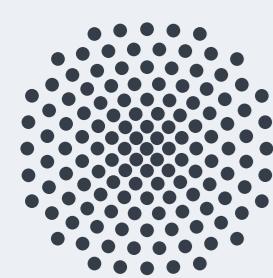


LVF-lemon – Towards a Linked Data Representation of “Les Verbes français”



University of Stuttgart
Germany

Ingrid Falk & Achim Stein
Project B5 @ SFB 732: Incremental Specification in Context—Institut für Linguistik/Romanistik

About

How to convert traditional [lexical syntactic-semantic resource](#) – “Les Verbes français” (LVF, [Dubois and Dubois-Charlier, 1997, François et al., 2007]) into standardised and normalised [linked open data model](#)?

Objective

- ▶ explicit and make accessible [linguistic knowledge](#)
- ▶ [interoperability](#) with other linguistic resources (e.g. corpus data and subcat lexicons).

To do

RDF conversion ≈ straightforward

Data modeling: model content in terms of established vocabularies:

general: RDFS, OWL, SKOS, ...

linguistic: lemon, LexInfo, OLIA, LMF, ...

LVF specific

Linking the data

Focus: [Syntax](#), [semantics](#) and [syntax-semantics interface](#).

The Lexicon – LVF, Les Verbes français [Dubois and Dubois-Charlier, 1997, François et al., 2007]

- ▶ ≈12 300 verbs, ≈25 610 readings, ≈15 readings/verb

- ▶ elaborate morpho-syntactic and semantic description.

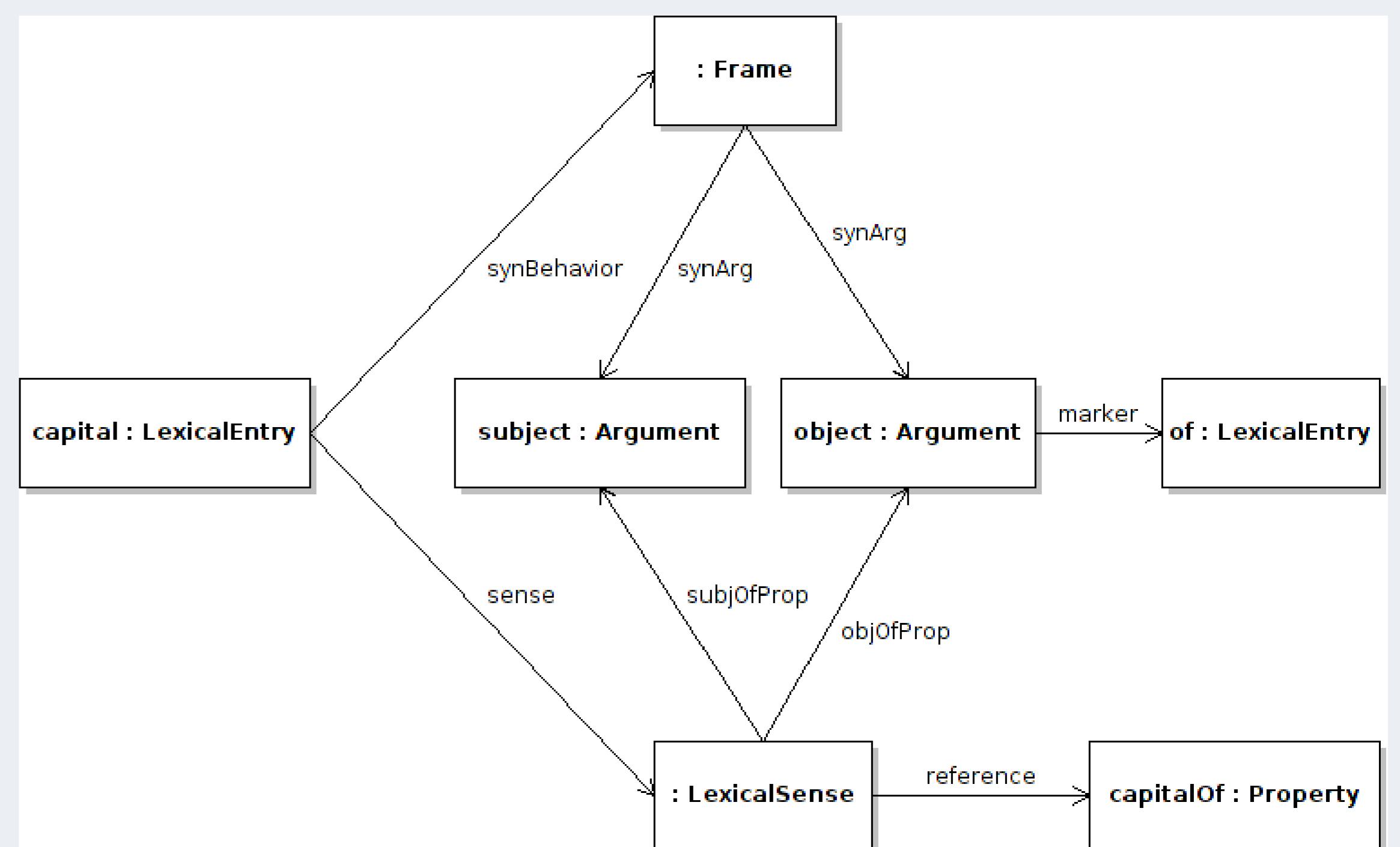
Sample uses, semantic descriptions			Syntactic descriptions			
id example	semantic decomposition	sem. primitive	sem. class	id schema	encoded information	
01 On élargit une route/ La route s'élargit.	r/d+qt large	(make) become	Transformation	01 A30	intransitive with adjunct, inanimate subject	
02 Cette veste élargit Paul aux épaules.	d large	become	Transformation	T1308	transitive, human subject, inanimate direct object, instrumental adjunct	
03 On élargit ses connaissances.	r/d large abs	(make) become, figurative	Transformation	P3008	reflexive, inanimate subject, instrumental adjunct	
04 On élargit le débat à la politique étrangère. f.ire abs VRS		directed move, figurative	Enter/Leave	02 N1i	intransitive, animate subject, prep. phrase headed by <i>de</i> (<i>of</i>)	
				A90	intransitive with adjunct, subject human or thing	
				T3900	transitive, inanimate subject, object human or thing	

Table: Sample readings for élargir/enlarge

Lemon: the linguistic LLOD model [McCrae et al., 2012]

- ▶ links lexical entries to knowledge repositories (ontology)
- ▶ lexicon and knowledge (ontology) layers separated
- ▶ linguistically sound structure based on Linguistic Markup Framework – LMF [Francopoulo et al., 2006]
- ▶ linking to data categories: allows for arbitrarily complex linguistic description, integration with annotated corpora
- ▶ RDF-native form: enables leverage of existing Semantic Web technologies
- ▶ [syntax-semantics interface](#) based on LexInfo ontology
 - ▷ lexical entries ↔ ontology sense
 - ▷ link between subcat frame ↔ sense not represented

Problem in LVF: [syntax & semantics intricately interleaved!](#)



lemonUBY – [Eckle-Kohler et al., 2012]

- ▶ representation of UBY, network of multilingual, interlinked lexical semantic resources
- ▶ based on LMF, ISOcat data categories
- ▶ conversion by mapping LMF representations
- ▶ [synsem mapping](#) implicit:
 - ▷ via subcats, syntactic arguments ↔ thematic roles

PDEV-lemon – [Maarouf et al., 2014]

- ▶ representation of PDEV, the Pattern Dictionary of English verbs
 - ▷ verbs associated with attested patterns of use.
- ▶ [synsem mapping](#): patterns ↔ PDEV extension frame sense
 - ▷ to model phraseological or idiomatic expressions
- ▶ ontology of semantic types: *human*, *animate*, etc.

LVF-lemon: road-map

- ▶ [morphology](#) can be appropriately represented by [lemon](#) core model (morphology module)
- ▶ [syntax](#): schemes can be mapped to [lemon](#) or LMF syntactic arguments
- ▶ syntactic/semantic types assigned to syntactic arguments:
 - ▷ LVF specific, related to PDEV semantic type ontology
 - ▷ 8 LVF specific semantic roles, related to VerbNet/FrameNet role inventories

semantics: need LVF specific representation

semantic classes: based on linguistic theories

semantic primitives in *opérateur*: [Pinker, 1989, Jackendoff, 1983] among others
⇒ relate to VerbNet/FrameNet representations

References

- [Dubois and Dubois-Charlier, 1997] Dubois, J. and Dubois-Charlier, F. (1997). *Les Verbes français*.
- [Eckle-Kohler et al., 2012] Eckle-Kohler, J., McCrae, J., and Chiarcos, C. (2012). lemonUby - A large, interlinked, syntactically-rich resource for ontologies. *SWJ (Semantic Web Journal)*.
- [Francopoulo et al., 2006] Francopoulo, G., Bel, N., George, M., Calzolari, N., Monachini, M., Pet, M., and Soria, C. (2006). Lexical Markup Framework (LMF). In *Proceedings of Fifth LREC*.
- [François et al., 2007] François, J., Le Pesant, D., and Leeman, D. (2007). Présentation de la classification des Verbes Français de Jean Dubois et Françoise Dubois-Charlier.
- [Jackendoff, 1983] Jackendoff, R. (1983). *Semantics and Cognition*.
- [Maarouf et al., 2014] Maarouf, I. E., Bradbury, J., and Hanks, P. (2014). PDEV-lemon: a Linked Data implementation of the Pattern Dictionary of English Verbs based on the Lemon model. In *Proceedings of the 3rd Workshop on Linked Data at LREC'14*.
- [McCrae et al., 2012] McCrae, J., Aguado-de Cea, G., Buitelaar, P., Cimiano, P., Declerck, T., Gómez-Pérez, A., Gracia, J., Hollink, L., Montiel-Ponsoda, E., Spohr, D., and Wunner, T. (2012). Interchanging lexical resources on the Semantic Web. *Language Resources and Evaluation*.
- [Pinker, 1989] Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*.