



## Towards TreeLex++: Syntactico-Semantic Lexical Resource for French

Anna Kupść, Pauline Haas, Rafael Marin, Antonio Balvet

### ► To cite this version:

Anna Kupść, Pauline Haas, Rafael Marin, Antonio Balvet. Towards TreeLex++: Syntactico-Semantic Lexical Resource for French. Language & Technology Conference, May 2019, Poznan, Poland. hal-02120183

**HAL Id: hal-02120183**

**<https://hal.archives-ouvertes.fr/hal-02120183>**

Submitted on 5 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Towards TreeLex++

## Syntactico-Semantic Lexical Resource for French

Anna Kupść, Pauline Haas, Rafael Marin and Antonio Balvet

akupsc@u-bordeaux-montaigne.fr, pauline.haas@univ-paris13.fr, rafael.marin@univ-lille.fr, antonio.balvet@univ-lille.fr

### Abstract

The paper presents a syntactico-semantic lexicon of over a thousand French verbs. It has been created by manually adding lexical aspect to verb frames gathered in TreeLex (Kupść and Abeillé, 2008). We present how the original syntactic resource has been adapted to the current project, our aspect assignment procedure and an overview of the resulting database.

**Keywords:** computational semantics, corpora-based methods in language engineering, electronic language resources and tools, formalization of natural languages

## 1. Introduction

For practical Natural Language Processing applications (e.g. Information Extraction, Syntactic Parsing, Text Generation), machine-tractable as well as human-readable large-scale lexical resources are still a very valuable asset, even in a scene which appears dominated by robust Machine-Learning algorithms and giga-word corpora. For instance, even though syntactic parsing has seen great advances in the past 10 years, thanks to the development of Treebanks and dependency-annotated corpora, even the best parser fails to capture such an intuitive linguistic notion as transitivity. In this sense, (semi)manually constructed lexica are an indispensable complementary resource to corpus-driven resources, such as word embeddings and lexical conditional probabilities databases. Each type of resource captures a portion of the problem (i.e. Language), and thus the challenge contemporary NLP systems are facing today is more how to integrate different knowledge sources than proving that one source is better – or more consistent – than the other.

In this paper, we present TreeLex++, an extension of TreeLex (Kupść and Abeillé, 2008), a syntactic lexicon for French, based on the French Treebank, enriched here with aspectual semantics information. Different lexica have been devised over several decades for the automatic processing of French texts, in different theoretical frameworks: from the manually-encoded Lexicon-Grammar tables (Gross, 1975) couched in a distributionalist framework, to contemporary large-scale, semi-automatically induced lexica such as the Lefff (Sagot *et al.*, 2006 ; Sagot, 2010). Most of those lexical resources have focused on providing a formalized syntactic description of the main syntactic categories, with an emphasis on verbal predicates. In extending TreeLex with aspectual semantics information, our goal is primarily to set up a large-scale aspectual semantics characterization process of lexical units. Secondly, we wish to provide the NLP and Natural Language Engineering communities with a resource which combines corpus-induced syntactic characterizations (as opposed to theory-driven ones) as well as basic aspectual distinctions, based on Vendler's hierarchy. We believe that such a resource will make it possible to capture selectional restrictions that are not accessible to 'purely' syntactic descriptions.

In the first sections, we present how TreeLex++ derives from the original FTB-induced TreeLex resource (Section

2 and 3). Then we move on to the presentation of our aspectual semantics characterization process (Section 4). In section 5 we give a general overview of the present state of the resource. Section 6 is dedicated to conclusions and perspectives.

## 2. TreeLex

TreeLex is a syntactic lexicon automatically extracted from the French Treebank (FTB), cf. Abeillé *et al.* (2003). The lexicon contains ca. 2000 contemporary French verbs with their syntactic realizations and frequencies found in the FTB.

The FTB is a corpus of newspaper texts (*Le Monde* journal, 1990-1993) encoded in XML format. In addition to lexical information for every word (category, lemma, person, number, gender etc.), the corpus provides a syntactic structure for each sentence: both syntactic groups and functions are indicated, see Fig.1 below. The part of the corpus where both syntactic groups and functions are annotated covers ca. 21500 sentences (about 525000 words).

```
<SENT argument="ETR" author="MINANGOY
ROBERT" date="1990-01-19" nb="1006"
textID="456">
  <NP fct="SUJ">
    <w cat="D" ee="D-def-ms" ei="Dms"
lemma="le" mph="ms" subcat="def">Le</w>
    <w cat="A" ee="A-ord-ms" ei="Ams"
lemma="deuxième" mph="ms"
subcat="ord">deuxième</w>
    <w cat="N" ee="N-C-ms" ei="NCms"
lemma="problème" mph="ms"
subcat="C">problème</w>
  </NP>
  <VN>
    <w cat="V" ee="V--P3s" ei="VP3s"
lemma="être" mph="P3s" subcat="">est</w>
  </VN>
  <NP fct="ATS">
    <w cat="D" ee="D-def-fs" ei="Dfs"
lemma="le" mph="fs" subcat="def">la</w>
    <w cat="N" ee="N-C-fs" ei="NCfs"
lemma="nourriture" mph="fs"
subcat="C">nourriture</w>
  </NP>
  <w cat="PONCT" ee="PONCT-S"
ei="PONCTS" lemma="." subcat="S">.</w>
</SENT>
```

Fig1. A sample of FTB sentence annotation

The FTB annotation schema is centered around the verbal nucleus (VN) which makes syntactic dependents easily accessible. This corpus organization is exploited by Kupść and Abeillé (2008) in order to obtain obligatory arguments and provide syntactic frames for verbs present in the FTB. The resulting lexicon, called TreeLex<sup>1</sup>, provides a rich syntactic representation of each argument since both functions and their phrasal realizations are encoded. Example (1) shows a lexical entry for the transitive verb *entraver* ‘to impede’ which takes a nominal subject (SUJ:NP) and a nominal direct object (OBJ:NP).

(1) *entraver*: SUJ:NP, OBJ:NP

Names of functions and phrases are adopted directly from FTB notation with two additions (*ref* et *obj*) for obligatory clitics, cf. Tab.1. Arguments with clitic realizations are used to indicate reflexive verbs (ex. *se réjouir* ‘to rejoice’: SUJ:NP, *ref*:CL), idiomatic expressions (ex. *s’en sortir* ‘to manage’: SUJ:NP, *obj*:en, *ref*:CL) or an impersonal subject (ex. *falloir* ‘to have to’: SUJ:il, OBJ:VPinf).

function	meaning	possible phrasal realizations
SUJ	subject	NP, VPinf, Ssub
OBJ	direct object	NP, VPinf, Ssub
A-OBJ	indirect object introduced by <i>à</i>	VPinf, PP
DE-OBJ	indirect object introduced by <i>de</i>	VPinf, PP
P-OBJ	indirect prepositional object (other than <i>de</i> and <i>à</i> )	PP
ATS	subject’s complement	AP, NP, VPpart, VPinf, Ssub
ATO	direct object’s complement	AP, NP, VPpart, VPinf, Ssub
ref	obligatory reflexive clitic	CL
obj	other obligatory clitic	en, y

Table 1. TreeLex functions with syntactic realizations

If a verb allows for different syntactic combinations (i.e., either a list of functions or their realizations differ), every frame is listed separately. Therefore, a single verb (more precisely, its lemma) can be found several times in the lexicon, see (2). As no semantic disambiguation is made, this strategy aims at distinguishing potentially different senses associated with each frame. Indeed, in (2a-b), *voler* has the meaning of ‘to steal’ whereas in (2c) it can be translated as ‘to fly’.

- (2) a. *voler*: SUJ:NP, OBJ:NP, A-OBJ:NP  
 b. *voler*: SUJ:NP, DE-OBJ:NP  
 c. *voler*: SUJ:NP

As noted on TreeLex’s website, an optional realization of specific arguments has been added manually, (3).

- (3) *détruire*: SUJ:NP, (OBJ:NP)

Finally, since multi-word units are indicated in FTB, TreeLex lists 465 multi-word verbs, ex. *courir le risque* ‘to take a risk’, *donner lieu* ‘to result/take place’.

### 3. Beyond TreeLex: Towards TreeLex++

TreeLex contains 1912 verbs and 3229 entries, i.e., verb-frame couples, which correspond to 24660 verb occurrences<sup>2</sup> attested in the FTB corpus. The resource provides a rich syntactic information and, as stated in Kupść and Abeillé (2008,p.38), it can be easily integrated with other resources to be employed in NLP tasks such as parsing or text generation. However, its relatively small size makes large-scale applications problematic.

On the other hand, TreeLex’s size makes an in-depth qualitative linguistic study feasible. For example, it could be extended with semantic information to investigate interactions between semantic and syntactic properties of verbs. For French, several projects regroup syntactic and semantic verbal properties, incorporating different levels of semantic information, e.g., verbal semantic classes (LVF, cf. François *et al.* (2007)), thematic roles (French FrameNet, cf. Djemaa *et al.* (2016)) or lexical aspect (Nomage, cf. Balvet *et al.* (2012) or Falk and Martin, 2016). In the current project, we decided to focus on high-level syntax-semantics relationships and thus we augmented the syntactic frames in TreeLex with manually encoded aspectual information. Our approach differs from Balvet *et al.* (2012) or Falk and Martin (2016), as verbal<sup>3</sup> aspect assignment is guided by corpus examples rather than by elicited sentences. Similarly to Falk and Martin (2016), aspect is assigned to a verb-frame couple rather than to a verb alone. Nevertheless the level of detail of our aspectual classes is distinct both from Balvet *et al.* (2012) and Falk and Martin (2016): we use only the four major Vendlerian classes (see sec.4 for details).

In order to prepare the TreeLex data for aspect assignment, several modifications have been adopted. First, all frames had to be represented in a uniform way. Therefore all syntactic arguments, whether optional or not, have been treated equally and indications of optional realizations have been removed. In particular, verbs such as *détruire* ‘to destroy’ in (3) were transformed into (4):

- (4) *détruire*: SUJ:NP, OBJ:NP

Second, we had to address the ambiguity in TreeLex entries. As shown in (2), TreeLex verbs may appear with several frames. According to Kupść and Abeillé (2008), this affects about 40% of TreeLex verbs. Such multiple frames may indicate a polysemous and/or a polyaspectual verb. However, different syntactic realizations of a single argument structure (the same sequence of functions) are also listed as separate frames in TreeLex, see (5). This representation usually introduces an artificial syntactic (frame) and semantic (meaning) ambiguity. The direct object (OBJ) of the verb *déplore* ‘to regret/deplore’ in (5) has two syntactic realizations (a nominal phrase, NP, or a subordinate phrase, Ssub) but this syntactic variation does not imply a difference in meaning.

- (5) *déplore*: SUJ:NP, OBJ:Ssub  
*déplore*: SUJ:NP, OBJ:NP

<sup>1</sup> [http://redac.univ-tlse2.fr/lexiques/treelex\\_en.html](http://redac.univ-tlse2.fr/lexiques/treelex_en.html)

<sup>2</sup> We present here counts calculated from the on-line TreeLex version, [http://redac.univ-tlse2.fr/lexiques/treelex/treelex\\_verbs.csv](http://redac.univ-tlse2.fr/lexiques/treelex/treelex_verbs.csv)

<sup>3</sup> Balvet *et al.* (2012) use corpus examples to assign aspectual properties to nouns but not to verbs.

In order to avoid such an artificial ambiguity, we grouped all frames which differed only in type of phrasal realization. Therefore, the double nature of OBJ in (5) is currently represented as in (6).

(6) *déplorer*: SUJ:NP, OBJ:NP/Ssub

In order to reduce the semantic ambiguity, we decided to consider only verbs which, after syntactic grouping, appeared with a single syntactic frame. As a consequence, verbs such as *voler* in (2) have been excluded.<sup>4</sup> Multi-word verbal units have been omitted as well as their meaning is usually idiomatic.

Finally, all remaining 1161 verbs have been coupled with examples extracted from the FTB. We collected corpus examples in order to illustrate how each frame is instantiated and provide a real context for aspect assignment.

#### 4. Incorporating Lexical Aspect

Aspectual information has been added manually to TreeLex verbs. Unlike grammatical aspect, lexical aspect refers to inherent semantic properties indicating the way in which predicates are structured in relation to time.

In the most general terms, the properties in question have to do with the presence (or lack thereof) of some end (limit or boundary) or dynamism in the lexical structure of certain classes of verbs. Thus, for instance, the presence of a limit distinguishes between telic (i.e., a time-limited situation) and atelic verbs.

These semantic properties give rise to four major aspectual classes (cf. Vendler, 1967): STATE, ACTIVITY (ACT), ACCOMPLISHMENT (ACC) and ACHIEVEMENT (ACH). Their semantic features are listed in Tab.2 below.

CLASS	dynamic	durative	telic
STATE	-	+	-
ACT	+	+	-
ACC	+	+	+
ACH	+	-	+

Table 2. The four situation types, based on Vendler (1967)

Our aspect assignment procedure consisted in a double manual annotation by two experts in semantics. Strictly speaking, our method is not an annotation process as each expert had access to the decision taken by the other annotator and the final result has been commonly discussed. Each verb has been considered along with its syntactic frame and the corresponding examples found in FTB. The assignment task consisted in choosing one of the four classes (tags) in Tab.2. Each decision was made after applying the usual tests presented in the literature on verb lexical aspect (see Garey 1957, Kenny 1963/[1994], Vendler 1967, Dowty 1979, Wilmet 1980, Lamiroy 1987,

Daladier 1996, Rothstein 2004, among others). We have used the following six tests:

T1: progressive form of *être en train de* 'to be V-ing'

T2: question related to dynamicity *Que s'est-il passé hier?* 'What happened yesterday?'

T3: use of aspectual semi-auxiliaries *commencer à* 'to start doing something', *continuer de* 'to keep on doing something', *finir de* 'to finish doing something', *arrêter de* 'to stop doing something'

T4: duration complement *en x temps* 'in x time'

T5: duration complement *pendant x temps* 'during x time'

T6: imperfective paradox *V[temps inaccompli] IMPLIQUE V [temps accompli]* 'V[imperfect tense] IMPLIES V [perfect tense]'

Situation type	T1	T2	T3	T4	T5	T6
STATE	no	no	no	yes/no	yes/no	yes
ACT	yes	yes/no	yes/no	no	yes	yes
ACC	yes	yes	yes	yes	yes/no	no
ACH	no	yes	no	no	no	no

Table 3. A grid for the allocation of aspectual classes to TreeLex verbs

In order to illustrate our procedure, let us take the verb *invoker* 'to invoke' in one of the sentences where it appears in the corpus:

(7) Pour justifier cette décision, la direction **invoque** la déprime du marché automobile.

'To justify this decision, the management invokes the depression of the automobile market.'

T1: This verb cannot appear in a progressive form: \**La direction est en train d'invoker la déprime du marché automobile.*

T2: *La direction a invoqué la déprime du marché automobile* is an acceptable answer to the question *Que s'est-il passé hier?*

T3: This verb cannot appear as a complement of *commencer*, *continuer*, etc.: \**La direction a commencé/continué à invoquer la déprime du marché automobile.*

T4: *Invoker* is not compatible with *en x temps*: \**La direction a invoqué la déprime du marché automobile en deux heures.*

T5: It is not compatible with *pendant x temps* either: \**La direction a invoqué la déprime du marché automobile en deux heures.* This sentence is only acceptable in an iterative reading.

T6: *La direction invoquait la déprime du marché automobile* does not imply *La direction a invoqué la déprime du marché automobile.*

<sup>4</sup> This strategy does not replace a real semantic disambiguation since verbs which allow for a single

syntactic frame may still be polysemous. This issue will be addressed in further sections.

This way, according to the battery of tests summarized in Table 4, *invoquer* in (7) should be assigned to the ACHIEVEMENT class:

	T1	T2	T3	T4	T5	T6
<i>invoquer</i>	no	yes	no	no	no	no

Table 4. Test results for (7)

It is important to mention that verbs were annotated according to their meaning in the sentences found in the FTB corpus. Verbal polysemy was addressed only if different meanings appeared in the corpus, but the effects of coercion were not taken into account. When applying the tests, subjects and direct objects were transformed into the singular number (to avoid iterative readings, which can coerce an ACC V into an ACT V). We have used imperfect tenses in order to avoid habitual readings. Finally, frequency adverbs triggering iterative readings were not taken into account either.

We obtain an aspectual characterization limited to the meanings appearing in the corpus. It is not an annotation of the verbs as lemmas, neither verbs in sentences, but rather an annotation of verbal structures (verb + arguments) in a discursive context, which allowed us to identify verbal meaning and to avoid polysemy as much as possible.

## 5. Data in TreeLex++

The resulting resource, which we call TreeLex++, contains 1161 verbs enriched with syntactic (frame) and semantic (lexical aspect) properties. It is available in a text format as a CSV file (comma separated value). Each verb is accompanied by its frame, the lexical aspect, the number of examples found in FTB and their full list<sup>5</sup>. To simplify the search of the inflected form in the example text, the corresponding verb is indicated between <b> and </b> tags, as below:

- (8) Quant à moi , je trouve qu' on se <b>fiche</b> du monde en n' expliquant pas les choses en langage courant .

‘As for me, I think that they don’t give a toss about anybody by not giving an explanation in the common language.’

To make linguistic generalizations easier, information encoded in syntactic frames has been translated into several representations:

- number of syntactic arguments (clitic arguments are excluded from this count)
- whether a verb is reflexive or not
- a general frame (a list of syntactic functions and obligatory clitics)
- a simplified frame (a list of syntactic functions alone)
- the full frame including syntactic realizations (types of phrases)

The corresponding syntactic information for *déplorer* in (6) and the reflexive verb *se fiche* ‘not caring about’ presented in TreeLex++ format is given in Tab.5.

verb	Nr Arg	r e f l	general frame	simplified frame	full frame
<i>déplorer</i>	2	n o	SUJ.OBJ	SUJ.OBJ	SUJ:NP. OBJ: NP/Ssub
<i>ficher</i>	2	y e s	SUJ.DE-OBJ.refl	SUJ.DE-OBJ	SUJ:NP. DE-OBJ:PP. refl:CL

Table 5. Syntactic information in TreeLex++

A brief summary of syntactic realizations (the number of syntactic arguments) of TreeLex++ verbs is given in Tab.6 below. The number of arguments in TreeLex++ does not exceed three and the vast majority of verbs (74.24%) have two arguments. However, as indicated in Tab. 5, this does not necessarily correspond to a transitive structure (SUJ.OBJ) as the second argument may have a different function than a direct object (see Tab.1).

NRARGS	NUMBER OF VERBS	%
1	183	15.76%
2	862	74.24%
3	116	9.99%

Table 6. The distribution of verbs with respect to the number of arguments

The distribution of verbal aspectual classes found in TreeLex++ is given in Tab.7.

ASPECT	NUMBER OF VERBS	%
ACH	576	49.61%
ACC	260	22.39%
ACT	219	18.86%
ETAT (STATE)	103	8.87%
ETAT ACT	1	0.08%
ETAT ACH	1	0.08%
ACH ACT	1	0.08%

Table 7. Aspect distribution in TreeLex++

The majority of verbs in TreeLex++ are telic (ACH or ACC) or even dynamic (ACH, ACC or ACT). However, the distribution of durative (STATE, ACT, ACC) and non-durative (ACH) verbs is almost equal.

<sup>5</sup> Individual examples are separated by a vertical bar ‘|’.

The resource is neither syntactically nor semantically balanced, which is probably due to the content of the FTB corpus (newspaper texts).

As shown in Tab.7, most verbs are assigned a single aspect. Hence, it seems that our approximate disambiguation technique is quite efficient. There are 3 verbs, however, which have a double aspect: *excéder*, *observer*, and *traverser*. Indeed, judging from their context, these verbs are truly polysemous in FTB: *excéder* is ambiguous between ‘to exceed’ and ‘to infuriate’, *observer* is used as either ‘to observe’ or ‘to respect/keep’ and *traverser* corresponds to ‘to cross’ or ‘to experience’. Therefore, even when syntactic properties are restricted to a single frame and all coercion factors from FTB examples are eliminated, semantic and aspectual ambiguity are still present.

## 6. Conclusions and Perspectives

TreeLex++ is a lexicon which groups both syntactic and semantic properties for over a thousand verbs illustrated with attested examples. Such a database offers a valuable resource both for a linguistic study and NLP applications. From a research perspective, it will be interesting to verify correlations between syntactic frames and aspect values. For instance, intuitively, the accomplishment verbs (ACC) should be associated with transitive verbs (2-argument predicates). TreeLex++ provides an opportunity to verify this hypothesis empirically: not only can it be confirmed or infirmed but we can also calculate its strength. As for NLP applications, a number of practical uses of aspectual information is cited in Falk and Martin (2016): the assessment of event factuality, text summarization, machine translation or automatic detection of temporal relations. It would be interesting to test our resource in this kind of tasks.

In the current version, TreeLex++ contains only single-frame verbs, which roughly covers a half of the entries in TreeLex. In order to include the remaining half in TreeLex++, we have to employ a true semantic disambiguation technique first. As mentioned in Section 5, a verb with a unique syntactic combination may still be polysemous and polyaspectual. In case of several frames, this potential ambiguity is multiplied and human disambiguation effort, already complex and time-consuming, increases considerably. A possible solution could be a lexical look-up of verb-frame couples in LVF (François *et al.*, 2007) in order to identify different verb senses. However, pairing the senses with corresponding FTB examples would require a separate technique.

We wish also to develop an evaluation methodology for our resource. For example, we could compare our results with aspect values attributed to verbs in the Nomage project (Balvet *et al.*, 2012). However, Nomage methodology (for verbs) differs from ours as aspect assignment is based on elicited examples rather than on verb uses in a corpus. Another comparison could be made with the syntactico-semantic resource described in Falk and Martin (2016) which served for training of an automatic classifier of verbal aspect. Unfortunately, this data does not seem to be publicly available. Moreover, both resources use different aspectual values from ours

thus the corresponding tagsets have to be converted first in order to provide the equivalent information.

The current version of TreeLex++ is freely available on-line: <http://redac.univ-tlse2.fr/lexiques/treelexPlusPlus.html>.

## References

- Abeillé, A., L. Clément, and F. Toussenen. 2003. Building a treebank for French, in A. Abeillé (ed.), *Treebanks*, Kluwer, Dordrecht, pp.165-187.
- Balvet A., L. Barque, M.-H. Condette, P. Haas, R. Huyghe, R. Marín, A. Merlo. 2012. La ressource Nomage, Confronter les attentes théoriques aux observations du comportement linguistique des nominalisations en corpus, *TAL* (52), n° 3/2012, pp.1-24.
- Daladier, A. 1996. Le rôle des verbes supports dans un système de conjugaison nominale et l’existence d’une voix nominale en français, *Langages* 121, pp.35-53.
- Djemaa M., Candito M., Muller P., Vieu L. 2016. Corpus annotation within the French FrameNet: a domain-by-domain methodology. In *Proceedings of LREC 2016*. Portoroz, Slovenia, May 2016
- Dowty, D. 1979. *Word Meaning and Montague Grammar: the Semantics of Verbs and Times in Generative Semantics and Montague's PTQ*, Dordrecht : Reidel.
- François, J., D. Le Pesant, D. Leeman. 2007. Présentation de la classification des *Verbes Français* de Jean Dubois et Françoise Dubois-Charlier, *Langue française*, 153, pp.3-19.
- Falk I., Martin F. 2016. Automatic Identification of Aspectual Classes across Verbal Readings. In *Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics, \*SEM@ACL 2016*, Berlin, Germany, 11-12 August 2016
- Gross, M., *Méthodes en syntaxe*. 1975. Paris: Hermann.
- Garey, H. B. 1957. Verbal aspect in French, *Language* 33, pp.91-110.
- Kenny, A. 1963/[1994]. *Action, Emotion and Will*, Bristol : Thoemmes Press.
- Kupść, A., Abeillé A. Growing TreeLex. 2008. In Gelbukh, A. (ed.), *9th International Conference (CICLing)*, Haifa (Israel), February 2008, pp. 28-39 *Lecture Notes in Computational Linguistics*, 4919
- Lamiroy, B. 1987. The complementation of Aspectual Verbs in French, *language* 63-2, pp. 278-298.
- Rothstein, S. 2004. *Structuring Events: A Study in the Semantics of Lexical Aspect*, Blackwell : Oxford.
- Sagot, B., Clément, L., de La Clergerie, E., & Boullier, P., The Lefff 2 syntactic lexicon for French: architecture, acquisition, use. 2006. In *5th international conference on Language Resources and Evaluation (LREC 2006)*.
- Sagot, B., The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French. 2010. In *7th international conference on Language Resources and Evaluation (LREC 2010)*.
- Vendler, Z. 1967. *Linguistics in Philosophy*, Ithaca N.Y.: Cornell University.
- Wilmet, M. 1980. Aspect grammatical, aspect sémantique, aspect lexical : un problème de limites, in J. David & R. Martin (dir.), *La notion d'aspect*, Metz : Centre d'Analyse Syntaxique de l'Université de Metz, pp. 51-68.