and compound adverbs (Gross, 1986; 1990). Various features are encoded in both LG tables and describe the relations of deletion, permutation, intensification and paraphrasing that associate, on the one hand, the simple and compound adverbs and, on the other hand, different types of compound adverbs. The resulting syntactic resource is manually evaluated and freely available under the LGPL-LR license in both plain text format and XML, at the following site:

<http://infolingu.univ-mlv.fr/english> > Language Resources > Lexicon-Grammar > Download

Syntactic-semantic classes of simple adverbs

Class	Syntacticosemantic subclass	Example	Size
ADVPC	Conjunct	finalement	51
ADVPS	Utterance-level	sincèrement	17
ADVP AE	Evaluative	curieusement	10
ADVP AH	Habit	habituellement	10
ADVP AM	Modal	certainement	28
ADVP AS	Subject oriented attitude	bêtement	30
ADVMS	Subject oriented manner	gentiment	883
ADV MV	Verbal manner	démocratiquement	1523
ADV MQI	Intensifier	excessivement	247
ADV MQC	Completion	complètement	13
ADV MQE	Qualitative extension	fondièrement	21
ADV MTQ	Time	actuellement	31
ADV MTD	Duration	temporairement	11
ADV MTF	Frequency	régulièrement	33
ADV MP	Viewpoint	linguistiquement	279
ADV MF	Focus	essentiellement	16
Total			3,203

Simple adverbs in *-ment*

ID	Adv	Class	Subclass	Example	Size
164	linguistiquement	ADV	ADV MQC	linguistique	1
165	littérairement	ADV	ADV MQC	littéraire	1

Molinier & Levrier, 2000

Notations: Adv (adverb), Adj (adjective), Adj-n (noun derived from an adjective), P (clause), V (verb), W (any range of verb complements), and Ddef (definite article)

Most features encoded in both LG tables of simple adverbs in *-ment* and compound adverbs produce variants that can be easily added to the syntactic lexicon *LGLex*

Type of feature	Example	Number
paraphrase	Adj-ment = (du + d'un) point de vue Adj	12
paraphrase embedded in a construction	N0 V W de (E + une) (façon + manière) Adj	8
deletion	Prep1 Det1 C1	3
permutation	Prep Det Modif pré-adj Adj C	2
possessivation	Prep1 Det1 C1 de N2 = Prep1 Poss2 C1	2
intensification	très Adv	6
Total		33

Fine-grained linguistic descriptions

Lexicon-Grammar framework
Gross, 1975; 1994

LG tables

electronic inventory of lexical & syntactic information

LADL since 1970s, LIGM since late 1990s

Conversion with *LGExtract*
Constant & Tolone, 2010

LGLex

NLP oriented syntactic lexicon

Tolone, 2011

Morphosyntactic classes of compound adverbs

Class	Morphosyntactic structure	Example	Size
PADV	C	enfin	520
PC	Prep C	par exemple	664
PDETC	Prep Det C	de nos jours	848
PAC	Prep DetAdj C	à la dernière minute	776
PCA	Prep Det C Modif pré-adj Adj	à la nuit tombante	840
PCDC	Prep Det1 C1 de Det2 C2	dans la limite du possible	750
PCPC	Prep1 Det1 C1 Prep2 Det2 C2	des pieds à la tête	287
PCONJ	Prep1 Det1 C1 Conj Prep2 Det2 C2	tôt ou tard	333
PCDN	Prep Det1 C1 de N2	à l'insu de N	555
PCPN	Prep1 Det1 C1 Prep2 N2	en comparaison avec N	151
PV	Prep V Prepv Det Cv	à dire vrai	285
PF	ConjS (Det0 C0 + N0) V Prep1 (Det1 C1 + N1)	jusqu'à ce que mort s'ensuive	396
PECO	(Adj) comme Det C	(fidèle) comme un chien	305
PVCO	(V) comme Det C	(travailler) comme un chien	338
PPCO	(V) comme Prep Det C	(disparaître) comme par enchantement	50
PJC	ConjC Det C1 Prep C2	mais aussi et surtout	185
Total			7,283

Compound adverbs

ID	NO = Nhum	NO = N-hum	Neg obl	Ppv	Prédicat type	<ENT>-Prep	<ENT>-Det	<ENT>-C	<ENT>-Modif pré-adj	<ENT>-Adj	Prep Det C	Prep Det Modif pré-adj Adj C	Prep Modif pré-adj Adj C	Conjunction
424	-	+	-	-	:se produire	<E>	le	matin	<E>	dit	-	+	-	-
425	+	-	-	-	:faire N	<E>	le	maximum	<E>	possible	+	-	-	-

Gross, 1986; 1990

Encoding of features: + (binary, valid), - (binary, not valid), <E> (lexical, empty)

Notations: Prép (preposition), Det (determiner), Adj (adjective), Modif pré-adj (pre-adjectival modifier), N (noun), V (verb), Conj (conjunction), and C (noun tied with the rest of the adverb)

Adverbial entry in *LGLex*

```

<entry>
- <entry status="completed" id="P_advmp_164">
- <lexical-info cat="adv">
- <lexical-value>
- <complete value="linguistiquement"/>
- <lexical-value>
- <adv value="en linguistique"/>
- <adv value="d'un point de vue linguistique"/>
- <adv value="du point de vue linguistique"/>
- <adv value="du point de vue de la linguistique"/>
- <adv value="au point de vue linguistique"/>
- <adv value="au niveau linguistique"/>
- <adv value="linguistiquement parlant"/>
- <adv value="sur le plan linguistique"/>
- <adv value="sur le plan linguistique"/>
- <lexical-info>
- <arguments>
- <call-constructions>
- <absolute-constructions>
- <construction value="0:Adv, NO V W?"/>
- <construction value="0:est Adv Qui P"/>
- <construction value="true:NO V Adv W"/>
- <construction value="0:Adv, NO ne V pas W"/>
- <construction value="0:Adv parlant, P"/>
- <construction value="true:Adv, NO V W"/>
- <structureAdv>
- <construction value="0:en Ndomaine"/>
- <construction value="0:un point de vue Adj"/>
- <construction value="0:du point de vue Adj"/>
- <construction value="base:Adv"/>
- <construction value="0:du point de vue de Ddef Ndomaine"/>
- <construction value="0:au point de vue Adj"/>
- <construction value="0:au niveau Adj"/>
- <construction value="0:Adv parlant, P"/>
- <construction value="0:au plan Adj"/>
- <construction value="0:sur le plan Adj"/>
- <structureAdv>
- <relative-constructions>
- <call-constructions>
- </call-constructions>
</entry>

```

Identifier of the adverbial entry and subcategorization class

Adverbial entry and its generated variants

Constructions in which a given adverb can occur, i.e. at the beginning of a negative clause

Various morpho-syntactic structures of a given adverb

RESULTS

Adverbial entries in <i>LGLex</i>	Enhancement of <i>LGLex</i>
Initial entries	10,487
Paraphrases	9,208 (88%)
Deletions	1519 (14%)
Permutations	103 (1%)
Possessivations	288 (3%)
Intensifications	210 (2%)
Final entries	21,815 (108%)

FUTURE WORK

Conversion into the *Lefff* format
Sagot et al., 2006; 2010

Conversion into the *DELA* format
Courtois & Silbertztein, 1990

FRMG
French Meta-Grammar

WOLF
French WordNet

DELAS/DELAC
French Morphological Dictionary

RELEX network

More than twenty research teams in Computational Linguistics have been working on the construction of LRs using identical methods within the Lexicon-Grammar framework. LG tables of adverbs are available in French, English, Portuguese, Spanish, German, Italian, Korean and Modern Greek, at: <http://infolingu.univ-mlv.fr>