Annotating Events, Temporal Expressions and Relations in Italian: the It-TimeML Experience for the Ita-TimeBank

Tommaso Caselli
ILC "A.Zampolli" - CNR
Via G. Moruzzi, 1
56124 Pisa
caselli@ilc.cnt.it

Valentina Bartalesi Lenzi
CELCT
Via della Cascata 56/c
38123 Povo (TN)
bartalesi@celct.it

Rachele Sprugnoli
CELCT
Via della Cascata 56/c
38123 Povo (TN)
sprugnoli@celct.it

Emanuele Pianta
CELCT
Via della Cascata 56/c
38123 Povo (TN)
pianta@fbk.eu

Irina Prodanof
ILC "A.Zampolli" - CNR
Via G. Moruzzi, 1
56124 Pisa
prodanof@ilc.cnr.it

Abstract

This paper presents the *annotation guidelines* and *specifications* which have been developed for the creation of the Italian TimeBank, a language resource composed of two corpora manually annotated with temporal and event information. In particular, the adaptation of the TimeML scheme to Italian is described, and a special attention is given to the methodology used for the realization of the annotation specifications, which are strategic in order to create good quality annotated resources and to justify the annotated items. The reliability of the It-TimeML guidelines and specifications is evaluated on the basis of the results of the inter-coder agreement performed during the annotation of the two corpora.

1 Introduction

In recent years a renewed interest in temporal processing has spread in the NLP community, thanks to the success of the TimeML annotation scheme (Pustejovsky et al., 2003a) and to the availability of annotated resources, such as the English and French TimeBanks (Pustejovsky et al., 2003b; Bittar, 2010) and the TempEval corpora (Verhagen et al., 2010).

The ISO TC 37 / SC 4 initiative ("Terminology and other language and content resources") and the TempEval-2 contest have contributed to the development of TimeML-compliant annotation schemes in languages other than English, namely Spanish, Korean, Chinese, French and Italian. Once the corresponding corpora will be completed and made available, the NLP community will benefit from having access to different language resources with a common layer of annotation which could boost studies in multilingual temporal processing and improve the performance of complex multilingual NLP systems, such as Question-Answering and Textual Entailment.

This paper focuses on the annotation guidelines and specifications which have been developed for the creation of the Italian TimeBank (hereafter, Ita-TimeBank). The distinction between *annotation guidelines* and *annotation specifications* is of utmost importance in order to distinguish between the abstract, formal definition of an annotation scheme and the actual realization of the annotated language resource. In addition to this, documenting the annotation specification facilitates the reduplication of annotations and justify the annotated items.

The paper is organized as follows: Section 2 will describe in detail specific issues related to the temporal annotation of Italian for the two main tags of the TimeML annotation scheme,

Proceedings of the Fifth Law Workshop (LAW V), pages 143–151,
Portland, Oregon, 23-24 June 2011.

namely <EVENT> and <TIMEX3>. Section 3 will present the realization of the annotation specifications and will document them. Section 4 focuses on the evaluation of the annotation scheme on the Ita-TimeBank, formed by two corpora independently realized by applying the annotation specifications. Finally, in Section 5 conclusions and extensions to the current annotation effort will be reported.

Notice that, for clarity's sake, in this paper the examples will focus only on the tag (or attribute or link) under discussion.

2 It-TimeML: Extensions and Language Specific Issues

Applying an annotation scheme to a language other than the one for which it was initially developed, requires a careful study of the language specific issues related to the linguistic phenomena taken into account (Im et al., 2009; Bittar, 2008).

TimeML focuses on *Events* (i.e. actions, states, and processes - <EVENT> tag), *Temporal Expressions* (i.e. durations, calendar dates, times of day and sets of time - <TIMEX3> tag), *Signals* (e.g. temporal prepositions and subordinators - <SIGNAL> tag) and various kind of *dependencies between Events and/or Temporal Expressions* (i.e. temporal, aspectual and subordination relations - <TLINK>, <ALINK> and <SLINK> tags respectively).

An ISO language-independent specification of TimeML is under development but it is still in the *enquiry stage*[1]. For this reason, in the following subsections we will mostly compare the Italian annotation guidelines with the latest version of the English annotation guidelines (TimeML Working group, 2010), focusing on the two main tags, i.e <EVENT> and <TIMEX3>, in Italian.

2.1 The <EVENT> tag

The <EVENT> tag is used to mark-up instances of eventualities (Bach, 1986). This category comprises all types of actions (punctual or durative) and states as well. With respect to previous annotations schemes (Katz and Arosio, 2001, Filatova and Hovy, 2001, Setzer and Gaizauskas, 2001 among other), TimeML allows for annotating as Events not only verbs but also nouns, adjectives and prepositional phrases.

In the adaptation to Italian, two annotation principles adopted for English, that is an orientation towards surface linguistic phenomena and the notion of minimal chunk for the tag extent, have been preserved without major modifications. The main differences with respect to the English version rely i.) in the attribute list; and ii.) in the attributes values.

In Italian 12 core attributes apply with respect to the 10 attributes in English. The newly introduced attributes are MOOD and VFORM which capture key distinctions of the Tense-Mood-Aspect (TMA) system of the Italian language. These two attributes are common to other languages, such as Spanish, Catalan, French and Korean.

The MOOD attribute captures the contrastive grammatical expression of different modalities of presentation of an Event when realized by a verb. Annotating this attribute is important since grammatical modality has an impact on the identification of temporal and subordinating relations, and on the assessment of veridicity/factivity values. Mood in Italian is expressed as part of the verb morphology and not by means of modal auxiliary verbs as in English (e.g. through the auxiliary “would”),. Thus, the solution to deal with this phenomenon adopted for English TimeML (where the main verb is annotated with the attribute MODALITY=”would”, see below) is not applicable in Italian unless relevant information is lost. The values of the MOOD attribute, as listed below, have been adapted to Italian and extended with respect to those proposed in the ISO-TimeML specification:

- NONE: it is used as the default value and corresponds to the Indicative mood:
  (1.) Le forze dell’ordine hanno <EVENT … mood="NONE"> schierato </EVENT> 3.000 agenti. [The police has deployed 3,000 agents.]

[1] http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37331

- CONDITIONAL: it signals the conditional mood which is used to speak of an Event whose realization is dependent on a certain condition, or to signal the future-in-the-past:
  (2.) <EVENT ... mood="COND"> Mangerei </EVENT> del pesce. [I would eat fish.]

- SUBJUNCTIVE: it has several uses in independent clauses and is required for certain types of dependent clauses.
  (3.) Voglio che tu te ne <EVENT … mood="SUBJUNCTIVE">vada</EVENT> [I want you to go.]

- IMPERATIVE: it is used to express direct commands or requests, to signal a prohibition, permission or any other kind of exhortation.

The attribute VFORM is responsible for distinguishing between non-finite and finite forms of verbal Events. Its values are:

- NONE: it is the default value and signals finite verb forms:
  (4.) Le forze dell'ordine hanno <EVENT … vForm="NONE">schierato</EVENT> 3.000 agenti. [The police has deployed 3,000 agents.]

- INFINITIVE: for infinitive verb forms:
  (5.) Non è possibile <EVENT … vForm="INFINITIVE">viaggiare</EVENT>. [It's not possible to travel.]

- GERUND: for gerundive verb forms:
  (6.) Ha evitato l'incidente <EVENT … vForm="GERUND"> andando </EVENT> piano. [Driving slowly, he avoided the incident.]

- PARTICIPLE: for participle verb forms:
  (7.) <EVENT … vForm="PARTICIPLE"> Vista </EVENT> Maria, se ne andò. [Having seen Maria, he left.]

As for attribute values, the most important changes introduced for Italian in comparison with the English TimeML, are related to the ASPECT and MODALITY attributes.

The ASPECT attribute captures standard distinctions in the grammatical category of aspect or Event viewpoint (Smith, 1991). In English TimeML it has the following values: i.) PROGRESSIVE; ii.) PERFECTIVE; iii.) PERFECTIVE_PROGRESSIVE, or iv.) NONE. The main differences with respect to the English guidelines concern the following points:

i.) the absence of the value PERFECTIVE_PROGRESSIVE and

ii.) the presence of the value IMPERFECTIVE, which is part of the ISO TimeML current definition.

These differences are due to language specific phenomena related to the expression of the grammatical aspect in Italian and English and to the application of the TimeML surface oriented annotation philosophy. In particular, the assignment of the aspectual values is strictly determined by the verb surface forms. For instance, in English the verb form "is teaching" requires the PROGRESSIVE value. On the other hand, the Italian counterpart of "is teaching" can be realized in two ways: either by means of the simple present (*insegna* [s/he teaches]) or by means of a specific verbal periphrasis (*sta insegnando* [s/he is teaching]). In order to distinguish between these two verb forms, and to account also for other typical Romance languages tense forms, such as the Italian *Imperfetto*, the use of the additional IMPERFECTIVE value is necessary. Thus, *insegna* [s/he teaches], as well as the Imperfetto *insegnava* [s/he was teaching] are annotated as IMPERFECTIVE, whereas *sta insegnando* [s/he is teaching] is annotated as PROGRESSIVE. On the other hand, the absence of the PERFECTIVE_PROGRESSIVE value, used for English tense forms of the kind "he has been teaching", is due to the lack of Italian verb surface forms which may require its use.

In English, modal verbs are not annotated as Events and the MODALITY attribute is associated to the main verb (the value of the attribute is the token corresponding to the modal verb). Unlike English modals, Italian modal verbs, such as *potere* [can/could; may/might], *volere* [want; will/would] and *dovere* [must/have to; ought to; shall/should], are to be

considered similar to other lexical verbs in that it is possible to assign them values for tense and aspect. Consequently, each instance of Italian modal verbs will be annotated with the tag <EVENT>. The value of the MODALITY attribute is the lemma of the verb (e.g. *dovere*).

A further language specific aspect concerns the annotation of verbal periphrases, that is special constructions with at least two verbs (and sometimes other words) that behave as a group like a single verb would. In Italian, it is possible to identify different instances of verbal periphrases, namely:

- aspectual periphrases (example 8 below), which encode progressive or habitual aspect;
- modal periphrases (example 9), which encode modality not realized by proper modal verbs;
- phasal periphrases (example 10), which encode information on a particular phase in the description of an Event.

Following Bertinetto (1991), in the last two cases, i.e. modal periphrases and phasal periphrases, both verbal elements involved should be annotated, while in the case of the aspectual periphrasis only the main verb (verb head) has to be marked; e.g.:

(8.) Maria stava <EVENT … ASPECT="PROGRESSIVE"> mangiando. [Maria was eating]

(9.) Il compito di matematica <EVENT ... MODALITY="ANDARE"> va </EVENT> <EVENT ... > svolto </EVENT> per domani. [Maths exercises must be done for tomorrow]

(10.) I contestatori hanno <EVENT ... CLASS="ASPECTUAL"> iniziato </EVENT> a <EVENT> lanciare </EVENT> pietre. [Demonstrators started to throw stones.]

Similarly to what proposed for English, in presence of multi-tokens realization of Events, two main annotation strategies have been followed:

- in case the multi-token Event expression corresponds to an instance of a collocation or of an idiomatic expression, then only the head (verbal, nominal or other) of the expression is marked up;
- in case the multi-token Event is realized by light verb expressions, then two separate <EVENT> tags are to be created both for the verb and the nominal/prepositional complement.

2.2 The <TIMEX3> tag

The TIMEX3 tag relies on and is as much compliant as possible with the TIDES TIMEX2 annotation. The Italian adaptation of this annotation scheme is presented in Magnini et al. (2006). The only difference concerns the annotation of articulated prepositions which are annotated as signals, while in the TIMEX2 specifications they are considered as part of the textual realization of Temporal Expressions:

(11a.) <TIMEX2 …> nel 2011 </TIMEX2> [in 2011]

(11b.) <SIGNAL …> nel </SIGNAL> <TIMEX3…>2011</TIMEX3> [in 2011]

On the other hand, with respect to the TIMEX3 annotation of other languages such as English, we decided to follow the TIMEX2 specification by annotating many adjectives as Temporal Expressions (e.g. *recente* [recent], *ex* [former]) and including modifiers like *che rimane* in *l'anno che rimane* [the remaining year] into the extent of the TIMEX3 tag since it is essential for the normalization of temporal expressions.

3 From Annotation Guidelines to Specifications

As already stated, the annotation guidelines represent an abstract, formal level of description which, in this case, is mainly based on a detailed study of the relevant linguistic levels. Once the guidelines are applied to real language data, further issues arise and need to be tackled. This section focuses on a method for developing annotation specifications. Annotation specifications are to be seen as the actual realization of the annotation guidelines. The identification and distinction of annotation guidelines from annotation specification is of major importance as it is to be conceived as a new level of Best Practice for the creation of

semantically annotated Language Resources (Calzolari and Caselli, 2009).

The process of realization of the annotation specifications is strategic both to realize good quality annotated resources and to justify why certain textual items have to be annotated. As for the It-TimeML experience we will illustrate this process by making reference and reporting examples for two tags, namely for the <EVENT> and the <TLINK> tags.

As a general procedure for the development of the annotation specifications, we have taken inspiration from the DAMSL Manual (Core and Allen, 1997). Different decision trees have been created for each task. For instance, for the annotation of the <EVENT> tag, four different decision trees have been designed for each POS (i.e. nouns, verbs, adjectives and prepositional phrases) which could be involved in the realization of an Event. In particular, the most complex decision tree is that developed for noun annotation. The identification of the eventive reading of nouns has been formalized into a discrimination process of different properties: firstly superficial properties are taken into consideration, i.e. whether a morphologically related verb exists or not, and whether the noun co-occurs with special verb predicates (for instance aspectual verbs such as *iniziare* [to start] or light verbs such as *fare* [to do]); then, deeper semantic properties are analyzed, which involve other levels such as word sense disambiguation and noun classification (e.g. whether the noun is a functional or an incremental one).

Other decision trees have been improved to avoid inconsistencies in Event classification. For instance, the identification of *Reporting Events* showed to be problematic because of the vague definition adopted in the guidelines. A Reporting Event is a giving information speech act in which a communicator conveys a message to an addressee. To help annotators in deciding whether an event is a Reporting one, the annotation specifications suggest to rely on FrameNet as a starting point (Baker, et al. 1998). More specifically, an Italian lexical unit has been classified as Reporting if it is the translation equivalent of one of the lexical units assigned to the Communication frame, which has Message as a core element. Among the frames using and inherited from the Communication frame, only the ones having the Message as a core element and conveying a giving information speech act have been selected and the lexical units belonging to them have been classified as Reporting Events: e.g. *urlare* [to scream] from the Communication_noise frame, *sottolineare* [to stress] from the Convey_importance frame, *dichiarare* [to declare] from the Statement frame.

Similarly, for the identification of TLINKs, a set of decision trees has been developed to identify the conditions under which a temporal relation is to be annotated and a method to decide the value of the *reltype* attribute. For instance, the annotation of temporal relations between nominal Events and Temporal Expressions in the same sentence is allowed only when the Temporal Expression is realized either by an adjective or a prepositional phrase of the form "*di (of)* + TEMPORAL EXPRESSION" e.g.:

(12.) La <EVENT eid="e1" ... > riunione </EVENT> <SIGNAL sid="s1" ... > di </SIGNAL> <TIMEX3 tid="t1" ... > ieri </TIMEX3> [yesterday meeting]
<TLINK lid="l1" eventInstanceID="e01" relatedToTime="t01" signalID="s1" relType="IS_INCLUDED"/>

In addition, decision trees based on the idea that signals provide useful information to TLINK classification have been used to assign the *reltype* value to TLINKs holding between a duration and an Event. For example, the pattern "EVENT + *tra (in)* + DURATION" identifies the value AFTER, while the pattern "EVENT + *per (for)* + DURATION" is associated with the value MEASURE.

(13.) Il pacco <EVENT eid="e1" ... >arriverà </EVENT> <SIGNAL sid="s1" ... > tra </SIGNAL> <TIMEX3 tid="t1" ... > due giorni </TIMEX3> [the package will arrive in two days]
<TLINK lid="l1" eventInstanceID="e1" relatedToTime="t1" signalID="s1" relType="AFTER"/>

(14.) Sono stati <EVENT eid="e1" ... > sposati </EVENT> <SIGNAL sid="s1" ... > per </SIGNAL> <TIMEX3 tid="t1" ... > dieci anni

</TIMEX3> [they have been married for ten years]
<TLINK lid="l1" eventInstanceID="e1" relatedToTime="t1" signalID="s1" relType="MEASURE"/>

The advantages of this formalization are many. The impact of the annotators' subjectivity is limited, thus reducing the risk of disagreement. Moreover, trees can then be easily used either as features for the development of a automatic learner or as instructions in a rule-based automatic annotation system.

4 Evaluating Annotations

Two corpora have been developed in parallel following the It-TimeML annotation scheme, namely the CELCT corpus and the ILC corpus. Once these two corpora will be completed and released, they will form the Italian TimeBank providing the NLP community with the largest resource annotated with temporal and event information (more than 150K tokens).

In this section, the two corpora are briefly described and the results of the inter-coder agreement (Artstein and Poesio, 2008) achieved during their annotation are compared in order to evaluate the quality of the guidelines and of the resources.

The *CELCT corpus* has been created within the LiveMemories project[2] and it consists of news stories taken from the Italian Content Annotation Bank (I-CAB, Magnini et al., 2006). More than 180,000 tokens have been annotated with Temporal Expressions and more than 90,000 tokens have been annotated also with Events, Signals and Links. The Brandeis Annotation Tool[3] (BAT) has been used for the pilot annotation and for the automatic computation of the inter-coder agreement on the extent and the attributes of Temporal Expressions, Events and Signals. After the pilot annotation, the first prototype of the CELCT Annotation Tool (CAT) has been used to perform the annotation and to compute the inter-coder agreement on Links. For what concern the annotation effort, the work on Temporal Expressions, Events and Signals involved 2 annotators while 3 annotators have been engaged in the annotation of Links. The annotation started in January 2010 and required a total of 1.3 person/years. Table 1 shows the total number of annotated markables together with the results of the inter-coder agreement on tag extent performed by two annotators on a subset of the corpus of about four thousand tokens. For the annotation of Event and Signal extents, statistics include average precision and recall and Cohen' kappa, while the Dice Coefficient has been computed for the extent of Links and Temporal Expressions.

Markable	#	Agreement
TIMEX3	4,852	Dice=0.94
EVENT	17,554	K=0.93 P&R=0.94
SIGNAL	2,045	K=0.88 P&R=0.88
TLINK	3,373	Dice=0.86
SLINK	3,985	Dice=0.93
ALINK	238	Dice=0.90

Table 1: Annotated markables and results of the inter-coder agreement on tag extent[4]

Table 2 provides the value of Fleiss' kappa computed for the annotation of Temporal Expression, Event and Link attributes.

Tag and attribute	Agreement-Kappa
TIMEX3.type	1.00
TIMEX3.value	0.92
TIMEX3.mod	0.89
EVENT.aspect	0.96
EVENT.class	0.87
EVENT.modality	1.00
EVENT.mood	0.90
EVENT.polarity	1.00
EVENT.pos	1.00
EVENT.tense	0.94
EVENT.vform	0.98
TLINK.relType	0.88
SLINK.relType	0.93
ALINK.relType	1.00

Table 2: Inter-coder agreement on attributes

[2] http://www.livememories.org

[3] http://www.timeml.org/site/bat/

[4] Please note that the number of annotated Temporal Expressions is calculated on a total of 180,000 tokens, while the number of Events, Signals and Links is calculated on more than 90,000 tokens.

The *ILC corpus* is composed of 171 newspaper stories collected from the Italian Syntactic-Semantic Treebank, the PAROLE corpus and the web for a total of 68,000 tokens (40,398 tokens are freely available, the remaining are available with restrictions). The news reports were selected to be comparable in content and size to the English TimeBank and they are mainly about international and national affairs, political and financial subject. The annotation of Temporal Expressions, Event extents and Signals has been completed while the annotation of Event attributes and LINKs is a work in progress. A subset of the corpus has been used as data set in the TempEval-2 evaluation campaign organized within SemEval-2 in 2010. So far the annotation has been performed thanks to eight voluntary students under the supervision of two judges using BAT. The annotation started in March 2009 and is requiring a total of 3 person/years. Table 3 reports the total number of Temporal Expressions, Events, Signals and TLINKs together with the results of the inter-coder agreement on tag extent performed on about 30,000 tokens. To measure the agreement on tag extents, average precision and recall and Cohen' kappa have been calculated. The annotation of Temporal Links has been divided into three subtasks: the first subtask is the relation between two Temporal Expressions, the second is the relation between an Event and a Temporal Expression, the third regards the relation between two Events.

Markable		#	Agreement
TIMEX3		2,314	K=0.95 P&R= 0.95
EVENT		10,633	K=0.87 P&R= 0.86
SIGNAL		1,704	K=0.83 P&R= 0.84
TLINK	TIMEX3–TIMEX3	353	K=0.95
	EVENT–TIMEX3	512	K=0.87
	EVENT–EVENT	1,014	in progress

Table 3: Annotated markables and results of the inter-coder agreement on tag extent

The values of Fleiss' kappa computed for the assignment of attribute values are illustrated in Table 4.

Tag and attribute	Agreement – Kappa
TIMEX3.type	0.96
TIMEX3.value	0.96
TIMEX3.mod	0.97
EVENT.aspect	0.93
EVENT.class	0.82
EVENT.modality	0.92
EVENT.mood	0.89
EVENT.polarity	0.75
EVENT.pos	0.95
EVENT.tense	0.97
EVENT.vform	0.94
TLINK.relType	in progress

Table 4: Annotated TLINKs and results of the inter-coder agreement

Given the data reported in the above tables, it is possible to claim that the results of the inter-coder agreement are good and comparable beyond the different annotation method used to develop the two corpora. So far, the ILC corpus has been annotated without time constraints by several annotators with varying backgrounds in linguistics using BAT. With this web-based tool, each file has been assigned to many annotators and an adjudication phase on discrepancies has been performed by an expert judge. As required by BAT, the annotation has been divided into many annotation layers so each annotator focused only on a specific set of It-TimeML tags. On the other hand, few expert annotators have been involved in the development of the CELCT corpus interacting and negotiating common solutions to controversial annotations. With respect to BAT, the CELCT Annotation Tool is stand-alone and it does not require neither the parallel annotation of the same text, nor the decomposition of annotation tasks allowing to have flexibility in the annotation process and a unitary view of all annotation layers. These features are helpful when working with strict project deadlines.

A comparison with the inter-coder agreement achieved during the annotation of the English TimeBank 1.2 (Pustejovsky et al., 2006a), shows that the scores obtained for the CELCT

and the ILC corpora are substantially higher in the following results: (i) average precision and recall on the identification of tag extent (e.g. 0.83 *vs*. 0.95 of ILC Corpus and 0.94 of CELCT Corpus for TIMEX3; 0.78 *vs*. 0.87 of ILC Corpus and 0.93 of CECLT Corpus); (ii) kappa score on Event classification (0.67 vs. 0.82 of ILC Corpus and 0.87 of the CELCT Corpus); (iii) kappa score on TLINK classification (0.77 *vs*. 0.86 of CELCT Corpus).

The similarity of the agreement results among the three resources and the improvement of the scores obtained on the CELCT and the ILC corpora with respect to the English TimeBank 1.2, can be taken as an indication of the quality and coverage of the It-TimeML annotation guidelines and specifications. Annotators showed to perform consistently demonstrating the reliability of the annotation scheme.

5 Conclusions and Future Works

This paper reports on the creation of a new semantic resource for Italian which has been developed independently but with a joint effort between two different research institutions. The Ita-TimeBank will represent a large corpus annotated with information for temporal processing which can boost the multilingual research in this field and represent a case study for the creation of semantic annotated resources.

One of the most interesting point of this work is represented by the methodology followed for the development of the corpora: in addition to the guidelines, annotation specifications have been created in order to report in detail the actual choices done during the annotation. This element should be pushed forward in the community as a new best practice for the creation of good quality semantically annotated resources.

The results obtained show the reliability of the adaptation of the annotation guidelines to Italian and of the methodology used for the creation of the resources.

Future works will concentrate in different directions, mainly due to the research interests of the two groups which have taken part to this effort but they will be coordinated.

An interesting aspect which could be investigated is the annotation of the anaphoric relations between Events. This effort could be done in a more reliable way since the primary linguistic items have been already annotated. Moreover, this should boost research in the development of annotation schemes which could be easily integrated with each other without losing descriptive and representational information for other language phenomena.

Another topic to deepen regards the definition of the appropriate argument structure in It-TimeML in order to annotate relations between entities (e.g. persons and organizations) and Events in which they are involved (Pustejovsky et al., 2006b).

As regards the distribution of the Ita-TimeBank, the resource will soon be available in an in-line format. In order to integrate the temporal annotation with other linguistic annotations, a standoff version of the Ita-TimeBank needs to be developed. When this is made available, we plan to merge the manual annotation of temporal and event information with other types of linguistic stand-off annotations (i.e. tokenization, lemma, PoS, multi-words, various kinds of named entities) which are already available for the I-CAB corpus.

In order to encourage research on systems capable of temporal inference and event-based reasoning, the Ita-TimeBank could be used as gold standard within specific evaluation campaigns as the next TempEval initiative.

Finally, the use of crowdsourcing will be explored to reduce annotation effort in terms of financial cost and time. The most difficult challenge to face will be the splitting of a complicated annotation scheme as It-TimeML into simple tasks which can be effectively performed by not expert contributors.

Acknowledgments

The development of the CELCT corpus has been supported by the LiveMemories project (Active Digital Memories of Collective Life), funded by the Autonomous Province of Trento under the Major Projects 2006 research program. We would like to thank Alessandro Marchetti, Giovanni Moretti and Marc Verhagen who collaborated with us in processing and annotating the CELCT corpus.

References

André Bittar. 2008. Annotation des informations temporelles dans des textes en français,. In Proceedings of RECITAL 2008, Avignon, France.

André Bittar. 2010. Building a TimeBank for French: A Reference Corpus Annotated According to the ISO-TimeML Standard. PhD Thesis.

Andrea Setzer and Robert Gaizauskas.2001. A Pilot Study On Annotating Temporal Relations In Text. In: Proceedings of the ACL 2001 Workshop on Temporal and Spatial Information Processing.

Bernardo Magnini, Emanuele Pianta, Christian Girardi, Matteo Negri, Lorenza Romano, Manuela Speranza, Valentina Bartalesi Lenzi and Rachele Sprugnoli. 2006. I-CAB: the Italian Content Annotation Bank. In Proceedings of LREC 2006, Genova, Italy.

Bernardo Magnini, Matteo Negri, Emanuele Pianta, Manuela Speranza, Valentina Bartalesi Lenzi, and Rachele Sprugnoli. 2006. Italian Content Annotation Bank (I-CAB): Temporal Expressions (V.2.0). Technical Report, FBK-irst.

Carlota S. Smith. 1991. The Parameter of Aspect. Kluwer, Dordrecht.

Collin F., Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In: Proceedings of the COLING-ACL, pages 86-90. Montreal, Canada.

Elena Filatova and Eduard Hovy. 2001. Assigning Time-Stamps To Event-Clauses. In: Proceedings of the ACL 2001 Workshop on Temporal and Spatial Information Processing.

Emmon Bach. 1986. The algebra of events. Linguistics and Philosophy, 9, 5–16.

Graham Katz and Fabrizio Arosio. 2001. The Annotation Of Temporal Information In Natural Language Sentences. In: Proceedings of the ACL 2001 Workshop on Temporal and Spatial Information Processing.

ISO: Language Resource Management – Semantic Annotation Framework (SemAF) - Part 1: Time and Events. Secretariat KATS, August 2007. ISO Report ISO/TC37/SC4 N269 version 19 (ISO/WD 24617-1).

James Pustejovsky, Jessica Littman and Roser Saurí. 2006b. Argument Structure in TimeML. In: Graham Katz, James Pustejovsky and Frank Schilder (eds.) Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum (IB-FI), Schloss Dagstuhl, Germany.

James Pustejovsky, Jessica Littman, Roser Saurí, and Marc Verhagen. 2006a. TimeBank 1.2 Documentation. http://timeml.org/site/timebank/documentation-1.2.html

James Pustejovsky, José Castaño, Robert Ingria, Roser Saurí, Robert Gaizauskas, Andrea Setzer and Graham Katz. 2003a. TimeML: Robust Specification of Event and Temporal Expressions in Text. In: Proceedings of IWCS-5, Fifth International Workshop on Computational Semantics.

James Pustejovsky, Patrick Hanks, Roser, Saurì, Andrew See, Robert Gaizauskas, Andrea Setzer, Dragomir Radev, Beth Sundheim, David Day,Lisa Ferro, and Marcia Lazo. 2003b. The TIMEBANK corpus. In: Proceedings of Corpus Linguistics 2003, pages 647-656.

Marc Verhagen, Roser Saurí, Tommaso Caselli and James Pustejovsky. 2010. SemEval-2010 Task 13: TempEval-2. In: Proceedings of the 5th International Workshop on Semantic Evaluation.

Mark G. Core and James F. Allen. 1997. Coding Dialogs with the DAMSL Annotation Scheme. In: Working Notes of AAAI Fall Symposium on Communicative Action in Humans and Machines.

Nicoletta Calzolari, and Tommaso Caselli 2009. Short Report on the FLaReNet / SILT Workshop and Panel on Semantic Annotation, TR-ILC-CNR.

Pier Marco Bertinetto. 1991. Il verbo. In: R. L. and G. Salvi (eds.) Grande Grammatica Italiana di Consultazione, volume II, pages 13-161. Il Mulino.

Ron Artstein and Massimo Poesio. Inter-coder agreement for computational linguistics. Computational Linguistics, pages 555–596, 2008.

Seohyun Im, Hyunjo You, Hayun Jang, Seungho Nam, and Hyopil Shin. 2009. KTimeML: Specification of Temporal and Event Expressions in Korean Text. In: Proceedings of the 7th workshop on Asian Language Resources in conjunction with ACL-IJCNLP 2009, Suntec City, Singapore.

TimeML Working Group. 2010. TimeML Annotation Guidelines version 1.3.Manuscript, Brandeis University.