Extending the adverbial coverage of a French morphological lexicon

Elsa Tolone, Stavroula Voyatzi, Claude Martineau & Matthieu Constant
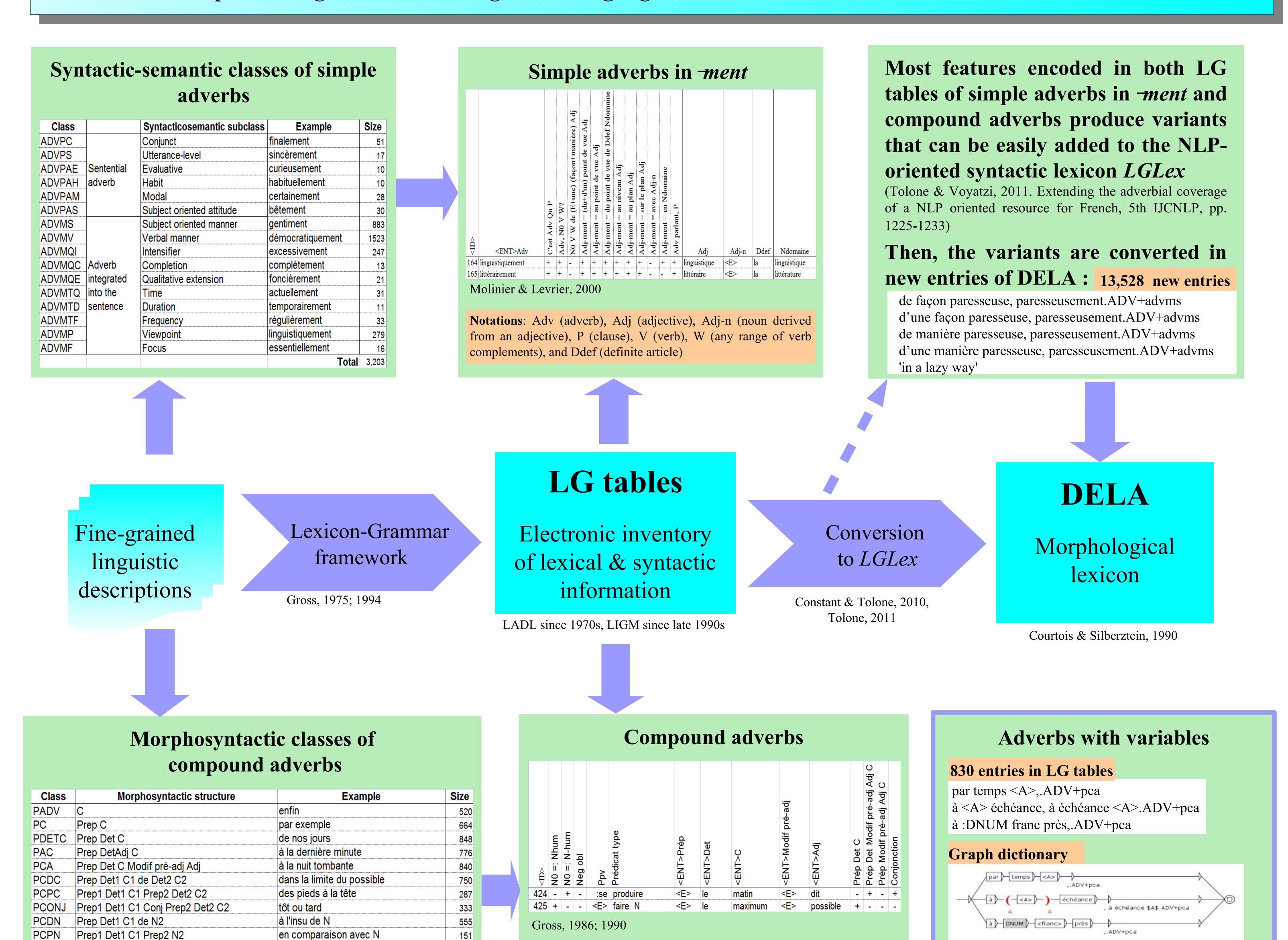
LIGM, Université Paris-Est, France
PLN, FaMAF, Universidad Nacional de Córdoba, Argentina
{tolone;voyatzi;martinea;mconstant}@univ-paris-est.fr

SUMMARY

We present an extension of the adverbial entries of the French morphological lexicon DELA (Dictionnaires Electroniques du LADL / LADL electronic dictionaries). Adverbs were extracted from *LGLex*, a NLP-oriented syntactic resource for French, which in its turn contains all adverbs extracted from the Lexicon-Grammar tables of both simple adverbs ending in ¬*ment* ('ɫy') (Molinier and Levrier, 2000) and compound adverbs (Gross, 1986; 1990). This work exploits fine-grained linguistic information provided in existing resources. We obtained 23,389 final entries by merging 9,038 initial entries and 21,593 new entries in lexicon DELA. The resulting resource is reviewed in order to delete duplicates and is freely available under the LGPL-LR license, 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
ADVPAE	Sentential	Evaluative	curieusement	10
ADVPAH	adverb	Habit	habituellement	10
ADVPAM		Modal	certainement	28
ADVPAS		Subject oriented attitude	bêtement	30
ADVMS		Subject oriented manner	gentiment	883
ADVMV		Verbal manner	démocratiquement	1523
ADVMQI		Intensifier	excessivement	247
ADVMQC	Adverb	Completion	complètement	13
ADVMQE	integrated	Qualitative extension	foncièrement	21
ADVMTQ	into the	Time	actuellement	31
ADVMTD	sentence	Duration	temporairement	11
ADVMTF		Frequency	régulièrement	33
ADVMP		Viewpoint	linguistiquement	279
ADVMF		Focus	essentiellement	16
			Total	**3,203**

Simple adverbs in ¬*ment*

<D>	<ENT>Adv	C'est Adv Qu P	Adv, N0 V W?	N0 V W de (E/une) (façon/manière) Adj	Adj-ment = (du+d'une) point de vue Adj	Adj-ment = au point de vue Adj	Adj-ment = du point de vue de Ddef Ndomaine	Adj-ment = au niveau Adj	Adj-ment = au plan Adj	Adj-ment = sur le plan Adj	Adj-ment = avec Adj-n	Adj-ment = en Ndomaine	Adv parlant, P		Adj	Adj-n	Ddef	Ndomaine
164	linguistiquement	+	+	.	+	+	+	+	+	+	+	+		linguistique	<E>	la	linguistique
165	littérairement	+	+	.	+	+	+	+	+	+	.	.	+		littéraire	<E>	la	littérature

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 NLP-oriented syntactic lexicon *LGLex*

(Tolone & Voyatzi, 2011. Extending the adverbial coverage of a NLP oriented resource for French, 5th IJCNLP, pp. 1225-1233)

Then, the variants are converted in new entries of DELA : `13,528 new entries`

de façon paresseuse, paresseusement.ADV+advms
d'une façon paresseuse, paresseusement.ADV+advms
de manière paresseuse, paresseusement.ADV+advms
d'une manière paresseuse, paresseusement.ADV+advms
'in a lazy way'

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 to *LGLex*
Constant & Tolone, 2010, Tolone, 2011

DELA
Morphological lexicon
Courtois & Silberztein, 1990

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	Prep Det1 C1 Conj Prep2 Det2 C2	tôt ou tard	333
PCDN	Prep Det1 C1 de N2	à l'insu de N	555
PCPN	Prep Det1 C1 Prep2 N2	en comparaison avec N	151
PV	Prep V Prepv Detv 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

<D>	N0 = Nhum	N0 = N-hum	Neg obl	Ppv	Predicat type	<ENT>Prep	<ENT>Det	<ENT>C	<ENT>Modif pré-adj	<ENT>Adj	Prép Det C	Prép Det Modif pré-adj Adj C	Prep Modif pré-adj Adj C	Conjction
424	+	-	:se	produire	<E>	le	matin	<E>	dit	-	+	-	+	
425	+	-	-	<E>	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)

Adverbs with variables

830 entries in LG tables
par temps <A>,.ADV+pca
à <A> échéance, à échéance <A>.ADV+pca
à :DNUM franc près,.ADV+pca

Graph dictionary

Entries in DELA format
par temps sec,.ADV+pca
par temps pluvieux,.ADV+pca
à brève échéance,à échéance brève.ADV+pca
à cinq francs près,.ADV+pca

RESULTS

	without variables	with variables
Initial entries in DELA	9,036	/
New entries from *LGLex*	21,898	952
New entries from *LGLex* (without duplicates entries)	20,757	830
All entries	29,793	/
All entries (without duplicates entries)	22,564	/

FUTURE WORK

EVALUATION

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

FRMG
French Meta-Grammar

WOLF
French WordNet

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