

Using Semantic Networks to Identify Temporal Expressions from Semantic Roles

Hector Llorens, Borja Navarro, Estela Saquete
Natural Language Processing Research Group
University of Alicante, Spain
{*hlllorens, borja, stela*}@*dlsi.ua.es*

Abstract

Nowadays, the temporal aspects of natural language are receiving a great research interest. TimeML has been adopted as a standard for temporal information annotation by a large number of researchers. Available TimeML resources are very limited in size and in diversity of languages. This paper analyzes a combination of semantic roles and semantic networks information for improving this situation. An automatic approach using semantic networks to convert temporal semantic roles into TimeML TIMEX3 elements is presented. This approach has been quantitatively evaluated for English and Spanish. The results point out that the presented approach can help in a semi-automatic creation of TimeML resources for the evaluated languages and could be also valid for other European languages.

Keywords

TimeML, TimeBank, TE identification, Semantic Roles

1 Introduction

In recent years, the research interest on automatic treatment of temporal information of natural language (NL) text has experienced an important growth [8]. One of the main reasons for that are the benefits that temporal information brings to Question answering (QA), summarization and many other natural language processing (NLP) areas [17]. Specialized workshops and conferences [12, 15], and evaluation forums [20, 21] reflect the importance of this field. Furthermore, the development of language independent systems has become an important issue among NLP community. This has been reflected in many conferences such as CLEF¹, as well as in works specific to Temporal Expression (TE) recognition field [24, 10]. In this paper, we present an approach to temporal expression identification from a multilingual point of view.

There are different ways to represent temporal information in NL. One of them is TimeML [13], which has been recently adopted as *de facto* standard annotation scheme by a large number of researchers [17].

A different way to represent time is defined in Semantic role labeling (SRL). SRL consists of determining basic event structures in a sentence, detecting semantic relations among entities and events. The tem-

poral information of the events is represented by the temporal semantic role. SRL field has achieved important results in the last years [5].

Currently, the major problem of TimeML lies on the lack of resources, specially the lack of corpora for languages other than English. Specifically, this work is focused on the benefits that available semantic networks and semantic roles corpora can introduce to TE identification. To achieve the proposed objective, we present an automatic system that identifies TimeML TEs from semantic roles using semantic networks as validation method. Furthermore, this system is designed to handle the task multilingually, provided that there are semantic networks and semantic roles resources available for the target language. To measure the performance and the possibilities of the presented proposal, an evaluation for English and Spanish is carried out, as well as an in-depth analysis of the results.

The paper is structured as follows: Section 2 focuses on the background of temporal information processing and SRL fields and Section 3 provides detailed information about our proposal to obtain TimeML TEs from semantic roles and semantic networks. Section 4 includes the evaluation and error analysis and, finally, conclusions and further work lines are presented.

2 Background

The importance of temporal aspects of NL is not a new issue in artificial intelligence (AI) [1]. Several efforts have been done in order to define standard ways to represent temporal information in NL. Since temporal information extraction was included in Message Understanding Conference context, there have been three important annotation schemes for temporal information: STAG [18], TIDES [4] and TimeML [13]. TimeML is a rich specification language for events and TEs in NL that combines and extends features of both preceding schemes. It was designed to address time stamping, ordering and reasoning about TEs and events of NL. Fig. 1 illustrates an example annotation. In the example, “*came*” (EVENT) represents an event which is linked to the temporal expression “*Monday*” (TIMEX3) through a temporal link (TLINK), in which the temporal signal “*on*” (SIGNAL) is involved.

An English corpus illustrating TimeML annotation, TimeBank [14], was created together with the first version of this annotation scheme. The last version of the corpus, TimeBank 1.2, is considered a gold standard and has been published by Linguistic Data Consor-

¹ European Cross-Language Evaluation Forum

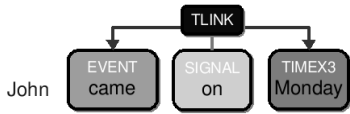


Fig. 1: TimeML example

tium. An in-depth analysis of TimeBank corpus can be found in [2]. Unfortunately, there are not TimeML corpora available for other languages like Spanish.

There have been different works on developing systems for automatically tagging NL text following TIMEX3 specifications. On the one hand, the work of Boguraev and Ando [2] presents an evaluation on automatic TimeML annotation over TimeBank using machine learning techniques. The results for TIMEX3 recognition using 5-fold cross validation were 89.6% and 81.7% $F_{\beta=1}$ for relaxed and strict span. On the other hand, TTK [22] accomplishes this task using the GUTime module. TTK has not been evaluated for TIMEX3. However, it was benchmarked on training data from TERN 2004 [20] at 85% and 78% $F_{\beta=1}$ for TIMEX2 relaxed and strict span respectively.

As introduced in previous section, another way of representing temporal information in NL texts, is through semantic roles. They represent temporality from a different perspective. Essentially, Semantic role labeling consists of determining basic event structures in a sentence, detecting semantic relations among entities and events. The temporal semantic role (TSR) represents “when” an event takes place. Fig. 2 illustrates how semantic roles represent temporal information through the temporal semantic role.

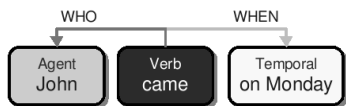


Fig. 2: Semantic roles example

Only one reference about using semantic roles for temporal information processing has been found in literature [6]. That work used them as complementary information to identify temporal relations.

Semantic networks have been used in many NLP fields for different purposes. WordNet [3] and EuroWordNet [23] represent the most used semantic networks for English and European languages respectively. Specifically, in temporal expression identification task, semantic networks have been used in the following works: Negri et al. [11] used WordNet to create a list of temporal named entities such as Bastille Day, Hanukkah, etc., by collecting all hyponyms tree of “calendar.day” synset. Also, in [16], semantic networks were used to expand a list of temporal triggers by adding all the synonyms. These works show that the information contained in semantic networks can be useful for temporal information extraction task.

3 Proposal

In order to study the benefits that semantic networks and semantic roles can introduce to temporal expressions identification task, this section presents an automatic system that identifies TimeML TEs using such resources. Two versions are described, firstly, TIPSem, which uses morphosyntactic information to transform temporal semantic role (TSR) into TIMEX3 element, and secondly, TIPSem+WN, which uses semantic networks to validate TIMEX3 elements identified by TIPSem system.

3.1 TIPSem

Temporal role is not defined exactly as a TIMEX3. A TSR represents a complete semantic predicate with a temporal function. However, the full extent of a TIMEX3 tag must correspond to one of the following categories: noun phrase (“yesterday” NP), adjective phrase (“3-day” ADJP) or adverbial phrase (“3 days ago” ADVP). As shown in example 1, both representations are not equivalent.

- (1) She was born [in 1999 TSR]
She was born in <TIMEX3>1999</TIMEX3>

TIPSem (Temporal Information Processing based on Semantic roles) implements the following set of transformation rules from TSR to TIMEX3 solving the main differences between them.

1. **Removing TSR overlapping:** Due to the fact that each verb has its own roles, it is possible to find overlapped TSRs. In such cases, TIPSem system keeps only the TSR representing the minimum syntactic unit (NP, ADJP or ADVP).
2. **Removing subordination of TSR:** If a TSR corresponds to a subordination clause it does not correspond to a TIMEX3. The system detects and removes it using the syntactic tree.
3. **Splitting TSR:** A TSR composed of more than one NP can contain a set of related TIMEX3, linked by a temporal preposition or a coordination conjunction. There are two exceptions for this rule. Times “[ten minutes to four]”, where the “to” preposition is denoting a specification relation, and the preposition “of” (“the end of 1999”), which is usually part of the expression. Our system looks for prepositions or coordination conjunctions in every TSR containing more than one NP. If they are found and do not represent an exception, the TSR is split in n TIMEX3 corresponding to each NP.
4. **TSR syntactic reduction:** As described above, a TSR generally differs from a TIMEX3 on its boundaries. If a TSR has any element out of the minimum syntactic unit (NP, ADJP or ADVP), this element is not included as part of the TIMEX3. The most common cases are the ones in which the TSR consists of a prepositional phrase

(PP). This PP normally contains some preposition (before, at, etc.) or a combination adverb-preposition (later in, ahead of, etc.) followed by an NP which represents the TIMEX3 element.

5. **Tagging resulting TSR as TIMEX3:** Finally, after the application of all the previous rules, resulting TSR are directly tagged as TIMEX3.

Furthermore, due to the fact that SRL relies on verbs, nominal sentences can not be labeled. These sentences are commonly found in titles, brackets, notes, etc. Hence, as a post-processing step, a TE tagger capable of identifying basic explicit TEs (only times and dates) is executed for these sentences.

3.2 TIPSem+WN

There are cases in which TSR does not contain a TIMEX3. These cases represent one of the main problems of the TIPSem approach. Example 2 illustrates the problem showing a sentence annotated with the TSR, the correct TIMEX3 annotation and the incorrect TIMEX3 annotation obtained by TIPSem.

- (2) **TSR:** She ate [before the meeting TSR]
Correct TIMEX3: She ate before the meeting
Incorrect TIMEX3 (TIPSem):
 She ate before <TIMEX3>the meeting</TIMEX3>

As shown in the example 2, “*the meeting*” is incorrectly tagged as TIMEX3 by TIPSem approach. In this case the temporal information provided by the TSR corresponds to a TimeML EVENT instead. The difficulty arises on how to differentiate this kind of events from real TEs. The following example illustrates the reasons why this is not an easy issue.

- (3) (S(NP (PRP She))(VP (VBD ate)
 (PP (IN before)(NP (DT the) (NN night))))))

 (S(NP (PRP She))(VP (VBD ate)
 (PP (IN before)(NP (DT the) (NN meeting))))))

In example 3, “*before the night*” and “*before the meeting*” are represented by a TSR at semantic roles level, and are identical at morphosyntactic level. However, “*the night*” corresponds to a TIMEX3, but not “*the meeting*”. In this manner, it is not trivial to distinguish between them using only morphosyntactic and semantic roles information.

One possible solution would be to manually encode a list of temporal triggers. This solution is costly and language dependent. For that reason, we propose an automatic multilingual solution to problem using the multilingual temporal information encoded in different languages semantic networks such as WordNet [3], EuroWordNet [23]. A list of different languages “*WordNets*” can be found at Global WordNet site².

For each word sense (synset), semantic networks bring, among other things, the complete hypernyms hierarchy. Our hypothesis is that all words related to time should have a general time concept among their

hypernyms. Example 4 shows two words related to a general time concept.

- (4) hour (hypernyms hierarchy)
 => **time_unit** => measure => abstraction => entity

Monday (hypernyms hierarchy)
 => day_of.the_week => calendar_day => **time_period**
 => measure => abstraction => entity

The unique exception we include in this hypothesis are purely numeric dates and times such as “*1999, 12-12-2001 and 18:25*”.

Taking this hypothesis into consideration, we define a TE validation algorithm based on semantic networks. It is defined as follows:

- A TSR is validated to be a TIMEX3 if at least one of its words has a hypernym that matches a general temporal concept or a numeric date/time.
- To handle polysemous words, the word part-of-speech (PoS) is used to query the semantic networks. If the word has different senses with the same PoS, if at least one of them is related to a time concept the system validates it, because if the word is contained by a temporal role, this is probably the correct sense.
- The algorithm also searches for multiword expressions to handle compound temporal concepts like “*Corpus Christi*” and “*Saint Joseph*”.

The described algorithm has been implemented for English and Spanish. WordNet has been used for English, taking as general time concepts: *time_period*, *time_unit* and *time*. EuroWordNet has been used for Spanish, taking as general time concepts the same as in English: *periodo*, *unidad_de_tiempo* and *tiempo*³. Fig. 3 illustrates the TIPSem+WN system architecture.

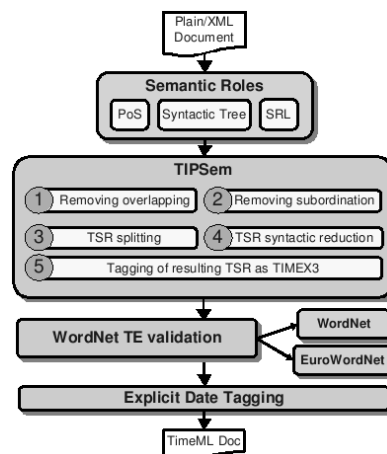


Fig. 3: TIPSem+WN Architecture

² <http://www.globalwordnet.org/>

³ time_period, time_unit and time

4 Evaluation

The objective of the evaluation is to provide a quantitative study on how well does TIPSem and TIPSem+WN approaches perform in TIMEX3 identification task and which are the effects of the usage of semantic networks. It covers English and Spanish, and also includes a Baseline implementation, that tags every TSR as TIMEX3, to measure how accurate are temporal roles by their own on representing TIMEX3.

4.1 Evaluation Environment

4.1.1 Corpora

The presented approaches have been evaluated using TimeBank 1.2 corpus [14] for English, and a manually annotated sample of AnCora [19, 9] for Spanish.

- **English (TimeBank):** TimeBank 1.2 consists of 183 news articles tagged following the TimeML 1.2.1 specification. For this evaluation, this corpus has been automatically annotated using the SRL tool developed by University of Illinois CCG group [7], which uses PropBank role set. This tool obtained a 77.44% $F_{\beta=1}$ in TSR (AM-TMP PropBank role) labeling in CoNLL 2005.
- **Spanish (AnCora TimeML Sample):** Due to the lack of TimeML corpus for Spanish, we have developed a Spanish TimeML TIMEX3 corpus sample annotating manually 30 docs of AnCora. AnCora is the largest corpus annotated at different linguistic levels in Spanish and Catalan. It consists of 500K words in each language, mainly taken from newspaper texts. The corpus is annotated and manually reviewed at: morphological level, syntactic level, and semantic level.

Both corpora statistics are shown in Table 1. In the table, the *in TEXT* value indicates the TIMEX3 tags found in corpus text (between *TEXT* tags), ignoring explicit dates in documents headers.

Corpus	docs	words	TIMEX3 (in TEXT)
TimeBank	183	61.8K	1414 (1228)
AnCora Sample	30	7.3K	155 (125)

Table 1: Corpora statistics

4.1.2 Criteria

The presented approaches have been tested in TE identification within the previously described corpora and the results have been compared to the original TIMEX3 annotation. The explicit dates of document headers have been ignored to make a more reliable test. We applied the criteria used in TERN-2004. The measures, inherited from it, are:

- **ACT:** TIMEX3 tags returned by the system.
- **Correct (corr):** Correct instances
- **Incorrect (inco):** Wrongly bounded instances
- **Missing (miss):** Not detected instances
- **Spurious (spur):** False positives

- **Precision (prec):** corr/ACT
- **Recall (rec):** corr/POS
- **$F_{\beta=1}$:** $(2*prec*rec)/(prec+rec)$

An adaptation to TIMEX3 of the TERN-2004 scorer⁴ has been used to calculate these measures.

4.2 Results

Tables 2 and 3 show the obtained results for English and tables 4 and 5 the ones obtained for Spanish. For each system, span relaxed *R* and span strict *S* results are indicated. *S* refers to strict match of both boundaries of a TIMEX3 expression (exact extent) while *R* results consider as correct every tag including a TIMEX3 even if it is wrongly bounded.

System		ACT	corr	inco	miss	spur
Baseline	R	1410	764	0	464	646
	S	1410	368	396	464	646
TIPSem	R	1245	908	0	320	337
	S	1245	817	91	320	337
TIPSem+WN	R	1020	905	0	323	115
	S	1020	815	90	323	115

Table 2: TIMEX3 results for English (1)

System		prec %	rec %	$F_{\beta=1}$ %
Baseline	R	54.2	62.2	57.9
	S	26.1	30.0	27.9
TIPSem	R	72.9	73.9	73.4
	S	65.6	66.5	66.1
TIPSem+WN	R	88.7	73.7	80.5
	S	79.9	66.4	72.5

Table 3: TIMEX3 results for English (2)

For English, the Baseline obtains a 57.9% $F_{\beta=1}$ for R, but it falls to 27.9% for S. Nevertheless, TIPSem achieves a 66.1% $F_{\beta=1}$ for S, and TIPSem+WN outperforms the previous two obtaining a 72.5%.

System		ACT	corr	inco	miss	spur
Baseline	R	147	93	0	32	54
	S	147	44	49	32	54
TIPSem	R	144	108	0	17	36
	S	144	102	6	17	36
TIPSem+WN	R	114	107	0	18	7
	S	114	101	6	18	7

Table 4: TIMEX3 results for Spanish (1)

For Spanish, the Baseline obtains a 68.4% $F_{\beta=1}$ for R, but it falls to 32.4% for S. However, TIPSem achieves a 75.8% $F_{\beta=1}$ for S, and TIPSem+WN surpasses them obtaining an 84.5%.

Although both corpora consist of news articles and have a similar TE distribution, English and Spanish results are not strictly comparable due to the difference in size of the corpora. Thus, prior to analyzing the results obtained for different languages, we studied the comparability of the results. The English corpus is approximately 10 times greater in size than the Spanish corpus. For that reason, we created a TimeBank

⁴ <http://fofoca.mitre.org/tern.html#scorer>

System		prec %	rec %	$F_{\beta=1}$ %
Baseline	R	63.3	74.4	68.4
	S	29.9	35.2	32.4
TIPSem	R	75.0	86.4	80.3
	S	70.8	81.6	75.8
TIPSem+WN	R	93.9	85.6	89.5
	S	88.6	80.8	84.5

Table 5: *TIMEX3* results for Spanish (2)

normalized corpus dividing TimeBank corpus into 10 parts whose average statistics are closer to the Spanish (18 docs, 6.1 K words, 140 TIMEX3 and 122 TIMEX3 in TEXT). The approaches have been evaluated over each part and the results have been averaged. The normalized corpus and the complete TimeBank corpus show same quality results with an average difference of 0.47% $F_{\beta=1}$. Therefore, the results can be compared taking into account these numbers. Fig. 4, illustrates the strict span $F_{\beta=1}$ results for the three presented approaches in both evaluated languages.

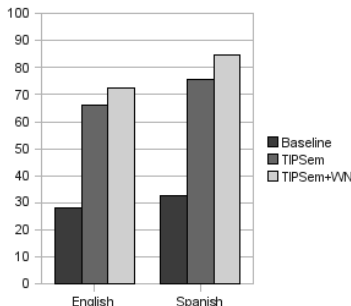


Fig. 4: Strict span $F_{\beta=1}$ results comparison

As shown in the Fig. 4, the results for both languages follow the same pattern and offer similar quality. The Spanish evaluation achieved better results than the English evaluation. This may be because contrary to English, Spanish SRL has been done manually in AnCora corpus.

Results show that, taking TSR as TIMEX3 without any post processing (Baseline), they are reasonably good in the span relaxed identification case, but not in the strict case. However, TIPSem approach obtains much higher results. It indicates that the transformation rules of TIPSem approach have resolved several differences between TSR and TIMEX3.

Focusing on the benefits that the usage of semantic networks introduced to TIPSem approach, we can observe that $F_{\beta=1}$ results have been increased in both languages. The improvement in strict span $F_{\beta=1}$ is a 9.68% for English and a 11.48% for Spanish. This fact indicates that the method defined in this paper accomplishes the objectives for which it was created. TIPSem+WN improves the TIPSem precision via reducing spurious errors produced by the problem described in section 3.2. Moreover, TIPSem+WN does not sacrifice the recall (-0.1% English S, -0.8% Spanish S), see next section (4.3) for details.

There are no strictly comparable results in the literature. Currently, there are no published results for

TIMEX3 identification in Spanish. The closest evaluation is the one done by Boguraev and Ando [2] for English using TimeBank, which is described in section 2. Our approach obtains similar quality results specially in the case of Spanish which has been done over a corpus manually labeled with semantic roles.

4.3 Error analysis

The aim of this section is to show in which aspects is TIPSem+WN failing and analyze the error reduction introduced by semantic networks method.

- **Spurious** (8% EN / 6% ES): False positives have been reduced drastically by the application of the method based on semantic networks defined in this paper, which confirms that the proposed hypothesis is valid for this task. Specifically, it decreases TIPSem spurious errors from 27% to 8% for English and from 25% to 6% for Spanish.

The few errors that remain spurious, apart from SRL errors, are indefinite TEs⁵ (see example 5). The problem is that, although they are indeed TEs, they do not correspond to TIMEX3 elements following the TimeML specifications.

- (5) **EN:** in just a moment
ES: en ese momento⁶

- **Missing** (27% EN / 14% ES): This problem appears because semantic roles not always cover all possibilities of TE in NL.

- The major problem appear in nominal sentences, parenthesis, titles and, in general, all kinds of NL text where verbs are not present. Due to the fact that semantic roles are mainly related to verbs, and semantic networks method is only applied to TSR, TIPSem+WN is not applicable in these sentences (see example 6).

- (6) **EN:** The 1999 results
ES: Tres años en Francia⁷

- Also, cases in which a TE has no temporal function in the sentence (i.e., *Agent* role) but it is a TIMEX3 (see example 7). Semantic networks have not been applied to roles other than TSR, because the ambiguity would introduce noise, for example, in proper nouns like “*Doris Day*”.

- (7) **EN:** [He A0][spent V][6 days A3]
ES: [Estas semanas A0][fueron V][nefastas A1]⁸

- Very few correct TEs obtained by TIPSem have been missed by TIPSem+WN approach, which indicates that the used semantic networks are enough complete in temporal information relations to satisfy this task needs. Example 8 shows the unique cases.

⁵ TEs with an indefinite temporal value (a moment, a while,...)
⁶ at that time

⁷ Three years in France

⁸ These weeks were terrible

- (8) **EN1**: nineteen seventy-nine
- EN2**: as soon as possible
- EN3**: second (adjective)
- ES**: el primer cuatrimestre⁹

– Minor problems are caused by SRL errors.

- **Incorrect** (6% EN / 5% ES): Span errors are mostly produced by SRL errors. For example, in “[10 p.m]. [Wednesday]” the error is produced because the last period of “10 p.m.” has been interpreted as a sentence separation mark.

5 Conclusions

This paper studied the application of semantic networks to the identification of temporal expressions from semantic roles following TimeML specifications. For this purpose, two approaches have been defined (1) TIPSem, which does not use semantic networks, and (2) TIPSem+WN using them. They both, together with a Baseline, have been evaluated in TIMEX3 identification for English and Spanish.

The TIPSem+WN approach obtained a 80.5% and 89.5% $F_{\beta=1}$ for English and Spanish respectively. This means a significant improvement over the Baseline, and an important improvement over TIPSem approach (S $F_{\beta=1}$: +9.68% English and +11.48% Spanish).

The results and errors analysis have confirmed that semantic networks usage produces a reduction of spurious values, but not an increment of missing ones. In this manner, the precision has been increased and the recall maintained, producing a final $F_{\beta=1}$ increase.

The results for both languages follow the same pattern and offer similar quality, facing equivalent error percentages and types. Hence, we can confirm that the approach is valid for English and Spanish. Due to the fact the presented approach is based on semantic roles and multilingual semantic networks information, it could be valid also for other European languages that share several features at this level.

The results lead us to propose potential applications as further work. On the one hand, taking into account that same quality results have been obtained for English and Spanish using the same approach, this study will be extended to other languages to confirm if the analyzed hypothesis could be considered multilingual. On the other hand, due to the lack of TimeML corpora, it will be analyzed if the presented study could be exploited as part of a semi-automatic process of building TimeML corpora for other languages.

Acknowledgments. This paper has been supported by the Spanish Government, project TIN-2006-15265-C06-01, where Llorens is funded (BES-2007-16256).

References

[1] J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of ACM*, 26(11):832–843, 1983.

[2] B. Boguraev and R. K. Ando. Effective Use of TimeBank for TimeML Analysis. In F. Schilder, G. Katz, and J. Pustejovsky, editors, *Annotating, Extracting and Reasoning about Time and Events, International Seminar*, volume 4795 of *LNAI, LNCS*, pages 41–58. Springer, 2007.

[3] C. Fellbaum. *WordNet: An Electronic Lexical Database*. The MIT Press, May 1998.

[4] L. Ferro, L. Gerber, I. Mani, B. Sundheim, and G. Wilson. TIDES 2005 Standard for the Annotation of Temporal Expressions. Technical report, MITRE, 2005.

[5] D. Gildea and D. Jurafsky. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288, 2002.

[6] C. Hagège and X. Tannier. XRCE-T: XIP temporal module for TempEval campaign. In *TempEval (SemEval)*, pages 492–495, Prague, Czech Republic, 2007. ACL.

[7] P. Koomen, V. Punyakanok, D. Roth, and W.-t. Yih. Generalized inference with multiple semantic role labeling systems (shared task paper). In *CoNLL-2004*, pages 181–184, 2005.

[8] I. Mani, J. Pustejovsky, and R. J. Gaizauskas, editors. *The Language of Time: A Reader*. Oxford University Press, 2005.

[9] M. A. Martí, M. Taulé, L. Márquez, and M. Bertran. Anotación semiautomática con papeles temáticos de los corpus CESS-ECE. In *Sociedad Española para el Procesamiento del Lenguaje Natural*, volume 38, 2007.

[10] T. Moia. Telling apart temporal locating adverbials and time-denoting expressions. In *Proceedings of the workshop on Temporal and Spatial information processing*, pages 1–8, NJ, USA, 2001. ACL.

[11] M. Negri and L. Marseglia. Recognition and Normalization of Time Expressions: ITC-irst at TERN 2004. Technical report, Information Society Technologies, 2004.

[12] J. Pustejovsky. TERQAS: Time and Event Recognition for Question Answering Systems. In *ARDA Workshop*, 2002.

[13] J. Pustejovsky, J. M. Castaño, R. Ingria, R. Saurí, R. J. Gaizauskas, A. Setzer, and G. Katz. TimeML: Robust Specification of Event and Temporal Expressions in Text. In *IWCS-5, 5th Int. Workshop on Computational Semantics*, 2003.

[14] J. Pustejovsky, P. Hanks, R. Saurí, A. See, R. J. Gaizauskas, A. Setzer, D. R. Radev, B. Sundheim, D. Day, L. Ferro, and M. Lazo. The TIMEBANK Corpus. In *Corpus Linguistics*, pages 647–656, 2003.

[15] J. Pustejovsky, I. Mani, Belanger, B. Boguraev, Knippen, Litman, Rumshisky, See, Symonen, and Guilder. ARDA summer workshop on graphical annotation toolkit for TimeML. Technical report, MITRE, 2003.

[16] E. Saquete, P. Martínez-Barco, and R. Muñoz. Automatic Multilinguality for Time Expression Resolution. In *MICAI*, volume 2972 of *LNCS*, pages 458–467, 2004.

[17] F. Schilder, G. Katz, and J. Pustejovsky. *Annotating, Extracting and Reasoning About Time and Events*, volume 4795 of *LNCS*. Springer, 2007.

[18] A. Setzer and R. Gaizauskas. Annotating Events and Temporal Information in Newswire Texts. In *LREC 2000*, pages 1287–1294, Athens, 2000.

[19] M. Taulé, M. A. Martí, and M. Recasens. AnCorra: Multilevel Annotated Corpora for Catalan and Spanish. In ELRA, editor, *LREC*, Marrakech, Morocco, 2008.

[20] TERN-2004. Time Expression Recognition and Normalization Evaluation Workshop (<http://fofoca.mitre.org/>), 2004.

[21] M. Verhagen, R. J. Gaizauskas, M. Hepple, F. Schilder, 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, Prague, 2007. ACL.

[22] M. Verhagen, I. Mani, R. Saurí, R. Knippen, S. B. Jang, J. Littman, A. Rumshisky, J. Phillips, and J. Pustejovsky. Automating temporal annotation with TARSQI. In *ACL*, pages 81–84, NJ, USA, 2005. ACL.

[23] P. Vossen, editor. *EuroWordNet: a multilingual database with lexical semantic networks*. Kluwer Academic Publishers, MA, USA, 1998.

[24] G. Wilson, I. Mani, B. Sundheim, and L. Ferro. A multilingual approach to annotating and extracting temporal information. In *Proceedings of the workshop on Temporal and Spatial information processing*, pages 1–7, NJ, USA, 2001. ACL.

⁹ the first four-month period