

Annotating Time Expressions in Catalan

TimeML Annotation Guidelines

(Version TempEval-2010)

Barcelona Media
Technical Report BM 2010-03

Roser Saurí, James Pustejovsky

Contents

1	Introduction	2
2	Time expressions in TimeML	3
2.1	Dates	3
2.2	Times of day	5
2.3	Durations	5
2.4	Sets	8
3	What to annotate as time expressions	9
3.1	Markable expressions	9
3.1.1	Lexical triggers	9
3.1.2	Fuzzy expressions	10
3.1.3	Metonymic expressions	10
3.2	Non-Markable expressions	10
3.2.1	Parts of speech: Prepositions and subordinating conjunctions	10
3.2.2	Expressions of sequencing	10
3.2.3	Expressions of speed (or celerity)	11
3.2.4	Expressions of frequency	12
3.2.5	Expressions of non-quantifiable duration	13
3.2.6	The word <i>temps</i> ('time') when it means 'situation' or 'occasion' . . .	13
3.2.7	Time expressions in proper names	13

4	Timex3 extents	14
4.1	Grammatical criteria	14
4.2	Relational criteria	16
4.2.1	Specification relation	16
4.2.2	Anchoring relation: Involving a duration expression	18
4.2.3	Anchoring relation: Involving a date expression	20
4.2.4	Range relation	22
4.2.5	Frame relation	23
4.2.6	Conjunction relation	23
5	Timex3 attributes	24
5.1	Attribute <code>type</code>	24
5.2	Attribute <code>value</code>	25
5.2.1	Dates	25
5.2.2	Time of day	29
5.2.3	Durations	30
5.2.4	Sets	31
5.3	Attribute <code>mod</code>	32
5.4	Attribute <code>temporalFunction</code>	33
5.5	Attribute <code>anchorTimeID</code>	34
5.6	Attribute <code>valueFromFunction</code>	35
5.7	Attribute <code>functionInDocument</code>	35
5.8	Attributes <code>beginPoint</code> and <code>endPoint</code>	35
5.9	Attributes <code>quant</code> and <code>freq</code>	36
A	Complex time expressions	37
A.1	Dates	37
A.2	Durations	38
B	TimeML attributes and values for <code>timex3</code>	39

1 Introduction

This document describes the annotation guidelines for marking up instances of *time expressions* in Catalan text, according to the TimeML language. TimeML (Pustejovsky et al., 2005) is a specification language for events and time expressions. It was first developed in 2002 in an extended workshop called TERQAS (Time and Event Recognition for Question Answering Systems),¹ which focused on the issue of answering temporally based questions regarding events and entities in news articles. In 2003, TimeML was further developed

¹<http://www.timeml.org/site/terqas/index.html>

in the context of the TANGO workshop (TimeML Annotation Graphical Organizer).² In addition, TimeML has been consolidated as an international cross-language ISO standard (ISO WD 24617-1:2007), and has been approved as the annotation language for TempEval, one of the tasks in the SemEval International Workshop on Semantic Evaluations (Verhagen et al., 2007, 2009).

In the context of the current document, time expressions are referred to as *timexes*. *Timex* stands for TIME EXpression. In TimeML, they are marked up with the tag `timex3`. Early annotation schemes for time expressions included the tags `timex` and `timex2`. TimeML’s `timex3` tag was developed with these earlier schemes in mind, as will be appreciated by the experienced reader in specific areas of the current annotation guidelines. The tag `timex3`, however, has some significant differences with respect to its predecessors –hence, the different name.

The current annotation guidelines parallel those for Spanish time expressions (Saurí et al., 2010), while focusing on the specifics of time expressions as expressed in the Catalan language. The annotation process will be split into two sequential subtasks. The first one is devoted to identifying the time expressions in text, while the second explains how to characterize these with their appropriate attributes (e.g., `type`, `value`, `mod`, or `anchorTimeID`).

The following section analyzes the notion of temporal expression (or *timex*) as understood in TimeML. Then, sections 3 and 4 address the task of **time expression identification**, laying out first what to annotate as such and then describing how much text to mark up –i.e., its extent. Finally, section 5 focuses on the task of **attribute annotation**.

2 Time expressions in TimeML

Time expressions in TimeML (*aka* `timex3`) are those constructions referring to points or intervals on the timeline. They can express: (a) **times of day**, e.g., *migdia* (‘noon’), *les 3h de la tarda* (‘3:00pm’), *aquest matí* (‘this morning’); (b) **dates**, e.g., *20 d’abril del 2008* (‘April 28th, 2008’), *ahir* (‘yesterday’), *la setmana que ve* (‘next week’), *tres mesos després* (‘three months later’), *l’any passat* (‘last year’); (c) **durations**, e.g., *dos mesos* (‘two months’), *cinc hores* (‘five hours’), *els 9 propers anys* (‘the coming 9 years’); and (d) **sets**, e.g., *una vegada al mes* (‘once a month’), *cada dimarts* (‘every Tuesday’). The following subsections present these types in some detail.

2.1 Dates

Dates include all sorts of calendar dates. They can refer to a point in time (e.g., *14 d’abril del 1931*, ‘April 14, 1931’) or to a full interval (*els seixanta*, ‘the 60s’).

²<http://www.timeml.org/site/tango/index.html>

Different degrees of granularity. Dates can have different degrees of granularity:

- **Days:** *ahir, 8 de gener del 2001, el proper divendres, aquest dissabte*, etc.
- **Weeks:** *la pròxima setmana, la segona setmana del mes*, etc.
- **Months:** *tres mesos després, el proper mes, agost del 2000*.
- **Seasons or business quarters:** *la pròxima primavera, el primer trimestre*, etc.
- **Years and decades:** *1980, l'any que ve, els setanta*.
- **Centuries, millennia:** *el segle passat, el nou mil.lenni*, etc.

Precise vs. fuzzy dates. Dates can be precise or fuzzy. A *precise date* refers to a specific calendar point or duration. For example:

- (1) a. La primera trobada va tenir lloc el 20 de maig del 2009.
- b. La seva oferta consisteix en suprimir els descomptes que l'empresa aplica a les nòmines per la crisi del 1994.

Sometimes, the value of a precise date is not fully specified but needs to be interpreted based on contextual information.

- (2) a. El concurs públic que es va convocar el novembre ha quedat desert.
- b. Castelló va arribar ahir a 12.000 habitants de 76 nacionalitats diferents.

Additional examples of (fully or partially specified) precise dates follow below.

- (3) a. Clinton estudià com a becari a Oxford a finals dels seixanta.
- b. Per a la comunitat científica, el llegat de l'encefalopatia bovina marcarà el pròxim segle.
- c. A finals d'aquest mil.lenni, existeixen unes 100 mil ONG a tot el món que reben prop de 10 mil milions de dòlars.

Fuzzy dates, on the other hand, express vague calendar points, or calendar intervals with imprecise boundaries. This is particularly the case of time expressions referring to the past, the present, or the future in general, imprecise terms, as exemplified in table 1.

Table 1: Fuzzy dates in TimeML

Present reference:	<i>Ara, avui</i> (both of them when interpreted as 'currently', 'nowadays'), <i>actualment, en aquest moment, en el moment present,</i> <i>en el moment actual.</i>
Future reference:	<i>En el/un futur,</i> <i>el dia de demà, demà</i> (when referring to a future time)
Past reference:	<i>Passat, en el passat, ahir</i> (referring to a past time), <i>llavors</i> (as in: <i>des de llavors, de llavors fins ara</i>), <i>recentement, últimament, antigament</i>

Anchored dates. Some dates are explicitly anchored to a second date by means of a modifier of sequence, such as *anterior*, *següent*, *que segueix*, *posterior*, *abans*, *després* (‘preceding’, ‘following’, ‘that follows’, ‘next’, ‘before’, ‘after’), etc.

- (4) a. el següent dimarts
 b. el dimarts passat
 c. el dimarts que ve
 d. el dimarts vinent

In the examples above, the anchored date is expressed by *dimarts* (‘Tuesday’), whereas the anchoring date is implicit. Other constructions, however, have the anchoring date explicit. Note that this anchoring element can denote either a date (e.g., *el diumenge de Pasqua* ‘Easter’) or an event (e.g., *el seu casament* ‘their wedding’).

$$(5) \text{ el dimarts } \left\{ \begin{array}{l} \textit{següent} \\ \textit{que segueix} \\ \textit{posterior} \\ \textit{anterior} \end{array} \right\} a \left\{ \begin{array}{l} \textit{el diumenge de Pasqua} \\ \textit{el seu casament} \end{array} \right\}$$

$$(6) \text{ el dimarts } \left\{ \begin{array}{l} \textit{abans} \\ \textit{després} \end{array} \right\} de \left\{ \begin{array}{l} \textit{el diumenge de Pasqua} \\ \textit{el seu casament} \end{array} \right\}$$

See section 4.2.3 for a complete description of these elements.

2.2 Times of day

Times of day include expressions referring to points or intervals of time smaller than a day. In the following examples, the time of day is underlined.

- (7) a. El patrocinador va arribar a les 3h menys 10.
 b. El rellotge marcava la una del migdia.
 c. El dia 19 tinc classe fins a les 11h del matí.
 d. Va estar mirant la tele fins a la mitjanit/matinada/...
 e. Enviat: 12:27 PM EDT.

Time expressions that have a reference to a calendar date but which also include a reference to a time of day will, in any case, be considered a `time3` of type time of day.

- (8) Enviat l'11 d'abril del 1996, 11:13 GMT.

2.3 Durations

An expression of duration indicates how long something lasts. In TimeML, duration expressions are those that denote a quantification over time.

Precise vs. fuzzy durations. Like dates, durations can be precise or fuzzy. *Precise durations* include a *quantity expression*, namely, a numeral (*dos*, 52) or a quantifier (*uns*, *pocs*, *molts* –‘some’, ‘few’, ‘many’), plus a *time unit* (e.g., *dia*, *setmana*, *mes*, *any* –‘day’, ‘week’, ‘month’, ‘year’):

- (9) *38 setmanes* (‘38 weeks’)
vuit anys i mig (‘eight years and a half’)
molts anys (‘many years’)
alguns dies (‘several days’)

Some examples:

- (10) a. Alliberen de Guantànamo el càmera d’Al-Jazira empresonat durant sis anys.
 b. Joan Saura va tardar tres dies a informar l’alcalde de Mataró.
 c. El país va entrar en un llarg període d’instabilitat política que va durar tota una dècada.

Precise durations can also be expressed by means of lexical items conveying a quantification over a time unit. For example, partitive nouns like *meitat* (‘half’) or *terç* (‘third’), and collective nouns like *quinzena* (‘fortnight’).

- (11) *la segona quinzena* (‘the second fortnight’)
la meitat dels dies (‘half of the days’)

Other expressions denote *fuzzy durations*. For example, nouns like *temporada* (‘season, period’), *període* (‘period’), and *estona* (‘while’) in examples such as (12), quantifying adverbs such as *molt* (‘long, a lot’) and *poc* (‘short, a little’) in (13), or constructions such as *molt de temps* (‘a long time’), *poca estona* (‘a short time’). Here, they are considered as *timex3s* as well.

- (12) a. L’espectacle va durar una estona llarga.
 b. Tamek va passar una temporada a la capital catalana per tractar-se de les seqüeles dels anys que va passar empresonat per les autoritats marroquines.
- (13) a. L’equip va durar poc, ja que una sanció obligà a clausurar el camp.
 b. Mingote no va tardar gaire a posar al servei de la música popular la seva vàlua.

Finally, in Catalan there are other expressions that can denote both precise and fuzzy durations. They tend to have complex patterns, such as:

- (14) *el que* $\left\{ \begin{array}{l} \textit{queda/quedava/quedarà}/... \\ \textit{segueix/seguí}/... \\ \textit{va/anava}/... \end{array} \right\}$

followed by the PP: *de N_{time}*. For instance:

- (15) La ministra preveu que les caigudes de preus es vagin moderant en el que queda d’any.

Finally, in Catalan there are other expressions that can denote both precise and fuzzy durations. They tend to have complex patterns, such as:

- (16) *el que* $\left\{ \begin{array}{l} \textit{queda/quedava/quedarà}/... \\ \textit{va/anava}/... \end{array} \right\}$

followed by the PP: *de N_{time}*. For instance:

- (17) Iberia perd 182 milions d'euros en el que va d'any.

Durations vs. dates. It is important to distinguish between durations (like those just exemplified) and dates that can be interpreted as referring to intervals, such as:

- (18) Durant el 1992 es van produir nombroses escissions dins el Front Popular de Letònia.

Deciding whether a time expression refers to a point or an interval is, to some extent, an arbitrary issue related to the granularity of the time system. Bigger time units (e.g., years, decades) can be more easily considered as intervals or durations than smaller ones (e.g., seconds). Here, we will consider as durations those expressions defined above as precise or fuzzy durations. For example: *alguns dies* ('some days'), *4 mesos* ('4 months'), *una dècada* ('one decade'), *38 setmanes* ('38 weeks'), *diversos anys* ('several years'), *molta estona* ('a while'), *poc* ('a little'), etc., as exemplified above.

These expressions give no indication of their position in the calendar, contrary to, e.g., *1992* in (18), although in some cases the context places them in the timeline. For instance:

- (19) a. Les obres de la nova catenària rígida de la Plaça de Catalunya han afectat durant les dues setmanes passades un total de 800.000 viatgers.
b. Els propers 5 anys seran els millors de la seva vida.

In TimeML, these expressions will still be considered durations. Their placement in the timeline will be codified by means of temporal links (or TLINKs). Refer to the appropriate guidelines for that.

Anchored durations. Special attention should be paid to duration expressions which, due to the constructions they appear in, refer in fact to points in time and hence can be seen as denoting *dates*. We call them *anchored durations* because they are anchored by a second time reference. For instance:

- (20) a. Mèxic va deixar passar fa 2 anys el tren dels bioenergètics.
b. Avui fa 5 mesos de l'inici de Dharmakaya Barcelona.

In (20a), the duration expressions *fa 2 anys* ('2 years ago') is employed to determine a point in time two years earlier. Similarly, in (20b) the expression *fa 5 mesos* ('it is 5

months’) uses the value provided by *avui* (‘today’) in order to compute the point in the calendar 5 months previous to today.

Additional examples include:

- (21) a. [dues setmanes] a partir [del proper dijous] (‘two weeks from next Thursday’)
 b. [dos dies] abans de [Cap d’any] (‘two days before New Year’s Day’)

Framed durations. Some duration expressions appear in constructions here referred to as *framed durations*, given that they locate (or frame) a duration within the scope of a temporal unit which has a precise reference in the calendar. The following are examples of this kind of constructions. The extent of the duration timex is underlined, whereas the extent of the framing date is in bold face.

- (22) a. la segona quinzena de **novembre** (‘the second fortnight in November’)
 b. la primera meitat de **la setmana** (‘the first half of the week’)
 c. els últims dies de **l’any** (‘the last days of the year’)
 d. les tres primeres dècades **del segle XX** (‘the first three decades of the 20th century’)

In this kind of constructions, we will also accept as durations those expressions that are compliant with the pattern below, whenever followed by the PP: *de N_{time}* (which will be expressing the framing date).

- (23) *el que* $\left\{ \begin{array}{l} \textit{queda/quedava/quedarà}/\dots \text{ (lit., ‘what remains/remained/...’)} \\ \textit{va/anava}/\dots \text{ (lit., ‘what goes/went/...’)} \end{array} \right\} + \textit{de N}_{time}$

For instance:

- (24) Iberia perd 182 milions d’euros en el que va d’any.

2.4 Sets

Whereas time expressions of type *time of date* and *date* refer to when something happened, and *durations* indicate how long something lasted, *sets* tell how often something happened. Typical examples of set-denoting time expressions are:

- (25) *cada dia/mes/dilluns/dimarts/...* (‘each day/month/Monday/Tuesday/...’)
tots els dies/mesos/dilluns/dimarts/... (‘every day/month/Monday/Tuesday/...’)
diariament, a diari, mensualment, anualment,... (‘daily, monthly, yearly’)
x vegades al/cada dia/mes/any/... (‘x times a/every day/month/year/...’)

3 What to annotate as time expressions

3.1 Markable expressions

3.1.1 Lexical triggers

In TimeML, we will mark as `timex3` those constructions which have an appropriate lexical trigger as their syntactic head. Lexical triggers are words:

- whose meaning conveys a temporal unit or concept, such as *dia* ('day') o *mensualment* ('monthly'), or
- whose referent can be oriented on a timeline, or at least oriented with relation to a time (past, present, future).

Table 2 contains a sampling of lexical triggers. In the case of nouns (and adjectives), only the singular (and masculine) forms are presented, but other forms are possible as well.

Table 2: Lexical triggers sampling

Part of Speech	Lexical triggers
Noun	<i>Matí, migdia, tarda, nit.</i> <i>Dia, setmana, cap de setmana, mes, any, dècada, segle, mil.lenni.</i> <i>Semestre, trimestre, quadrimestre.</i> <i>Solstici, equinocci.</i> <i>Temporada, etapa, període, instant.</i>
Noun/Proper name	<i>Gener, febrer, març, abril, maig, juny, juliol, agost,...</i> <i>Nadal, Setmana santa, Sant Joan,...</i>
Time noun/adverb	<i>Ahir, avui, demà.</i> <i>Dilluns, dimarts, dimecres, dijous, divendres, dissabte, diumenge.</i>
Adjective	<i>Diari, mensual, anual.</i> <i>Semestral, bimensual, trimestral.</i> <i>Passat, recent; present, actual; proper, futur.</i>
Adverb	<i>Diàriament, mensualment, anualment.</i> <i>Semestralment, bimensualment, trimestralment.</i> <i>Recentement; ara, actualment; properament.</i> <i>Llavors</i> (in constructions like: ' <i>des de llavors fins ara</i> ', ' <i>de llavors ençà</i> '). <i>Aquí</i> (in constructions like: ' <i>d'aquí a Nadal</i> ') <i>Ençà</i> (in constructions like: ' <i>de Nadal ençà</i> ', but not: ' <i>d'ençà de Nadal</i> ', where ' <i>d'ençà de</i> ' will be treated as a signal).
Number	<i>3, tres, tercer, etc., as in: el 3 de febrer, el tercer de cada mes.</i>
Quantifiers (of duration)	<i>molt, poc, força.</i>
Patterns	<i>08:00, 19:47, 31/11/2008, 2007, ...</i>

Sometimes, the trigger word is not the syntactic but the semantic head of the expression, such as in partitives like *la major part de la setmana* ('the main part of the week'). These

constructions will, in any case, be considered `timex3`. In other cases, however, these (or similar) lexical triggers are non-referring, and thus impossible to relate to a timeline. For example, the use of *dia* ('day') in the idiomatic phrase *estar a l'ordre del dia* ('be the order of the day'). These cases will not be marked up as `timex3s` expressions in TimeML.

3.1.2 Fuzzy expressions

Expressions denoting fuzzy dates, presented in table 1, are `timex3` markables as well.

3.1.3 Metonymic expressions

Some time expressions such as *11-S*, *11-M*, etc., have been recategorized into proper nouns and refer to specific events (in these cases, the terrorist attacks that occurred on September 11th, 2001 and March 11th, 2004, respectively), and not simply to the original date. Regardless of that, in TimeML these expressions will be annotated as `timex3`. Other time expressions that will receive the same treatment are: *Sant Joan* ('Saint John', *Dia internacional de la infància* ('International Day of the Childhood'), and *Primer de maig* ('Labor Day').

3.2 Non-Markable expressions

On the other hand, there are a set of lexical items that, although they are temporal in their semantics, are less amenable to being pinned down to a timeline. Broadly speaking, they belong to one of the classes presented in the following subsections.

3.2.1 Parts of speech: Prepositions and subordinating conjunctions

Prepositions (which introduce noun phrases) and *subordinating conjunctions* (which introduce clauses) are two parts of speech that are never triggers; that is, they never appear as the syntactic head of an annotated expression. Table 3 provides some examples. It includes both simple (e.g., *de*, *quan*) and complex (*abans de*, *tan aviat com*) expressions.

Table 3: Non-Triggers. Prepositions and Subordinating conjunctions

Part of Speech	Non-triggers
Subordinating conjunctions:	<i>Ara que, abans que, cada vegada que, quan, després que, mentres, tan aviat com, ...</i>
Prepositions:	<i>A, abans de, de, durant, en, des de, fins, ...</i>

3.2.2 Expressions of sequencing

Expressions establishing temporal sequencing between two or more events will not be marked up as `timexes`. The following examples illustrate some of them. The two events

are in bold face, whereas the sequencing expression appears underlined.

- (26) Quatre anys de **confrontació** entre ZP i el PP i mentrestant augmenta el cost de la vida, l'atur i cada cop costa més arribar a final de mes.
- (27) Només cal recordar els **fets** de Nova York i Washington del passat setembre i la **situació** subsegüent d'estat de conflicte internacional.

Expressions of sequencing belong mainly to the parts of speech of adjectives and adverbs. Table 4 lists the most frequent ones.

Table 4: Non-Triggers. Sequencing expressions

Part of Speech	Non-triggers
Adjectives	<i>Abans, després</i> (e.g., <i>El dia després.</i>) <i>Primerenc; tardà, tardívol.</i> <i>Primer; últim, darrer, final.</i> <i>Anterior, previ.</i> <i>Simultani, coetani.</i> <i>Posterior, següent, subsegüent, ulterior.</i>
Adverbs and adverbial constructions	<i>Abans, després</i> (e.g., <i>Arribà després.</i>) <i>D'hora, aviat; tardanament.</i> <i>Primerament; eventualment, finalment, al final.</i> <i>Anteriorment, previament.</i> <i>Simultàniament, coetàniament, mentres, mentrestant, entretant.</i> <i>Posteriorment, seguidament, subsegüentment, ulteriorment, en qual-sevol moment.</i> <i>Encara, novament, de nou.</i>

Non-trigger adjectives are permitted within the extent of a markable expression, as in *l'any **passat** anys* ('the preceding year'), *el **següent** dia* ('next day'), or *les hores **prèvies*** ('the previous hours'). They are not markable on their own, as in *la **següent** reunió* ('the next meeting').

Note that some adjectives and adverbs presented here (e.g., *previ, següent* –'previous', 'subsequent') are semantically very similar to those presented as lexical triggers in Table 2 (e.g., *recent, proper* –'recent', 'coming'). Notwithstanding, there is a distinctive feature between the two groups. Namely, lexical triggers are necessarily anchored to the time of the speech (or utterance) act, whereas adjectives and adverbs presented as non-triggers can have any time point as their anchoring reference. In case of doubt, use this test to decide whether to mark up the expression.

3.2.3 Expressions of speed (or celerity)

Some manner adverbs (or adverbial constructions) express how soon or how quickly something happens. They are known as speed (or celerity) adverbs. They will be not marked up as **timex3s**. Table 5 provides some examples.

(28) L'impacte dels projectils va crear unes bombolles de tres metres d'alçada quasi instantàniament a la base del salt d'aigua.

(29) Laporta espera que Márquez renovi ben aviat pel club.

Table 5: Non-Triggers. Expressions signaling speed

Part of Speech	Non-triggers
Adjectives	<i>Instantani, immediato.</i> <i>Lent.</i>
Adverbs and adverbial constructions	<i>Instantàniament, immediatament, ràpidament.</i> <i>A l'acte, a l'instant, d'immediat, aviat, de seguida.</i> <i>Amb preses, ràpidament.</i> <i>Lentament.</i>

3.2.4 Expressions of frequency

Bare frequencies (that is, frequency expressions which do not convey a time unit) are not markable. Some of them are presented in Table 6.

Table 6: Not markable frequency expressions

Part of Speech	Expressions
Adjectives:	<i>Freqüent, normal, habitual, usual, comú.</i>
Adverbs:	<i>Normalment, freqüentment, habitualment, usualment, típicament.</i> <i>Repetidament.</i> <i>Rarament, difícilment.</i> <i>Sempre; mai.</i>
Constructions:	<i>Una vegada, tres cops, en cinc ocasions,</i> <i>Rara vegada, alguna vegada.</i>

(30) Va passar només una vegada.

(31) Hi va haver denúncies freqüents de tortura i maltractament de detinguts polítics i penals.

(32) El 66% dels tarragonins utilitza Internet habitualment.

(33) Els nazis sempre van insistir en una entesa amb la Gran Bretanya.

By contrast, frequencies whose semantics include a temporal unit will be marked up. Table 7 illustrates some of them.

Table 7: Markable frequency expressions

Part of Speech	Expressions
Adjectives:	<i>Diari, mensual, anual, trimestral, semestral.</i>
Adverbs:	<i>Diariament, mensualment, anualment, trimestralment, semestralment.</i>
Constructions:	<i>Dues vegades al dia, cinc cops a la setmana, ... A diari, cada dia/mes/any/... Tots els dies/mesos/anys/...</i>

3.2.5 Expressions of non-quantifiable duration

Non-quantifiable durations are expressed by means of items such as those presented in Table 8. They will not be marked up.

- (34) L'ordre d'arrest domiciliari contra Augusto Pinochet ha quedat temporalment suspesa.

Table 8: Non-Triggers. Non-quantifiable durations

Part of Speech	Non-triggers
Adjectives:	<i>Permanent, perpetu, etern. Temporal, provisional.</i>
Adverbs:	<i>Permanentment, perpètuament, eternament. Temporalment, provisionalment.</i>

3.2.6 The word *temps* ('time') when it means 'situation' or 'occasion'

The word *temps* ('time') is one of the most difficult expressions in terms of markability. In many cases, it is not referring to a time interval but it means 'situation' or 'occasion'. This use of the word is not markable. Hence, *temps* ('time') in (35b) will be marked up (since it expresses a vague duration), but not in (35a).

- (35) a. L'estiu es temps de vacances, inclús per a les estrelles de la ràdio.
b. LoQUo va estar molt de temps funcionant des de casa meva al carrer Pintor Fortuny.

3.2.7 Time expressions in proper names

Proper names that designate something other than a temporal entity (e.g., films, books, organizations) but happen to correspond to lexical triggers are not markable. For example:

- (36) La pel·lícula Quatre dies de setembre està dirigida por Bruno Barreto.
(37) *El somni d'una nit d'estiu* es una comedia escrita per Shakespeare al voltant del 1595.
(38) La companyia DIA va obtenir unes vendes brutes el 2008 de 10.560 milions d'euros.

However, triggers that are functioning as temporal modifiers within titles are markable. Examples include titles of conferences (39a) and awards (39b).

- (39) a. TERQAS 2002.
b. L'entrenador de l'any.

4 Timex3 extents

This section provides the criteria for determining where each time expression begins and ends. We refer to this as the **extent** of the expression.

The criteria to be used here can be grammatical (section 4.1) or relational (4.2). Grammatical criteria are applied first, whereas relational criteria are employed in case of complex expressions, for which it is not clear whether one or more timexes must be distinguished.

4.1 Grammatical criteria

For most timexes, the full extent of the tag corresponds to one of the following categories:

Noun phrase, including also noun phrases with proper nouns. Some examples: *la tarda*, *Nadal*, *l'estiu passat*, *ahir*, *diumenge*, *aproximadament mitja hora* ('the afternoon', 'last summer', 'yesterday', 'Sunday', 'approximately half an hour'). With regard to annotating timex3-referring NPs, we will apply the following considerations:

- **Prepositions.** Any preposition preceding the temporal expression will not be included as part of the tag.³ For example, the following expressions will have only the underlined part marked up as the timex3 element: *durant la tarda* ('during the afternoon'), *abans d'aquell dimarts* ('before that Tuesday'), *d'aquí a mitja hora* ('in half an hour').

However, some corpora do not have the **preposition + determinant** contractions split, such as *del* (*de*, 'of' + *el*, 'the'), *al* (*a*, 'to' + *el*, 'the'), or *pel* (*per*, 'for' + *el*, 'the'). In this situation, the preposition will be included in the span of the temporal expression. For example: *abans del dimarts, previ al diumenge de Pasqua*.

- **Postnominal modifiers.** Some modifiers of temporal nouns express an event. In the following examples, the temporal noun is in italics while the event is in bold face.

- (40) a. El dia de la seva **mort**.
b. El millor quatrimestre que ha hagi **tingut** mai l'empresa.

³Temporally relevant prepositions will be annotated as **signals**, as established in the annotation guidelines for this other TimeML entity.

c. Aproximadament quatre dècades de renovada **hostilitat**.

Postnominal modifiers like these will not be included as part of the `timex3` tag, which will span as indicated by the square brackets in the preceding examples.

Adjective phrase; e.g., *recent* (‘recent’), as in *un viatge recent* (‘a recent trip’).

Adverbial phrase; e.g., *molt recentement* (‘very recently’), *d’aquí a un any* (lit. ‘from here in a year’).

Special pattern expressions. Constructions like *fa un mes* (‘a month ago’ –lit., ‘it makes one month’), *falta un mes* (‘in one month’ –lit.: ‘it misses one month’), or *la setmana que ve* (‘next week’ –lit. ‘the week that is coming’), are characteristic time expressions in Catalan which do not correspond to any of the previous categories. They will be considered as time expressions as well. Specifically, the following patterns apply here:

1. $\text{NP}_{duration} + \left\{ \begin{array}{l} abans \\ després \\ \text{[adv]} \text{ més tard} \end{array} \right\}$ (e.g., *tres dies més tard*)
2. $\left\{ \begin{array}{l} fa/feia/... \\ passen/passaven/... \end{array} \right\} + \text{NP}_{duration}$ (e.g., *fa una hora, feia molta estona*)
3. $\left\{ \begin{array}{l} queden/quedaven/... \\ falten/faltaven/... \end{array} \right\} + \text{NP}_{duration}$ (e.g., *falten deu anys, falta una mica*)
4. $\left\{ \begin{array}{l} el \\ la \end{array} \right\} + \text{N}_{time} + que \left\{ \begin{array}{l} ve \\ segueix \end{array} \right\}$ (e.g., *el dimarts que ve*)
5. $\left\{ \begin{array}{l} el \\ la \end{array} \right\} + \left\{ \begin{array}{l} següent \\ anterior \\ passat \\ proper \end{array} \right\} + \text{N}_{time}$ (e.g., *el següent dimarts*)

In items 1-3, $\text{NP}_{duration}$ simbolizes an NP denoting a duration (as defined in section 2.3), such as *temporada* (‘season, period’), *període* (‘period’), or *estona* (‘while’), or alternatively, an NP consisting of a numeral (*dos, 52*) or quantifier (*uns, pocs, molts* –‘some’, ‘few’, ‘many’), plus a time unit (e.g., *dia, setmana, mes, any* –‘day’, ‘week’, ‘month’, ‘year’).

On the other hand, N_{time} in items 4-5 represents temporal units such as: *dia, setmana, mes, any, gener...desembre, dilluns...diumenge*, etc. (‘day’, ‘week’, ‘month’, ‘year’, ‘January, ..., December’, ‘Monday, ..., Sunday’, etc.).

Refer to the tables in appendix A, for an overview of the treatment for these and similar constructions.

4.2 Relational criteria

There are time expressions that present a more complex structure, which makes it difficult to determine the extent of the `timex3` tag. For example, does an expression like *el dimarts 30 de juny* ('Tuesday, June 30th') contain one or two `timex3`s? That is, one for the NP *el dimarts*, and the other for the NP *30 de juny*.

In cases like this, the `timex3` extent will be determined according to the relation that holds between the two time expression candidates. The following sections present the possible relations between two time expressions and details their treatment.

4.2.1 Specification relation

Definition. Involving two time expressions, one of which is helping to further specify the other. The specification relation between two elements is characterized based on the classification of temporal units shown in Table 9, which can be smaller than a day ($t < day$), as long as a day ($t = day$), bigger than a day but smaller than a year ($day < t < year$), as long as a year ($t = year$), and bigger than a year ($t > year$).

t<day	t=day	day<t<year	t=year	t>year
12 en punt	dimecres	la quarta setmana	1984	l'últim segle
2 menys 10	demà	mes	proper any	
mitjanit	2 de gener	semestre		
matinada	el dia 8	gener		
tarda	el dia de Nadal	temporada		
nit	el dia del treballador	tardor		

Table 9: Classification of time units.

Two temporal expressions in a specification relation can belong to the same type of time unit (according to Table 9). For example, in (41a), the referents of both *les dotze en punt* ('twelve o'clock') and *la mitjanit* ('midnight') correspond to the same unit type, namely, units smaller than a day ($t < day$).

- (41) a. [*les dotze en punt*] de [*la mitjanit*] ('12 o'clock midnight')
 b. [*les quatre*] de [*la tarda*] ('4pm in the afternoon')
 c. [*el dimarts*] [*18 de gener*] ('Tuesday, January 18th')

In other cases, one of the expressions (generally, the first one) refers to a temporal unit smaller than the second. For example:

- (42) a. [*l'estiu*] d'[*aquest any*] ('the summer of the current year')
 b. [*alguns dimarts*] de [*1984*] ('some Tuesdays of 1984')

Note that in these and the previous set of examples, the temporal expressions are related by the preposition *de* ('of'), or by its absence (41c).

A third possible type of constructions holding a specification relation expresses a duration, which is modified by another duration of smaller size. The modification can be additive (43a) or subtractive (43b-c).

- (43) a. [*tres anys*], [*tres mesos*] *i* [*tres dies*] ('three years, three months and three days')
 b. [*les 4*] *menys* [*10*] ('10 minutes to 4', lit. '4 (hours) minus 10 (minutes)')
 c. [*un mes*] *menys* [*3 dies*] ('one month minus 3 days')

Annotation. Two time expressions in a specification relation will be marked up with a single tag if:

- The two expressions belong to the same time unit, according to Table 9. For example, *les 12h en punt de la mitjanit* ('12 o'clock midnight'); *el dimarts, 18 de gener* ('Tuesday, January 18th'); *les 11h del matí* ('11am in the morning').
- The two expressions refer to durations of different magnitude, and the smaller one modifies (shortens or lengthens) the bigger one. For example: *tres mesos i tres dies* ('three months and three days'); *les 4 menys 10* ('10 minutes to 4'); *un mes menys 3 dies* ('one month minus 3 days').
- The two expressions belong to the same syntactic constituent and are (generally, but not always) connected with the preposition *de* ('of'). This preposition appears, for example, when expressing dates that include the month and/or the year.

Syntactic constituency can be checked using the fronting and clefting tests (sentences *b-c* and *d-e* in the examples below, respectively). Two different constituents will allow for fronting and clefting, but not two parts of the same constituent. For example, the two time constructions in square brackets in (44) are part of the same constituent, while those in (45) are not.⁴

- (44) a. Tots els grups es reuniran [a les 11h] [del dijous 18 de gener].
 b. *[Del dijous 18 de gener], tots els grups es reuniran [a les 11h].
 c. *[A les 11h], tots els grups es reuniran [del dijous 18 de gener].
 d. *Serà [a les 11h], quan tots els grups es reuniran [del dijous 18 de gener].
 e. *Serà [del dijous 18 de gener], quan tots els grups se reuniran [a les 11h].

- (45) a. Tots els grups es reuniran [a les 11h] [el dijous 18 de gener].

⁴Note that fronting and clefting tests must be applied to the whole PP containing the time expression, and not only to the NP part considered as the `timex3` extent. That is, it will be applied to: **a** *les 11h* ('**at** 11am'), instead of: *les 11h* ('11am'). And to: **del** *dijous 18 de gener* ('**of** Thursday, January 11th'), instead of: *el dijous 18 de gener* ('Thursday, January 18th').

- b. [El dijous 18 de gener], tots els grups es reuniran [a les 11h].
- c. [A les 11h], tots els grups es reuniran [el dijous 18 de gener].
- d. Serà [el dijous 18 de gener], quan tots els grups es reuniran [a les 11h].
- e. Serà [a les 11h], quan tots els grups es reuniran [el dijous 18 de gener].

Some other examples of one-constituent temporal expression are:

- (46) *la nit del 3 de gener del 2005* ('the night of January 3rd, 2005')
el dos de desembre ('December 2nd')
l'octubre del 1963 ('October 1963')
l'estiu de l'any passat ('last year's summer')
11 d'abril del 1996, 11:13 GMT. ('April 11th, 1996')

4.2.2 Anchoring relation: Involving a duration expression

Definition. Some of the constructions with anchoring relations are here referred to as *anchored durations*. Anchored durations contain a typical duration expression (see section 2.3), but refer in fact to a point in time (i.e., a date or time of day). They are so called because the duration is anchored to a further temporal reference. In Catalan, anchored durations are typically expressed by means of any of the following structures, which involve the use of temporal prepositions (simple or complex) denoting sequencing, like *a partir de*, *abans de*, *després de*, *seguint a*, *anterior/s a*, *des de*. Some constructions have the anchoring element explicit, whereas others have it implicit.

- **With an explicit anchoring reference:** In the examples by the side, the explicit anchoring element is in bold face, whereas the anchored duration extent is underlined.

1. $\text{NP}_{duration} + \left\{ \begin{array}{l} \textit{abans} \\ \textit{després} \\ \textit{a partir de} \end{array} \right\} + \textit{de} + \text{NP}_{time}$ (e.g., *tres dies* ***abans de Nadal***)
2. $\left\{ \begin{array}{l} \textit{queden}/\dots \\ \textit{falten}/\dots \end{array} \right\} + \text{NP}_{duration} + \textit{per} + \text{NP}_{time}$ (e.g., *queden tres dies* ***per Nadal***)
3. $\left\{ \textit{passen}/\dots \right\} + \text{NP}_{duration} + \textit{de} + \text{NP}_{time}$ (e.g., *passen tres dies* ***de Nadal***)
4. $\text{NP}_{time} + \left\{ \textit{fa}/\textit{feia}/\dots \right\} + \text{NP}_{duration} + \textit{de}$ (e.g., ***ahir*** *va fer cinc mesos*)
5. $\textit{de} + \left\{ \begin{array}{l} \textit{aquí} \\ \textit{avui} \end{array} \right\} + \left\{ \begin{array}{l} \textit{a} \\ \textit{en} \end{array} \right\} + \text{NP}_{duration}$ (e.g., ***d'aquí*** *a tres dies*)

- **With an implicit anchoring reference:** The anchoring element here is interpreted based on contextual information.

1. $\text{NP}_{duration} + \left\{ \begin{array}{l} \textit{abans} \\ \textit{després} \\ \textit{més tard} \end{array} \right\}_{[adv]}$ (e.g., *tres dies més tard*)

2. $\left\{ \begin{array}{l} \textit{queden/quedaven}/\dots \\ \textit{falt/en}/\textit{faltaven}/\dots \end{array} \right\} + \text{NP}_{\textit{duration}}$ (e.g., *queden dos anys*, *falta una mica*)
3. $\left\{ \begin{array}{l} \textit{fa/feia}/\dots \\ \textit{passen/passaven}/\dots \end{array} \right\} + \text{NP}_{\textit{duration}}$ (e.g., *fa una hora*, *feia molta estona*)

Annotation. Anchored durations will be annotated as indicated below:

- **With an explicit anchoring reference:** Two different treatments will be applied, depending on whether the resulting temporal reference denoted by the whole construction corresponds to a date or a time of day.

- a. If the resulting temporal entity refers to a **date**, such as:

- (47) a. Va arribar **avui** *fa un mes*.
 b. Va arribar quan *faltaven 3 dies* per **Nadal**.

The annotation of these constructions will include:

- A `timex3` tag of type `duration`, containing the expression denoting the anchored duration (underlined in the examples above).
- A `timex3` tag of type `date`, encoding the explicit anchoring reference (in bold face above).
- A non-consuming (or empty) `timex3` tag of type `date`, representing the resulting date that the whole construction is expressing. For instance, assuming that today is Friday, July 30, 2009, the construction *avui fa un mes* ('one month ago today') in example (47a) refers to the date June 30, 2009. This will be the value of such empty tag.⁵

- b. If the resulting temporal entity refers to a **time of day**, such as:

- (48) Va arribar quan *faltaven dos minuts per a les 3h de la tarda*.

The annotation of these constructions will include:

- A single `timex3` tag, of type `time`, will be marked up. Compare this treatment with that in construction (47b).

- **With an implicit anchoring reference**, such as:

- (49) a. Va arribar *fa un mes*.
 b. Va arribar *fa alguns mesos*.

These constructions will be marked up with the following set of tags:

⁵Furthermore, links will be used to express the relative ordering of the two time expressions. See the annotation guidelines for Link entities in TimeML.

- A `timex3` tag of type `duration`, corresponding to the span of the anchored duration in (49), underlined.
- A non-consuming (or empty) `timex3` tag of type `date`, representing the resulting date that the anchored duration is expressing. The value of this expression can be precise, as it would be the case in example (49a), or fuzzy, as in (49b), where we can interpret that the arrival took place in a past reference, few months earlier, but we do not know precisely how many.

4.2.3 Anchoring relation: Involving a date expression

Definition. The second type of constructions with anchoring relations are called here *anchored dates*. Anchored dates contain at least one date expression (see section 2.1), followed by an element of sequencing such as *anterior*, *següent*, *que segueix*, *posterior*, *abans*, *després* (‘preceding’, ‘following’, ‘that follows’, ‘next’, ‘before’, ‘after’), etc. Like anchored durations, anchored dates can have the anchoring reference explicit or implicit.

- **With an explicit anchoring reference:** The anchoring element can denote either a date (e.g., *Nadal* ‘Christmas’) or an event (e.g., *la reunió* ‘the meeting’). In the examples by the side, the explicit anchoring reference is in bold face, whereas the anchored date is underlined.

$$\begin{aligned}
 1. & \text{NP}_{time} + \left\{ \begin{array}{l} \textit{següent} \\ \textit{que segueix} \\ \textit{previ} \\ \textit{posterior} \\ \textit{anterior} \end{array} \right\} + a + \left\{ \begin{array}{l} \text{NP}_{time} \\ \text{NP}_{event} \end{array} \right\} \quad (\text{e.g., } \underline{\textit{el dimarts}} \textit{ següent a } \left\{ \begin{array}{l} \mathbf{Nadal} \\ \mathbf{la reunió} \end{array} \right\}) \\
 2. & \text{NP}_{time} + \left\{ \begin{array}{l} \textit{abans} \\ \textit{després} \end{array} \right\} + de + \left\{ \begin{array}{l} \text{NP}_{time} \\ \text{NP}_{event} \end{array} \right\} \quad (\text{e.g., } \underline{\textit{el dimarts}} \textit{ abans de } \left\{ \begin{array}{l} \mathbf{Nadal} \\ \mathbf{la reunió} \end{array} \right\})
 \end{aligned}$$

- **With an implicit anchoring reference:** The elements expressing the sequencing (or anchoring) relation have an adjectival use. Some of them can be found preceding the noun (as in pattern 1 below), although its most typical use is following it (pattern 2).⁶

$$1. \left\{ \begin{array}{l} \textit{el} \\ \textit{la} \end{array} \right\} + \left\{ \begin{array}{l} \textit{següent} \\ \textit{anterior} \\ \textit{passat} \end{array} \right\} + \text{N}_{time} \quad (\text{e.g., } \underline{\textit{el següent dimarts}})$$

⁶The expressions ‘*el N_{time} que segueix*’ (lit. ‘the *N_{time} which follows*’) and ‘*el N_{time} que ve*’ (lit. ‘the *N_{time} which is coming*’) are included here as well given their equivalence with the construction ‘*el N_{time} següent*’ (‘the coming *N_{time}*’).

$$2. \left\{ \begin{array}{l} el \\ la \end{array} \right\} + N_{time} + \left\{ \begin{array}{l} següent \\ que segueix \\ vinent \\ que ve \\ passat \\ anterior \\ posterior \\ abans \\ després \end{array} \right\} \quad (\text{e.g., } \underline{\text{el dimarts següent}})$$

Annotation. The annotation of these constructions is as follows:

- **With an explicit anchoring reference**, such as:

(50) La festa dels Ous Pintos se celebra a la localitat asturiana de Pola de Siero el dimarts següent al diumenge de Pasqua.

The annotation of these constructions will include:

- A `timex3` tag of type `date`, containing the expression denoting the anchored date (underlined in the example above).
- A `timex3` tag of type `date`, encoding the explicit anchoring date (in bold face above).

- **With an implicit anchoring reference**, such as:

(51) a. Em van donar hora per al dimarts següent a les 9 del matí.
 b. L'any que ve cada persona menjarà uns 40 Kg de carn.

The annotation of these constructions will include only:

- A `timex3` tag of type `date`, containing the expression that denotes the anchored date (underlined in the example above). Note that, given the adjectival behaviour of the sequencing elements in these constructions (e.g., *següent*, *que ve*, etc.), the span of the `timex3` tag will include them as well. This contrasts with their treatment in the case of anchored durations with explicit anchoring elements, where these are disregarded as part of any `timex3` tag extent and, instead, annotated as `signals`.

4.2.4 Range relation

Definition. A range relation involves two time expressions of type **date**, which respectively denote the begin and end points of an interval. For example:

- (52) a. entre el 2005 i el 2012
b. del 5 de desembre fins el 30 de març
c. la temporada 92-93

Typical patterns in Catalan for range relations of this type are:

1. *entre* + NP_{time} + *i* + NP_{time} (e.g., *entre* dilluns *i* dijous)
2. (*des*) *de* + NP_{time} + $\left\{ \begin{array}{l} (fins) (a) + NP_{time} \\ encà \end{array} \right\}$ (e.g., *de* dilluns/llavors/ara (*fins*) *a* dijous)
3. NP_{time} + $\left\{ \begin{array}{l} - \\ / \end{array} \right\}$ + NP_{time} (e.g., 1992/93, 92-93)

Annotation. Some examples illustrating this relation are:

- (53) Des de Nadal encà, s'han intensificat els assajos.
- (54) Però des del 2005 fins al 2008, amb l'explosió del baix cost, el Prat ha aconseguit repetidament situar-se com el segon aeroport de l'Estat en passatgers.

Their annotation includes:

- Two tags of type **date**, each spanning over the two date expressions (underlined above).
- An additional empty tag of type **duration**, encoding the span of the interval delimited by these two dates. In example (54), for instance, it will encode an interval of 3 years.

It must be pointed out, however, that these annotation guidelines may not be applicable in the case of pattern 3. In particular, when it is not possible to annotate only token fragments, be it given a limitation in the annotation tool or because the text has not previously tokenized these elements into separate tokens, as is the case here. Hence, in the current annotation effort, pattern 3 will be annotated as follows:

- A single **timex3** tag of type **duration**, scoping over the extent of the 2 date expressions.
- Two empty tags of type **date**, each referring to one of the date expressions.

Example (110) illustrates this marking-up in detail.

4.2.5 Frame relation

Definition. Similar to range relations, constructions conveying framing relations also denote a time interval. However, they contain a date (in bold face below) and a duration expression (underlined), instead of the two date expressions observed before. The date refers to a particular temporal frame within which the duration is located.

- (55) a. tres setmanes d'**octubre** ('three weeks in October ')
 b. els dos últims mesos d'**any** ('the last two months of the year')
- (56) a. la segona quinzena d'**octubre** ('the second fortnight of October ')
 b. el primer ter c de **segle** ('the first third of this century')
- (57) a. el que va d'**any** (lit., 'what is gone from this year')
 b. el que queda d'**any** (lit., 'what is left from this year')

The date expresses one of the boundaries of the interval. For instance, in (57b), the boundary is set by the end of the current day. In Catalan, the typical patterns for constructions presenting a framing relation is:

1. $N_{duration} + de + N_{date}$ (as in examples (55-56))
2. $el\ que \left\{ \begin{array}{l} queda \\ segueix \\ va \end{array} \right\} de + N_{date}$ (e.g., *el que queda d'any*)

Annotation. The following tags will be introduced:

- A tag of type **duration**, spanning over the expression with the pattern *el que queda/va/ segueix/...* (underlined above).
- A tag of type **date**, spanning over the N_{time} (in bold face above).
- Two (potentially empty) tags of type **date**, expressing the begin and end points from which the length of the duration is computed. For example, if the expression in (??) is uttered in April 2009, the values of these two tags will be, respectively, January of 2009, and April of 2009.

4.2.6 Conjunction relation

Definition. Involving two time expressions related by a coordination conjunction, mainly, *i* and *o* ('and' and 'or'). For example: [*sis mesos*] *o* [*un any*] ('six months or a year').

Annotation. The two coordinated expressions will be marked up as two different tags if they refer to two independent points or intervals of time. For example, the conjunction in (59) is expressing an alternative between two different durations. Therefore, the two time expressions will be marked up independently, as indicated by the underlining.⁷

(58) Saddam jugarà les mateixes cartes de nou d'aquí a sis mesos o un any a partir d'avui.

On the other hand, the conjunction in (59) is expressing a relation of specification between the three time expressions. That is, *tres anys* ('three years') is modified by *tres mesos* ('three months') and *tres dies* ('three days'). The three constructions will be marked up as one single `timex3` entity.

(59) A inicis de març del 2009 va començar, a Dag Shang Kagyu, el primer retir tradicional de tres anys, tres mesos i tres dies.

5 Timex3 attributes

The attributes for `timex3` entities in TimeML are listed below. Those between parentheses are not relevant for the current annotation edition.

1. `type`
2. `value`
3. `mod`
4. `(temporalFunction)`
5. `anchorTimeID`
6. `(valueFromFunction)`
7. `(functionInDocument)`
8. `beginPoint`
9. `endPoint`
10. `quant`
11. `freq`

5.1 Attribute type

Each `timex3` is assigned one of the following types: `DATE`, `TIME` (for times of day), `DURATION`, or `SET`. Section 2 provides a detailed description of the different time expressions corresponding to each of these values.

⁷In example (59), the time expression *avui* ('today') is in an anchoring relation with both *sis mesos* ('six months') and *un any* ('one year'). Hence, *avui* ('today') needs to be marked up as an additional `timex3` entity. We omit it here for the sake of simplicity.

5.2 Attribute value

The attribute `value` indicates the temporal reference expressed by the `timex3` expression, that is, the date, time of day, length of the duration that it is being denoted, etc. This value is given in an extended ISO 8601 format, which applies a different pattern depending on the `type` attribute. For instance, a `DURATION` must have a value that begins with 'P' (standing for 'period of time'), and a `TIME` a value beginning with the letter 'T' (standing for 'time'). The main guidelines for annotating this attribute are presented in the following subsections, based on the `timex3` types.

5.2.1 Dates

Days. They are expressed in the form of 8 digits separated by means of hyphens: YYYY-MM-DD. The first 4 digits represent the year, the next 2 express the month, and the last 2 convey the day. For example, the time expression in (60a) will be represented as (60b).

- (60) a. El 29 de juny del 2009
b. `value`: 2009-06-29

The annotator will introduce as much information as is available. For example, given the sentence in (61a), and assuming that the document creation time is Friday, July 12, 2002, then the `value` attribute of the underlined `timex3` must specify the full date that can be computed from the document creation time, that is (61b).

- (61) a. La reunió de l'últim divendres va ser un èxit
b. `value`: 2002-07-05

Unknown information is left underspecified by means of the placeholder 'X'. In the next example, the year is unknown:

- (62) a. El 20 de maig
b. `value`: XXXX-05-20

Week and weekend dates. The format for weeks is YYYY-Www, where YYYY indicates the year, and Www is the week number prefixed by the letter 'W'. W01 refers to the first week of the year and W53 to the last one.

- (63) a. La setmana passada.
b. `value`: 2009-W30

References to the day of the week (i.e., Monday to Sunday) will be expressed with the more complete format: YYYY-Www-D, where D is the weekday number, from 1 (Monday) to 7 (Sunday). This format will be applied if the text presents the trigger expression *semana* ('week'), or if the `timex` expresses a generic reference, as in (64). Otherwise, dates will be represented in the year-month-date format introduced earlier (YYYY-MM-DD).

- (64) a. El dilluns no és un bon dia per a reunions. ??
b. value: XXXX-WXX-1

If the timex refers to a weekend, the token WE will be appended at the weekday position:

- (65) a. El primer cap de setmana d'aquest any.
b. value: 2009-W01-WE

As before, the annotator should include as much information as possible. Unknown or underspecified information is left underspecified with the placeholder 'X'.

Months and years. References to months are specified as: YYYY-MM. References to years are expressed as: YYYY.

- (66) a. Va néixer el març del 1963.
b. value: 1963-03

- (67) a. Va néixer el 1963.
b. value: 1963

Given that ISO assumes that years have four digits, years before the year 1000 (or after 1000 BCE) are written with leading 0s:

- (68) a. L'any 878, Guifré rebé els comptats de Barcelona, Girona i Besalú .
b. value: 0878

Decades, Centuries, Millenia. In order to represent these units, we will follow TIDES 2005, which introduces a flexible solution on top of the fix-width, four-place annotation for years in ISO (i.e., YYYY). Specifically, we consider the first position to represent the “millenium” element, the second the “century” element, the third the “decade” element, and the fourth the “year” component.

Thus, according to this version, a well-formed value can consist of only three digits, two, or one. **Decades** will be expressed with the format **YYY** (69), **centuries** will follow the format **YY** (70), and **millenia** will apply the format **Y** (73).

- (69) a. els anys seixanta
b. value: 196

- (70) a. el proper segle
b. value: 21⁸

- (71) a. el proper mil.lenni
b. value: 3

⁸Note that the value is not in accordance with the name of the century (twenty-second, and thus 22, but it agrees with the prefix of the years during that century (21XX), hence 21.

Before Current Era (BCE). The ISO standard has no way to capture Before Current Era (BCE) dates, and so we follow the solution adopted by TIDES 2005 and others in the community. Namely, we will use the token `BC` as a prefix to the year portion of the ISO value.

- (72) a. Buda va néixer cap a l'any 563 abans de la nostra era.
 b. `value:` BC0536

We also follow the convention of pretending that there is a year 0 in the Judeo-Christian calendar, to make the arithmetic straightforward for computers.

- (73) a. Aquest rei va viure fa 4000 anys. [Stated in 2009]
 b. `value:` BC1990 [2009+1990+1 = 4000]

Note that in TimeML, the timex *fa 4000 anys* above is considered an anchored duration. Hence, the value showed in (73b) corresponds to the value of the non-consuming `timex3` tag of type `date`, which is associated to the duration. Review section 4.2.2 for further details.

Seasons. Seasons have different meanings to different people. For example, some might interpret “summer” as the hot part of the year, while others interpret it more literally, from summer solstice to fall equinox. This is the reason why tokens, rather than precise numerical values, are used in the `value` attribute. Seasons will be represented using the tokens in Table 10.

Table 10: Season tokens

Period	Token	Position in value	Example expression	Example annotation
Spring	SP	YYYY-SP	<i>La primavera del 2009</i> <i>El període primaveral</i>	<code>value:2008-SP</code> <code>value:XXXX-SP</code>
Summer	SU	YYYY-SU	<i>L'estiu passat</i> <i>L'època d'estiu</i>	<code>value:2008-SU</code> <code>value:XXXX-SU</code>
Fall	FA	YYYY-FA	<i>La tardor que ve</i> <i>A la tardor</i>	<code>value:2009-FA</code> <code>value:XXXX-FA</code>
Winter	WI	YYYY-WI	<i>L'hivern del 1962</i> <i>Durant l'hivern</i>	<code>value:1962-WI</code> <code>value:XXXX-WI</code>

Yearly quarters, trimesters, and halves (or semesters). Quarters (equivalent to *trimestres* in Catalan), quatrimesters, and halves (equivalent to *semestres*) will be expressed using the tokens in Table 11. Quatrimesters are units not contemplated by the ISO 8601 standards, but given that it is a time unit quite frequent in (at least some parts of) the Catalan-speaking culture, we include here the notation `TX` for the sake of convenience.

Table 11: Season tokens

Period	Token	Position in val.	Example expression	Example
First quarter	Q1	YYYY-Q1	<i>el primer trimestre del 2001</i>	2001-Q1
Second quarter	Q2	YYYY-Q2	<i>el segon trimestre del 2001</i>	2001-Q2
Third quarter	Q3	YYYY-Q3	<i>el tercer trimestre del 2001</i>	2001-Q3
Fourth quarter	Q4	YYYY-Q4	<i>el quart trimestre del 2001</i>	2001-Q4
First quadrimester	T1	YYYY-T1	<i>el primer quadrimestre del 2001</i>	2001-T1
Second quadrimester	T2	YYYY-T2	<i>el segon quadrimestre del 2001</i>	2001-T2
Third quadrimester	T3	YYYY-T3	<i>el tercer quadrimestre del 2001</i>	2001-T3
First half	H1	YYYY-H1	<i>el primer semestre del 2001</i>	2001-H1
Second half	H2	YYYY-H2	<i>el segon semestre del 2001</i>	2001-H2

Geologic eras. In referring to the far distant past, scientific convention does not use the standard calendar and instead gives the time as offset. We adopt the same two-letter abbreviations used by geologists, and place them at the initial position in the value attribute:

- KA Some number of thousand years ago.
- MA Some number of million years ago.
- GA Some number of billion years ago.

For example:

- (74) a. El Parc Juràssic va començar fa 210 milions d’anys.
b. value: MA210

Fiscal years. They are expressed using the token FY prefixed to the year number.

- (75) a. L’any fiscal actual.
b. value: FY2009

Fuzzy time expressions. Table 1 provides some of the expressions in Catalan bearing fuzzy semantics. The tokens in (103) below will be used for annotating this type of expressions.

- (76) PAST_REF: The time expression denotes a past reference.
PRESENT_REF: The time expression denotes a present reference.
FUTURE_REF: The time expression denotes a future reference

5.2.2 Time of day

Precise day times. They have the format: THH:MM:SS, where the digits at the H positions express the hour (conceived in cycles of 24 hours), those at the M positions express the minutes, and those at the S positions express the seconds. The positions for minutes and seconds are optional. Thus, alternative formats in case that minutes and/or seconds are not specified are: THH:MM and THH. For example:

- (77) a. Se'n va anar a les 6h de la tarda.
b. value: T18:00

Note that when the time referred to is at the top of the hour (e.g., *les dues en punt*, '2 o'clock'), the minute portion is given as :00 to ensure that it is not misinterpreted as unknown (expressed with the placeholder X).

If the text includes some reference to the specific date in which the time is anchored (78a), then the **value** attribute must also contain the date. This is exemplified in the sentence below. Note that the date expression (*el 29 de juny del 2009*, 'June 29th, 2009') is annotated independently from the time expression (*les 6h de la tarda*, '6pm in the afternoon'). The former receives the **value** in (78b), while the second receives that in (78c). However, the information in the former is used to complete the **value** attribute of the latter.

- (78) a. Se'n va anar **el 29 de juny del 2009** a les 6h de la tarda.
b. value: 2009-06-29
c. value: 2009-06-29T18:00

For purposes of annotation consistency, '12:00am midnight' will always be represented as hour 24 of the earlier day:

- (79) a. Ahir van estar desperts fins a les 12h de la mitjanit. [Uttered on July 30, 2009]
b. value: 2009-07-29T24:00

If the time is given in Universal Coordinated Time (UTC) or Greenwich Meridian Time (GMT), the letter Z will be added at the end of the value:

- (80) a. 8 de gener del 2001, 11:13 GMT
b. value: 2001-01-08T11:13Z

If the time is given relative to UTC or GMT, as e.g. UTC+2 or GMT-4, the value will then be normalized to the corresponding UTC or GMT value. For example:

- (81) a. 8 de gener de 2001, 10:13 GMT+4
b. value: 2001-01-08T06:13Z

Table 12: Season tokens

Period	Token	Example expression	Example annotation
Morning	MO	<i>avui al matí</i>	2009-07-30TMO
Noon	MI	<i>avui al migdia</i>	2009-07-30TMI
Afternoon	AF	<i>avui a la tarda</i>	2009-07-30TAF
Evening	EV	<i>avui al vespre</i>	2009-07-30TEV
Night	NI	<i>avui a la nit</i>	2009-07-30TNI
Day time (i.e., matí + tarda)	DT	<i>en hores de treball</i>	2009-07-30TDT

Periods of the day. Day periods, such as *matí*, *tarda*, *nit* (‘morning’, ‘afternoon’, ‘night’), will be represented by the tokens in Table 12, placed at the hour position in the standard ISO representation, namely, at the position signaled with \$\$ in: YYYY-MM-DDT\$\$.

These tokens will be used only if the precise time is not present in the expression. For instance, *les 11h del matí* (‘11am in the morning’) will be given the day time value of T11:00.

5.2.3 Durations

They are represented by the format P[n]Y[n]M[n]DT[n]H[n]M[n]S or P[n]W. In these representations, the [n] is replaced by the value for each of the date and time elements that follow the [n]. Letter P is the duration designator (historically called *period*) placed at the start of the duration value, and letter T is the time designator, preceding the time components of the representation. Furthermore, the other capital letters in the formula express:

- Y Year designator that follows the value for the number of years.
- M Month designator.
- W Week designator.
- D Day designator.
- H Hour designator.
- M Minute designator (always needs to be preceded by the T)
- S Seconds designator.

For example, P1Y2M3DT4H5M6S represents a duration of *one year, two months, three days, four hours, five minutes, and six seconds*. Date and time elements including their designator may be omitted if their value is zero. Similarly, lower order elements may also be omitted for reduced precision. For example, P11DT22H (i.e., 11 days and 22 hours) and P33Y (i.e., 33 years) are both acceptable duration representations.

The value may also have a decimal fraction, as in P0.5Y to indicate half a year. Note that this same duration can be expressed as P6M as well. However, the format should follow the linguistic expression wherever possible. Hence:

- (82) a. ... durant mig any value:P0.5Y
 b. ... durant sis mesos value:P6M

To avoid ambiguity between references to months and references to minutes, P1M is a one-month duration and PT1M is a one-minute duration (note the time designator, T, that precedes the minute value).

Durations can also be expressed using the tokens to refer to: periods of the day (MO, MI, AF, EV, NI, DT), weekends (WE), yearly seasons (SP, SU, FA, WI), quarters (Q), quatrimesters (T), year halves (H), and fiscal years (FY). For example:

- (83) a. Les sessions d'introducció al Linux duren dues tardes.
 b. value:PT2AF
- (84) a. El període de reproducció d'aquesta au dura tota la primavera.
 b. value:P1SP
- (85) a. El curs d'interpretació de nivell avançat ocupa tres trimestres.
 b. value: P3Q

5.2.4 Sets

To fully annotate sets, we need, in addition to the attribute **value**, either the **quant** or **freq** attributes, if not both. The meaning of these two attributes will be presented later (section 5.9). For the time being, we focus on the **value** attribute.

For sets, the **value** attribute expresses the time interval in which the iteration (of events or times) takes place. For example, in the expression *dues vegades a la setmana* ('twice a week'), the interval is a duration of one week –hence, **value=P1W**. Additional examples follow:

- (86) a. cada dos dies
 b. value:P2D
- (87) a. cada octubre
 b. value:XXXX-10
- (88) a. tots els dimarts
 b. value:XXXX-WXX-2

5.3 Attribute mod

This attribute is inherited directly from TIDES 2005 (Ferro et al., 2005). Hence, the section is highly based on the description offered there.

Time expressions can be qualified or modified in some ways. For instance, *2009* is an unmodified expression, whereas *principis del 2009* ('beginning of 2009') is modified. In general, we want the annotation to capture the basic semantics of quantity modifiers, such as *aproximadament*, *no més de* ('approximately', 'no more than'), and lexicalized aspect markers, e.g., *principis de*, *mitjans de* ('the beginning of', 'mid').

The attribute `mod` captures the semantics of those modifiers within the scope of the `timex3` expression, but not the semantics of leading prepositions or other terms that are outside the extent of the tagged `timex3`. For example, the expression *abans del dimarts* ('before Tuesday') is not considered a modified expression for our purposes, because the preposition *abans de* ('before') is not included within the tag extent.

Some of the `mod` values apply only to points in time (i.e., dates and times of day), other values apply only to durations, and finally other values apply to either, as presented in Table 13.

Table 13: Values for the `mod` attribute

	Token	Example
Points	BEFORE	<i>fa més d'un any</i>
	AFTER	<i>fa menys d'un any</i>
	ON_OR_BEFORE	<i>fa almenys un any</i>
	ON_OR_AFTER	<i>no fa més d'un any</i>
Intervals	LESS_THAN	<i>menys de dues hores, quasi dues hores.</i>
	MORE_THAN	<i>més de 5 minuts</i>
	EQUAL_OR_LESS	<i>no més de tres dies</i>
	EQUAL_OR_MORE	<i>almenys tres dies</i>
Both	START	<i>al principi de, els primers (dies, mesos, etc.), començant (el dia, mes, etc.)</i>
	MID	<i>a meitat de, a mitjans de</i>
	END	<i>acabant, a finals de</i>
	APPROX	<i>al voltant de, aproximadament, més o menys</i>

Note that the examples in the first 4 rows seem to be referring to durations instead of points –e.g., *més d'un any*, *al menys un any* ('more than one year', 'at least one year'). Recall, however, the notion of *anchored durations* (end of section 2.3): they are time constructions denoting a duration anchored to a point in time in order to express a second point in time. The duration corresponds, in fact, to the temporal distance between the two points. Hence, the complete annotation of the temporal expressions in (89) and (91) is as shown in (90) and (92), respectively.

(89) Microsoft ha admès que coneixia des de feia més d'un any un error de seguretat a l'Internet Explorer. [Published: August 4, 2008]

(90) a. `extent`: *fa més d'un any* `type`:DURATION `value`:P1Y `mod`:MORE_THAN
b. `extent`: [empty tag] `type`:DATE `value`:2007-08-04 `mod`:BEFORE

(91) Ara fa quasi vint anys de la caiguda del mur de Berlin. [Published: September 24, 2008]

(92) a. `extent`: *ara* `type`:DATE `value`:PRESENT_REF `mod`:--
b. `extent`: *fa quasi vint anys* `type`:DURATION `value`:P20Y `mod`:LESS_THAN
c. `extent`: [empty tag] `type`:DATE `value`:1988-09-24 `mod`:AFTER

5.4 Attribute `temporalFunction`

This attribute will not be used for TempEval-2010.⁹

`TemporalFunction` is a binary attribute which expresses whether the value of the temporal expression needs to be determined via evaluation of a temporal function. The value of this attribute will be positive for those cases that do not contain all the information necessary to fill the higher-order (or left-hand) positions in the `value` attribute. Consider for example, the time expression *les 10h del matí* ('10am in the morning') in (104). Given that the part of information concerning the date in its `value` attribute has been provided by another time expression (i.e., *el 30 de juny del 2009*, 'June 30th, 2009'), the `temporalFunction` value for *les 10h. del matí* will be `True`:

(93) Va arribar el 30 de juny del 2009, a les 10h. del matí.

(94) a. `ext`: *el 30 de juny del 2009* `type`:DATE `value`:2009-06-30 `temporalFunction`:False
b. `ext`: *les 10h. del matí* `type`:TIME `value`:2009-06-30T10:00 `temporalFunction`:True

Even if the context is not providing the complete information, all expressions that have some higher-order position underspecified will have `temporalFunction: True`. Some examples are:

(95) *el 31 de gener* ('January 31st')
la nit d'ahir ('yesterday's night')
la setmana passada ('last week')

On the other hand, for cases in which the higher-order position of the values are filled, the `temporalFunction` attribute should be assigned a `False` value. Such cases include:

⁹The presence of the attributes `anchorTimeID` (see section 5.5), `beginPoint` or `endPoint` (both in section 5.8) will suffice to set the `temporalFunction` attribute to `True`.

- (96) *les 4h. de l'1 de març del 1980* ('4am on March 1st, 1980')
l'estiu del 2008 ('summer of 2008')
dimecres, 20 de març del 1963 ('Wednesday, March 20th 1963')

Time expressions of type `DURATION` will receive `True` in case of underspecification. Examples of expressions in this situation are:

- (97) *diversos mesos* ('several months')
alguns anys ('some years')
les hores següents ('the following hours')

Examples of durations that will receive `False` as the value of their `temporalFunction` attribute are:

- (98) *dues hores* ('two hours')
31 dies ('31 days')
9 mesos ('9 months')

5.5 Attribute `anchorTimeID`

Optional attribute. It introduces the ID of the time expression to which the `timex3` in question is anchored. That is, the time expression that is needed in order to compute the `value` attribute of the current `timex3`.

In (99), for instance, the two underlined `timex3` expressions are anchored to another time expression. In order to know the particular year in the `value` attribute of *el dilluns 15 de juliol* ('Monday, July 15th'), this expression needs to be evaluated with respect to the document creation time (DCT), with ID number `ID: t0`. And the `value` of the expression *les 9h. del matí* ('9am in the morning') needs to be evaluated relative to *el dilluns 15 de juliol* (Monday, July 15th), with ID: `t1`.

- (99) El taller TERQAS es reemplenirà el dilluns 15 de juliol. La sessió començarà a les 9h del matí.

- (100) a. `ext: 2002-07-12` (DCT) ID:t0 value:2002-07-12
b. `ext: el dilluns 15 de juliol` ID:t1 value:2002-07-15 anchorTimeID:t0
c. `ext: les 9h. del matí` ID:t2 value:2002-07-15T09:00 anchorTimeID:t1

Similarly, some duration `timexes` need information provided by other duration expressions in order to have its `value` attribute filled, as is the case with expression *tres* in the following example:

- (101) El viatge dura entre tres i cinc hores.

- (102) a. `ext: tres` ID:t1 value:PT3H anchorTimeID:t2
b. `ext: cinc hores` ID:t2 value:PT5H

5.6 Attribute valueFromFunction

Attribute not relevant for the purpose of manual annotation. Omitted for TempEval 2010.

5.7 Attribute functionInDocument

Omitted for TempEval 2010. It indicates the function of a `timex3` in providing a temporal anchor for other temporal expressions in the document. There are several times that mark the major milestones in the life of a textual document:

(103) CREATION_TIME:	The time the text is created.
MODIFICATION_TIME:	The time the text is modified.
PUBLICATION_TIME:	The time the text is published.
RELEASE_TIME:	The time it may be released (if not immediately).
RECEPTION_TIME:	The time it is received by a reader.
EXPIRATION_TIME:	The time that the text expires (if any).
NONE:	This attribute is not explicitly supplied.

5.8 Attributes beginPoint and endPoint

Optional attributes. These two attributes are present in tags of type `DURATION` used for annotating *anchored durations* and *range* expressions. In particular, they encode the ID of the temporal expressions referring to the begin and end points of such durations.

Examples (104)-(105) illustrate their use in the case of anchored durations, whereas examples (106)-(107) show it for range expressions. Focus on the tags of type `DURATION`, in (105a) and (107c). The `timex3` Ids given as values of the attributes `beginPoint` and `endPoint` refer to the Ids of the two other `timex3` tags (of type `DATE`) in each annotation.

(104) La tercera reunió tindrà lloc una setmana després del 10 de juliol.

(105) a. ext: <i>una setmana despres</i>	ID:t1	type:DURATION	value:P1W	
	beginPoint:t2		endPoint:t3	
b. ext: <i>del 10 de juliol</i>	ID:t2	type:DATE	value:2009-07-10	
	temporalFunction:True		anchorTimeID:t0	
c. ext: [empty tag]	ID:t3	type:DATE	value:2009-07-17	
	temporalFunction:True		anchorTimeID:t1	

(106) Els falciots arriben entre principis de maig i finals de juny.

(107) a. ext: <i>principis de maig</i>	ID:t1	type:DATE	value:XXXX-05	mod:START
	temporalFunction:False			
b. ext: <i>finals de juny</i>	ID:t2	type:DATE	value:XXXX-06	mod:END
	temporalFunction:False			
c. ext: [empty tag]	ID:t3	type:DURATION	value:P2M	
	beginPoint:t1		endPoint:t2	

Finally, the next example illustrates the use of these two attributes in the case of range expressions following pattern: $NP_{time} + \left\{ \begin{array}{c} - \\ / \end{array} \right\} + NP_{time}$.

Two alternative annotations are provided for the same expression in sentence (108). In the first case (109), we see what would be the annotation in optimal conditions, that is, when it is possible to mark up only token fragments. The second annotation variation (110), on the other hand, illustrates the annotation for when, due to constraints from the annotation tool, the two date expressions will need to be annotated as a unique `timex3` tag of type `duration` (refer to section 4.2.4).

(108) A principis de la temporada 94-95 va ser traspasat a Golden State Warriors.

(109) a. ext: <i>94</i>	ID:t1	type:DATE	value:1994-XX		
		temporalFunction:	False		
b. ext: <i>95</i>	ID:t2	type:DATE	value:1995-XX		
		temporalFunction:	False		
c. ext: [empty tag]	ID:t3	type:DURATION	value:P1Y	mod: APPROX	
	beginPoint:t1		endPoint:t2		

(110) a. ext: [empty tag]	ID:t1	type:DATE	value:1994-XX		
b. ext: [empty tag]	ID:t2	type:DATE	value:1995-XX		
c. ext: <i>94-95</i>	ID:t3	type:DURATION	value:P1Y	mod: APPROX	
	beginPoint:t1		endPoint:t2		

5.9 Attributes `quant` and `freq`

These attributes are used for a temporal expressions of type `SET`. `quant` is generally a literal from the text that quantifies over the expression. `freq` contains an integer value and a time granularity that represent the frequency at which the temporal expression regularly reoccurs. These attributes are only used if their values are supplied by the temporal expression (or by a temporal anchor). Though it seems on occasions that values for these attributes can be inferred, they will not be for purposes of manual annotation. Although, if there is no specified `quant`, one imagines that the set is universally quantified. The following examples complete the annotations of the sets listed earlier in this section:

(111) a. <i>dos cops per setmana</i>	type:SET	value:P1W	freq:2X		
b. <i>cada dues hores</i>	type:SET	value:P2H	quant:EVERY		
c. <i>3 dies cada setmana</i>	type:SET	value:P1W	quant:EVERY	freq:3D	
d. <i>cada octubre</i>	type:SET	value:XXXX-10	quant:EVERY		

A Complex time expressions

A.1 Dates

Pattern	Example	Annotation
Range expressions		
1. $entre + NP_{time} + i + NP_{time}$	<i>entre <u>dilluns</u> i <u>dijous</u></i>	2 tags of type DATE (underlined)
2. $(des) de + NP_{time} + \left\{ \begin{array}{l} (fins) (a) + NP_{time} \\ encà \end{array} \right\}$	<i>de <u>dilluns</u> <u>encà</u></i> <i>des de <u>dilluns</u> fins <u>avui</u></i>	1 empty ag of type DURATION (for the resulting interval value)
3. $NP_{time} + \left\{ \begin{array}{l} - \\ / \end{array} \right\} + NP_{time}$	<i><u>1992-93</u></i>	1 tag of type DURATION 2 empty tags of type DATE
Anchored dates		
– With an implicit anchoring reference:		
1. $\left\{ \begin{array}{l} el \\ la \end{array} \right\} + \left\{ \begin{array}{l} següent \\ anterior \\ passat \end{array} \right\} + N_{time}$	<i><u>el següent dimarts</u></i>	1 tag of type DATE (underlined)
2. $\left\{ \begin{array}{l} el \\ la \end{array} \right\} + N_{time} + \left\{ \begin{array}{l} següent \\ que segueix, ve \\ vinent \\ passat \\ previ \\ anterior \\ posterior \\ abans \\ després \end{array} \right\}$	<i><u>el dimarts següent</u></i>	
– With an explicit anchoring reference:		
3. $NP_{time} + \left\{ \begin{array}{l} següent \\ que segueix \\ posterior \\ anterior \\ previ \end{array} \right\} + a + \left\{ \begin{array}{l} NP_{time} \\ (NP_{event}) \end{array} \right\}$	<i><u>el dia</u> previ a <u>cap d'any</u></i>	2 tags of type DATE (underlined)
4. $NP_{time} + \left\{ \begin{array}{l} abans \\ després \end{array} \right\} + de + \left\{ \begin{array}{l} NP_{time} \\ (NP_{event}) \end{array} \right\}$	<i><u>el dimarts</u> abans de <u>Nadal</u></i>	

A.2 Durations

Pattern	Example	Annotation
Frame durations		
1. $el\ que + \left\{ \begin{array}{l} queda \\ va \\ \dots \end{array} \right\} de + N_{time}$ 2. $la\ resta\ de + N_{time}$ 3. $el/la + \left\{ \begin{array}{l} primer/a \\ \dots \end{array} \right\} + N_{part/coll} + de\ N_{time}$	$\underline{el\ que\ queda\ d'any}$ $\underline{la\ resta\ de\ l'any}$ $\underline{el\ primer\ terç\ de\ mes}$	1 tag of type DURATION (underlined) 1 tag of type DATE (bold face) 2 (possibly empty) tags of type DATE (for the begin and end points of the duration)
Anchored durations		
– With an implicit anchoring reference:		
1. $NP_{duration} + \left\{ \begin{array}{l} abans \\ després \\ més\ tard \end{array} \right\}$ 2. $\left\{ \begin{array}{l} queden/\dots \\ falten/\dots \end{array} \right\} + NP_{duration}$ 3. $\left\{ \begin{array}{l} fa/\dots \\ passen/\dots \end{array} \right\} + NP_{duration}$	$\underline{tres\ dies\ més\ tard}$ $\underline{queden\ dos\ anys}$ $\underline{fa\ una\ hora}$	1 tag of type DURATION (underlined) 1 empty tag of type DATE (for the resulting date value)
– With an explicit anchoring reference:		
4. $NP_{dur} + \left\{ \begin{array}{l} abans \\ després \\ a\ partir\ de \end{array} \right\} + de + NP_{time}$ 5. $\left\{ \begin{array}{l} queden/\dots \\ falten/\dots \end{array} \right\} + NP_{dur} + per + NP_{time}$ 6. $\left\{ \begin{array}{l} passen/\dots \end{array} \right\} + NP_{dur} + de + NP_{time}$ 7. $NP_{time} + \left\{ \begin{array}{l} fa/\dots \end{array} \right\} + NP_{dur} + de$ 8. $de + \left\{ \begin{array}{l} aquí \\ avui \end{array} \right\} + \left\{ \begin{array}{l} a \\ en \end{array} \right\} + NP_{dur}$	$\underline{tres\ dies\ abans\ de\ Nadal}$ $\underline{queden\ tres\ dies\ per\ Nadal}$ $\underline{passen\ tres\ dies\ de\ Nadal}$ $\underline{ahir\ va\ fer\ cinc\ mesos}$ $d'aquí\ a\ \underline{tres\ dies}$	a. Referring to a DATE: 1 tag of type DURATION (underlined) 1 tag of type DATE (bold) 1 empty tag of type DATE (for the resulting date value) b. Referring to a TIME: 1 tag of type TIME

B TimeML attributes and values for `timex3`

Attribute:	Possible values:
<code>type</code>	DATE TIME DURATION SET
<code>value</code>	Value compliant with the extended 8601 format.
<code>mod</code>	For points: BEFORE AFTER ON_OR_BEFORE ON_OR_AFTER
	For durations: LESS_THAN MORE_THAN EQUAL_OR_LESS EQUAL_OR_MORE
	For both: START MID END APPROX
<code>anchorTimeId</code>	ID of another <code>timex3</code> (e.g., <code>t11</code>).
<code>beginPoint</code>	ID of another <code>timex3</code> (e.g., <code>t12</code>).
<code>endPoint</code>	ID of another <code>timex3</code> (e.g., <code>t13</code>).
<code>quant</code>	Literal from the text quantifying over the time expression (e.g., <code>EVERY</code>).
<code>freq</code>	Integer value and time granularity that represent the frequency at which the time expression reoccurs (e.g., <code>3D</code>).

Acknowledgments. This work has been supported by an EU Marie Curie International Reintegration Grant (PIRG04-GA-2008-239414) to the first author.

References

- Ferro, L., Gerber, L., Mani, I., Sundheim, B., & Wilson, G. (2005). *TIDES 2005 Standard for the Annotation of Temporal Expressions*. The MITRE Corporation.
- Pustejovsky, J., Knippen, B., Littman, J., & Saurí, R. (2005). Temporal and event information in natural language text. *Language Resources and Evaluation*, 39(2), 123–164.
- Saurí, R., Saquete, E., & Pustejovsky, J. (2010). *Annotating Time Expressions in Spanish. TimeML Annotation Guidelines. Version TempEval-2010*. Barcelona Media Technical Report BM 2010-02.
- Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Katz, G., & Pustejovsky, J. (2007). SemEval-2007 task 15: TempEval temporal relation identification. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, (pp. 75–80), ACL.
- Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Moszkowicz, J., & Pustejovsky, J. (2009). The TempEval challenge: identifying temporal relations in text. *Language Resources and Evaluation*.