Automatically Generated Noun Lexicons for Event Extraction

Béatrice Arnulphy, Xavier Tannier, Anne Vilnat

LIMSI-CNRS, Univ. Paris-Sud
91403 Orsay, France
firstname.surname@limsi.fr

Abstract. In this paper, we propose a method for creating automatically weighted lexicons of event names. Almost all names of events are ambiguous in context (*i.e.*, they can be interpreted in an eventive or non-eventive reading). Therefore, weights representing the relative "eventiveness" of a noun can help for disambiguating event detection in texts.
We applied our method on both French and English corpora. Our method has been applied to both French and English corpora. We performed an evaluation based upon a machine-learning approach that shows that using weighted lexicons can be a good way to improve event extraction. We also propose a study concerning the necessary size of corpus to be used for creating a valuable lexicon.

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 (*e.g.,* temporal ordering of events on a chronological axis). However, events are, in open-domain information extraction, less studied than general named entities like location and person names.

We focused our study on nominal forms of events[1]. Lexicons provide lists of nouns that can be considered as events in context. These lexicons only contain common nouns, but the events are not only named with common nouns or with words that are in the existing lexicons. Indeed, almost all nouns are highly dependent on context to assign those nouns an event property. In this paper, we propose a method using patterns and shallow parsing to automatically build a lexicon for nouns event extraction. We apply this method on two languages (French and English). Our work is close to Bel et. al [5], which present cues for the disambiguation of non-deverbal event nouns. Contrary to Bel et al. [5], our lexicon provides quantitative information concerning the "eventiveness" of the words. Such a lexicon would help disambiguation of noun class in context.

First, we present our observations about the way we name events and we propose a brief survey of works dealing with nominal forms of events. Then we

[1] This work has been partially funded by OSEO under the Quaero program, as well as French National Research Agency (ANR) under project Chronolines (ANR- 10-CORD-010)

present the resources we used in our study, before introducing our method for the automatic creation of the weighted lexicons in order to extract names of events. To conclude, we evaluate the performances of our weighted lexicons in comparison with other classical lexicons, based on annotated corpora.

2 The Event

From our point of view, 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 occur in the past, the present or the future.

2.1 Construction of Event Names

In the Humanities, studies about events usually deal with single events or only few events (*e.g.*, *Jasmin Revolution* or *H1N1 flu* [8]), and do not offer generalization hints. We do not consider events in the same way. According to those studies and based upon our corpus analysis, we propose a description of the lexical construction of names of events, organized into three types, according to their construction.

1. **Nominalization related to an action verb.** A lot of event names are formed from words morphologically related to action verbs. They can be supported by deverbal nouns, nouns derived from action or event verbs by a process of nominalization. For example:
   - the verb *fêter* ("to celebrate") is morphologically linked to *fête* ("party", "celebration"): *la fête de la musique* ("the music festival").
   - the verb *to assign* is nominalized into *assignment*.

   In all languages, this nominalization is often ambiguous (nominalization could refer, either to the process or to the result of the process, the location or the object). Here, *assignment* can be the act of assigning something, as well as the result of this action.
2. **Nominalization not related to verbs.** Some names of events are introduced by nouns that intrinsically denote events, such as *festival* or *match*. Then a disambiguation is needed: in French, *salon* can be either a lounge or an exhibition show (e.g., *salon de l'automobile* — "motor exhibition").
3. **Metonymic nominalization.** Other nouns or noun phrases become name of events in specific contexts, often by metonymy: a location name (*Tchernobyl* refers to the 1986 nuclear accident that occurred in this town [18]) or a date (*September-11* stands for the 2001 attacks [7]).

For each of those three classes, we could use resources is a first approach that must be refined in context; context must be used to decide whether nouns or noun phrases are events.

2.2 Event Nouns in NLP

In NLP, the definition of events seems to be quite *ad hoc* to the application they are meant to describe. We will focus here on works dealing with events nouns in temporal extraction project and those more specifically oriented towards nominal event extraction.

Events in Temporal Extraction. *TimeML* [21] is a specification language for events and temporal expressions, originally developed to improve the performance of question answering systems. In TimeML, an event is defined as "a cover term for situations that happen or occur". An event is based upon punctuality or duration properties and it can describe states. The TimeML specification language is used for the annotation of numerous corpora in several languages.

In our work, we consider all kinds of events, being proper names or not, taking place in the past, the present or the future. We do not consider states (even if they can be nominalized) and we focus on events based upon a nominalization, not on verbs or predicative clauses, which are the main interest of TimeML.

We must also pay attention to the few Named Entity Recognition campaigns which considered events in their frameworks. Automatic Content Extraction (**ACE**) [11] proposed an event extraction project [1] in which the classification of events is detailed (arguments are related to particular events) and precise, but it only concerns a very limited number of domains (the category "life" is composed of "be-born", "be-injured" sub-domains, etc.). The objective of ACE is to detect thematic events. We are interested in all mentions of nouns describing events without any thematical predefined class. In the continuation of MUC [15] and ACE, SemEval[2] paid interest to events within the framework of a semantic role labelling approach and detection of eventive verbs in Chinese news. French **ESTER** campaigns [14] provide a very different classification of events as named entities: the aim is to produce an open-domain named entity tagging. For this purpose, event typology is quite simple: *historical and unique* events on the one hand, *repetitive* events on the other hand. Even if this typology is not detailed, it corresponds to our point of view on events.[3]

Nominal Event Extraction. Little research has been fully dedicated to automatic extraction of nominal events. We described here some works that follow a comparable approach to ours, using lexicons and linguistic classed-based information. Evita [25] is an application recognizing verbal and nominal events in English texts. This work is based upon the TimeML definition. Disambiguation of nouns that have both eventive and non-eventive interpretations is based on a statistical module, using a lexical lookup in WordNet[4] and the use of a Bayesian Classifier trained on SemCor. Also for English, following the ACE definition of events, Creswell et al. [10] created a classifier that labels NPs as events or

[2] `http://semeval2.fbk.eu/semeval2.php`

[3] In our works, we developped a more detailed typology which takes into account modality (factual, abstract, etc.), frequency (unique, recurring, instanciation), and temporality of the event.

[4] `http://wordnet.princeton.edu/`

non-events. They worked on seed term lexicons from WordNet and the British National Corpus.[5] Eberle et al. [12] present a tool using cues for the disambiguation of readings of German *ung*-nominalizations within their sentential context. Russo et al. [24] focused on the eventive reading of deverbals in Italian, using syntagmatic and collocational cues. In a close approach, Resnik and Bel [23] worked on Spanish and Bel et al. [5] on Spanish and English. They tried to disambiguate result and event, as well as deverbal nouns and non deverbal nouns. In a machine-learning approach, they used cues which are assumed in the linguistic literature (aspectual verbs and prepositions, temporal quantifying expressions, etc.). Dealing with the classification of deverbals (result, event, underspecified or lexicalized nouns), Peris et al. [19] focused on Spanish. Several lexicons, as well as automatically or manually extracted features, are evaluated in a machine learning model.

3 Resources

In our study, we use several resources : corpora and existing lexicons. We worked with raw corpora for the lexicon extraction, manually annotated corpora for the evaluation, both type of corpus in French and English. Here is an overview of these resources for English and French.

3.1 Corpora

For the Lexicon Extraction. For the creation of our weighted lexicons in French and English, we used a corpus of newswires from the French News Agency *AFP*[6]. The *AFP* corpus is available on a same period in two languages, so we could have similar corpus. The English corpus is composed of 1.3 millions texts over the 2004-2011 period (120 million tokens). The French corpus is of 1 million texts over 2005-2011. In French, we also used a corpus of 120,246 newspaper articles from *Le Monde* (two years, other 2001-2002, 61 million tokens): this corpus is of similar size to the French *AFP* corpus ; these two corpora are also similar according to the realities they deal with, even if they are evoked differently (newspaper articles and short news). We thus created a weighted lexicon from this corpus in order to complete our French weighted lexicon.

For the Evaluation.

The two TimeML annotated corpora we used are based on newswires (cf. 2.2). In English, *TimeBank 1.2* [20] contains 1,722 non-stative nominal events. The annotated texts are extracted from news media (*Wall Street Journal*, *ABC*, *CNN*, *Voice Of America*) over the 1989-1998 period. In French, *FR-TimeBank* [6] contains 663 nominal mentions of event. The annotated texts come from the newspaper *L'Est Républicain* over the period 1999-2003.

[5] http://www.natcorp.ox.ac.uk/

[6] We thank the French News Agency (AFP) for providing us with the corpus.

Our French Manually Annotated Corpus is composed of 192 French newspaper articles from *Le Monde* and *L'Est Républicain* for a total amount of 48 thousand words. Our corpus contains 1,844 events, which is comparable to TimeBank 1.2, FR-TimeBank, as well as the Italian IT-TimeBank [24] (3,695 event nouns) and the English corpus from [10] (1,579). We defined and followed precise annotation guidelines: they detail a typology of events, as well as instructions for deciding whether a noun or a noun phrase is an event or not. Among these instructions:

- Try to imagine some non-ambiguous valuable substitutes for the noun. This proves to be very effective.
- Take inspiration from examples of eventive and not eventive uses of the same word, that can be found in dictionaries, together with their proper definition.
- Remember that enumeration items are often (not always) of the same class.
- When decision is impossible, choose to annotate as non-event.

Delimiting the event boundaries is also a difficult issue and the guidelines provide instructions for this other problem. Following the guidelines, the two annotators (the authors of the guidelines) obtained a good agreement for the annotation of the heads of noun phrases (kappa[7]=0.808). Among the corpus, the 109 documents from *L'Est Républicain* are common with FR-TimeBank [6]. The two annotations have a different purpose, but seem quite similar according to the good inter-annotator agreement (kappa=0.704).

3.2 Lexicons

In French, two lexicons can be useful to find nominal mentions of events: VerbAction [26] and Bittar's alternative lexicon [6]. In English, we used nouns of events and actions from WordNet [13].

VerbAction is a deverbal noun lexicon. It 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 contains 804 complementary event nouns.[8] 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* ("apéritif/cocktail") and *feu* ("fire"). Some of these nouns describe a state and do not match our definition of events (*e.g.*,

[7] We used the Carletta's Kappa coefficient [9]. This measure compares the agreement against what might be expected by chance. According to Landis and Koch [17], from 0.6 to 0.8 is what we consider a good agreement. Up to 0.8 is a very good agreement.

[8] We are thankful to André Bittar for providing us this list.

absence — "non-attendance"). Lots of these nouns (like *anticoagulothérapie* — "anticoagulation therapy") belong to language of speciality, such as the medical one. This lexicon has been used for TimeML manual annotation in French.

The Action and Eventive Nouns in WordNet contains 5903 nouns tagged as "act" (for action) or events . This list of English words can be considered as comparable with the French lexicons (VerbAction and Bittar). It contains words describing events in almost all cases (*war*, *election*, *show*, *carnival*), expressions which are very ambiguous (*arts and crafts*, *bet*, *coloration*), multi-word expressions (*a cappella singing*), name of events (*Arab-Israeli War*, *Battle of Britain*, *laser trabecular surgery*), but also expressions that do not seem to fit with any event definition (*Attorney General*, *judo*, *industry*).

4 Automatic Lexicon Creation

We showed in a previous work [3] that a lexicon of event nominals can be created by applying extraction rules. These experiments demonstrated that the French automatically generated lexicon (created from *Le Monde*) is as precise as manually-validated lists, and weights can be used to improve the classification of nouns. This work was only conducted on French. In the present study, we extend these experiments to English and evaluate the process. We also generated a new lexicon for French from the *AFP* corpus in order to obtain comparable multilingual results. From the corpora of AFP news, we extracted two lexicons of nouns describing events according to our extraction rules, the first one in French and the second one in English. Our extraction rules depend on the use of a syntactic parser. For this purpose, we chose a robust parser, XIP.

XIP [2] is a robust parser for French and English which provides dependency relations and "classical" named entities (like persons or locations). But events are not identified. 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 in order to enrich the representation. It is what we have done.The parser is language-dependent, but the extraction rules are commutable to other languages with a minimal cost. We developped the same type of rules for French and English. We performed a corpus analysis to evaluate the meaningful of those rules for event extraction.

4.1 Extraction Rules

Temporal Rules. Because events are anchored to time, they are often linked to temporal prepositions and used in temporal context. Using these temporal markers is a good way to extract event noun phrases. In this way, we focused on the more unambiguous prepositions. These prepositions or trigger-words show:

	(FR)	(EN)
the occurrence of an event:	*à l'occasion de*	*at the time/moment of*
	au moment de	*on the occasion of*
a referential use of the event:	*avant/après*	*the morning of*
	le lendemain de	*the day before*
	au matin de	*at the morning of*
	à la suite de	*following (temporal)*
	lors de	*during*
an internal moment of the event:	*à l'issue de*	*the beginning of*

However, few of these triggers are unambiguously temporal triggers. Some like *avant* ("before"), *après* ("after"), *au commencement de* ("at the beginning") can be either temporal or locative, while *à l'occasion de* ("when") or *la veille* ("the day before") have only a temporal interpretation.

Verbal Rules. A previous study on French [4] shows which verbs are the most meaningful for event extraction and in which configuration (subject and/or object) it would be greatful to use them. We took this information into account in the following rules:

	(FR)	(EN)
in a subject position:	*avoir lieu*, *se tenir*	*to take place*,*to come about*
in an argument position:	*entraîner*	*to be the result of*

We focus on three types of verbs. The first type concerns verbs which explicitly introduce events (occurrence predicates):
(FR) *se produire*,*avoir lieu*
(EN) *to befall*, *to occur*

Le ***sommet du G8*** *est organisé* à Deauville.
The ***G8 Summit*** *is organized* in Deauville.

The second type of verbs introduce a relation of cause and/or effect for events. Indeed as we can see in the following examples, a causal action or event provokes another event.
(FR) *occasionner*
(EN) *to ensure*

La ***crise économique*** *entraînera* la ***famine*** dans les pays sous-développés.
The ***economic crisis*** *will lead to* ***famine*** in underdeveloped countries.
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 [...]

And the last one is for verbs which present a moment of an event (aspectual predicates):
(EN) *to begin*, *to last*

'The Event' will *end* like all successfull US TV shows.
Let the ***spectacle*** *begin*.

We used verbs which are quite always meaningful for event extraction, according to the observation from a corpus analysis. The verbs we selected introduce events in more than 90% of the cases.

4.2 Calculating the Eventiveness Relative Weight

The extraction rules based on contextual clues gives precise results (P>0.80) but a low recall (R<0.10). Therefore, to be representative, the lexicon has to be extracted from a large corpus (see Section 5.3). The application of the extraction rules allows the extraction of a list of eventive nouns. From this list and our corpus, we can fetch 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. This prediction is achieved by computing the Eventiveness Relative Weight (ERW): after applying the rules on the corpus, we calculate a weight for each noun extracted as an event at least twice. $ERW(w)$ is the number of occurrences $e(w)$ of the word w tagged by the rules, divided by the total number of its occurrences $t(w)$:

$$ERW(w) = \frac{e(w)}{t(w)} \tag{1}$$

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

Potential triggers *French*	Translation	Nb. detected / total occ	**ERW**	Potential triggers *English*	Nb. detected / total occ	**ERW**
chute	fall	434 / 2620	**0.166**	overthrow	383 / 448	**0.855**
clôture	closing	63 / 470	**0.134**	intifada	7 / 11	**0.636**
élection	election	1243 / 9713	**0.128**	bombardement	6 / 12	**0.500**
bousculade	jostle	12/115	**0.104**	testimony	426 / 13109	**0.032**
crise	crisis	286 / 6185	**0.046**	sleepover	3 / 27	**0.111**
tension	tension	16 / 1595	**0.001**	publication	154 / 9337	**0.016**
subvention	subvention	2 / 867	**0.002**	marathon	52 / 8070	**0.006**
Anschluss	Anschluss	3 / 4	**0.750**	play-off	73 / 75	**0.973**
méchoui	mechoui	3 / 5	**0.600**	breastfeeding	3 / 4	**0.750**
krach	krach	20 / 169	**0.118**	overheat	3 / 7	**0.428**
RTT	∼ day off	14 / 166	**0.084**	stopover	372 / 1345	**0.276**
demi-finale	semifinal	35 / 553	**0.063**	cross-examination	53 / 416	**0.127**
cessez-le-feu	cease-fire	15 / 440	**0.034**	distillery	4 / 126	**0.032**
accès	access	9 / 2828	**0.003**	welcome	66 / 3884	**0.017**
11 septembre	September-11	12 / 4354	**0.003**	influenza	37 / 6019	**0.006**

Table 1. Examples of trigger words extracted by the extraction rules

This interest is illustrated by examples given in Table 1: the upper part of the tables presents words which are found in the English or French standard lexicons while the lower part presents words fetched by the extraction rules which are not in the standard lexicons. We created three weighted lexicons: one based on the two years *Le Monde* French corpus, and two from the whole *AFP* corpora (one in English and one in French). The lemmas present in the weighted lexicons must be extracted by our rules at least twice. See Table 2.

Corpus used for the lexicon creation	Number of tokens total size	extracted	differents	Number of lemmas in the weighted lexicon
(FR) AFP (2005-2011)		166,077	8,053	3,538
(EN) AFP (2004-2011)	120,091,099	543,394	14,619	3,452
(FR) LM (2001-2002)	61,920,573	19,767	4,843	1,559

Table 2. From corpora to weighted lexicons: Size in number of tokens

5 A Machine-Learning Evaluation

We applied the French and English automatically-built weighted lexicons using a machine-learning approach and conducted an evaluation. We added the *ERW* value as a feature in the rule-based classifier J48, an implementation of C4.5 algorithm [22], as implemented in the software Weka [16]. The manually annotated corpus was split into a training set (75% of the annotated corpus) and a test set (the remaining 25% of the annotated corpus). The training set contains the same number of event entries than non event entries (see Table 3).

	Training Set			*Test Set*		
	total	YES	NO	total	YES	NO
English	2,182	1,092	1,092	3,246	453	2,793
French	5,226	1,263	1,263	2,700	566	2,134

Table 3. Number of tokens in the training and test corpus

For each language, we implemented three very basic models, allowing us to show the trade-off introduced by the *ERW*, without any suspicion of side effect due to other features:

- M_l uses only the standard manually validated lexicons:
  (FR) VerbAction and Bittar (EN) WordNet action and event nouns
- M_r uses only the *ERW*, as a real value. As we have two weighted lexicons in French, they are called:
  - M_r^{LM}, based on an extraction of the lexicon from two years of *Le Monde* corpus.
  - M_r^{AFP}, our new weighted lexicon based on the *AFP* corpus.
- M_{rl} uses both existing and weighted lexicons.

Our models are evaluated using the classical measures of precision (P), recall (R) and F-measure (F1)[9]

[9] 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 *ERW* lexicons vs. Standard Lexicons Comparison

Table 5.1 presents the evaluation of the French *LM* and *AFP* weighted lexicons in comparison to standard lexicons (Bittar's and VerbAction lexicons) on our annotated corpus. Table 4 presents the evaluation of the English AFP weighted lexicon in comparison to the standard lexicon extracted from WordNet on the TimeBank 1.2 corpus.

	Our EWR lexicons		*Standard*	*Mixed*	
	M_r^{LM}	M_r^{AFP}	M_l	M_{lr}^{LM}	M_{lr}^{AFP}
P	0.49	0.55	0.53	0.54	**0.60**
R	0.89	0.77	0.88	0.89	0.84
F1	**0.63**	**0.64**	**0.66**	**0.67**	**0.70**

	Our EWR lexicons M_r^{AFP}	*Standard* M_l	*Mixed* M_{lr}^{AFP}
P	0.36	0.30	**0.36**
R	0.71	0.64	0.77
F1	0.476	0.414	**0.493**

Table 4. Evaluation of the weighted lexicon in French (left) and in English (right)

First of all, in both French and English, we notice that:

- Using only our weighted lexicons (M_r) leads to similar results than using standard manually validated lexicons (M_l).
- Combining all information leads to a small but substantial improvement of precision and recall.

From these observations, we confirm that our automatically created weighted lexicons are as precise as the standard manually validated lexicons in French and in English. In French, we also notice that the weighted lexicon from *AFP* corpus is more precise than both the standard lexicon (P=0.53) and that the *LM* one (P=0.49). Besides, as a point of comparison, we applied the M_r^{AFP} model on the FR-TimeBank and our annotated corpus. The performances of the *AFP* weighted lexicon are similar on the two annotated corpora, even if the corpora were not annotated with the same aim or guidelines. Precision reaches 0.56 on FR-TimeBank and 0.55 on our annotations, recall is of 0.77 on both corpora and F1 is 0.648 and 0.642. Moreover, we observe that results for English are much lower than results for French. However, this difference is not due to the lexicons quality. Indeed, the trade-off between standard lexicons (VerbAction + Bittar in French, WordNet in English) and our ratio lexicon is similar. This means that their quality are similar as well. Our initial guess that a direct translation of French rules was enough is then confirmed. The fact that lexicons perform so poorly in English rather tends to prove that the problem is just more difficult in English. Studying this difference is one of our prespectives.

5.2 ML-Evaluation vs. Threshold Based Approach Comparison

As a comparison to the ML-Evaluation and in order to observe the evolution of performances, we tested different "slices" of the lexicon in a threshold based approach. According to the value of the *ERW*: all words with an *ERW* higher than 10%, then all those with an *ERW* greater than 8%, 6%, etc. The results are presented in the Table 5. 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% ERW) is 0.63, a value similar to the F-measure of the VerbAction and Bittar's lexicons combined (0.61).

Words of $ERW >$	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 5. Results when applying "slices" of ERW on the corpus (French LM lexicon).

5.3 Impact of the Size of the Corpus

As the precision of our extraction rules is good and the recall is low, we stated that a large corpus was necessary. But how large must the corpus used for the lexicon extraction corpus be? We created several weighted lexicons from parts of our corpus, from one month to one year of news. We studied the performances of M_r^{AFP} models depending of the size of the corpus it was based on (cf. Table 6).

Lexicon created on		1 month 07 2005	6 months 07-12 2005	1 year 2005	all 2004-2011
French	P	0.665	0.539	0.512	0.56
	R	0.303	0.628	0.692	0.77
	F1	**0.416**	**0.58**	**0.588**	**0.648**
English	P	0.36	0.31	0.35	0.36
	R	0.35	0.7	0.76	0.71
	F1	**0.36**	**0.43**	**0.48**	**0.48**

Table 6. Evaluation of the weighted lexicons depending of the size of the corpus

Figure 1 shows that, in English and in French, the gain in terms of F-measure of a model trained on a one-year-learned lexicon is as good as for a whole-corpus-learned lexicon. The figures and the shape of the curves seem to show that more corpora would not increase significantly the performances.

However, even if global performances are not improved by adding more and more documents, it is still interesting to extract names of event in a much longer period or during a specific period of time. Indeed, events and their names are anchored to time, and very particular event names will be used only at a precise moment (*e.g. tsunami*, *Arab Spring*).

6 Conclusion

We automatically created lexicons of eventive nouns in French and English by using rules based on verbs and temporal clues. A relative weight of eventiveness

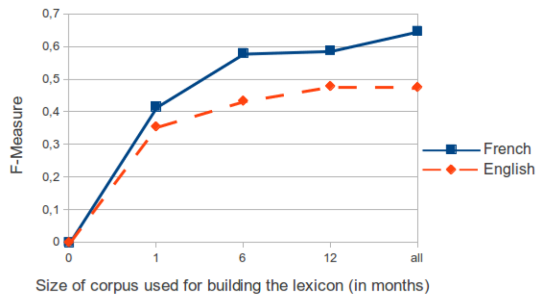

Fig. 1. Progression of the F-measure depending of the size of the corpus

(*ERW*) is added to the lexicon. The *ERW* a great information in order to help for the disambiguation of the words. In a machine-learning evaluation, we showed that our automatically generated weighted lexicons are competitive to the lexicons which were manually created. These experiments also prove that the transposition of the rules from a language to another one is possible. As well, we observed that a one-year corpus is significant enough to build a lexicon with our method and to obtain comparable result as those of classical lexicons.
According to our experiments on French, we conclude that the performance of the weighted lexicon is dependent on the corpus chosen to generate the lexicon. It would be interesting to apply our method on other domains. In English, as the result with the lexicon from WordNet is low, we plan to study this difference. However, because some words take an eventive meaning at a given moment (*e.g.*, *le nuage islandais* (literally "Icelandic cloud") refers to the blast of the Eyjafjöll volcano from March to October 2010), we would like to work on a new lexicon which would consider the date of the appearance of an event name.

References

1. ACE (Automatic Content Extraction) - English Annotation Guidelines for Events, V 5.4.3 2005.07.01. Tech. rep., LDC (2005)
2. Aït-Mokhtar, S., Chanod, J.P., Roux, C.: Robustness beyond Shallowness: Incremental Deep Parsing. Natural Language Engineering 8 (2002)
3. Arnulphy, B.: A weighted Lexicon of French Names. In: Proc. of RANLP Student Workshop (2011)
4. Arnulphy, B., Tannier, X., Vilnat, A.: Les entités nommées événement et les verbes de cause-conséquence. In: Actes de TALN (2010)
5. Bel, N., Coll, M., Resnik, G.: Automatic Detection of Non-deverbal Event Nouns for Quick Lexicon Production. In: Proc. of COLING (2010)
6. Bittar, A.: Building a TimeBank for French: A Reference Corpus Annotated According to the ISO-TimeML Standard. Ph.D. thesis, Univ. Paris Diderot (2010)
7. Calabrese Steimberg, L.: Les héméronymes. Ces évènements qui font date, ces dates qui deviennent évènements. Mots. Les langages du politique 3 (2008)
8. Calabrese Steimberg, L.: La nomination d'événements dans le discours d'information : entre activité collective et déférence épistémiologique. In: Colloque Langage, discours, événements (2011)

9. Carletta, J.: Assessing Agreement on Classification Tasks: the Kappa Statistic. Computational Linguistics 22 (1996)
10. Creswell, C., Beal, M.J., Chen, J., Cornell, T.L., Nilsson, L., Srihari, R.K.: Automatically Extracting Nominal Mentions of Events with a Bootstrapped Probabilistic Classifier. In: Proc. of the COLING/ACL (2006)
11. Doddington, G., Mitchell, A., Przybocki, M., Ramshaw, L., Strassel, S., Weischedel, R.: The Automatic Content Extraction (ACE) Program - Tasks, Data, and Evaluation. In: Proc. of LREC (2004)
12. Eberle, K., Faaß, G., Heid, U.: Corpus-based identification and disambiguation of reading indicators for German nominalizations. In: Proc. of Corpus Linguistics (2009)
13. Fellbaum, C.: WordNet: An Electronic Lexical Database. Bradford Books (1998)
14. Gravier, G., Bonastre, J.F., Geoffrois, E., Galliano, S., McTait, K., Choukri, K.: Ester, une campagne d'évaluation des systèmes d'indexation automatique d'émissions radiophoniques en français. In: Proc. of JEP (2004)
15. Grishman, R., Sundheim, B.: Message Understanding Conference: A Brief History. In: Proc. of COLING (1996)
16. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA Data Mining Software: An Update. SIGKDD Explorations 11(1) (2009)
17. Landis, J.R., Koch, G.G.: The Measurement of Observer Agreement for Categorical Data. Biometrics 33 (1977)
18. Lecolle, M.: Éléments pour la caractérisation des toponymes en emploi événementiel. In: Evrard, I., Pierrard, M., Rosier, L., Raemdonck, D.V. (eds.) Les sens en marge - Représentations linguistiques et observables discursifs : actes du colloque international de Bruxelles, 3-5 novembre 2005. L'Harmattan (2009)
19. Peris, A., Taulé, M., Boleda, G., Rodriguez, H.: ADN-classifier: Automatically assigning denotation types to nominalizations. In: Proc. of LREC (2010)
20. Pustejovsky, J., Verhagen, M., Saurí, R., Littman, J., Gaizauskas, R., Katz, G., Mani, I., Knippen, R., Setzer, A.: TimeBank 1.2. LDC (2006)
21. Pustejovsky, J., Castaño, J., Ingria, R., Saurí, R., Gaizauskas, R., Setzer, A., Katz, G.: TimeML: Robust Specification of Event and Temporal Expressions in Text. In: IWCS-5, Fifth International Workshop on Computational Semantics. (2003)
22. Quinlan, R.: C4.5: Programs for Machine Learning. Morgan Kaufman Publishers (1993)
23. Resnik, G., Bel, N.: Automatic detection of non-deverbal event nouns in spanish. In: di Linguistica Computazionale, I. (ed.) Proc. of the 5th Int. Conference on Generative Approaches to the Lexicon (2009)
24. Russo, I., Caselli, T., Rubino, F.: Recognizing deverbal events in context. In: Proc. of CICLing. Springer (2011)
25. Saurí, R., Knippen, R., Verhagen, M., Pustejovsky, J.: Evita: A Robust Event Recognizer for QA Systems. In: Proc. of HLT/EMNLP (2005)
26. Tanguy, L., Hathout, N.: Webaffix : un outil d'acquisition morphologique dérivationnelle à partir du Web. In: Pierrel, J.M. (ed.) Actes de TALN. ATILF (2002)