

# TimeLine: Cross-Document Event Ordering

## SemEval 2015 - Task 4

### Manual Annotation Guidelines

Anne-Lyse Minard, Alessandro Marchetti, Manuela Speranza, Bernardo Magnini  
Fondazione Bruno Kessler

Marieke van Erp  
VU University Amsterdam

Itziar Aldabe, Rubén Urizar, Eneko Agirre, German Rigau  
The University of the Basque Country

[Events](#)  
[Time Anchor](#)  
[Event Ordering](#)  
[TimeLine](#)

#### 1. Events

The annotation of event mentions is based on the intra-document annotation guidelines defined under the NewsReader project:

<http://www.newsreader-project.eu/files/2013/01/NWR-2014-2.pdf>.

*“Event is used as a cover term to identify “something that can be said to obtain or hold true, to happen or to occur” [ISO TimeML Working Group, 2008]. This notion can also be referred to as eventuality [Bach, 1986] including all types of actions (punctuals or duratives) and states as well.”* (section 1, NewsReader Guidelines)

Some events annotated following the NewsReader guidelines could not go on a timeline, for example because they didn't happen (counter-factual events) or they are uncertain. In order to annotate only events potentially candidates to participate to a timeline, we have defined criteria based on the NewsReader Guidelines.

We annotate verbs, except if they are modified by a modal word, nouns and pronouns.

Adjectives generally express a property or attribute of an entity, and anchoring them in time is not simple. So adjectival events will not be placed on a timeline.

Events are classified according to semantic features. Those classified as “grammatical” are dependent to a content verb/noun and don’t have a time span, so they will not be annotated. We have also decided to leave out cognitive events (i.e. events that describe mental states or mental acts).

The last criteria is based on the factuality and certainty of events. Counter-factual events will not be part of a timeline because they did not take place. Non-factual events are speculative events, so we don’t know if they happen or not. If it’s certain that they will happen (e.g. “the conference will take place on Monday”), they will be annotated. But if they are uncertain (e.g. “the conference may take place later”), we will not annotate them.

Following are the rules applied to select events which are potentially candidates of a timeline. Into brackets we specify the section in the NewsReader Guidelines where annotators can find more details. For each criteria we provide some examples, in which selected events are in bold.

1. Part-of-speech (section 5.2.2)
  - a. verb (except if it is modified by a modal word)
    - i. Apple has **claimed** that the 3G iPhone will **provide** internet access at double the speed of the internet access **provided** by the previous versions.
  - b. noun
    - i. Apple Inc. today has introduced the much-anticipated iPhone at the Macworld **Conference** in San Francisco.
  - c. pronoun
    - i. Apple Corps claims that Apple Computer's iTunes Music Store violates an agreement reached between the two companies in 1991, **which** barred Apple Computer from using the "Apple" brand.
2. Event Classes (section 4.2)
  - a. SPEECH\_COGNITIVE: only events that describe the action of a person or an organization declaring something, narrating an event, informing about an event
    - i. Apple Inc. **announced** Friday that it will give owners of its new iPhone 4 a free case in response to mounting concerns over the device's antenna placement.
    - ii. Jobs also **said** that the phones can be returned for a refund as well.
  - b. OTHER
    - i. Apple Computer today **launched** a 3G version of its iPhone device.
    - ii. He **plans** a meeting on Friday.
3. Factuality - certainty (section 5.2.3 and 5.2.4)
  - a. FACTUAL events

- i. Apple Corps **claims** that Apple Computer's iTunes Music Store violates an **agreement** reached between the two companies in 1991, which barred Apple Computer from using the "Apple" brand in certain uses in the music business.
- b. NON-FACTUAL + CERTAIN (except if they are part of a conditional construction)
  - i. Apple has claimed that the 3G iPhone will **provide** internet access at double the speed of the internet access provided by the previous versions.
  - ii. The former Chief Operating Officer, Tim Cook will be **succeeding** Jobs as CEO.

Following we define criteria to leave some events out of the timeline (in the examples events not to be annotated are in underlined).

1. POS
  - a. adjective
    - i. The new phone will also come pre-loaded with the new iPhone 3.0 software, which will be available on June 17.
    - ii. The iPod nano is now smaller and squarer.
  - b. prepositional events
    - i. They had 150 passengers on board.
2. Event Classes
  - a. SPEECH\_COGNITIVE: events that describe mental states and mental acts that involve mental or cognitive processes
    - i. He plans to go to Roma.
    - ii. "We love the Beatles, and it has been painful being at odds with them over these trademarks," said Apple Inc. CEO Steve Jobs.
  - b. GRAMMATICAL
    - i. The announcement came in a rare press conference held by Apple on Friday morning.
3. Factuality - certainty
  - a. NON-FACTUAL + UNCERTAIN
    - i. Apple CEO Steve Jobs believes that this *may* have a large impact on the film industry.
  - b. COUNTER-FACTUAL
    - i. Papermaster *did not* comment on the situation.
4. Specific cases:
  - a. conditional constructions
    - i. The iPhone 4 was plagued by highly publicized reports that abnormalities in its new antenna design caused the device to lose its cellular signal *if* held in a certain way.
    - ii. The iPhone 4 antenna is actually a metal strip that wraps around the side of the device, which has caused dropped calls *when* held in a certain way.

- b. verbs modified by modal verbs (note: according to the NewsReader guidelines “will” is considered as an auxiliary and not as a modal verb)
  - i. Any members of the iPhone Developer Program *can add* applications to the store for free.
- c. conjunction of events:
  - i. Developers will be able to **set** the price of the applications or **release** them for free . → “set” and “release” are annotated but not the discontinuous event mention “set (or) release”
- d. causality (section 10.3): ENABLE-type verbs (e.g. enable, aid, allow, permit) and verbs in their complements
  - i. Features new to the iPhone include an upgraded camera, which also *allows* users to record video and sound, as well as " voice control", which will *allow* users to control most features of the iPhone with their voice.

## 2. Time Anchor

Each event is associated to a time anchor and the annotation of time anchors is based on TIMEML.

A time anchor is always a DATE (as defined in TIMEML) and it’s format follows the ISO-8601 standard: YYYY-MM-DD (that is Year, Month, and Day), the maximum granularity admitted being DAY.

As for anchors with a lower granularity, we admit only months and years: references to months are specified as: YYYY-MM, whereas references to years are expressed as: YYYY.

The place-holder character, X, is used for each unfilled position in the value of a component.

Examples:

- February 6, 2007 → 2007-02-06
- April 2010 → 2010-04
- in 2009 → 2009
- May 23 → XXXX-05-23

A time anchor takes as value the point in time when the event occurred (in case of punctual events) or began (in case of durative events).

### Durative events

The anchor time is the beginning of the period.

(1) *He has been **fighting** pancreatic cancer since 2004.*

event: fighting

anchor time: 2004

(2) *I travelled from november 20th to December 15th* (Document creation time: 2007-01-18)

event: travelled

anchor time: 2006-11-20

### **Punctual events**

The anchor time is the DATE when it occurred.

(3) *Jobs, was **born** in San Francisco on February 24, 1955.*

event: born

anchor time: 1955-02-24

(4) *In April, the technology blog Gizmodo **obtained** a prototype of the new phone.* (Document

creation time: 2010-06-08)

event: obtained

anchor time: 2010-04

### **Special cases**

**Present.** If an event is at the present tense, the value of the anchor time should be the document creation time.

(5) *Apple **has** 109 retail stores.* (Document creation time: 2005-06-06)

event: has

anchor time: 2005-06-06

Note: If it's a "general truth" event, it may have a corefered event in another file with a different document creation date. In this case the two events will have different time anchors and won't be considered as coreferent events.

**Narrative present.** If the present tense is employed when narrating past events, the value of the anchor time should be a past date.

(6) *NPD said according to those figures, Apple's iTunes store **passes** U.S. electronics retailer Best Buy for the No. 2 U.S. music retailer in 2007.*

event: passes

anchor time: 2007

**Direct speech.** Events occurring in direct speech are anchored taking into consideration the time of utterance.

(7) *On May 23rd she said "Today I am moving".* (Document creation time: 2005-06-06)

event: moving

anchor time: 2005-05-23

**Undefined value for anchor time.** If it's not possible to anchor a past event, future event or an event at the narrative present tense, the time anchor value should be XXXX-XX-XX.

(8) *Cook **held** positions at IBM and Compaq.*

anchor time: XXXX-XX-XX

### 3. Event Ordering

Event ordering is based on event time anchors and on temporal relations between events (more specifically on the before/after and includes/simultaneous relations as defined by ISO-TimeML).

#### **Ordering of events associated to time anchors with the same granularity.**

Ordering of events associated to time anchors with the same granularity is based on the value of the time anchors.

“In April 2010 he moved to Italy and three months later he bought a house”

ordering	time_anchor	event(s)
1	2010-04	moved
2	2010-07	bought

When it is not possible to order events based on the time anchor, ordering is based on textual information.

“In April 2010 he graduated and moved to Italy”

ordering	time_anchor	event(s)
1	2010-04	graduated
2	2010-04	moved

When it is not possible to order events based either on the time anchor or on textual information, they will be considered **simultaneous**.

“He graduated in April 2010.”

‘He left Italy last month.’ (DCT: 2010-05-20)

ordering	time_anchor	event(s)
1	2010-04	graduated
1	2010-04	left

#### **Ordering of events associated to time anchors with different granularity.**

Ordering of events associated to time anchors with different granularity is based on the value of the time anchors.

“He met her in April 2010 and married her on April 6, 2012.”

“In 2011, John went to the USA.”

ordering	time_anchor	event(s)
1	2010-04	bought
2	2011	went
3	2012-04-06	gave

When it is not possible to order events based on the time anchor, ordering is based on textual information.

“He met her in April 2010 and married her before the end of the year”

ordering	time_anchor	event(s)
1	2010-04	met
2	2010	married

When it is not possible to order events based either on the time anchor or on textual information, **precedence** should be given **to events with lower granularity**.

“John bought a new car in 2010.”

“He met Mary in April 2010.”

“He graduated on April 6, 2010.”

ordering	time_anchor	event(s)
1	2010	bought
2	2010-04	met
3	2010-04-06	graduated

### **Ordering of events associated to time anchors with the place-holder character X**

They should be ordered based on the value of the anchor time and on information available in the text as precisely as possible.

“John bought a house in this region in 2009.”

“John got married on April 6, 2011.”

“In 2010, around her birthday Kathy went<sup>1</sup> to the USA. The 13th of the same month John left Kathy’s apartment.”

ordering	time_anchor	event(s)
1	2009	bought
2	2010-XX-13	left
3	2011-04-06	married

---

<sup>1</sup> The event *went* is not in the timeline about John because it does not involve the target entity John as defined in the Timeline task guidelines.

“John bought a house in this region in 2009.”

“After he met Mary, he left the region.”

“John get married on April 6, 2011.”

ordering	time_anchor	event(s)
1	2009	bought
2	XXXX-XX-XX	met
3	XXXX-XX-XX	left
4	2011-04-06	married

When in doubt where to position an event with anchor containing a place holder, annotators give precedence to this events wrt events with more specific anchors. So, if no information is available, they should be put in the first place of the timeline, and their ordering number will be 0.

“John bought a house in this region in 2009.”

“He met Mary on February 2, 2010 and one year after, on April 6, he married her.”

“In 2010, around her birthday Kathy went to the USA. The 13th of the same month John left Kathy’s apartment.”

“John travelled around Europe.”

ordering	time_anchor	event(s)
0	XXXX-XX-XX	travelled
1	2009	bought
2	2010-XX-13	left
3	2010-02-02	met
4	2011-04-06	married

## 4. TimeLine

### Target Entities

One TimeLine is associated to one target entity. The entity can be of type *organization*, *person* or *product*.

The TimeLine would contain events in which the target entity **explicitly** participates with the semantic role ARG0 (i.e. agent) or ARG1 (i.e. patient). In the sentence (1) *iPhone 4* is ARG0 of the verb *use*, and in sentence (2) it is ARG1 of the verb *unveil*.

(1) **The iPhone 4** will use iOS.

(2) Yesterday, Steve Jobs unveiled **iPhone 4**.

Entity coreference must be resolved. A TimeLine should contain events involving besides the target entity its coreferences (including pronominal coreferences).



The **is\_part\_of** relations are not considered as coreferences. For example “The 16GB version” in the sentence (3) is not a coreference of iPhone4.

(3) **The 16GB version is priced at US\$ 199 and the 32GB version at US\$ 299.**

In sentence (4) “the parties” refers to the two companies “Apple Inc.” and “Apple Corps”, but “the parties” doesn’t corefer with neither “Apple Inc.” or “Apple Corps”.

(4) *On September 21, 2004 **the parties** agreed to have the case heard by the UK court.*

### Event coreference

Event coreference must be resolved. If two event mentions corefer, only one event must appear in the TimeLine.

The sentence (3) and (4) contain two event mentions which corefer: “introduced” and “introducing”.

(3) *The newest iPhone, [iPhone 4] was **introduced** by [Apple CEO Steve Jobs] at the company's 2010 Worldwide Developer's Conference less than two weeks ago.*

(4) *While **introducing** [iPhone 4], at the annual conference, [Jobs] [...]*

### TimeLine format

A TimeLine is represented in a simple tab format:

ordering      time\_anchor    event(s)

The first column (*ordering*) contains a cardinal number which indicates the position of the event in the TimeLine (two events can be associated to the same number if they are simultaneous).

The second column (*time\_anchor*) contains a time anchor. The third column (*event*) consists of one event or a list of corefered events separated by a tab. Each event is represented by the id of the file (<DOCID>), the id of the sentence and **the extent of the event mention** (i.e. tokens that compose the event mention) in the following format: 11778-2-launch (docid-sentid-event)

In the case of multi-words event, tokens are separated by an underscore:

16844-12-showed\_off

One file by TimeLine must be created. The first line contains the target entity.

iTunes

1	2003	11778-3-launch	11778-4-launch
2	2007	11778-4-pass	
3	2008-01	11778-7-hold	
4	2008-02	11778-2-pass	11778-5-pass
4	2008-02	11778-3-accounts_for	