Multilingual Extension of a Temporal Expression Normalizer using Annotated Corpora

E. Saquete   P. Martínez-Barco   R. Muñoz
gPLSI
DLSI. UA
Alicante, Spain
{stela,patricio,rafael}@dlsi.ua.es

M. Negri   M. Speranza
ITC-irst
Povo (TN), Italy
{negri,manspera}@itc.it

R. Sprugnoli
CELCT
Trento, Italy
sprugnoli@celct.it

Abstract

This paper presents the automatic extension to other languages of TERSEO, a knowledge-based system for the recognition and normalization of temporal expressions originally developed for Spanish[1]. TERSEO was first extended to English through the automatic translation of the temporal expressions. Then, an improved porting process was applied to Italian, where the automatic translation of the temporal expressions from English and from Spanish was combined with the extraction of new expressions from an Italian annotated corpus. Experimental results demonstrate how, while still adhering to the rule-based paradigm, the development of automatic rule translation procedures allowed us to minimize the effort required for porting to new languages. Relying on such procedures, and without any manual effort or previous knowledge of the target language, TERSEO recognizes and normalizes temporal expressions in Italian with good results (72% precision and 83% recall for recognition).

1   Introduction

Recently, the Natural Language Processing community has become more and more interested in developing language independent systems, in the effort of breaking the language barrier hampering their application in real use scenarios. Such a strong interest in multilinguality is demonstrated by the growing number of international conferences and initiatives placing systems' multilingual/cross-language capabilities among the hottest research topics, such as the European Cross-Language Evaluation Forum[2] (CLEF), a successful evaluation campaign which aims at fostering research in different areas of multilingual information retrieval. At the same time, in the temporal expressions recognition and normalization field, systems featuring multilingual capabilities have been proposed. Among others, (Moia, 2001; Wilson et al., 2001; Negri and Marseglia, 2004) emphasized the potentialities of such applications for different information retrieval related tasks.

As many other NLP areas, research in automated temporal reasoning has recently seen the emergence of machine learning approaches trying to overcome the difficulties of extending a language model to other languages (Carpenter, 2004; Ittycheriah et al., 2003). In this direction, the outcomes of the first Time Expression Recognition and Normalization Workshop (TERN 2004[3]) provide a clear indication of the state of the field. In spite of the good results obtained in the *recognition* task, *normalization* by means of machine learning techniques still shows relatively poor results with respect to rule-based approaches, and still remains an unresolved problem.

The difficulty of porting systems to new languages (or domains) affects both rule-based and machine learning approaches. With rule-based approaches (Schilder and Habel, 2001; Filatova and Hovy, 2001), the main problems are related to the fact that the porting process requires rewriting from scratch, or adapting to each new language, large numbers of rules, which is costly and time-

---

[2]http://www.clef-campaign.org/
[3]http://timex2.mitre.org/tern.html

consuming work. Machine learning approaches (Setzer and Gaizauskas, 2002; Katz and Arosio, 2001), on the other hand, can be extended with little human intervention through the use of language corpora. However, the large annotated corpora that are necessary to obtain high performance are not always available. In this paper we describe a new procedure to build temporal models for new languages, starting from previously defined ones. While still adhering to the rule-based paradigm, its main contribution is the proposal of a simple, but effective, methodology to automate the porting of a system from one language to another. In this procedure, we take advantage of the architecture of an existing system developed for Spanish (TERSEO, see (Saquete et al., 2005)), where the recognition model is language-dependent but the normalizing procedure is completely independent. In this way, the approach is capable of automatically learning the recognition model by adjusting the set of normalization rules.

The paper is structured as follows: Section 2 provides a short overview of TERSEO; Section 3 describes the automatic extension of the system to Italian; Section 4 presents the results of our evaluation experiments, comparing the performance of Ita-TERSEO (i.e. our extended system) with the performance of a state of the art system for Italian.

2 The TERSEO system architecture

TERSEO has been developed in order to automatically recognize temporal expressions (TEs) appearing in a Spanish written text, and normalize them according to the temporal model proposed in (Saquete, 2005), which is compatible with the ACE annotation standards for temporal expressions (Ferro et al., 2005). As shown in Figure 1, the first step (recognition) includes pre-processing of the input texts, which are tagged with lexical and morphological information that will be used as input to the temporal parser. The temporal parser is implemented using an ascending technique (chart parser) and is based on a temporal grammar. Once the parser has recognized the TEs in an input text, these are passed to the normalization unit, which updates the value of the reference according to the date they refer to, and generates the XML tags for each expression.

As TEs can be categorized as explicit and implicit, the grammar used by the parser is tuned for discriminating between the two groups. On the
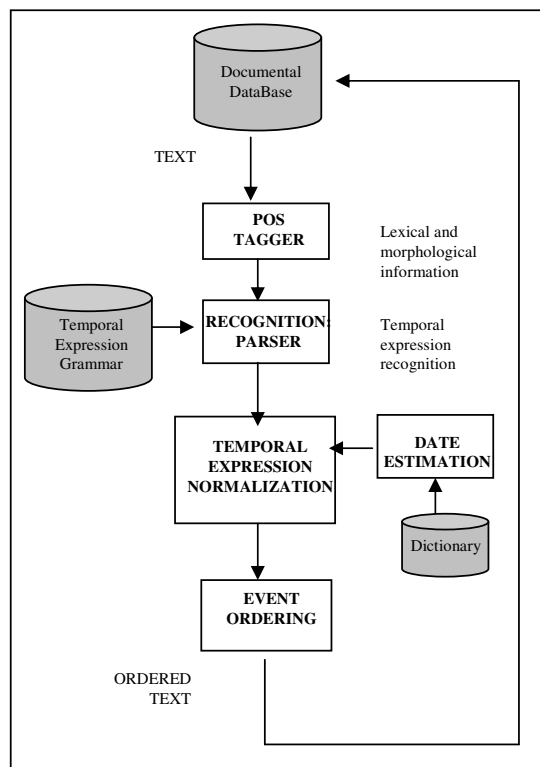

Figure 1: Graphic representation of the TERSEO architecture.

one hand, *explicit* temporal expressions directly provide and fully describe a date which does not require any further reasoning process to be interpreted (e.g. "1st May 2005", "05/01/2005"). On the other hand, *implicit* (or anaphoric) time expressions (e.g. "yesterday", "three years later") require some degree of reasoning (as in the case of anaphora resolution). In order to translate such expressions into explicit dates, such reasoning capabilities consider the information provided by the lexical context in which they occur (see (Saquete, 2005) for a thorough description of the reasoning techniques used by TERSEO).

2.1 Recognition using a temporal expression parser

The parser uses a grammar based on two different sets of rules. The first set of rules is in charge of date and time recognition (i.e. explicit dates, such as "05/01/2005"). For this type of TEs, the grammar adopted by TERSEO recognizes a large number of date and time formats (see Table 1 for some examples).

The second set of rules is in charge of the recognition of the temporal reference for implicit TEs,

```
fecha→ dd+'/'+mm+'/'+(yy)yy          (12/06/1975)
                                     (06/12/1975)
fecha→ dd+'-'+mes+'-'+(yy)yy         (12-junio-1975)
                                     (12-Jun.-1975)
fecha→ dd+'de'+mm+'de'+(yy)yy (12 de junio de 1975)
```

Table 1: Sample of rules for Explicit Dates Recognition.

Implicit dates referring to Document Date Concrete	reference→ 'ayer' *(yesterday)* reference→ 'mañana' *(tomorrow)* reference→ 'anteayer' *(the day before yesterday)* reference→ 'el próximo día' *(the next day)*
Implicit Dates Previous Date Period	reference→ 'un mes después' *(a month later)* reference→ num+'años después'*(num years later)*
Imp. Dates Prev.Date Concrete	reference→ 'un día antes' *(a day before)*
Implicit Dates Previous Date Fuzzy	reference→ 'días después' *(some days later)* reference→ 'días antes' *(some days before)*

Table 2: Sample of rules for Implicit Dates recognition.

i.e. TEs that need to be related to an explicit TE to be interpreted. These can be divided into time adverbs (e.g. "yesterday", "tomorrow"), and nominal phrases that are referring to temporal relationships (e.g. "three years later", "the day before"). Table 2 shows some of the rules used for the detection of these kinds of references.

2.2 Normalization

When the system finds an explicit temporal expression, the normalization process is direct as no resolution of the expression is necessary. For implicit expressions, an inference engine that interprets every reference previously found in the input text is used. In some cases references are solved using the newspaper's date (FechaP). Other TEs have to be interpreted by referring to a date named before in the text that is being analyzed (FechaA). In these cases, a temporal model that allows the system to determine the reference date over which the dictionary operations are going to be done, has been defined. This model is based on the following two rules:

1. The newspaper's date, when available, is used as a base temporal referent by default; otherwise, the current date is used as anchor.

2. In case a non-anaphoric TE is found, it is stored as FechaA. This value is updated every time a non-anaphoric TE appears in the text.

Table 3 shows some of the entries of the dictionary used in the inference engine.

3   Extending TERSEO: from Spanish and English to Italian

As stated before, the main purpose of this paper is to describe a new procedure to automatically build temporal models for new languages, starting from previously defined models. In our case, an English model has been automatically obtained from the Spanish one through the automatic translation of the Spanish temporal expressions to English. The resulting system for the recognition and normalization of English TEs obtains good results both in terms of precision (P) and recall (R) (Saquete et al., 2004). The comparison of the results between the Spanish and the English system is shown in Table 4.

	SPANISH	ENGLISH
DOCS	50	100
POS	199	634
ACT	156	511
CORRECT	138	393
INCORRECT	18	118
MISSED	43	123
P	88%	77%
R	69%	62%
F	77%	69%

Table 4: Comparison between Spanish TERSEO and English TERSEO.

This section presents the procedure we followed to extend our system to Italian, starting from the Spanish and English models already available, and a manually annotated corpus. In this case, both models have been considered as they can be complemented reciprocally. The Spanish model was

REFERENCE	DICTIONARY ENTRY
'ayer' *(yesterday)*	`Day(FechaP)-1/Month(FechaP)/Year(FechaP)`
'mañana' *(tomorrow)*	`Day(FechaP)+1/Month(FechaP)/Year(FechaP)`
'anteayer' *(the day before yesterday)*	`Day(FechaP)-2/Month(FechaP)/Year(FechaP)`
'el próximo día' *(the next day)*	`Day(FechaP)+1/Month(FechaP)/Year(FechaP)`
'un mes después' *(a month later)*	`[DayI/Month(FechaA)+1/Year(FechaA)--` `    DayF/Month(FechaA)+1/Year(FechaA)]`
num+'años después' *(num years later)*	`[01/01/Year(FechaA)+num --` `              31/12/Year(FechaA)+num]`
'un día antes' *(a day before)*	`Day(FechaA)-1/Month(FechaA)/Year(FechaA)`
'días después' *(some days later)*	`>>>>FechaA`
'días antes' *(some days before)*	`<<<<FechaA`

Table 3: Normalization rules

manually obtained and evaluated showing high scores for precision (88%), so better results could be expected when it is used. However, in spite of the fact that the English model has shown lower results on precision (77%), the on-line translators between Italian and English have better results than Spanish to Italian translators. As a result, both models are considered in the following steps for the multilingual extension:

- Firstly, a set of Italian temporal expressions is extracted from an Italian annotated corpus and stored in a database. The selected corpus is the training part of I-CAB, the Italian Content Annotation Bank (Lavelli et al., 2005). More detailed information about I-CAB is provided in Section 4.

- Secondly, the resulting set of Italian TEs must be related to the appropriate normalization rule. In order to do that, a double translation procedure has been developed. We first translate all the expressions into English and Spanish simultaneously; then, the normalization rules related to the translated expressions are obtained. If both the Spanish and English expressions are found in their respective models in agreement with the same normalization rule, then this rule is also assigned to the Italian expression. Also, when only one of the translated expressions is found in the existing models, the normalization rule is assigned. In case of discrepancies, i.e. if both expressions are found, but not coinciding in the same normalization rule, then one of the languages must be prioritized. As the Spanish model was manually obtained and has shown a higher precision, Spanish rules are preferred. In other cases, the expression is reserved for a manual assignment.

- Finally, the set is automatically augmented using the Spanish and English sets of temporal expressions. These expressions were also translated into Italian by on-line machine translation systems (Spanish-Italian[4], English-Italian[5]). In this case, a filtering module is used to guarantee that all the expressions were correctly translated. This module searches the web with Google[6] for the translated expression. If the expression is not frequently found, then the translation is abandoned. After that, the new Italian expression is included in the model, and related to the same normalization rule assigned to the Spanish or English temporal expression.

The entire translation process has been completed with an automatic generalization process, oriented to obtain generalized rules from the concrete cases that have been collected from the cor-

---

[4] http://www.tranexp.com:2000/Translate/result.shtml
[5] http://world.altavista.com/
[6] http://www.google.com/

pus. This generalization process has a double effect. On the one hand, it reduces the number of recognition rules. On the other hand, it allows the system to identify new expressions that were not previously learned. For instance, the expression *"Dieci mesi dopo"* (i.e. "Ten months later") could be recognized if the expression *"Nove mesi dopo"* (i.e. Nine months later) was learned.

The multilingual extension procedure (Figure 3) is carried out in three phases:
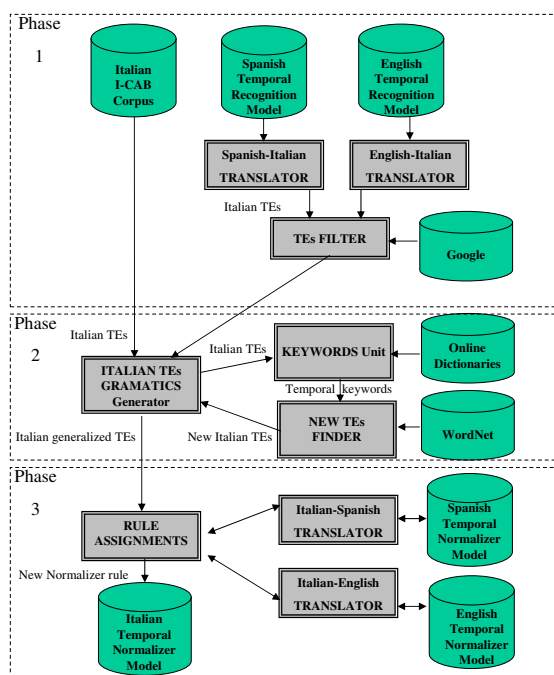

Figure 2: Multilingual extension procedure.

- Phase 1: TE Collection. During this phase, the Italian temporal expressions are collected from I-CAB (Italian Content Annotation Bank), and the automatically translated Italian TEs are derived from the set of Spanish and English TEs. In this case, the TEs are filtered removing those not being found by Google.

- Phase 2: TE Generalization. In this phase, the TEs Gramatics Generator uses the morphological and syntactical information from the collected TEs to generate the grammatical rules that generalize the recognition of the TEs. Moreover, the keyword unit is able to extract the temporal keywords that will be used to build new TEs. These keywords are augmented with their synonyms in WordNet (Vossen, 2000) to generate new TEs.

- Phase 3: TE Normalizing Rule Assignment. In the last phase, the translators are used to relate the recognizing rule to the appropriate normalization rule. For this purpose, the system takes advantage of the previously defined Spanish and English temporal models.

4 Evaluation

The automatic extension of the system to Italian (Ita-TERSEO) has been evaluated using I-CAB, which has been divided in two parts: training and test. The training part has been used, first of all, in order to automatically extend the system. After this extension, the system was evaluated both against the training and the test corpora. The purpose of this double evaluation experiment was to compare the recall obtained over the training corpus with the value obtained over the test corpus.

An additional evaluation experiment has also been carried out in order to compare the performance of the automatically developed system with a state of the art system specifically developed for Italian and English, i.e. the Chronos system described in (Negri and Marseglia, 2004).

In the following sections, more details about I-CAB and the evaluation process are presented, together with the evaluation results.

4.1 The I-CAB Corpus

The evaluation has been performed on the temporal annotations of I-CAB (I-CAB-temp) created as part of the three-year project ONTOTEXT[7] funded by the Provincia Autonoma di Trento.

I-CAB consists of 525 news documents taken from the local newspaper L'Adige (http://www.adige.it). The selected news stories belong to four different days (September, 7th and 8th 2004 and October, 7th and 8th 2004) and are grouped into five categories: News Stories, Cultural News, Economic News, Sports News and Local News. The corpus consists of around 182,500 words (on average 347 words per file). The total number of annotated temporal expressions is 4,553; the average length of a temporal expression is 1.9 words.

The annotation of I-CAB has been carried out adopting the standards developed within the ACE program (Automatic Content Extraction[8]) for the Time Expressions Recognition and Normalization

---

tasks, which allows for a semantically rich and normalized annotation of different types of temporal expressions (for further details on the TIMEX2 annotation standard for English see (Ferro et al., 2005)).

The ACE guidelines have been adapted to the specific morpho-syntactic features of Italian, which has a far richer morphology than English. In particular, some changes concerning the extension of the temporal expressions have been introduced. According to the English guidelines, in fact, definite and indefinite articles are considered as part of the textual realization of an entity, while prepositions are not. As the annotation is word-based, this does not account for Italian articulated prepositions, where a definite article and a preposition are merged. Within I-CAB, this type of preposition has been included as possible constituents of an entity, so as to consistently include all the articles.

An assessment of the inter-annotator agreement based on the Dice coefficient has shown that the task is a well-defined one, as the agreement is 95.5% for the recognition of temporal expressions.

4.2 Evaluation process

The evaluation of the automatic extension of TERSEO to Italian has been performed in three steps. First of all, the system has been evaluated both against the training and the test corpora with two main purposes:

- Determining if the recall obtained in the evaluation of the training part of the corpus is a bit higher than the one obtained in the evaluation of the test part of I-CAB, due to the fact that in the TE collection phase of the extension, temporal expressions were extracted from this part of the corpus.

- Determining the performance of the automatically extended system without any manual revision of both the Italian translations and the resolution rules automatically related to the expressions.

Secondly, we were also interested in verifying if the performance of the system in terms of precision could be improved through a manual revision of the automatically translated temporal expressions.

Finally, a comparison with a state of the art system for Italian has been carried out in order to es-timate the real potentialities of the proposed approach. All the evaluation results are compared and presented in the following sections using the same metrics adopted at the TERN2004 conference.

4.2.1 Evaluation of Ita-TERSEO

In the automatic extension of the system, a total of 1,183 Italian temporal expressions have been stored in the database. As shown in Table 5, these expressions have been obtained from the different resources available:

- ENG_ITA: This group of expressions has been obtained from the automatic translation into Italian of the English Temporal Expressions stored in the knowledge DB.

- ESP_ITA: This group of expressions has been obtained from the automatic translation into Italian of the Spanish Temporal Expressions stored in the knowledge DB.

- CORPUS: This group of expressions has been extracted directly from the training part of the I-CAB corpus.

Source	N	%
ENG_ITA	593	50.1
ESP_ITA	358	30.3
CORPUS	232	19.6
TOTAL TEs	1183	100.0

Table 5: Italian TEs in the Knowledge DB.

Both the training part and the test part of I-CAB have been used for evaluation. The results of precision (P), recall (R) and F-Measure (F) are presented in Table 6, which provides details about the system performance over the general recognition task (timex2), and the different normalization attributes used by the TIMEX2 annotation standard.

As expected, recall performance over the training corpus is slightly higher. However, although the temporal expressions have been extracted from such corpus, in the automatic process of obtaining the normalization rules for these expressions, some errors could have been introduced.

Comparing these results with those obtained by the automatic extension of TERSEO to English and taking into account the recognition task (see Table 4), precision (P) is slightly better for English (77% Vs. 72%) whereas recall (R) is better in the Italian extension (62% Vs. 83%). This is

Tag	Ita-TERSEO: TRAINING			Ita-TERSEO: TEST			Chronos: TEST		
	P	R	F	P	R	F	P	R	F
timex2	0.694	0.848	0.763	0.726	0.834	0.776	0.925	0.908	0.917
anchor_dir	0.495	0.562	0.526	0.578	0.475	0.521	0.733	0.636	0.681
anchor_val	0.464	0.527	0.493	0.516	0.424	0.465	0.495	0.462	0.478
set	0.308	0.903	0.459	0.182	1.000	0.308	0.616	0.5	0.552
text	0.265	0.324	0.292	0.258	0.296	0.276	0.859	0.843	0.851
val	0.581	0.564	0.573	0.564	0.545	0.555	0.636	0.673	0.654

Table 6: Results obtained over I-CAB by Ita-TERSEO and Chronos.

due to the fact that in the Italian extension, more temporal expressions have been covered with respect to the English extension. In this case, in fact, Ita-TERSEO is not only using the temporal expressions translated from the English or Spanish knowledge database, but also the temporal expressions extracted from the training part of I-CAB.

4.2.2 Manual revision of the acquired TEs

A manual revision of the Italian TEs stored in the Knowledge DB has been done in two steps. First of all, the incorrectly translated expressions (from Spanish and English to Italian) were removed from the database. A total of 334 expressions were detected as wrong translated expressions. After this, another revision was performed. In this case, some expressions were modified because the expressions have some minor errors in the translation. 213 expressions were modified in this second revision cycle. Moreover, since pattern constituents in Italian might have different orthographical features (e.g. masculine/feminine, initial vowel/consonant, etc.), new patterns had to be introduced to capture such variants. For example, as months' names in Italian could start with a vowel, the temporal expression pattern "nell'-MONTH" has been inserted in the Knowledge DB. After these changes, the total amount of expressions stored in the DB are shown in Table 7.

Source	N	%
ING_ITA	416	47.9
ESP_ITA	201	23.1
CORPUS	232	26.7
REV_MAN	20	2.3
TOTAL TEs	869	100.0

Table 7: Italian TEs in the Knowledge DB after manual revision.

In order to evaluate the system after this manual revision, the training and the test part of I-CAB have been used. However, the results of precision (PREC), recall (REC) and F-Measure were exactly the same as presented in Table 6. That is not really surprising. The existence of wrong expressions in the knowledge database does not affect the final results of the system, as they will never be used for recognition or resolution. This is because these expressions will not appear in real documents, and are redundant as the correct expression is also stored in the Knowledge DB.

4.2.3 Comparing Italian TERSEO with a language-specific system

Finally, in order to compare Ita-TERSEO with a state of the art system specifically designed for Italian, we chose Chronos (Negri and Marseglia, 2004), a multilingual system for the recognition and normalization of TEs in Italian and English. Like all the other state of the art systems addressing the recognition/normalization task, Chronos is a rule-based system. From a design point of view, it shares with TERSEO a rather similar architecture which relies on different sets of rules. These are regular expressions that check for specific features of the input text, such as the presence of particular word senses, lemmas, parts of speech, symbols, or strings satisfying specific predicates. Each set of rules is in charge of dealing with different aspects of the problem. In particular, a set of around 350 rules is designed for TE recognition and is capable of recognizing with high Precision/Recall rates both explicit and implicit TEs. Other sets of regular expressions, for a total of around 700 rules, are used in the normalization phase, and are in charge of handling a specific TIMEX2 attribute (i.e. VAL, SET, ANCHOR_VAL, and ANCHOR_DIR). The results obtained by the Italian version of Chronos over the test part of I-CAB are shown in the last three columns of Table 6.

As expected, the distance between the results obtained by the two systems is considerable. However, the following considerations should be taken into account. First, there is a great difference, both

in terms of the required time and effort, in the development of the two systems. While the implementation of the manual one took several months, the porting procedure of TERSEO to Italian is a very fast process that can be accomplished in less than an hour. Second, even if an annotated corpus for a new language is not available, the automatic porting procedure we present still remains feasible. In fact, most of the TEs for a new language that are stored in the Knowledge DB are the result of the translation of the Spanish/English TEs into such a target language. In our case, as shown in Table 5, more than 80% of the acquired Italian TEs result from the automatic translation of the expressions already stored in the DB. This makes the proposed approach a viable solution which allows for a rapid porting of the system to other languages, while just requiring an on-line translator (note that the Altavista Babel Fish translator[9] provides translations from English to 12 target languages). In light of these considerations, the results obtained by Ita-TERSEO are encouraging.

5 Conclusions

In this paper we have presented an automatic extension of a rule-based approach to TEs recognition and normalization. The procedure is based on building temporal models for new languages starting from previously defined ones. This procedure is able to fill the gap left by machine learning systems that, up to date, are still far from providing acceptable performance on this task. As results illustrate, the proposed methodology (even though with a lower performance with respect to language-specific systems) is a viable and effective solution for a rapid and automatic porting of an existing system to new languages.

References

B. Carpenter. 2004. Phrasal Queries with LingPipe and Lucene. In *13th Text REtrieval Conference*, NIST Special Publication. National Institute of Standards and Technology.

L. Ferro, L. Gerber, I. Mani, B. Sundheim, and G. Wilson. 2005. TIDES 2005 Standard for the annotation of temporal expressions. Technical report, MITRE.

E. Filatova and E. Hovy. 2001. Assigning time-stamps to event-clauses. In ACL, editor, *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*, pages 88–95, Toulouse, France.

A. Ittycheriah, L.V. Lita, N. Kambhatla, N. Nicolov, S. Roukos, and M. Stys. 2003. Identifying and Tracking Entity Mentions in a Maximum Entropy Framework. In ACL, editor, *Proceedings of the NAACL Workshop WordNet and Other Lexical Resources: Applications, Extensions and Customizations*.

G. Katz and F. Arosio. 2001. The annotation of temporal information in natural language sentences. In ACL, editor, *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*, pages 104–111, Toulouse, France.

A. Lavelli, B. Magnini, M. Negri, E. Pianta, M. Speranza, and R. Sprugnoli. 2005. Italian Content Annotation Bank (I-CAB): Temporal expressions (v. 1.0.): T-0505-12. Technical report, ITC-irst, Trento.

T. Moia. 2001. Telling apart temporal locating adverbials and time-denoting expressions. In ACL, editor, *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*, Toulouse, France.

M. Negri and L. Marseglia. 2004. Recognition and normalization of time expressions: Itc-irst at TERN 2004. Technical report, ITC-irst, Trento.

E. Saquete, P. Martínez-Barco, and R. Muñoz. 2004. Evaluation of the automatic multilinguality for time expression resolution. In *DEXA Workshops*, pages 25–30. IEEE Computer Society.

E. Saquete, R. Muñoz, and P. Martínez-Barco. 2005. Event ordering using terseo system. *Data and Knowledge Engineering Journal*, page (To be published).

E. Saquete. 2005. *Temporal information Resolution and its application to Temporal Question Answering*. Phd, Departamento de Lenguages y Sistemas Informáticos. Universidad de Alicante, June.

F. Schilder and C. Habel. 2001. From temporal expressions to temporal information: Semantic tagging of news messages. In ACL, editor, *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*, pages 65–72, Toulouse, France.

A. Setzer and R. Gaizauskas. 2002. On the importance of annotating event-event temporal relations in text. In LREC, editor, *Proceedings of the LREC Workshop on Temporal Annotation Standards, 2002*, pages 52–60, Las Palmas de Gran Canaria,Spain.

P. Vossen. 2000. EuroWordNet: Building a Multilingual Database with WordNets in 8 European Languages. *The ELRA Newsletter*, 5(1):9–10.

G. Wilson, I. Mani, B. Sundheim, and L. Ferro. 2001. A multilingual approach to annotating and extracting temporal information. In ACL, editor, *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*, pages 81–87, Toulouse, France.

---

[9]http://world.altavista.com/