



**HAL**  
open science

## TempoWordNet for Sentence Time Tagging

Gaël Dias, Mohammed Hasanuzzaman, Stéphane Ferrari, Yann Mathet

► **To cite this version:**

Gaël Dias, Mohammed Hasanuzzaman, Stéphane Ferrari, Yann Mathet. TempoWordNet for Sentence Time Tagging. 23rd international conference on World wide web companion, Apr 2014, Seoul, South Korea. pp.Pages 833-838, 10.1145/2567948.2579042 . hal-01074964

**HAL Id: hal-01074964**

**<https://hal.science/hal-01074964>**

Submitted on 23 Oct 2014

**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.

# TempoWordNet for Sentence Time Tagging

Gaël Dias  
Normandie University  
GREYC UMR 6072  
Caen, France  
first.last@unicaen.fr

Mohammed  
Hasanuzzaman  
Normandie University  
GREYC UMR 6072  
Caen, France  
first.last@unicaen.fr

Stéphane Ferrari  
Normandie University  
GREYC UMR 6072  
Caen, France  
first.last@unicaen.fr

Yann Mathet  
Normandie University  
GREYC UMR 6072  
Caen, France  
first.last@unicaen.fr

## ABSTRACT

In this paper, we propose to build a temporal ontology, which may contribute to the success of time-related applications. Temporal classifiers are learned from a set of time-sensitive synsets and then applied to the whole WordNet to give rise to TempoWordNet. So, each synset is augmented with its intrinsic temporal value. To evaluate TempoWordNet, we use a semantic vector space representation for sentence temporal classification, which shows that improvements may be achieved with the time-augmented knowledge base against a bag-of-ngrams representation.

## Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Content Analysis and Indexing—*Dictionaries; Thesauruses; Linguistic processing*

## General Terms

Algorithms, Experimentation

## Keywords

TempoWordNet, Temporal Ontology, Sentence Temporal Classification

## 1. INTRODUCTION

Temporality has recently received increased attention in Natural Language Processing (NLP) and Information Retrieval (IR). Initial works have been proposed in NLP and are exhaustively summarized in [12]. More recently, the introduction of the TempEval task [19] in the Semantic Evaluation workshop series has clearly established the importance

of time to deal with different NLP tasks such as syntactic analysis [3].

In IR, the time dimension has also received particular attention for the past few years. According to [13], time is one of the key five aspects that determine a document credibility besides relevance, accuracy, objectivity and coverage. So, the value of information or its quality is intrinsically time-dependent. As a consequence, a new research field called Temporal Information Retrieval (T-IR) has emerged [1] and deals with all classical IR tasks such as crawling, indexing, ranking from the time viewpoint.

However, NLP and IR evidence different grains of analysis for temporality. NLP aims to understand time at a fine-grained level. For example, automatic temporal ordering of events in text is usually performed via various linguistic mechanisms including the use of time expressions such as “before”, “after” or “during” that explicitly assert a temporal relation. In particular, [4] investigate the role of temporal signals in temporal relation extraction.

From the IR viewpoint, temporality has been studied at a coarse-grained level. Different models have been proposed at the year level (or large periods of time) for crawling [11], indexing [2] or ranking [10]. As such, most methodologies rely on the presence of explicit timexes and hardly bridge the gap when no explicit mention of time is available. One recent exception is proposed in [9] where text time-tagging is seen as a classification task.

In this paper, we propose to build a time-sensitive WordNet (TempoWordNet) where each synset is associated to its intrinsic temporal value. As such, we expect to provide a better understanding of time in language, which may benefit both fine-grained (NLP) and coarse-grained (IR) temporal studies. TempoWordNet is evaluated based on the sentence temporal classification task. Overall results show that TempoWordNet allows 13.9% improvements of  $F_1$ -measure against the vector space model representation and 14.7% against the semantic vector space model obtained with the existing WordNet time subtree.

## 2. RELATED WORK

As expressed in [17], time taggers usually contain pattern files with words and phrases, which are typically used to express temporal expressions in a given language (e.g. names

of months). In fact, most temporal NLP tasks rely on a time-sensitive vocabulary. On the contrary, T-IR systems usually do not use information about time in language although they could benefit from it when facing the recurrent problem of missing explicit timexes.

WordNet [14] is a good place to start to find time-sensitive concepts. Indeed, one can list a set of 21 temporal synsets by iteratively following the hyponymy relation from the concept of time (synset # 00028270) represented by the following gloss: *the continuum of experience in which events pass from the future through the present to the past*. However, likewise the tennis problem evidenced in [7], most temporal words are not under the concept of time. For example, concepts such as “prediction”, “remember”, “ancient”, “fresh” clearly have a time dimension although they are not listed under the time subtree of WordNet.

Based on the initial ideas of [15] on temporal ontologies and inspired by SentiWordNet [6], we propose to enrich all WordNet synsets with their temporal dimensions.

### 3. BUILDING TEMPOWORDNET

We build TempoWordNet based on WordNet such that each synset is automatically time-tagged with four dimensions: *atemporal*, *past*, *present* and *future*. This can be achieved in different ways. In this paper, we focus on two strategies following the ideas of [5] and [6]: a one-step strategy (baseline) and a two-steps process. We will start by detailing the two-steps strategy, which embodies most of the relevant concepts and then straightforwardly define the one-step process.

#### 3.1 Two-Steps Classification

The overall idea of the two-steps strategy can be described as follows. First, a three-class temporal classifier is built over a set of manually selected seed synsets defined by their corresponding glosses. The underlying idea is that temporal synsets should embody temporality in their definition in a similar way. The classification process is iterated based on the repetitive semantic expansion of the initial seed list until cross-validation accuracy drops. By semantic expansion, we mean that different lexico-semantic relations are used to encounter temporality in WordNet. This first step results in a *past*, *present* and *future* classifier and an expanded list of temporal synset candidates.

A second temporal classifier is then learned to time-tag synsets as *atemporal* or *temporal*. This process is obtained by taking the final list of expanded seed synsets from the previous learning problem and randomly choosing a balanced number *atemporal* synsets. A 10-fold cross-validation is then used to learn the model.

TempoWordNet is finally obtained by first classifying all WordNet synsets as *atemporal* or *temporal* with the second classifier and then the resulting temporal synsets are tagged as *past*, *present* and *future* by the first classifier.

##### 3.1.1 Past, Present, Future Classification

The first step to build TempoWordNet is based on a classification model, which aims to distinguish between *past*, *present* and *future* synsets. This first step is defined in Algorithm 1 and all subtasks are explained as follows.

**Initial Seeds Lists Selection:** In SentiWordNet, [5] start by selecting words that are relevant to express positive or

---

#### Algorithm 1 Past, present, future classification.

---

```

Selection of the initial seeds lists
repeat
  Expansion of the seeds lists
  Learning the model Past, Present, Future
  Measure Accuracy by 10-fold cross-validation
until accuracy drops

```

---

negative opinions. Similarly, we need to select seeds used as good paradigms for *past*, *present* and *future* categories. For example, words like “yesterday”, “previously”, “remember” are good paradigmatic words for the *past* category, “current”, “existing”, “presently” for *present* and “prophecy”, “predict”, “tomorrow” for *future*. The selection of the initial set of seed synsets is a crucial step in the process as their properties must be preserved along the expansion process.

In order to catch the most relevant synsets for each time category, a first selection was made by several individuals through intensive and freewheeling group discussion. Every participant was encouraged to think aloud and to suggest as many words as possible. We preferred to use this process as choosing all words from the WordNet time subtree would have resulted in a biased sample as almost all synsets are nouns. Indeed, we wanted to make sure that each grammatical category existing in WordNet (i.e. Noun, Adjective, Adverb and Verb) would be present in the sets of seeds for *past*, *present* and *future* categories.

As each synset in WordNet contains one or more words, the synsets expressing the temporal connotations listed by the individuals were selected.

Finally, we performed an inter-annotator agreement process over the three seeds lists with four different annotators who were presented with the synsets and their respective glosses. The results of the multirater agreement evaluation are presented in Table 1. In particular, we processed the free-marginal multirater kappa values [16] as Fleiss’ popular multirater kappa [8] is known to be influenced by prevalence and bias, which can lead to the paradox of high agreement but low kappa. Overall figures assess adequate agreement.

Metric	Past	Present	Future
% of overall agreement	0.85	0.83	0.90
Free-marginal $\kappa$	0.70	0.66	0.80

**Table 1: Inter-annotator agreement.**

The initial lists of *past*, *present* and *future* seeds are given in Table 2 for reproducibility of the experiments.

**Expansion Process:** The guiding idea behind the expansion process is that the temporal properties of the initial hand-crafted seeds lists should be preserved as we strategically travel through WordNet. Depending on the morpho-syntactic class of an initial temporal synset, choosing an appropriate set of conceptual relations may allow to expand the notion of time in WordNet. As a consequence, we propose to exploit the following conceptual relations for the expansion process depending on the morpho-syntactic class:

- **synonymy:** for each morpho-syntactic class (e.g. “past” vs. “yesteryear” for noun),

Words	Sense	Category	Class
past	1,2	n.	past
past	1,2	adj.	past
yesterday	1,2	n.	past
yesterday	1,2	adv.	past
commemorate	2	v.	past
previously	1	adv.	past
present	1	n.	present
present	1,2	adj.	present
now	1	n.	present
now	3	adv.	present
nowadays	1	adv.	present
today	1	n.	present
ongoing	1	adj.	present
existing	1	adj.	present
current	1	adj.	present
future	1	n.	future
future	1,2	adj.	future
tomorrow	1,2	n.	future
tomorrow	1	adv.	future
predict	1	v.	future
expected	1	adj.	future
prophecy	1	v.	future
aforethought	1	adj.	future

Table 2: List of initial *temporal* seeds.

- **hyponymy:** for nouns<sup>1</sup> (e.g. “future” vs. “tomorrow”),
- **troponymy:** for verbs (e.g. “will” vs. “plan”),
- **related nouns:** for adjectives (e.g. “future” vs. “approaching”),
- **root adjectives:** for adverbs (e.g. “recently” vs. “recent”).

**Classification:** Finally, a semi-supervised learning strategy is used to learn the temporal (*past*, *present* and *future*) classifier. At each semantic expansion step (or iteration), a three-class text classifier is trained over the glosses of each synset contained in the seeds lists. After each iteration, the accuracy of the learned model is measured through a 10-fold cross-validation process. The expansion process continues until the classifier accuracy steadily drops.

**Results:** In our experiments, we used the initial seeds lists containing 30 synsets and then performed the semi-supervised learning process using different classifiers and representations. As for classifiers, we used Support Vector Machines (SVM), Multinomial Naïve Bayes models (MNB) and Decision Trees (C4.5) from the Weka platform<sup>2</sup> and performed all the experiments with default parameters set by Weka. As for the representation space, each synset was represented by its gloss encoded as a vector of word unigrams weighted by their frequency in the gloss<sup>3</sup>.

Overall results are presented in Table 3 and show that the “optimal” expansion is obtained after three iterations using

<sup>1</sup>Following the hypernymy relation leads to the classical semantic shift problem.

<sup>2</sup><http://www.cs.waikato.ac.nz/ml/weka/> [Last access: 12/02/2014].

<sup>3</sup>Stop words removal has been performed using the Weka database.

Steps		1	2	3	4	5
SVM	Precision	81.7	84.4	86.1	86.0	85.4
	Recall	79.9	82.6	83.9	83.5	83.2
	$F_1$ -measure	79.8	82.1	<b>83.5</b>	83.1	82.9
MNB	Precision	83.8	76.9	78.2	77.4	78.1
	Recall	82.7	76.7	77.5	76.3	77.0
	$F_1$ -measure	<b>83.2</b>	76.8	77.8	76.8	77.5
C4.5	Precision	76.3	73.5	71.1	72.4	73.5
	Recall	70.8	63.5	63.5	63.8	62.5
	$F_1$ -measure	<b>74.4</b>	68.1	67.1	68.4	67.6

Table 3: SVM, naïve bayes and decision trees accuracy results for Past, Present, Future classification at each iteration step.

SVM. In the cases of MNB and C4.5, accuracy immediately drops as the introduction of possible noisy synsets is hard to handle for such simple models<sup>4</sup>.

Finally, we end up with a list of 632 temporal synsets distributed as follows: 210 synsets marked as *past*, 291 as *present* and 131 as *future*. In Table 4, we provide the top 10 synsets and the bottom 10 synsets classified as temporal by the SVM at iteration 3. Likewise the distribution of the number of extracted synsets, the distribution of morpho-syntactic categories depends on the temporal class. For instance, *future* is mainly referred to by nouns, while *present* evidences a high number of action verbs and *past* is represented by ancient animals or adjectives and adverbs.

These results were expected. However, when digging up results, the temporal issue of synsets is sometimes difficult to guess. In fact, it is important to remember that classification is made over glosses. As such, the temporal values of concepts are given by their definition. For instance, “here”, which is classified as a *present* adjective has an unclear temporal connotation. However, its gloss “*being here now*” clearly refers to a *present* situation.

Also, within the expansion process, noisy synsets may be introduced and difficult the learning process. For instance, “augur”, which is automatically defined as a noun with a *future* connotation is incorrectly classified. Indeed, its gloss “*a religious official who interpreted omens to guide public policy*” does not embody any *future* issue. This situation is discussed in section 5 but has mainly to deal with the fact that the temporal connotation is not always present for a same denotation.

### 3.1.2 Atemporal vs. Temporal Classification

Once the *past*, *present* and *future* classifier has been learned, we end up with a list of 632 temporal synsets, which “abusively” embody the notion of time in WordNet. Indeed, there are more temporal categories than just *past*, *present* and *future* as there are more than positive and negative classes in the expression of sentiments. Although, as a first step towards building the first temporal ontology so far<sup>5</sup>, we found wise to refer to the common sense connotations of time.

Continuing with the analogy proposed by [5] where any word is objective if it is not negative or positive, any concept, which is not associated to the notions of *past*, *present* or *future* is called *atemporal*. So, in order to learn the second

<sup>4</sup>Remind that the set of synsets is small.

<sup>5</sup>As far as we know.

Past			
Top 10		Bottom 10	
Word (Sense)	Cat.	Word (Sense)	Cat.
by (1)	adv.	iguanodon (1)	n.
recently (1)	adv.	ground-shaker (1)	n.
in the first place (1)	adv.	diplococus (1)	n.
remember (3)	v.	saurischian (1)	n.
old (1)	n.	argentinosauro (1)	n.
old (1)	adj.	ornithischian (1)	n.
old (2)	adj.	titanosaur (1)	n.
old (6)	adj.	mellowing (1)	n.
erstwhile (1)	adj.	appearance (4)	n.
honest to god (1)	adj.	psychosexuality (1)	n.
Present			
Top 10		Bottom 10	
Word (Sense)	Cat.	Word (Sense)	Cat.
immediately (1)	adv.	overstay (1)	v.
now (3)	adv.	visit (7)	v.
presently (2)	adv.	run (30)	v.
present (3)	n.	drag on (1)	v.
immediate (3)	adj.	wear (6)	v.
instant (2)	adj.	crawl (3)	v.
attendant (1)	adj.	bond (3)	v.
ever-present (1)	adj.	ramp (5)	v.
here (1)	adj.	stand back (2)	v.
omnipresent (1)	adj.	line up (3)	v.
Future			
Top 10		Bottom 10	
Word (Sense)	Cat.	Word (Sense)	Cat.
prophecy (1)	n.	example (4)	n.
prefiguration (2)	n.	referral (1)	n.
prognosis (1)	n.	palmist (1)	n.
prophecy (2)	n.	sibyl (1)	n.
meteorology (1)	n.	prophetess (1)	n.
fortunetelling (1)	n.	augur (1)	n.
extropy (1)	n.	sibyl (2)	n.
horoscope (1)	n.	onomanancy (1)	n.
guess (2)	n.	arithmancy (1)	n.
credit rating (1)	n.	lithomancy (1)	n.

**Table 4: List of automatically retrieved temporal synsets.**

classifier, we randomly chose a set 632 *atemporal* synset candidates within WordNet each one being outside the time subtree of WordNet and the list of pre-computed 632 temporal synsets. We performed a manual cross-annotation process to ensure the atemporality of the candidates. For that purpose, we randomly selected a subset of 10 synsets and asked four annotators to decide upon their atemporality. The results of the free-marginal multirater kappa evidence a substantial agreement with 0.73.

So, based on the set of 632 temporal synsets and 632 atemporal ones, a SVM was learned for a two-class problem reaching 85.6% accuracy over a 10-fold cross-validation process. Similarly to the first classification task, we used a linear kernel and represented each synset by its gloss based on the vector space model with each word feature being represented by its frequency.

Finally, TempoWordNet is obtained by a two-step process: (1) all synsets in WordNet are classified as *temporal* or *atemporal* based on the classifier mentioned in subsection 3.1.2 and (2) each temporal synset is associated to its tem-

poral values (*past*, *present* and *future* summing up to one) using the classifier built in subsection 3.1.1.

### 3.2 One-Step Classification

One direct comparison has been experimented with a one-step classification strategy. Instead of expanding temporal synsets in a first step to finally execute two-stage classification, propagation can be executed in a single step.

So, we propose to expand both *temporal* and *atemporal* synsets at the same time and directly produce a four-class temporal classifier: *past*, *present*, *future* and *atemporal*. For that purpose, we presented a set of 30 *atemporal* synsets to four annotators who agreed with a free-marginal multirater kappa value over 0.8 indicating almost perfect agreement. The list of *atemporal* synsets is given in Table 5 for reproducibility of the results.

Words	Sense	Category	Class
mixing bowl	1	n.	atemporal
freshen	2	v.	atemporal
carnation	2	n.	atemporal
chadian	1	adj.	atemporal
wren warbler	1	n.	atemporal
brainsick	1	adj.	atemporal
estriol	1	n.	atemporal
theology	2	n.	atemporal
unexpectedly	1	adv.	atemporal
jabber	1	n.	atemporal
human waste	1	n.	atemporal
cruciferous	1	adj.	atemporal
pet sitter	1	n.	atemporal
trombicula	1	n.	atemporal
drum	1	v.	atemporal
dateline	1	n.	atemporal
shot	11	n.	atemporal
okinawa	1	adv.	atemporal
chatter	1	v.	atemporal
polecat	2	n.	atemporal
foster home	1	n.	atemporal
lymph node	1	n.	atemporal
arabian sea	1	n.	atemporal
semanticist	1	n.	atemporal
strauss	3	n.	atemporal
doric order	1	n.	atemporal
reptantia	1	n.	atemporal
belt	2	v.	atemporal
half dollar	1	n.	atemporal
staggered board of directors	1	n.	atemporal

**Table 5: List of initial *atemporal* seeds.**

The same semi-supervised learning strategy is used to learn the four-class (*past*, *present*, *future* and *atemporal*) classifier. At each iteration, the classifier is trained over the glosses of each synset contained in the seeds lists. After each iteration, accuracy is measured through a 10-fold cross-validation process and the expansion process stops when accuracy drops. Results are presented in Table 6 for the same experimental setups as for the two-step strategy.

Unlike the two-step strategy, the one-step process shows incapacity to solve the temporality issue. Indeed, introducing the *atemporal* synsets in the propagation process since the first iteration evidences different problems: (1) atemporality is difficult to define unless when opposed to temporality as it embodies many denotations and no connotation

Steps		1	2	3	4	5
SVM	Precision	81.3	68.0	71.0	71.3	71.7
	Recall	80.3	63.0	66.8	67.8	68.6
	$F_1$ -measure	<b>80.8</b>	65.4	68.8	69.5	72.8
BNB	Precision	75.2	67.0	67.1	67.8	77.4
	Recall	74.3	64.6	63.6	64.0	76.3
	$F_1$ -measure	74.7	65.8	65.3	65.8	<b>76.8</b>
C4.5	Precision	73.3	68.6	59.2	61.4	72.4
	Recall	70.6	52.4	48.6	48.0	63.8
	$F_1$ -measure	<b>71.9</b>	59.4	53.3	68.4	69.9

**Table 6: SVM, naïve bayes and decision trees accuracy results for Past, Present, Future, Atemporal classification at each iteration step.**

and (2) temporality only spreads over a small proportion of WordNet while atemporality covers most of WordNet, and as such, as iterations grow, the set of atemporal candidate synsets gets predominant (i.e. more unbalanced datasets are obtained after each iteration).

#### 4. CLASSIFICATION OF TEMPORAL SENTENCES

In order to evaluate TempoWordNet, we propose to evaluate its usefulness based on an external task: sentence temporal classification. The underlying idea is that a temporal knowledge base can help to classify sentences into three different categories: *past*, *present* and *future*.

For that purpose, we automatically selected a set of *past*, *present* and *future* sentences from the well-known SemEval-2007 corpus developed for task 15 [18]. This corpus is a version of TimeBank containing approximately 2500 sentences with TimeML annotations. So, all sentences exclusively containing *past* (resp. *present*) expressions were marked as *past* (resp. *present*). As for *future*, all sentences containing *future* expressions combined or not with *present* timexes were tagged as *future*. The final corpus consists of 1455 sentences distributed as follows: 724 for *past*, 385 for *present* and 346 for *future*. Some examples are given as follows:

1. *In New York Stock Exchange composite trading yesterday, Oneida’s shares closed at \$18.375 a share, unchanged (Past),*
2. *Currently, Avon, based in Santa Monica, Calif., has 3.3 million common shares outstanding (Present),*
3. *A TOP-LEVEL investigation into Mark Thatcher’s alleged arms deals with Iraq is to be launched by Swiss government officials in the New Year (Future).*

With respect to the classification task, different sentence representations have been used. First, we propose to represent each sentence as a bag of unigrams or bigrams. Second, we propose a semantic vector space representation where each sentence is augmented with the synonyms of any temporal word contained in it. In particular, we propose that the words are matched directly from the WordNet time subtree (Wn) or from TempoWordNet (TWn). For the experimental setups, we used the SVM implementation of Weka and report the results for 10-fold cross-validation for tf.idf weighting scheme. Note that word sense disambiguation has

not been performed and the most frequent sense is used instead. Results are reported in Table 7 where stop words removal has been performed.

The results clearly evidence that the WordNet time subtree does not embody enough time-related information and the process of automatically time-tagging WordNet can greatly improve the task of sentence temporal classification. Moreover, the results suggest that bigrams do not provide extra information for the given task.

In order to better understand the process of temporal classification of english sentences, we propose to compare the unigram representations with and without stop words. Results are shown in Table 8.

	Uni.+SW	Uni.+TWn+SW
Precision	85.8	87.8
Recall	85.7	87.8
$F_1$ -measure	85.6	87.8

**Table 8: Evaluation results for sentence classification with and without stop words. Balanced corpus: 346 sentences for *past*, 346 sentences for *present* and 346 sentences for *future*,**

Stop words indeed play an important role for sentence classification for the english language. In the list of stop words there are auxiliary verbs such as “will” or “did”, which are evident clues for sentence temporal classification. The improvement of TempoWordNet on this small dataset is therefore residual reaching 2.2% increased performance. However, its importance may not be neglected as complex temporal classification tasks are not likely to depend on auxiliary verbs. A clear example is given in sentences (1), (2) and (3). Another interesting example is shown in the Temporal Query Intent Classification subtask of the pilot task Temporalia of NTCIR-11<sup>6</sup>, where queries such as “*History of Coca-Cola*”, “*time in london*”, “*long term weather forecast*” or “*lose weight quickly*” must respectively be classified as *past*, *recency*, *future* or *atemporal*.

#### 5. DISCUSSION

One important remark must be expressed here for future related works. Although the process of time tagging WordNet has been inspired by the initial idea of [5], it can not be compared to the one of opinion tagging. Indeed, the subjective information linked to a word is both about connotation and denotation. Thus, hyponymy is a privileged relation for propagating opinions. Indeed, both the main sense and the same semantic orientation are kept along the hyponymy relation. On the contrary, we observed that the temporal information is more associated to denotation than connotation. As such, although hyponyms of a given temporal synset still have the same denotation, the temporal connotation may be lost. For instance, the present information linked to the concept “present” is already lost for its direct hyponym “date”. As a consequence, enhanced propagation strategies, instead of just lexico-semantic relations, should be proposed in order to improve the intrinsic quality of TempoWordNet.

<sup>6</sup><https://sites.google.com/site/ntcirtemporalia/home> [Last access: 11/02/2014].

Representation	Unigrams	Bigrams	Unigrams+Wn	Bigrams+Wn	Unigrams+TWn	Bigrams+TWn
Precision	64.2	63.9	63.6	63.3	<b>78.3</b>	78.2
Recall	64.2	63.4	63.3	62.9	77.8	<b>78.0</b>
$F_1$ -measure	64.2	63.6	63.4	63.1	78.0	<b>78.1</b>

**Table 7: Evaluation results for sentence classification over the Past, Present, Future version of the Semval-2007 task 15 Corpus. Balanced corpus: 346 sentences for *past*, 346 sentences for *present* and 346 sentences for *future*.**

## 6. CONCLUSION

In this paper, we proposed the first steps towards the automatic construction of a temporal ontology. In particular, we augmented WordNet with temporal information by following a two-step process. First, synsets are classified as *atemporal* or *temporal* and then, all temporal synsets are associated to *past*, *present* and *future* probabilities. The experiments made for sentence temporal classification showed that TempoWordNet allows 13.9% improvements of  $F_1$ -measure against the vector space model representation and 14.7% against the semantic vector space model obtained with the existing WordNet time subtree. We also evidenced the importance played by stop words in sentence temporal classification where improvements with TempoWordNet are less expressive (2.2%). Nevertheless, we deeply believe that TempoWordNet can be an important resource for time related applications both in NLP and IR. As a consequence, we provide free access to this resource as well as all developing materials at the following url <http://tempowordnet.greyc.fr>.

## 7. REFERENCES

- [1] O. Alonso, J. Strötgen, R. Baeza-Yates, and M. Gertz. Temporal information retrieval: Challenges and opportunities. In *Proceedings of the 1st International Temporal Web Analytics Workshop (TAWAW)*, pages 1–8, 2011.
- [2] A. Anand, S. Bedathur, K. Berberich, and R. Schenkel. Index maintenance for time-travel text search. In *Proceedings of the 35th International ACM Conference on Research and Development in Information Retrieval (SIGIR)*, pages 235–244, 2012.
- [3] F. Costa and A. Branco. Full-fledged temporal processing: Bridging the gap between deep linguistic processing and temporal extraction. *Journal of Language Modelling*, 1(1):97–154, 2013.
- [4] L. Derczynski and R. Gaizauskas. A corpus-based study of temporal signals. *arXiv:1203.5066*, 2012.
- [5] A. Esuli and F. Sebastiani. Determining the semantic orientation of terms through gloss analysis. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM)*, pages 617–624, 2005.
- [6] A. Esuli and F. Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC)*, pages 417–422, 2006.
- [7] C. Fellbaum. *WordNet: An Electronic Lexical Database*. Bradford Books, 1998.
- [8] J. Fleiss. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382, 1971.
- [9] A. Jatowt, C.-M. A. Yeung, and K. Tanaka. Estimating document focus time. In *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM)*, pages 2273–2278, 2013.
- [10] N. Kanhabua, R. Blanco, and M. Matthews. Ranking related news predictions. In *Proceedings of the 34th International ACM Conference on Research and Development in Information Retrieval (SIGIR)*, pages 755–764, 2011.
- [11] A. Kulkarni, J. Teevan, K. Svore, and S. Dumais. Understanding temporal query dynamics. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM)*, pages 167–176, 2011.
- [12] I. Mani, J. Pustejovsky, and R. Gaizauskas. *The language of time: a reader*, volume 126. Oxford University Press, 2005.
- [13] M. Metzger. Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13):2078–2091, 2007.
- [14] G. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
- [15] M. Moens and M. Steedman. Temporal ontology in natural language. In *Proceedings of the 25th Annual Meeting on Association for Computational Linguistics (ACL)*, pages 1–7, 1987.
- [16] J. Randolph. Free-marginal multirater kappa (multirater  $\kappa_{free}$ ): an alternative to fleiss’ fixed-marginal multirater kappa. *Joensuu Learning and Instruction Symposium*, 2005.
- [17] J. Strötgen and M. Gertz. Multilingual and cross-domain temporal tagging. *Language Resources and Evaluation (LRE)*, 47:269–298, 2013.
- [18] M. Verhagen, R. Gaizauskas, F. Schilder, M. Hepple, G. Katz, and J. Pustejovsky. Semeval-2007 task 15: Tempeval temporal relation identification. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 75–80, 2007.
- [19] M. Verhagen, R. Gaizauskas, F. Schilder, M. Hepple, J. Moszkowicz, and J. Pustejovsky. The tempeval challenge: Identifying temporal relations in text. *Language Resources and Evaluation (LRE)*, 43(2):161–179, 2009.