

A Weighted Lexicon of French Event Names

Béatrice Arnulphy

LIMSI-CNRS & Univ. Paris Sud 11

Orsay, France

{firstname.lastname}@limsi.fr

Abstract

This paper describes a study in the purpose of annotation of event names in French texts. It presents a theoretical study about the notion of Event and defines the types of event names under study. It then presents related works about Events in NLP. Afterwards, we first use manually supervised lexicons that provide lists of nouns representing events, and demonstrate the limitations of lexicons in the task of event recognition. Further experiments are presented to propose an automatic method for building a weighted lexicon of event names.¹

1 Introduction

Information extraction consists in a surface analysis of text dedicated to a specific application. Within this general purpose, detection of event descriptions is often an important clue. However, events are, in open-domain information extraction, less studied than general named entities like location and person names. Furthermore, other fields in NLP are concerned by the recognition of events.

Verbs vs. nouns. Most events are expressed in texts by verbs and nouns. (Vendler, 1967) described the events verbal forms in a formal way while (Pustejovsky et al., 2005) used natural language processing application to process this study. Verbs are also more frequent and easier than nouns to identify and to link to other temporal information than nouns, such as temporal expressions and signals. However, (especially in newswire articles and in all languages) verbs often express less meaningful events, especially in newswire articles, whatever the language observed is: the most frequent verbs in texts are common words like *say*, *accept*, *look*. Verbs are used to talk about “common” events, while important events are frequently nominalized. For this reason, studies on events in the humanities, like sociology and particularly linguistics, mainly focus on nominal events.

¹This work has been partially founded by OSEO under the Quaero program

Name of Events. An event is what happens, it corresponds to a change of state. It can be either recurring or unique, predicted or not. It may last a moment or be instantaneous. It can also occur indifferently in the past, the present or the future.

The name given to an event can be formed either from deverbal nouns, from nouns that intrinsically denote events, or words taking their eventness in context. These constructions are detailed in Section 2. For each of those three classes, we observed that using resources is a first approach that give results we have to refine in context; context must be used to decide whether nouns or noun phrases are events.

Objectives. Existing lexicons provide lists of nouns that can be considered as events in context. Indeed, almost all nouns are highly dependent on context to assign an event characteristic. In this paper, our aim is to present the interest to use lexicons in event recognition, and their limits. We then propose a lexicon of event nouns providing quantitative information concerning the “eventness” of the words. Such a lexicon would help disambiguation of noun class in context.

This paper is organized as follows: Section 2 introduces the notion of events and presents our vision of the construction of events names. Section 3 deals with events in NLP. In Section 4, we focus on the resources that we created or used: our manually-annotated corpus (used for evaluation), as well as existing lexicons and extraction rules that we identified. Section 5 is dedicated to our experiments, leading to the automatic elaboration of a weighted lexicon, presented in Section 6.

2 Events and their Names

Events have been studied for years in several fields and in different ways. Here is an overview of works dealing with the general definition of events. We also present our observation about the

formation of their names.

Event Entity There are some definitions of the event in philosophy, history, linguistics and also in media theory. The last two are of particular interest for our own work.

In the 70's, an important reflexion was conducted about the notion of **mediatical** event. Following Davidson's ideas of 1970 about Mental Events (Davidson, 1980), these works focus on "what makes the event" and how medias create it. More recently, Neveu and Quéré (1996) presents the notion of event, as a simple occurrence, unplanned, not repeatable, happened in a recent or distant past. We disagree with this definition and we consider planned or unplanned events, such as those taking place in the past, present or future. However, there is no information about the nominalization of event descriptions.

In **linguistics**, a few researches try to deal with the problem of events in its globality. Velde (2000) refers to the general notion of "triad" I-here-now, and notices that if persons and locations are considered as proper names, then "proper names of time" should exist as well. Moreover, location names and dates can, by the mean of metonymy, take an eventive reading. It is the case for the toponym *Tchernobyl* (Lecolle, 2004) that designates the nuclear explosion which happened in the city of Tchernobyl in 1986; or for the hemeronym *September-11* (Calabrese Steimberg, 2008) which names the terrorist attack on New York in 2001. We are interested in the detection of such metonymical event names.

How are they constructed? In the humanities, studies about events in humanities usually deal with one case among others. We do not consider events in the same way. This is why according to their studies and our corpus analysis we propose a description of the lexical construction of names of events.

We organize event names into three types, according to their construction:

- Events supported by **deverbal nouns**, derived from event verbs or verb phrases by a process of nominalization. For example, the verb *to assign* is nominalized into *assignment*. In all languages, this nominalization is often ambiguous (here, *assignment* can be the act of assigning something, but also the result of this action).

- Names introduced by **nouns that intrinsically denote events**, as *festival* or *match*. Once again, a disambiguation is needed: in French, *salon* can be either a lounge or an exhibition show – *salon de l'automobile* "motor exhibition").
- Nouns or noun phrases that become **events in specific contexts**, often by metonymy, as some location names (*Tchernobyl* designates the 1986 nuclear accident) or dates (*September-11* stands for the 2001 attacks).

In the litterature, we can find clues of definitions of the event, a challenge is to deal with them in a NLP approach.

3 Events in NLP

In NLP, the definition of events seems to be quite *ad hoc* to the application they are meant to.

Events in temporal extraction. TimeML (Pustejovsky et al., 2003) is a specification language for events and temporal expressions, originally developed to improve the performance of question answering systems. In TimeML, it is considered that an event is "a cover term for situations that happen or occur". Their qualities are punctuality or duration, and can describe states. In our own work, we consider all kinds of events, proper names or not, taking place in the past, the present or the future. We do not consider states (even if they can also be nominalized) and we focus only on nominalization of events, not on verbs or predicative clauses, which are the main interest of TimeML.

Events in Named Entity studies. The task of Named Entity recognition generally focuses on classical notions of location, organisation, person or date (e.g. the MUC campaigns). Events Named Entity are rarely considered, and only in a very specific, task-oriented type definition. However, events expressed by noun phrases have many common points with "traditional" named entities; in particular, applications are nearly the same (information extraction, relationship extraction, summarization, technology watch, etc.). Nevertheless, some aspects are different, for example event phrases are more subject to variations, and they are more frequently composed of several words with an internal structure (head, modifiers and arguments).

Only few named entity evaluation campaigns considered events in their frameworks. In the event extraction project **ACE** (Automatic Content Extraction) in (2005), the classification of events is detailed and precise, but concerns only a very limited number of domains. For example, the category “life” is composed of “be-born”, “die”, “be-injured”, etc. Specific arguments are related to particular events, such as the origin and destination for transportation. The objective of ACE is to detect thematic events, and the classification, precise but incomplete, is coherent from this point of view. We do not have the same objectives. In our work, we are interested in all mentions of nouns describing events without any thematical predefined class. In the continuation of MUC (Grishman and Sundheim, 1996) and ACE, SemEval² paid interest to events, but only in semantic role labelling approach and detection of eventive verbs in Chinese news.

French **ESTER** campaigns provide a very different classification of events as named entities: the aim is to produce an open-domain named entity tagging. For this reason, event typology is quite simple: *historical and unique* events on the one hand, *repetitive* events on the other hand. This typology is quite close to our point of view on events.

Nominal Event Extraction. Only a few researches have been fully dedicated to automatic extraction of nominal events. We described here some works that follow a comparable approach as ours, where clues can be used on various languages.

Evita (Saurí et al., 2005) is an application recognizing verbal and nominal events in natural language English texts. This work was achieved in a TimeML way. Disambiguation of nouns that have both eventive and non-eventive interpretations is based on a statistical module, using a lexical lookup in WordNet and the use of a Bayesian Classifier trained on SemCor.

Also for English, following the ACE definition of events, Creswell et al. (2006) created a classifier that labels NPs as events or non-events for English. They worked on seed term lexicons from WordNet and the British National Corpus.

Eberle et al. (2009) present a tool using cues for the disambiguation of readings of German *ung-nominalizations* within their sentential context.

²<http://semeval2.fbk.eu/semeval2.ph>

Russo et al. (2011) focused on the eventive reading of deverbals in Italian, using syntagmatic and collocational cues.

Dealing with the classification of deverbals (result, event, underspecified or lexicalized nouns), Peris et al. (2010) focus on Spanish. Several lexicons, as well as automatically or manually extracted features, are evaluated in a machine learning model. Using lexicons turned out to perform under a simple baseline (which is “all instances are *result*”).

4 Resources

In this section, we introduce the resources we used or developed to carry through the study proposed in this paper: corpora, trigger lexicons, extraction rules.

4.1 Corpora

Two types of corpora (one annotated and one text-only) have been used in this study.

Manually-Annotated Corpus. We annotated a corpus of 192 French newspaper articles from *Le Monde* and *L'Est Républicain*, for a total of 48K words and 1,844 nominal events. Our corpus is as large as those in other languages in term of number of tagged nouns.³

Among our annotated corpus, 109 documents are common with FR-TimeBank, the French manually TimeML-annotated corpus (Bittar, 2010). The annotations given in FR-TimeBank and ours are different, but seem quite similar according to the good inter-annotator agreement ($\kappa=0.704$).

We wrote a quite detailed document describing annotation guidelines: it details a typology of events, as well as instructions for deciding whether a noun or a noun phrase is an event or not. In this paper, we only focus on the heads of the noun phrases. Based upon this definition, the two annotators obtained a good agreement ($\kappa=0.808$). This score proves that guidelines are well defined.

In the whole manually-annotated corpus, there are 1,844 annotated events, among them 725 different occurrences of head nouns. 269 of these eventive nouns occur only once. Among the nouns

³For a comparison purpose : 3,695 event nouns in IT-TimeBank (Russo et al., 2011), 1,579 in the English corpus from (Creswell et al., 2006), 663 in the French FR-TimeBank (Bittar, 2010). TimeBank (Pustejovsky et al., 2003) contains 7,571 events in total, but the number of nouns among them is not specified.

that appear more than once in the corpus, only 31% denote events every time they occur (100% time event: disparition “disappearance”, démission “resignation”).

Non-annotated corpus. For an experimental purpose (see below Section 6), we also used a simple text corpus of 120,246 newswire articles from *Le Monde* (two years).

4.2 Lexicons

Two existing lexicons have been used for our experiments: VerbAction (Tanguy and Hathout, 2002) and Bittar’s alternative lexicon (Bittar, 2010).

VerbAction: a Deverbal Noun Lexicon. VerbAction lexicon contains a list of French verbs of action (e.g. *fêter* “to celebrate”) together with the deverbal nouns derived from these verbs (*la fête* “the feast/celebration”). However, deverbals’ eventive reading can be ambiguous, mainly because they can also refer to the result of the action. The *VerbAction* lexicon contains 9,393 noun-verb lemma pairs and 9,200 unique nominal lemmas. It was built by manually validating a list of candidate couples automatically composed from lexicographical resources and from the Web.

The Alternative Noun Lexicon of Bittar. This lexicon contains 804 complementary event nouns⁴. These nouns are not deverbals (e.g. *anniversaire* “birthday” and *grève* “strike”). They have at least only one eventive reading, and can be ambiguous, as for deverbals: they may denote the event or the object of the process, as it is the case for *apéro* “aperitif/cocktail” and *feu* “fire”. Some of these nouns describe a state and do not match our definition of events, e.g. *absence* “non-attendance”. Lots of these nouns (like *anticoagulothérapie* “anticoagulation therapy”) belong to specific language registers. This lexicon has been used for TimeML manual annotation in French.

4.3 Extraction Rules

Beside these reflections concerning lists of nouns having an eventive reading, we achieved a study concerning several contextual clues that can be used for nominal event extraction.

Trigger Verbs: VB Rules. In (Arnulphy et al., 2010), we focused on French verbs introducing

⁴We are thankful to André Bittar for providing us this list.

event names in at least one of their arguments. The NPs related to these verbs were manually annotated by three experts, by validating or not the eventive reading of nouns in context. The study showed which verbs are meaningful for event extraction and in which configuration it would be useful to use them. Two types of verbs are considered.

The first consideration is for the verbs which explicitly introduce events, such as *avoir lieu/se tenir* “to take place”, or:

- (1) *Le sommet du G8 est organisé à Deauville.*
(The **G8 Summit** is organized in Deauville)

The second type of verbs shows a relation of cause or consequence. The point of view is that a causal action or event provokes another event. It is the case of *entraîner* “to lead to/to entail” or *provoquer* “to provoke”.

- (2) *La crise économique entraînera la famine dans de nombreux pays sous-développés.*
(The **economic crisis** will lead to **famine** in many underdeveloped countries)
- (3) *Le feu provoqué par l’attaque-suicide, n’était pas encore éteint que [...]*
(The **fire** provoked by the **suicide attack**, was not extinguish yet that [...])

In sentence 2, syntactical subject and object of *entraîner* are both events. The “famine” is the eventual consequence of the event “economic crisis”. In 3, the verb *provoque* introduces the *fire* as an event, being a consequence of the agent *suicide attack*.

According to this former study, only a few verbs were quite always meaningful for event extraction, but these ones had a good precision. For example, five verbs have an eventive subject in 90 to 100% of the cases (*avoir lieu* “to take place” or *se traduire par* “to lead to”). Others introduce an event in argument position, such as *organiser* “to organize” in more than 94% of its occurrences (cf. Table 1). We called this list of verbs *VB90*.

Temporal Indications: IT Rules. Events are anchored to time, and this is why they are often used with temporal prepositions and in temporal context.

These prepositions can indicate the occurrence of an event (*à l’occasion de* “at the time/moment of”), a referential use of the event (*avant/après*

Lemma	Translation	Rate of events
<i>Subject Position</i>		
avoir lieu	to take place	100%
se produire	to happen	94%
s'expliquer par	to be the consequence of	92%
avoir pour origine	to originate from	100%
être entraîné	to be driven by	100%
<i>Argument Position</i>		
organiser	to organize	94%
déclencher	to trigger	100%
conduire à	to lead to	93%
assister à	to attend	93%
donner lieu à	to give rise to	100%

Table 1: Examples of *VB90*, verbs that lead to an eventive reading of their subject or argument in more than 90% of the cases.

“before/after” or *la veille / le lendemain de* “the day before/after”), an internal moment of the event (*à l’issue de* “at the close of”).

However, few of these prepositions are unambiguously temporal triggers. Some like *avant*, *après*, *au commencement de* can be either temporal or locative, while *à l’occasion de* or *la veille* have only a temporal interpretation.

Using these temporal markers (Table 2) is then a good way to extract event noun phrases. We call *IT rules* the rules using them in order to extract events.

Temporal indicator	Translation
à la suite de	following (only temporal)
lors de	during
à l’occasion de	on the occasion of
au moment de	at the moment of
au lendemain de	at the day after

Table 2: Examples of temporal indicators used as event triggers.

VB and IT rules will be used in next sections to build our weighted lexicon.

5 Experiments

This set of experiments concerns the whole manually-annotated corpus.

XIP (Aït-Mokhtar et al., 2002) is a robust parser for French and English which provides depen-

dency relations and named entity recognition. Syntactic relations, as well as “classical” named entities like persons or locations, are identified, but events are not. XIP is a product from XRCE (Xerox Research Centre Europe), distributed with encrypted grammars that cannot be changed by the users. However, it is possible to add resources and grammar rules to the existing ones in order to enrich the representation. The experiments described below are performed by this means.

Considering the resources described in this paper, two distinct runs can be performed:

1. Using the French lexicons *VerbAction* and *Bittar*, described in Section 4.2, in order to evaluate the performance of a system using only these lists of nouns for event extraction.
2. Using the clues introduced in Section 4.3, *i.e.* triggers verbs (*VB rules*) and temporal indicators (*IT rules*) in order to extract events independently from lexicons.

These experiments have been evaluated in terms of precision, recall and f-measure. Precision is defined as the observed probability for a hypothesized element to be correct, recall is the observed probability for a referenced element to have been found and F-measure is the weighted harmonic mean of precision and recall.

5.1 Existing Lexicons

VerbAction and *Bittar*’s lexicons are used to annotate the corpus. Results obtained by applying these lexicons on our corpus are presented in Table 3. They show that *VerbAction* obtained a precision of 48.7%, confirming that deverbals have more non-event than event reading. The recall is 66.8%; even if the lexicon does not contain *all* deverbal nouns, it is large enough (9,200 words) and we can conclude that about one third of the events do not come from a deverbalization.

Adding the nouns from *Bittar*’s lexicon increases the recall (from 66.8% to 84.1%) without affecting the precision (48.7% to 48.3%). However 15% of events are still missed, and the precision stays quite low.

5.2 Verbs and Temporal Clues.

We automatically annotated a noun as an event if XIP indicated that this noun was subject or argument of a verb from the *VB90* list. On the

	Precision	Recall	F-measure
<i>VerbAction</i>	48.7%	66.8%	0.56
<i>VA + Bittar</i>	48.3%	84.1%	0.61

Table 3: Results with *VerbAction* and *Bittar* lexicons on the whole manually-annotated corpus.

other hand, nouns introduced by a temporal context identified by the *IT* rules (*during, the day before, etc.*) were also marked as events.

	Precision	Recall	F-measure
<i>IT</i>	81.2%	6.1%	0.11
<i>VB90</i>	84.0%	1.1%	0.02
<i>VB90 + IT</i>	81.6%	7.2%	0.13

Table 4: Results with XIP rules on the whole corpus.

As expected, and contrary to the approach exclusively based on lexicon, our extraction rules obtained a good precision and a very bad recall (Table 4). As we already mentioned, the implemented rules are focused on precise event designations.

5.3 Combination of Lexicons and Rules.

When combining lexicons and rules, recall increases of 1.8 points (from 84.1% to 85.9%), precision decreases from 48.3% to 48%.

6 A Weighted Lexicon for Event Nominals

The experiments described in the previous section show that our rules lead to a quite good precision (higher than 80%). For this reason, they can be used in order to automatically build a lexicon. As the recall is low, the rules should be applied on a large corpus. We used the non-annotated corpus presented in Section 4.1 (120,246 articles from *Le Monde*).

This method allows the extraction of a list of eventive nouns, but also, and more interestingly, it provides information about the level of ambiguity (eventive or non-eventive reading) of each word in the corpus. Otherwise, we are able to predict how eventive the word is expected to be.

6.1 Building the Lexicon

This prediction is achieved as follows: after applying the rules on the corpus, we calculate a ratio for each noun extracted as an event at least twice.

This ratio $r(w)$ is the number of occurrences of the word w that are tagged by the rules, divided by the total number of its occurrences $t(w)$, then ratio $r(w) = e(w)/t(w)$.

As the recall of the rules is low, r is obviously not a rate or a probability of the eventive reading of this word. However, a relative comparison with other ratios allows us to estimate how ambiguous the noun is in a given corpus. This value is then interesting for noun classification. This interest is illustrated by examples given in Table 5.

Potential triggers		Nb. detected / total occurrences	Ratio
French	English		
chute	fall	434 / 2620	0.166
clôture	closing	63 / 470	0.134
élection	election	1243 / 9713	0.128
bousculade	jostle	12 / 115	0.104
crise	crisis	286 / 6185	0.046
tension	tension	16 / 1595	0.001
subvention	subvention	2 / 867	0.002
Anschluss	Anschluss	3 / 4	0.750
méchoui	mechoui	3 / 5	0.600
krach	krach	20 / 169	0.118
RTT	~ day off	14 / 166	0.084
demi-finale	semifinal	35 / 553	0.063
cessez-le-feu	cease-fire	15 / 440	0.034
accès	access	9 / 2828	0.003
11 septembre	September-11	12 / 4354	0.003

Table 5: Examples of trigger words extracted by the extraction rules.

Many of these words can be found in lexicons *VerbAction* or *Bittar* (first part of the list), while others are not (second part). Nouns that are non-ambiguous in their eventive reading have a quite high ratio (higher than the average recall described in previous section). It is the case of *fall*, *election* or *krach*. On the other hand, highly ambiguous words like *tension*, *subvention* or *access* get a low ratio. The date *September 11* is also in this latter case, but dates are very rare in these results, and this one has by far the best rate. The French *clôture*, that can be translated as *fencing* or *closing*, seems almost not ambiguous in newswire articles.

These rules helped us to discover 305 new names of events that were not present in the trigger lexicons, such as those shown in Table 5, but also *tollé* “hue and cry”, *mise en sourdine* “soften” or *couac* “false note”.

6.2 Evaluation

We evaluated our weighted lexicon by comparing its performances in event extraction with the two standard lexicons.

Direct Application of Lexicon In a first evaluation experiment, we applied this new weighted lexicon on our annotated corpus (see Section 4.1), as done for VerbAction and Bittar in Section 5. To observe the evolution of performances, we tested different “slices” of the lexicon, according to the ratios obtained: all words with a ratio higher than 10%, then all those with a ratio greater than 8%, 6%, etc. The results are presented in the Table 6.

Words of ratio >	Precision	Recall	F-measure
10%	84.1%	16.6%	0.28
8%	83.6%	24.3%	0.38
6%	79.8%	31.5%	0.45
1%	56.3%	71.0%	0.63
0.5%	43.4%	80.1%	0.56

Table 6: Results when applying “slices” of ratio on the corpus.

Precision and recall evolve in an opposite way: when the lexicon is less selective, the recall increases and the precision decreases. The best F-measure (for 1% ratio) is 0.63, a value similar to the F-measure of the VerbAction and Bittar’s lexicons combined (0.61).

Machine-Learning Evaluation As a second evaluation of the automatically-built weighted lexicon, we added the word ratios as a feature in the rule-based classifier J48, an implementation of C4.5 algorithm (Quinlan, 1993), as implemented in the software Weka (Hall et al., 2009).

Our corpus has been splitted into a test set (49 documents) and development set (143 documents, which is small but sufficient for our study). We implemented three very basic models, allowing us to show the trade-off introduced by the ratios, without any suspicion of side effect due to other features:

- M_l uses only the two standard lexicons VerbAction and Bittar.
- M_r uses only the ratios, as a real value.
- M_{rl} uses both existing and ratio lexicons. When a word is not in the ratio lexicon, this word is given ratio 1 if in standard lexicons and 0 otherwise.

Results are given at Table 7. Using only our automatic ratio lexicon M_r leads to similar results

than using standard, manually validated lexicons M_l . Combining all information leads to a small improvement of precision and recall.

	M_l	M_r	M_{rl}
Precision	0.51	0.49	0.54
Recall	0.86	0.89	0.89
F-measure	0.64	0.63	0.67

Table 7: Comparison of models using standard lexicons (M_l), ratio lexicon (M_r) and both (M_{rl}).

Discussion Those results show that our weighted lexicon, automatically built and without any manual validation, leads to obtain comparable results than manually-validated lexicons. Combining all of them improves both precision and recall.

Creating this weighted lexicon requires the implementation of language-specific rules, but these rules seem quite easy to adapt to another language, provided that a syntactic parser exists for this language. Building such a lexicon is then much less time-consuming than validating an entire lexicon.

Moreover, if applied on much larger and diverse corpora, this method should make possible the detection of more metonymic events, as *September-11*, in order to build a knowledge base of event candidates.

7 Conclusion

We presented in this article several experiments aiming at studying the use of event names in French texts, as well as an automatic method for building a weighted lexicon of event names.

We first defined the object we are dealing with: what is, from our point of view, an event. Then, we noticed that the existing lexicons which can be used in an event extraction perspective are not sufficient for a wide coverage. Some words are not intrinsically events, but take their eventness in the context. These words cannot be found in such lexicons.

We applied our rules based on verbs and temporal clues, which provide events in more than 80% of cases, that allows us to construct a new weighted lexicon. Our experiments show that the lexicon is as precise as manually-validated lists, and that weights can be used to improve the classification of nouns.

Because some words take an eventual meaning

at a given moment (eg. *le nuage islandais* – literally “Icelandic cloud” – refers to the blast of the Eyjafjöll volcano from march to october 2010), we are now working on a new lexicon which would consider the date of the apparition of an event name.

We are also working on a weighted lexicon in English.

References

2005. ACE (Automatic Content Extraction) - English Annotation Guidelines for Events, Version 5.4.3 2005.07.01. Technical report, Linguistic Data Consortium.
- Salah Aït-Mokhtar, Jean-Pierre Chanod, and Claude Roux. 2002. Robustness beyond Shallowness: Incremental Deep Parsing. *Natural Language Engineering*, 8.
- Béatrice Arnulphy, Xavier Tannier, and Anne Vilnat. 2010. Les entités nommées événement et les verbes de cause-conséquence. In *Actes de TALN 2010*.
- André Bittar. 2010. *Building a TimeBank for French: A Reference Corpus Annotated According to the ISO-TimeML Standard*. Ph.D. thesis.
- Laura Calabrese Steimberg. 2008. Les héméronymes. Ces évènements qui font date, ces dates qui deviennent évènements. *Mots. Les langages du politique*, 3.
- Cassandra Creswell, Matthew J. Beal, John Chen, Thomas L. Cornell, Lars Nilsson, and Rohini K. Srihari. 2006. Automatically extracting nominal mentions of events with a bootstrapped probabilistic classifier. In *Proceedings of the COLING-ACL '06*.
- Donald Davidson, 1980. *Essays on Actions and Events*, chapter 11 "Mental Events" (1970). Psychology as Philosophy. Calendron Press.
- Kurt Eberle, Gertrud Faaß, and Ulrich Heid. 2009. Corpus-based identification and disambiguation of reading indicators for German nominalizations. In *Proceedings of Corpus Linguistics 2009*.
- Ralph Grishman and Beth Sundheim. 1996. Message Understanding Conference: A Brief History. In *Proceedings of COLING 1996*.
- Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*.
- Michelle Lecolle. 2004. Toponymes en jeu : Diversité et mixage des emplois métonymiques de toponymes. In *Studii si cercetari filologice 3 / 2004. Universit  de Pitesti, Roumanie*.
- Erik Neveu and Louis Qu er . 1996. Pr sentation. *R seaux*, 14(75).
- Aina Peris, Mariona Taul , Gemma Boleda, and Horacio Rodriguez. 2010. ADN-classifier: Automatically assigning denotation types to nominalizations. In *Proceedings of LREC'2010*.
- James Pustejovsky, Jos  Casta o, Robert Ingria, Roser Saur , Robert Gaizauskas, Andrea Setzer, and Graham Katz. 2003. TimeML: Robust Specification of Event and Temporal Expressions in Text. In *Proceeding of IWCS-5*.
- James Pustejovsky, Robert Knippen, Jessica Littman, and Roser Saur . 2005. Temporal and Event Information in Natural Language Text. *Language Resources and Evaluation*, 39(2-3).
- R. Quinlan. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufman Publishers.
- Irene Russo, Tommaso Caselli, and Francesco Rubino. 2011. Recognizing deverbal events in context. In *Proceedings of CICLING 2011, poster session*. Springer.
- Roser Saur , Robert Knippen, Marc Verhagen, and James Pustejovsky. 2005. Evita: A Robust Event Recognizer for QA Systems. In *Proceedings of HLT/EMNLP*.
- Ludovic Tanguy and Nabil Hathout. 2002. Webafix : un outil d'acquisition morphologique d rivationnelle   partir du Web. In *Actes de TALN 2002*.
- Dani le Van De Velde. 2000. Existe-t-il des noms propres de temps ? *Lexique*, 15.
- Zeno Vendler, 1967. *Facts and events*, chapter Verbs and Times, pages 97–121. Cornell University Press.