

Two Layers of Annotation for Representing Event Mentions in News Stories

Maria Pia di Buono¹ Martin Tutek¹ Jan Šnajder¹ Goran Glavaš²
Bojana Dalbelo Bašić¹ Nataša Milić-Frayling³

¹ TakeLab, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
`firstname.lastname@fer.hr`

² Data and Web Science Group, University of Mannheim, Germany
`goran@informatik.uni-mannheim.de`

³ School of Computer Science, University of Nottingham, United Kingdom
`natasa.milic-frayling@nottingham.ac.uk`

Abstract

In this paper, we describe our preliminary study of methods for annotating event mentions as part of our research on high-precision models for event extraction from news. We propose a two-layer annotation scheme, designed to capture the functional and the conceptual aspects of event mentions separately. We hypothesize that the precision can be improved by modeling and extracting the different aspects of news events separately, and then combining the extracted information by leveraging the complementarities of the models. We carry out a preliminary annotation using the proposed scheme and analyze the annotation quality in terms of inter-annotator agreement.

1 Introduction

The task of representing events in news stories and the way in which they are formalized, namely their linguistic expressions (event mentions), is interesting from both a theoretical and practical perspective. Event mentions can be analyzed from various aspects; two aspects that emerge as particularly interesting are the linguistic aspect and the more practical information extraction (IE) aspect.

As far as the linguistic aspect is concerned, news reporting is characterized by specific mechanisms and requires a specific descriptive structure. Generally speaking, such mechanisms convey non-linear temporal information that complies with news values rather than narrative norms (Setzer and Gaizauskas, 2000). In fact, unlike traditional story telling, news writing follows the “in-

verted pyramid” mechanism that consists of introducing the main information at the beginning of an article and pushing other elements to the margin, as shown in Figure 1 (Ingram and Henshall, 2008). Besides, news texts use a mechanism of gradual specification of event-related information, entailing a widespread use of coreference relations among the textual elements.

On the other hand, the IE aspect is concerned with the information that can be automatically acquired from news story texts, to allow for more efficient processing, retrieval, and analysis of massive news data nowadays available in digital form.

In this paper, we describe our preliminary study on annotating event mention representations in news stories. Our work rests on two main assumptions. The first assumption is that event in news substantially differ from events in other texts, which warrants the use of a specific annotation scheme for news events. The second assumption is that, because news events can be analyzed from different aspects, it makes sense also to use different annotation layers for the different aspects. To this end, in this paper we propose a two-layer annotation scheme, designed to capture the functional and the conceptual aspects of event mentions separately. In addition, we carry out a preliminary annotation using the proposed scheme and analyze the annotation quality in terms of inter-annotator agreement.

The study presented in this paper is part of our research on high-precision models for event extraction from news. We hypothesize that the precision can be improved by modeling and extracting the different aspects of news events separately, and then combining the extracted information by leveraging the complementarities of the models. As a

Narrative	News
<p>When electricians wired the home of Mrs Mary Ume in Hohola, Port Moresby, some years ago they neglected to install sufficient insulation at a point in the laundry where a number of wires crossed.</p> <p>A short-circuit occurred early this morning.</p> <p>Contact between the wires is thought to have created a spark, which ignited the walls of the house.</p> <p>The flames quickly spread through the entire house.</p> <p>Mrs Ume, her daughter Peni (aged ten) and her son Jonah (aged five months) were asleep in a rear bedroom. They had no way of escape and all perished.</p>	<p>A Port Moresby woman and her two children died in a house fire in Hohola today.</p> <p>Mrs Mary Ume, her ten-year-old daughter Peni and baby son Jonah were trapped in a rear bedroom as flames swept through the house.</p> <p>The fire started in the laundry, where it is believed faulty electrical wiring caused a short-circuit. The family were asleep at the time.</p> <p>The flames quickly spread and soon the entire house was blazing.</p>

Table 1: An example of narrative and news styles (Ingram and Henshall, 2008).

first step towards that goal, in this paper we carry out a preliminary comparative analysis of the proposed annotation layers.

The rest of the paper is structured as follows. In the next section we briefly describe the related work on representing and annotating events. In Section 3 we present the annotation methodology. In Section 4 we describe the annotation task, while in Section 5 we discuss the results. In Section 6 we describe the comparative analysis. Section 7 concludes the paper.

2 Related Work

Several definitions of events have been proposed in the literature, including that from the Topic Detection and Tracking (TDT) community: “a TDT event is defined as a particular thing that happens at a specific time and place, along with all necessary preconditions and unavoidable consequences” (TDT, 2004). On the other hand, the ISO TimeML Working Group (Pustejovsky et al., 2003) defines an event as “something that can be said to obtain or hold true, to happen or to occur.”

On the basis of such definitions, different approaches have been developed to represent and extract events and those aspects considered representative of event factuality.

In recent years, several communities proposed different shared tasks aiming at evaluating event annotation systems, mainly devoted to recognize event factuality or specific aspects related to factuality representation (e.g., temporal annotation), or tasks devoted to annotate events in specific language, e.g., Event Factuality Annotation Task presented at EVALITA 2016, the first evaluation exercise for factuality profiling of events in Italian (Minard et al., 2016b).

Among the communities working in this field,

the TimeML community provides a rich specification language for event and temporal expressions aiming to capture different phenomena in event descriptions, namely “aspectual predication, modal subordination, and an initial treatment of lexical and constructional causation in text” (Pustejovsky et al., 2003).

Besides the work at these shared tasks, several authors proposed different schemes for event annotation, considering both the linguistic level and the conceptual one. The NewsReader Project (Vossen et al., 2016; Rospocher et al., 2016; Agerri and Rigau, 2016) is an initiative focused on extracting information about what happened to whom, when, and where, processing a large volume of financial and economic data. Within this project, in addition to description schemes (e.g., ECB+. (Cybulska and Vossen, 2014a)) and multilingual semantically annotated corpus of Wikinews articles (Minard et al., 2016a), van Son et al. (2016) propose a framework for annotating perspectives in texts using four different layers, i.e., events, attribution, factuality, and opinion. In the NewsReader Project the annotation is based on the guidelines to detect and annotate markables and relations among markables (Speranza and Minard, 2014). In the detection and annotation of markables, the authors distinguish among entities and entity mention in order to “handle both the annotation of single mentions and of the coreference chains that link several mentions to the same entity in a text” (Lösch and Nikitina, 2009). Entities and entity mention are then connected by the REFER TO link.

Another strand of research are the conceptual schemes, rooted in formal ontologies. Several upper ontologies for annotating events have been developed, e.g., the EVENT Model F (Scherp et al.,

2009). This ontology represents events solving two *competency questions*¹ about the participants in the events and the previous events that caused the event in question. EVENT Model F is based on the foundational ontology DOLCE+DnS Ultra-lite (DUL) (Gangemi et al., 2002) and it focuses on the participants involved in the event and on mereological, causal, and correlative relationships between events.

Most of the proposed ontologies are tailored for financial or economic domains. A case in point is The newsEvent Ontology, a conceptual scheme for describing events in business events (Lösch and Nikitina, 2009).

3 Methodology

Our methodology arises from the idea that events in news call for a representation that is different from event representations in other texts. We believe that a coherent and consistent description and, subsequently, extraction of event mentions in news stories should be dealt with conveying temporal information (When), but also distinguishing other information related to the action (What), the participants (Who), the location (Where), the motivation (Why) and the manner in which the event happened (How). This means that a meaningful news/event description should cover the proverbial 5Ws and one H, regarded basic in information gathering, providing a factual answer for all these aspects.

The above assumption implies that events cannot be considered black boxes or monolithic blocks describable merely by means of the temporal chain description. Instead, it is necessary to capture the *functional* and *conceptual* aspects of event mentions. Indeed, as previously claimed, language used in news stories is characterized by mechanisms that differ from the narrative one. Such differences may manifest themselves in both the syntactic structures and the patterns of discursive features that effect the sentence structure.

In line with the above, our approach aims at accomplishing a fine-grained description of event mentions in news stories applying a two-layer annotation scheme. The first layer conveys the different syntactic structures of sentences, accounting the functional aspects and the components in

¹Competency questions refer to natural language sentences that express the scope and content of an ontology. Answering these question represents a functional requirement for an ontology (Uschold and Gruninger, 1996).

events on the basis of their role. As noted by Papafragou (2015), “information about individual event components (e.g., the person being affected by an action) or relationships between event components that determine whether an event is coherent can be extracted rapidly by human viewers”. On the other hand, the second layer is suitable also to recognize the general topic or theme that underlies a news story, due to the fact that this layer concerns conceptual aspects. This theme can be described as a “semantic macro-proposition”, namely a proposition composed by the sequences of propositions retrievable in the text (Van Dijk, 1991). Thus, the conceptual scheme makes it possible to recognize these structures reducing the complexity of the information and guaranteeing a summarization process that is closer to users’ representation.

3.1 Functional Layer

Following the previously-mentioned broad definition of an event in news as something that happens or occurs, in the functional annotation layer we focus on the lower level representation of events, closer to the linguistic level.

We represent each event with an event action and a variable number of arguments of different (sub)categories. The event action is most commonly the verb associated with the event (e.g., “destroyed”, “awarded”), however it can also be other parts of speech (e.g., “explosion”) or a multiword expression (e.g., “give up”). The action defines the focus of the event and answers the “What happened” question, and is the main part of an event mention.

Along with the event action, we define four main categories of event arguments to be annotated, which are then split into fine-grained subcategories, as shown in Table 2. We subcategorize the standard Participant category into the AGENT, PATIENT, and OTHERPARTICIPANT subcategories. We further divide each of the aforementioned subcategory into HUMAN and NON-HUMAN subcategories. The AGENT subcategory pertains to the entities that perform an action either deliberately (usually in the case of human agents) or mindlessly (natural causes, such as earthquakes or hurricanes). The PATIENT is the entity that undergoes the action and, as a result of the action, changes its state. The TIME and LOCATION categories serve to further specify the event.

Category	Subcategory	
PARTICIPANT	AGENT	HUMAN NONHUMAN
	PATIENT	HUMAN NONHUMAN
	OTHERPARTICIPANT	HUMAN NONHUMAN
LOCATION	GEOLOCATION	
	OTHER	
TIME		
OTHERARGUMENT		

Table 2: Functional categories and subcategories.

Finally, the OTHERARGUMENT category covers themes and instruments (Baker, 1997; Jackendoff, 1985) of an action. Table 3 gives an example of sentence “Barcelona defeated Real Madrid yesterday at Camp Nou” annotated using the functional layer.

The action’s arguments focus on the specifics of the event that occurred. We depart from the standard arguments that can be found in schemes like ECB+ (Cybulska and Vossen, 2014a) or TimeML (Pustejovsky et al., 2003) in that we included the Other_argument category. Furthermore, in TimeML, predicates related to states or circumstances are considered as events, while in the scope of this work, sentences describing a state, e.g., “They live in Maine”, are not annotated. In fact, we argue that they do not represent the focus in news, but merely describe the situation surrounding the event.

Our functional annotation differs from PropBank (Palmer et al., 2005) definitions of semantic roles as we do not delineate our functional roles through a verb-by-verb analysis. More concretely, PropBank adds predicate-argument relations to the syntactic trees of the Penn Treebank, representing these relations as *framesets*, which describe the different sets of roles required for the different meanings of the verb. In contrast, our analysis aims to describe the focus of an event mention by means of identifying *actions*, which can involve also other lexical elements in addition to the verb. This is easily demonstrated through the example “fire broke out” from Figure 2a, where we annotate “fire broke out” as an action, since it fully specifies the nature of the event defining in a less general way the action.

Text span	Label
Barcelona	AGENT
defeated	ACTION
Real Madrid	PATIENT
yesterday	TIME
at Camp Nou	GEOLOCATION

Table 3: Sample of functional event annotation.

Entity	Property	
PERSON	IDENTITY	MOVEMENT
ORGANIZATION	ASSOCIATION	LOCATION
ANIMAL	PARTICIPATION	CAUSE
OBJECT	OCCURRENCE	PERFORMANCE
PLACE	ATTRIBUTION	INFLUENCE
TIME	CONSTRUCTION	SUPPORT
MANIFESTATION	CREATION	PURPOSE
	MODIFICATION	CLASSIFICATION
	DECLARATION	DEATH

Table 4: Main entity classes and properties.

3.2 Conceptual Layer

In order to represent semantically meaningful event mentions and, consequently, to develop an ontology of the considered domain, we define also a second layer of annotation, namely a conceptual model for news stories. This model, putting forward a classification of the main concepts retrievable in news stories, defines seven entity classes, six entity subclasses, and eighteen properties (Table 4).

Entities and properties. Entity classes are defined in order to represent a set of different individuals, sharing common characteristics. Thus, being representative of concepts in the domain, entities may be identified by noun phrases. On the other hand, properties describe the relations that link entity classes to each other and can be represented by the verb phrase. For this reason, each property is associated with some association rules that specify the constraints related to both its syntactic behaviors and the pertinence and the intension of the property itself. In other words, these association rules contribute to the description of the way in which entity classes can be combined through properties in sentence contexts. To formalize such rules in the form of a set of axioms, we take in consideration the possibility of combining semantic and lexical behaviors, suitable for identifying specific event patterns. Thus, for in-

stance, the property MOVEMENT may connect the entity class PERSON and the entity classes PLACE and TIME, but the same property cannot be used to describe the relation between MANIFESTATION and PLACE. The definition of these rules, and the corresponding axioms, relies on word combination principles that may occur in a language, derived from an analysis of work of Harris (1988), and conceptual considerations related to the domain.

Factuality. To represent the factuality in event descriptions, we specify three attributes for each property: polarity, speculation, and passive markers. The polarity refers to the presence of an explicit negation of the verb phrase or the property itself. The speculation attribute for the property identifies something that is characterized by speculation or uncertainty. Such an attribute is associated with the presence of some verbal aspects (e.g., the passive applied to specific verbs as in they were thought to be), some specific constructions/verbs (e.g., to suggest, to suppose, to hypothesize, to propose) or modality verbs. According to Hodge and Kress (1988), the “modality refers to the status, authority and reliability of a message, to its ontological status, or to its value as truth or fact”. Finally, we use an attribute for a passive marker due to the fact that passive voice is used mainly to indicate a process and can be applied to infer factual information. Note that, although the time marker is typically considered to be indicative of factuality, we prefer to avoid annotating time markers in our schema. Thus, we infer the temporal chain in event mentions by means of both temporal references in the sentence, e.g., the presence of adverbs of time, and the syntactic tense of the verb.

Coreference. To account for the coreference phenomenon among entities, we introduce a symmetric-transitive relation taking two entity classes as arguments. This allows for annotation of two types of coreference, identity and apposition, and can be used at inter-sentence level to annotate single or multiple mentions of the same entity; an example is shown in Table 5.

Complex events. In the description of event mentions in news stories we often encounter sentence structures expressing complex events, i.e., events characterized by the presence of more than one binary relation among their elements. Due to

Text span	Label
Blasts	MANMADEEVENT*
which	MANMADEEVENT*
killed	DEATH
38	COLLECTIVE
by stadium	PLACE
claimed by	DECLARATION (passive)
TAK	ORGANIZATION

Table 5: Sample of coreference and attribute annotation (* denotes coreferring elements).

the fact that properties generally express binary relationships between two entity classes, we introduce N-ary relations, namely reified relations, in order to describe these complex structures. The reified relations allow for the description of complex events composed by more than two entities and one property. According to the recommendation of the W3C,² these additional elements, which contribute to constitute complex events, can be formalized as a value of the property or as other arguments (entity classes) occurring in the sentence. In our scheme, we decide to deal with some of these reified relations creating three additional entity classes – MANNER, SCOPE, and INSTRUMENT – which may hold heterogeneous elements. Nevertheless, these elements present a shared intensive property defined by the main property they refer to.

4 Annotation Task

To calibrate the two annotation schemes, we performed two rounds of annotation on a set of news stories in English. We hired four annotators to work on each layer separately, to avoid interference between the layers.

We set up the annotation task as follows. First, we collected a corpus of news documents. Secondly, we gave each of four annotators per schema the same document set to annotate, along with the guidelines for that schema. We then examined the inter-annotator agreement for the documents, and discussed the major disagreements in person with the annotators. After the discussion, we revised the guidelines and again gave the annotators the

²We refer to W3C protocols for representing these structures to warrant the compliance of our schema with Semantic Web languages. More information can be found here: <https://www.w3.org/TR/swbp-n-aryRelations/>.

same set of documents. For the annotation tool, we used Brat (Stenetorp et al., 2012).

We collected the documents by compiling a list of recent events, then querying the web to find news articles about those events from various sources. We collected the articles from various sources to be invariant of the writing style of specific news sites. We aimed for approximately the same length of articles to keep the scale of agreement errors comparable. For this annotator calibration step, we used a set of five news documents, approximately 20 sentences in length each.

We computed the inter-annotator agreement between the documents on sentence level in order to determine the sources of annotator disagreement. We then organized discussion meetings with all of the annotators for each schema to determine whether the disagreement stems from the ambiguity of the source text or from the incomprehensiveness of the annotation schema.

After the meetings, we revised and refined the guidelines in a process which mostly included smaller changes such as adding explanatory samples of annotation for borderline cases as well as rephrasing and clarifying the text of the guidelines. However, we also made a couple of more substantial revisions such as adding label classes and determining what should or should not be included in the text spans for particular labels.

5 Inter-Annotator Agreement

We use two different metrics for calculating the inter-annotator agreement (IAA), namely Cohen’s kappa coefficient (Cohen, 1960) and the F1-score (van Rijsbergen, 1979). The former has been used in prior work on event annotations, e.g., in (Cybulska and Vossen, 2014b). On the other hand, F1-score is routinely used for evaluating annotations that involve variable-length text spans, e.g., named entity annotations (Tjong Kim Sang and De Meulder, 2003) used in named entity recognition (NER) tasks. In line with NER evaluations, we consider two F1-score calculations: strict F1-score (both the labels and the text spans have to match perfectly) and lenient F1-score (labels have to match, but text spans may only partially overlap). In both cases, we calculate the macro F1-score by averaging the F1-scores computed for each label.

The motivation for using the F1-score along with Cohen’s kappa coefficient lies in the fact that Cohen’s kappa treats the untagged tokens as true

	Kappa	F1-strict	F1-lenient
Functional layer			
Round 1	0.428 ± 0.08	0.383 ± 0.04	0.671 ± 0.10
Round 2	0.409 ± 0.07	0.424 ± 0.04	0.621 ± 0.07
Conceptual layer			
Round 1	0.280 ± 0.08	0.350 ± 0.06	0.680 ± 0.15
Round 2	0.476 ± 0.07	0.475 ± 0.03	0.778 ± 0.06

Table 6: Inter-annotator agreement scores for the two annotation layers and two annotation rounds, averaged across annotator pairs and documents.

negatives. If the majority of tokens is untagged, the agreement values will be inflated, as demonstrated by Cybulska and Vossen (2014b). In contrast, the F1-score disregards the untagged tokens, and is therefore a more suitable measure for sequence labeling tasks. In our case, the ratio of untagged vs. tagged tokens was less skewed (6:4 and 1:2 for the functional and conceptual layer, respectively), i.e., for both annotation layers a fair portion of text is covered by annotated text spans, which means that the discrepancy between kappa values and F1-scores is expected to be lower.

We compute the IAA across all annotator pairs working on the same document, separately for the same round of annotation, and separately for each annotation layer. We then calculate the IAA averaged across the five documents, along with standard deviations. Table 6 shows the IAA scores.

For the functional layer, the Cohen’s kappa coefficient is above 0.4, which, according to Landis and Koch (1977), is considered a borderline between fail and moderate agreement. Interestingly enough, the kappa agreement dropped between the first and the second round. We attribute this to the fact that the set of labels was refined (extended) between the two rounds, based on the discussion we had with the annotators after the first round of annotations. Apparently, the refinement made the annotation task more difficult, or we failed to cater for it in the guidelines. Conversely, for the conceptual layer, the agreement in first round was lower, but increased to a moderate level in the second round. The same observations hold for the F1-strict and F1-lenient measures. Furthermore, the IAA scores for the second round for the conceptual layer are higher than for the functional layer. A number of factors could be at play here: the annotators working on the conceptual layer were perhaps more skilled, the guidelines were more com-

prehensive, or the task is inherently less difficult or perhaps more intuitive.

While the IAA scores may seem moderate at first, one has to bear in mind the total number of different labels, which is 17 and 28 for the functional and conceptual layer, respectively. In view of this, and considering also the fact that this is a preliminary study, we consider the moderate agreement scores to be very satisfactory. Nonetheless, we believe the scores could be improved even further with additional calibration rounds.

6 Comparative analysis

In this section, we provide examples of a couple of sentences annotated in both layers, along with a brief discussion on why we believe that each layer compensates the shortcomings of the other.

Fig. 1 provides an example of a sentence annotated in the functional and conceptual layer. We observe that the last part of the sentence, “reading: Bye all!!!”, is not annotated in the functional layer (Fig. 1a). This is due to the fact that the last part is a modifier of the patient, and not the action. Even though we could argue that in this case the information provided by the modifier is unimportant for the event, we could conceive of a content of the note that would indeed be important. Along with that, any modifier of the event arguments that is not directly linked to the arguments is not annotated in the functional layer, leading to information loss. We argue that in such cases the conceptual layer (Fig. 1b) is more suited towards gathering the full picture of the event along with all the descriptions.

Fig. 2a exemplifies the case where, in the functional layer, the action is a noun phrase. Such cases are intentionally meant to be labeled as actions as they change the meaning of the verb itself. In the conceptual case (Fig. 2b), as the occurrence we label “broke out”, a phrase that, although clear, gives no indication of the true nature of the event, and the conceptual layer relies on the “natural event” argument for the full understanding of the event. We argue that having a noun phrase as an action, such as in the functional layer, is a more natural representation of an event as it fully answers the “What” question. We also argue that making a distinction between “fire broke out” and “broke out” as actions is beneficial for the training of the event extraction model as it emphasizes the distinction between a verb and an action.

7 Conclusions

We have presented a two-layered scheme for the annotation of event mentions in news, conveying different information aspects: the functional aspect and the conceptual aspect. The first one deals with a more general analysis of sentence structures in news and the lexical elements involved in events. The conceptual layer aims at describing event mentions in news focusing on the “semantic macro-propositions”, which compose the theme of the news story.

Our approach to event mentions in news is a part of a research project on high-precision news event extraction models. The main hypothesis, leading the development of our system, is that the precision of models can be improved by modeling and extracting separately the different aspects of news events, and then combining the extracted information by leveraging the complementarities of the models. As part of this examination, we have presented also a preliminary analysis of the inter-annotator agreement.

Acknowledgments

This work has been funded by the Unity Through Knowledge Fund of the Croatian Science Foundation, under the grant 19/15: “EVENT Retrieval Based on semantically Enriched Structures for Interactive user Tasks (EVERBEST)”.

References

- Rodrigo Agerri and German Rigau. 2016. Robust multilingual named entity recognition with shallow semi-supervised features. *Artificial Intelligence*, 238:63–82.
- Mark C. Baker. 1997. Thematic roles and syntactic structure. In *Elements of grammar*, pages 73–137. Springer.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.
- Agata Cybulska and Piek Vossen. 2014a. Guidelines for ECB+ annotation of events and their coreference. Technical report, Technical report, Technical Report NWR-2014-1, VU University Amsterdam.
- Agata Cybulska and Piek Vossen. 2014b. Using a sledgehammer to crack a nut? lexical diversity and event coreference resolution. In *LREC*, pages 4545–4552.

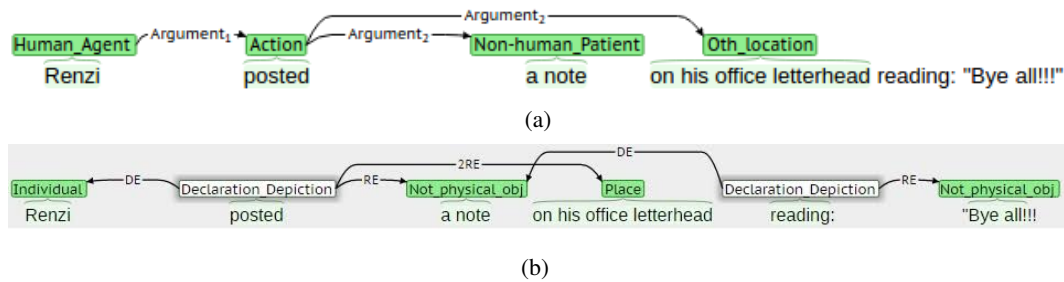


Figure 1: Example of sentence annotated in the (a) functional and the (b) conceptual layer.

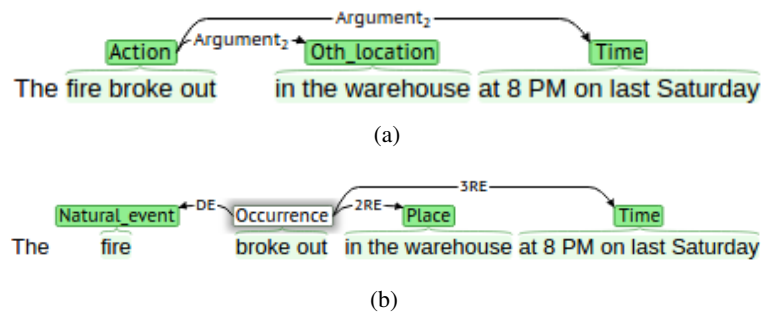


Figure 2: Example of sentence annotated in the (a) functional and the (b) conceptual layer.

Aldo Gangemi, Nicola Guarino, Claudio Masolo, Alessandro Oltramari, and Luc Schneider. 2002. Sweetening ontologies with dolce. In *International Conference on Knowledge Engineering and Knowledge Management*, pages 166–181. Springer.

Zellig Harris. 1988. *Language and information*. Columbia University Press.

Robert Hodge and Gunther R. Kress. 1988. *Social semiotics*. Cornell University Press.

D. Ingram and P. Henshall. 2008. *The news manual*.

Ray Jackendoff. 1985. Semantic structure and conceptual structure. *Semantics and Cognition*, pages 3–22.

J. Richard Landis and Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174.

Uta Lösch and Nadejda Nikitina. 2009. The new-sevents ontology: an ontology for describing business events. In *Proceedings of the 2009 International Conference on Ontology Patterns-Volume 516*, pages 187–193. CEUR-WS. org.

A-L Minard, Manuela Speranza, Ruben Urizar, Begona Altuna, MGJ van Erp, AM Schoen, CM van Son, et al. 2016a. MEANTIME, the NewsReader multilingual event and time corpus.

Anne-Lyse Minard, Manuela Speranza, Tommaso Caselli, and Fondazione Bruno Kessler. 2016b. The EVALITA 2016 event factuality annotation task

(FactA). In *Proceedings of Third Italian Conference on Computational Linguistics (CLiC-it 2016) & Fifth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2016)*.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.

Anna Papafragou. 2015. The representation of events in language and cognition. *The conceptual mind: New directions in the study of concepts*, page 327.

James Pustejovsky, José M Castano, Robert Ingria, Roser Sauri, Robert J Gaizauskas, Andrea Setzer, Graham Katz, and Dragomir R. Radev. 2003. TimeML: Robust specification of event and temporal expressions in text. *New directions in question answering*, 3:28–34.

Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard. 2016. Building event-centric knowledge graphs from news. *Web Semantics: Science, Services and Agents on the World Wide Web*, 37:132–151.

Ansgar Scherp, Thomas Franz, Carsten Saathoff, and Steffen Staab. 2009. F – a model of events based on the foundational ontology DOLCE+DnS ultra-lite. In *Proceedings of the fifth international conference on Knowledge capture*, pages 137–144. ACM.

- Andrea Setzer and Robert J. Gaizauskas. 2000. Annotating events and temporal information in newswire texts. In *LREC*, volume 2000, pages 1287–1294.
- Manuela Speranza and Anne-Lyse Minard. 2014. NewsReader guidelines for cross-document annotation NWR-2014-9.
- Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. BRAT: a web-based tool for NLP-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 102–107. Association for Computational Linguistics.
- Tdt TDT. 2004. Annotation manual version 1.2. *From knowledge accumulation to accommodation: cycles of collective cognition in work groups*.
- Erik F Tjong Kim Sang and Fien De Meulder. 2003. Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*, pages 142–147. Association for Computational Linguistics.
- Mike Uschold and Michael Gruninger. 1996. Ontologies: Principles, methods and applications. *The knowledge engineering review*, 11(02):93–136.
- Teun A. Van Dijk. 1991. Media contents: the interdisciplinary study of news as discourse, a handbook of qualitative methodologies for mass communication research.
- C.J. van Rijsbergen. 1979. *Information Retrieval*. Butterworths.
- Chantal van Son, Tommaso Caselli, Antske Fokkens, Isa Maks, Roser Morante, Lora Aroyo, and Piek Vossen. 2016. Grasp: A multilayered annotation scheme for perspectives. In *Proceedings of the 10th Edition of the Language Resources and Evaluation Conference (LREC)*.
- Piek Vossen, Rodrigo Agerri, Itziar Aldabe, Agata Cybulska, Marieke van Erp, Antske Fokkens, Egoitz Laparra, Anne-Lyse Minard, Alessio Palmero Aprosio, German Rigau, et al. 2016. NewsReader: Using knowledge resources in a cross-lingual reading machine to generate more knowledge from massive streams of news. *Knowledge-Based Systems*, 110:60–85.