

Event Detection between Literary Studies and NLP

A Survey, a Narratological Reflection, and a Case Study

Noa Visser Solissa¹ 
Andreas van Cranenburgh¹ 
Federico Pianzola¹ 

1. Center for Language and Cognition, University of Groningen , Groningen, The Netherlands.

Citation

Noa Visser Solissa, Andreas van Cranenburgh, and Federico Pianzola (2025). "Event Detection between Literary Studies and NLP. A Survey, a Narratological Reflection, and a Case Study". In: *Journal of Computational Literary Studies* 4 (1). [10.48694/jcls.4215](https://doi.org/10.48694/jcls.4215)

Date published 2025-12-01

Date accepted 2025-11-05

Date received 2025-02-07

Keywords

events, event detection, narratology, historical texts, news, literature

License

CC BY 4.0 

Reviewers

Andrew Piper, Haimo Stiemer

Note

This paper has passed through the conference track of JCLS. In addition to being peer reviewed, it was presented and discussed at the 4th Annual Conference of Computational Literary Studies at Krakow, Poland, in July 2025.

Abstract. Narrative structure in fiction relies on the strategic presentation of events, where the ordering and disclosure of information (*syuzhet*) shape reader engagement and tension. This study outlines a computational model for detecting *syuzhet* by formalizing event annotation in fictional texts across multiple languages. While automated event detection has been widely applied in domains like journalism and history, its theoretical foundations remain fragmented due to divergent definitions of "event" and domain-specific priorities. We critically synthesize prior approaches, highlighting their methodological and applicative distinctions, and position our model within this landscape. Additionally, we demonstrate its adaptability by comparing it to a storyline analysis framework designed for news, revealing cross-domain utility. Our work offers a flexible computational narratology framework for analyzing narrative progression in both fiction and non-fiction contexts.

1. Introduction

A reader does not only read a story to get to know what happens in a text, but also because of the manner in which this narrative is presented to them (Scheffel 2013). Information gaps created by the ordering and disordering of events according to logical and temporal links are what creates narrative tension and makes stories engaging (Baroni 2007; Sternberg 1992). Narrative organization is particularly important in fictional texts, as they are more likely to portray higher levels of non-linearity in comparison to non-fiction (Piper and Toubia 2023; van Cranenburgh et al. 2019). Our objective is to develop a computational model that detects the events in a fictional text the same way a reader learns about them (Genette 1980; Scheffel 2013). An intermediate goal is to create a theoretical model for the annotation of events in fiction in several languages, taking into account the many challenges posed by literary language and narrative strategies.

Automated event detection has been a task of interest in Natural Language Processing (NLP), linguistics, journalism, history, and literature (Caselli and Bos 2023; Norambuena et al. 2023; Santana et al. 2023; Sprugnoli and Tonelli 2017). However, despite this broad interest in automated event detection, the definitions of "event" differ greatly across scholarly works due to the different objectives for the task in the various fields and between different research projects.

In this article, we give an overview of related research and the manner in which these works are related to our theoretical model for the annotation of events in fiction. First, we elaborate on the definition of *syuzhet* – the concrete order in which events are *presented* (Scheffel 2013) – and provide a theoretical background on the different definitions of events in narratology, followed by an overview of related works in literary event detection (see section 3). Then we give an overview of research on automated event detection in news and historical texts (see section 4), concluding with a comparison of the different operationalizations of events used in related work (see Table 1). In section 5, we introduce our operationalization of literary events, as well as a comparison of our framework to a narratology-based framework developed for news (Vossen et al. 2021), to demonstrate how our theoretical model for fiction differs from frameworks in news (see section 6). In this review, we do not aim to be exhaustive. Instead, we have selected approaches from several domains to explore the comparability of NLP methods across different fields.

2. Literary Events

Our goal is a definition of narrative event that can be broadly operationalized (Pichler and Reiter 2022) for the automatic detection of events in literary texts. Thus, we aim at creating a domain-specific framework that contributes to bridging the gap between NLP research and its techniques to analyze events on the one hand, and our domain, computational literary studies, on the other. Additionally, it would be ideal to define narrative events in a way that is operationalizable across different languages. Many scholars in literary studies and narratology have addressed the concept of event, trying to define its constitutive properties and the role of events in stories. The main difference from NLP research is probably the conceptualization of different event categories (see subsection 2.1) and event sequences (see subsection 2.2).

2.1 Event Categories

Events can be considered the smallest units that make up a narrative. An event can also be seen as a change of state, i.e., any type of expressed change that contributes to the narration (Hühn 2013). To define what can be considered as a change of state, and therefore an event, Hühn (2013) distinguishes two types of events, based on the context in which the concept of event is used: (1) “a type of narration that can be described linguistically and manifests itself in predicates that express changes (event I), and (2) an interpretation- and context-dependent type of narration that implies changes of a special kind (event II), on the other.” Both *event I* and *event II* portray a basic type of narration and are characterized by a change of state, the transition from one situation to another, usually in relation to a character. *Event I* and *event II* are distinguished by the degree of specificity of the change of state. *Event I* changes of state consist of any change of state that contributes to the narrative, defining narrativity as the “relation of changes of any kind” (Hühn 2013). *Event I* concerns every type of change of state expressed in a text, whereas *event II* refers to specific changes of state that meet additional conditions, such as changes that are decisive, unpredictable turns in the narration, or a deviation from the norm of what is expected. The evaluation of the additional conditions of *event*

II is a matter of interpretation, and therefore *event II* is a hermeneutic category. On the contrary, *event I* can be evaluated rather objectively.

The definition of narrativity used in *event II* differs from the definition of narrativity used in *event I*. In *event II*, narration is considered to be the “representation of changes with certain qualities” (Hühn 2013). Whether these qualities are present is dependent on context and interpretation of the events in relation to the whole text. For example, “Mary stepped onto the ship” contains a type I event, namely the change of state of the character Mary by moving from the bank to the ship, resulting in a change of surroundings. However, in the context of a particular literary or cultural context, such as emigration, this can also be a type II event. Emigration can be seen as a new beginning and is therefore a deviation from what is expected. Therefore, this example can also be an *event II* change of state, depending on the literary and cultural context. *Event II* changes of state are considered to be more or less eventful, according to what extent they meet the following five criteria: relevance, unpredictability, effect, irreversibility, and non-iterativity (Hühn 2013). These additional criteria are also predominantly dependent on cultural, historical, or literary context. Therefore, the eventfulness of a change can be interpreted differently by different readers. Besides different event types, different event sequences have been conceptualized, too.

2.2 *Fabula and Syuzhet*

The Russian formalist Viktor Shklovsky introduced the terms *fabula* and *syuzhet* (Scheffel 2013) based on an analysis of the difference between chains of events in “actual life” and in art. Shklovsky argues that to understand the “aesthetic laws” of artistic narrative, the distinction between *fabula* and *syuzhet* is necessary. He defines *syuzhet* as “the material of the *fabula* in the artistic form.” In other words, the *fabula* represents what has happened or what was in the narrated world, whereas *syuzhet* is the artistic form in which the *fabula* is presented to the reader. *Fabula* is defined as “the material for *syuzhet* formation,” a chronological chain of events.

The *fabula/syuzhet* distinction is similar to the *story/plot* and *histoire/discourse* distinction (Pier 2003; Scheffel 2013). *Story* is “a narrative of events arranged in their time sequence” (Scheffel 2013). For instance, dinner comes after breakfast and Tuesday after Monday. *Plot* is a narrative of events focused on causality, for example, “The king died, and then the queen died of grief.” In the *plot* a causal relation between events is established, whereas in the *story*, the relationship is only chronological. More broadly, *plot* involves the transformation of “happenings” into a sequence of structured events that form a narrative (Xin 2022).

Similar to Shklovsky, Todorov identifies two aspects of literary works: *histoire* and *discourse*. A literary work is

at the same time a story [*histoire*] and a discourse [*discours*]. It is story, in the sense that it evokes a certain reality [...]. But the work is at the same time discourse [...]. At this level, it is not the events reported which count but the manner in which the narrator makes them known to us (Scheffel 2013).

The difference between *fabula/syuzhet* and *histoire/discours* is mainly found in the artistic value prescribed to the different terms. Todorov considers both *histoire* and *discours*

as important aspects of a literary work. *Histoire* is necessary to create a certain reality for the reader. *Discours* is important since literariness is not solely about the events reported, but also about the manner in which the narrator presents them to the reader. *Discours* also considers features such as perspective, style, and mode, whereas the *syuzhet* primarily focuses on the order of events represented in a text. Additionally, *histoire* contains the continuum of the narrated world, in contrast to *fabula* that only contains the parts of the narrated world that are relevant to the plot. Due to their broader definitions, *histoire* and *discours* are considered to be of equal literary value, whereas the *fabula* is considered not to be of literary value, and the artistic value of a text is represented solely in the *syuzhet*. Moreover, the interplay of the two sequences, with flashback and anticipations, generates a narrative tension, the narrativity that keeps readers engaged (Baroni 2007; Sternberg 1992). The automatic detection of both sequences is a difficult task, but computational literary studies have a unique interest in the way in which events are presented and can complement NLP efforts to detect the “histoire” of news. However, it is also worth noting that there are NLP works that consider aspects of the “syuzhet,” mostly in relation to the framing of events (Hamborg 2023; Minnema et al. 2022a).

3. Literary Event Detection

In this section, we provide an overview of event detection in literary texts, discussing whether and how the various approaches could be used for our objectives, namely to develop a computational model to detect the events of a fictional text in the way a reader learns about them, applicable in multiple languages.

3.1 Operationalizing the Narrativity of Event Representation

Given the specific interest of computational literary studies in the way in which events are presented, an operational model for the automatic detection of events in literary texts should enable the extraction of information not only about the semantics of events but also their rhetorical, narratological, and literary functions. To this end, Gius and Vauth (2022) started from operationalizing the narrativity of event representation at the level of discourse, using German prose as a case study.

Gius and Vauth (2022) define four different event categories that can be called *event I* in the context of Hühn:

1. Changes of state are physical or mental states’ changes of animate or inanimate entities
 “As Gregor Samsa awoke one morning from uneasy dreams;”
2. Process events are actions or happenings that do not result in a change of state, such as moving, thinking, feeling
 “he found himself transformed in his bed into a gigantic insect;”
3. Stative events are physical and mental states of animate or inanimate objects
 “His room lay quiet within its four familiar walls;”

4. Non-events do not relate to facts in the narrated world, such as general statements, questions or hypothetical situations

“She would have closed the door to the apartment.”

The four different types of events were specified in order to differentiate them for narrativity analysis and define events as “any change of state explicitly or implicitly represented in a text.” The events are ordered by degree of narrativity, with *changes of state* having the highest degree of narrativity and *stative events* the lowest. *Non-events* do not contribute to narrativity, but are included for comprehensive annotation. Gius and Vauth (2022) consider the whole text when annotating events. However, they aim to avoid “relatively strong interpretations necessary when primarily relating to the story world ‘behind’ its representation in the narrative” (Gius and Vauth 2022). The narrativity hierarchy of the four types of event categories ensures that the representation of eventfulness in discourse is reflected in the annotation. This indirect annotation of eventfulness is more aligned with the different types of eventfulness related to *event II* (Hühn 2013). One of the approaches of eventfulness discussed by Hühn (2013) requires that a change actually takes place in the narrated world (thus, it should be a fact in the narrated world) and that it reaches a conclusion (thus, the change cannot be described as only begun or in progress). This definition of eventfulness is similar to Gius and Vauth (2022)’s definition of *change of state*. However, as they annotate every event occurring in the text, and additionally non-events, their overall definition of event categories is broader than that allowed by *event II*, aligning more closely with that of *event I*.

The model with four event types has been used by Vauth et al. (2021) to annotate four German prose texts (Vauth and Gius 2021) and automatically classify events by following a two-step process. First, they extract verb phrases, which are then labeled with an event type in the second step. Since the annotation guidelines in Vauth and Gius (2021) focus on the finite verb of the sentence, the verb phrase extraction is done by selecting the finite verbs in each sentence using a pre-trained tagger. Then, for each verb, the dependency tree of a pre-trained parser is used to identify all tokens they cover, by traversing the tree. Relative clauses are not considered when moving down the dependency tree, and neither are conjunctions if their children consist of full verbs. On unseen data, the model reached a 0.71 F1 score in identifying the correct span and a 0.78 F1 score in classifying the event type. However, Vauth et al. (2021) only used German prose and therefore relied on a German pre-trained tagger and parser. Suitable pre-trained taggers and parsers will need to be selected for other languages to test this approach.

3.2 Literary Events as *realia*

Sims et al. (2019) define a literary event as an event that is actually happening in the story (*realis*), with the goal of analyzing the narrative plot. In this model, there are no stative events, they only consider activities, achievements, accomplishments, and changes of state, following Vendler (1957). A phrase is considered an event by Sims et al. (2019) if either (1) a change of state has occurred, (2) the cause of a state can be deduced, or (3) the phrase refers to an acute mental state, such as acute short-lasting responses like *shocked* or *astonished*. This specification of the types of events is in line with *event II*, in which an event is defined as a “representation of changes with certain

qualities” (Hühn 2013). As a consequence of these requirements, fewer phrases are annotated as events than in Vauth and Gius (2021). For example, the sentence “at the end of the path is a cave” is a stative event for Vauth and Gius (2021), but it would not be an event according to Sims et al. (2019).

Similar to the guidelines by Vauth and Gius (2021), events must have occurred, thus negations are not considered to be events, nor are possible future events. Generic phrases are also considered not to be events in both guidelines. However, Sims et al. (2019) do not treat wishes and desires as events, whereas Vauth et al. (2021) consider the act of wishing as stative events. Another difference is that Sims et al. (2019) consider single words as events, whereas for Vauth et al. (2021), all words that can be assigned to a finite verb are included in the annotation span of an event. Lastly, Sims et al. (2019) define their event triggers more broadly, including not only verbs but also adjectives and nouns. This approach has the advantage of being extendable to languages whose syntax does not rely on verbs as much as English does, but it also has the limitation that Vendler’s verb classes are not applicable to many languages.

For event detection, Sims et al. (2019) use an LSTM and five BiLSTMs that are trained on the annotated verbs only (baseline), as well as on a featurized model containing six extra token features: lemma, part of speech tag, context, syntax, WordNet synset and hyponymy information, word embeddings, and bare plurals as subjects. The differences between the BiLSTMs are the context included in the BiLSTM, including a sentence CNN, document context, and BERT contextual representations. The BiLSTM with BERT representations on the featurized model has the highest performance, with an F-score of 73.9.

3.3 Hylistic Analysis

Pannach (2023) analyses events in folktales using the hylistic theory (Zgoll 2020). Folktales are considered as sequences of hylemes. A hyleme is an individual statement containing events and states in chronological order. For example, the statement “Orpheus came to his end by being struck by a thunderbolt” results in the following hyleme sequence, which consists of three parts: (1) “Orpheus is struck by a thunderbolt,” (2) “Orpheus dies,” and (3) “Orpheus is dead.” This model does not include aspects related to how the events are presented, it rather focuses on achieving the best possible comparability between different variants of the same folktale, even across languages. That is why the events are translated into present-tense statements that describe precise actions or states. Additionally, Pannach uses four main categories in her hylistic analysis: single-point (punctual), durative-constant, durative-initial and durative-resultative, which are mainly associated to verbs in a phrase. Single-point hylemes consist of active actions, passive experiences, reactions, perceptions, or feelings. The beginning and end of the event are both included in the hyleme sequence. Durative hylemes are true for part of the sequence or the entire hyleme sequence. Durative-initial hylemes are true at the beginning of a sequence, durative-constant are true during the entire sequence, and durative-resultative at the end of the sequence.

Pannach (2023) compares this approach to the event model of Gius and Vauth (2022). The change-of-state event category of Gius and Vauth (2022) corresponds to the single-

point category used by Pannach (2023). Process events are also considered to be single-point. However, when the property of the event is iterative, such as “Charon works the sails,” the phrase would be considered to be durative-constant. Stative events correspond to the three durative hylestic classes, which class it belongs to depends on the context. Non-events are not annotated in the hylestic classes.

The vast majority of the annotated data consists of single-point statements. Due to the unequal distribution, as well as the similarity between the three different durative hylestic classes, a multinomial naive Bayes model was used, with a TF-IDF vectorizer. Three classifiers were implemented, one binary classifier distinguishing single-point and durative hylemes, one classifying durative-initial, durative-constant and durative-resultative hylemes, and one considering all four classes. The binary classifier reached a 0.79 F1 score for the durative hylemes and a 0.92 F1 score for the single-point hylemes. The second classifier obtained a 0.32 F1 score for the durative initial statements, a 0.85 F1 score for the durative constant, and a 0.56 F1 score for the durative-resultative. It is important to note that 69% of this test set consists of durative-constant hylemes, and 24% of durative-resultative hylemes. For the third classifier, the durative-initial hylemes obtained a 0.25 F1 score, the durative-constant hylemes a 0.69 F1 score, the durative-resultative hylemes a 0.43 F1 score, and a 0.93 F1 score on the single-point statements. In this test set, the distribution across the different classes is again unbalanced, as the test set only contains 30 durative initial hylemes and 1,151 single-point statements. As it is unclear whether the class imbalance in the test set of the second and the third classifier is reflected in the respective training sets, it is hard to determine how this imbalance has influenced the results, and whether this influences the strong preference for the single-point hylemes by the third classifier.

3.4 Analyzing Narrative Discourse with Large Language Models

Piper and Bagga (2024) use large language models (LLMs) to analyze narrative discourse within the framework of Genette (1980)’s narrative triangle concerning story, discourse, and narrating instance. They use three categories to analyze narrative discourse: (1) *POV* (point of view), focused on the experiencing agent; (2) *time*, including use of tense, anachrony, flashbacks, eventfulness, and event sequences; and (3) *setting*, including location and concreteness (realized and tangible space). Thus, they explicitly use event sequences and eventfulness as features to capture dimensions of time.

Piper and Bagga (2024) prompt LLMs to estimate the degree of presence of a given feature using a three-point scale. The dataset of Piper and Bagga (2022) is used to collect 13,543 passages from 18 genres, including contemporary novels, short stories, folktales, and non-fiction, such as memoirs and stories from AskReddit. The experiments were run on a subset of passages with a manually annotated narrativity score higher than 3.0. The evaluation consists of four steps: (1) replication, (2) honeypot, (3) inter-annotator agreement, and (4) model performance. First, 15 iterations are run on half of the validation data. For the best model, 95.6% replication occurs in all documents. Second, a nonsensical “honeypot” feature is used, for which the answer should never be positive. This feature is used to measure to what extent a model is randomly guessing. In the best model, all nonsensical prompts were answered negatively. Third, three annotators answered identical prompts the models received. The inter-annotator agreement is

fair, with a Fleiss's kappa of 0.38 and a universal agreement rate of 43%. Lastly, the model's accuracy is evaluated by comparing the model's results to the majority vote of the human annotators and the minimum match, where the results are compared with any human answer regardless of the majority vote.

There is a variance in the models' F1 score, from 0.28 to 0.79 of the majority vote, but a higher performance for the minimum match, with a maximum F1 score of 0.95, and four out of six models with an F1 score of 0.87 or higher. The annotator agreement correlates strongly with model performance. Thus, LLMs are a promising tool in the analysis of narrative discourse, specifically since the results show that the features "event sequences" and "eventfulness" can have different weights in classifying narrative. As the high variance across models is also seen between human annotators, the results emphasize the subjectivity and ambiguity in the task.

4. Related Work

In addition to computational literary studies, event detection has been a research topic in a multitude of domains, such as journalism and history (Lai 2022), using NLP and information extraction (IE) techniques (Santana et al. 2023). Despite the wide range of research conducted on events, adapting previous work to a new domain is complex – for example, due to the scarcity of corpora annotated with temporal information in historical texts (Sprugnoli and Tonelli 2017).

Another challenge is the lack of a general definition of events in Sprugnoli and Tonelli (2017) and across domains (Caselli and Bos 2023; Santana et al. 2023). In NLP, event detection is defined as the task of finding all pairs of linguistic expressions $(w_i, w_j) \in D$, in which D is a given document, w_i is an instance of an event trigger, and w_j is an instance of an event participant (Caselli and Bos 2023). The event triggers are defined as linguistic expressions that depict the happening of something, or a state. The event participants are expressions concerning the actors, location, and time of occurrence. Thus, by this definition, events represent complex relationships between actors, places, objects, actions, and states.

Because of the definition of events as complex relationships, events and storylines can be expressed as knowledge graphs (Kishore and He 2024; Wadhwa et al. 2024; Yan and Tang 2023). Yan and Tang (2023) introduce EventTKG, a narrative graph generation framework which can be used to generate storylines based on news and other media streams. They define an event as something that happens at a specific time and place, carried out by an individual or organization. Events are distinguished from complex events; complex events are clusters of events concerning the same topic that are also considered to be the basic elements of a storyline, with a storyline being a chronologically arranged sequence of events. Despite this broad definition of event, complex event, and storyline, Yan and Tang (2023) conclude that EventTKG can be applied only to a limited number of news datasets and that real-world events are also too complex for this framework. Therefore, the applicability of this framework to fiction appears limited.

Another approach is using LLMs to generate event sequences based on an event knowledge graph with partial causal relations (Wadhwa et al. 2024) or to track the context of

sentences and events (Miori and Petrov 2024); such information can then be incorporated in event knowledge graphs. Using LLMs in the development of event sequences and event knowledge graphs is promising, but the bias in an LLM can influence event extraction. For example, Kishore and He (2024) show that GPT-3.5 has a bias towards “AFTER” in a question-answer format concerning the chronological sequence of two events in a given text, whereas GPT-4 has a preference for “BEFORE.” When assessing truthfulness on the chronological order of events in a given text, GPT-3.5 has a bias towards “TRUE,” whereas GPT-4 tends towards “FALSE.”

In addition to these limitations, we need to consider that the definition of event in NLP, as a linguistic expression of a relationship between a happening or state and an actor, location, or time of occurrence (Caselli and Bos 2023), does not provide a way to distinguish different sequences of the same events, i.e., the *fabula* vs. the *syuzhet*. In NLP, the goal of event extraction is mainly to derive and represent the events occurring in a text so that the events and the text can be easily analyzed, visualized, and searched. The relationship between event triggers and event participants – as described by Caselli and Bos (2023) and applied in most NLP work – only links the *what* to actors (*who*), location (*where*), and time of occurrence (*when*). However, if we are interested in the way a reader learns about the events in a text, it becomes more important to focus on the how events relate to the fictional world (i.e., the *what* of narration) and their representation in the text (i.e., the *how* of narration) (Gius and Vauth 2022), for which using event categories distinguishing eventfulness are more suitable (Gius and Vauth 2022; Hühn 2009, 2013).

This basic theoretical difference makes it difficult to compare and relate previous work in NLP event detection to the goal of event detection in literary texts. However, in subsection 4.1 and subsection 4.2, we discuss different NLP techniques used in event extraction on news and historical texts, with the goal of showing in more detail to what extent these works can complement a narratological approach.

4.1 News

In this section, we discuss the issues identified concerning the comparability between different works and corpora in two recent surveys (Caselli and Bos 2023; Norambuena et al. 2023) and by discussing the data structure of events proposed in Vossen et al. (2021), since it is one of the most elaborate narratology-based frameworks.

Caselli and Bos (2023) find that variation in the definition of events and the annotation of linguistic realizations, and the assignment of events to specific semantic classes, make most of the event-labeled corpora incompatible with each other. They give an overview of six event-annotated news corpora, which all use a different event definition. The majority of these corpora restrict the annotation of events by solely annotating events that occur in given event classes. These restrictions make these frameworks unsuitable for literature. For example, ACE (Doddington et al. 2004) only annotates events in news articles that fall under one of eight semantic classes (*life, movement, conflict, business, contact, personnel, justice, transaction*). In contrast, TimeML (Pustejovsky et al. 2003) rejects restrictions on semantic classes and linguistic realizations of events, as annotations are based on the lexical aspect and their contextual syntactic structure.

As TimeML is aimed at portraying events as temporal expressions relative to each other, this approach is not applicable in the analysis of the *syuzhet*.

Norambuena et al. (2023) identify two fundamental units in news narrative extraction, *events* and *entities*, i.e., the actions and happenings in the text and the characters and other entities that are related to the events. Focusing on the former, they define *computational narrative representation* as a discrete story structure, such as a graph or a timeline of events. They observe that the most common and simple way to computationally represent a narrative is as a linear sequence of events, such as a timeline.

Since the survey only analyses research using news corpora, they assume that each text (news article) focuses on one single main event. Previous or secondary events, which can, for example, be used to link articles together, are not taken into account in this survey. As previous and secondary events are crucial in fiction, this assumption is not applicable to literary event detection. They identify three scopes: events as sentences, events as entire documents, and events as a cluster of documents. This is a broader view of events than in many other approaches. For example, TimeML defines events as more specific than an action, such as a perception.

Among these seemingly incompatible approaches, there are also two that leverage insight coming from narratological scholarship. The first one is Vossen et al. (2021), who propose a framework informed by narratology and argue that a plot structure is composed of three elements: (1) an *exposition*, in which the characters and the setting are introduced, (2) a *predicament*, which consists of a set of struggles or problems that an actor has to go through, and (3) the *extrication*, which is the end of the predicament. The predicament itself consists of three elements: (1) *rising action*, which consists of events that increase the tension, (2) *climax*, which consists of events where the tension reaches its maximum, and (3) *falling action*, which consists of events that resolve the climax and lower tension. Besides these dynamic patterns, they also define three data structures: the *timeline* based on the *fabula*, which they define as a chronological timeline; the *causeline*, related to the plot, which they define as a set of loose and strict causal relations; and the *storyline*, which they define as a set of (pairwise) relations between events according to the patterns mentioned above; it is associated to the plot structure. The storyline includes the explanatory causal relations between events related to a climax event by the strongest connection. The events in the storyline are chronologically ordered. In annotation, every event mention is associated with a temporal expression or is directly temporally related to other events in the timeline. In the causeline only events that express a loose causal relation are included. Based on the causeline, the storyline depicts explicit additional explanatory relations that may lead to a climax event. In section 6, we will compare this framework to our approach of analyzing narrative events.

Another NLP work looking at narratology – as well as at Critical Discourse Analysis (CDA) – is by Huang and Usbeck (2024), who propose a theoretical framework to construct new narratives from an author-focused perspective. CDA considers news narrators as a dominant group that shapes a narrated world encoded in language, in which real-world events are portrayed to the public. Therefore, the focus is on how real-world events are organized to shape a narrated world, using an adapted definition of *fabula* and *discourse* by Gervás and Calle (2024). They consider the information flow from a real-world event to a news item as follows: First, based on a real-world event, a

subset of an organized event sequence forms the *fabula*; then, the *discourse* is created through narrative composition, simplified as causal relations between the events in the *fabula*; and, lastly, the *discourse* is used to form textualized narratives in natural language. They define *fabula* as “the actual sequence of events, that is chronologically and causally ordered” and *discourse* as “the product of the telling, which reorganizes the chronological and causal order of this sequence.” They view the narrated world as event-event causal relations and narration as a function that shapes the narrated world. They consider events as the smallest unit in a narrative, but do not consider all events in a text to be part of the narrated world. Indeed, they make a distinction between constituent and supplementary events, of which only the former are represented as event-event causal relations. The proposed theoretical framework represents this information flow as the narrated world logic, which can be used to extract the core story of events told by a news narrator. As this is a proposal for a theoretical framework that has not been evaluated yet, it is unclear how effective it is and whether this framework is applicable to literature.

To conclude, in the task of event detection in news, there is no general consensus on the definition of event. This lack of consensus shows that relying on existing frameworks and corpora does not lead to broadly applicable annotations, as the different corpora are hard to compare and relate to each other (Caselli and Bos 2023). Moreover, most corpora restrict events to certain semantic event classes, but this is too restrictive for a comprehensive analysis of the *syuzhet*.

4.2 Historical Texts

The lack of a general consensus on the definition of events does not only occur with event extraction in news texts, but also with historical texts. Additionally, the aim of event extraction from historical texts is not focused on information extraction only, but also on the analysis and interpretation of events. To address the difference in objectives between fields, and to make NLP techniques applicable to historical texts in such a way that it will lead to a more homogeneous usage of event extraction in historical research, Sprugnoli and Tonelli (2017) suggest using the expertise of historians for the linguistic annotation of events.

Sprugnoli and Tonelli (2019) conclude from their discussions with historians that the semantic type of an event is the most relevant information for annotation, that multi-token annotation of event phrases should be possible, and that events can have different syntactical forms and grammatical classes. Accordingly, they define 22 relevant semantic classes, based on the semantic categories of the Historical Thesaurus of the Oxford English Dictionary (HTOED), aiming to avoid too much granularity while at the same time ensuring broad informativeness. The latter is important due to the diverse topics and genres in historical texts.

They consider three different types of event spans: (1) single-token, (2) multi-token, and (3) discontinuous expressions. Events can be verbs, past participles, present participles, adjectives, nouns, and pronouns. Multi-token events are restricted to seven types of linguistic construction, such as phrasal verb constructions, final and non-finite verbs, and nouns.

The resulting annotated corpus, the Histo Corpus (Sprugnoli and Tonelli 2019), is used to train two types of classifiers: CRF classifiers and a BiLSTM. Two CRF classifiers were implemented: one to identify the event span and the other to predict the correct event class on unseen text. The BiLSTM is used for sequence tagging as well as event detection and event classification. Overall, the BiLSTM outperforms the CRF classifiers in event classification, except for the event class ‘physical sensations.’

In another project (Verkijk and Vossen 2023), historians have been involved in the development of an ontology that can be used for event extraction from the archives of the Dutch East India Company (VOC). The ontology enables the extraction of implied events, as this is deemed to be important by experts. They use the CEO ontology (Segers et al. 2017), which models semantic circumstantial relations between event classes, as the basis for the definition of event classes, since they want to be able to annotate static events. They identify three relevant types of observable events: *ship movement*, *trade*, and *(geo)political/social relations*. More detailed classes, for example, whether an action is legal or illegal, depend on the context and the interpretation of an expert, and are therefore not considered as observable events. Building on FrameNet (Ruppenhofer et al. 2010) and CEO, they define participants specific to each event class. Other event arguments are spatial or temporal. Roles can be recycled from one event to another; for example, the agent in an *Attacking* event is a patient in the state *BeingInConflict*. Results show good agreement between human annotators for the labeling of event triggers, but poor performance of fine-tuned models for automated event detection (Verkijk et al. 2024), as the highest precision achieved is 0.55 using the model GysBERT (Manjavacas and Fonteyn 2022) and the highest recall is 0.43, obtained using the model XLM-R (Conneau et al. 2019).

From this type of research, we can observe that event annotation in historical texts differs greatly from approaches to annotate events in literature. Both Sprugnoli and Tonelli (2019) and Verkijk and Vossen (2023) use predefined semantic classes and themes to identify and analyze events, while considering a multitude of syntactical forms and grammatical classes. However, for research on literature, all events in the text are relevant because they can fulfill different functions that cannot be defined in advance (Pianzola 2018). Some events contribute to creating the setting for the story, other events contribute to the progression of the plot, and others contribute to show the personality of the fictional characters. All events potentially play a role in the cognitive and aesthetic processing of literary text by readers (Caracciolo 2014).

4.3 Review Summary

As can be seen in Table 1, multiple projects in different domains operationalize event detection without using a concrete definition of what should be considered to be an event. Instead, some use strictly defined semantic classes in their annotation guidelines (Sprugnoli and Tonelli 2017; Verkijk and Vossen 2023), some use narrative statements to detect events (Pannach 2023), and some use existing datasets and their respective definitions to analyze events (Piper and Bagga 2024; Vossen et al. 2021). These diverse approaches show that in the operationalization of events and related theoretical concepts, strict event definitions are not a prerequisite for implementing computational pipelines for event detection.

Authors	Granularity	Scope	Event definition	Notes
Gius and Vauth (2022)	Finite verbs	German literature	Any change of state explicitly or implicitly represented in a text.	Uses four event types: change of state, process event, stative event, and non-event.
Pannach (2023)	Finite verbs	Folktales	–	Uses a narrative statement instead of a definition on events. The categories used for hylistic analysis are: single-point (punctual), durative-constant, durative-initial and durative-resultative.
Piper and Bagga (2024)	Rate of non-helping verbs / rate of non-stative verbs (Piper and Bagga 2022)	Contemporary novels, short stories, folktales, and non-fiction such as memoirs and stories from AskReddit.		Focuses on event sequences (series of sequential actions) and eventfulness (how reliant the narrative discourse is on action rather than description, qualia, or dialogue).
Sims et al. (2019)	Verbs, adjectives, and nouns	English literature	They use three criteria to define event, which all need to occur within the context of the sentence: (1) an explicit change of state, (2) the cause of the state, and the cause and resulting state occur at the same location, (3) an acute mental state.	They tag event triggers, defined as the minimum extent of text capable of representing an event, including activities, achievements, accomplishments, and changes of state, as being events.
Huang and Usbeck (2024)		News	Smallest functional unit in the narrated world that causes a change of state. This state can be of a story world or of a mental world for a character or a reader.	They focus on constituent events, as they are the essential events that form the backbone of the narrative. They filter out the supplementary events that are not crucial to the plot, but add depth, richness, and complexity to the narrative.
Vossen et al. (2021)	Verbs, adjectives, and nouns	News	ECB+ event annotation (Cybulska and Vossen 2014).	ECB + models events from news data as a combination of four components: (1) an event action component, (2) an event time slot, (3) an event location component, (4) a participant component.
Yan and Tang (2023)	The arguments of event contain trigger, subject, object, time, and place	News	An event is something that happens at a specific time and place, and is carried out by an individual or organization. Complex events are blocks of events involving the same topic, represented by a group of non-overlapping clusters of events, and link multiple news with the same topic.	Framework is only applicable on a limited number of news datasets, as real-world events are too complex for this framework.
Sprugnoli and Tonelli (2019)	Verbs, verb constructions, adjectives, and nouns	Historical travel narratives and news	–	Does not explicitly define events, but uses the 22 semantic classes to identify events.
Verkijk and Vossen (2023)	Event classes defined by specialist historians, focusing on the circumstantial relationship between events	Historical texts with event classes focusing on ship movement, trade, and (geo)political/social relations	They do not give a general definition on events, but do define static (such as being at a location or being in conflict) and dynamic events (such as leaving a location or attacking).	The event classes are defined to represent observable events, steering clear of concepts that have an inherently subjective character.

Table 1: This table summarizes the research discussed in the sections above. As can be seen, not all studies use a definition of events; some only use an operationalization of events suitable for their task and computational pipelines.

5. An Operationalization of Literary Events for Multilingual Corpora

Aiming at a broader applicability of the model of Gius and Vauth (2022), we have modified their guidelines with extra examples and edge cases from English fiction. Our multilingual corpus consists of fiction, specifically fanfiction. Four of the added examples can be found below:

1. Change of state: [The baking sheet **sighed** a bit,]₁ [**beginning** to relax.]₂ PROCESS EVENT₁ + CHANGE OF STATE₂;
2. Process event: [you **are** on a path in the woods];
3. Stative event: [unsure what to make of a scene];
4. Non-event: [“You **need** to make friends, Ryeowook ah,”]₁ [he **had** said over the dinner table]₂ NON EVENT₁ + PROCESS EVENT₂.

The first example shows the importance of the duration of a motion, as the first part of the sentence, “the baking sheet sighed a bit,” is a process event, whereas the second part, “beginning to relax,” is a change of state. As sighing is a short-lasting motion, it is a process event. In the second part, the finite verb is *beginning*, which implies that this phrase marks a longer-lasting change in the character state, namely relaxation.

Our corpus also displays a great variety in the used type of narrators. For example, in a narrative told by a second person narrator, the sentence, “you are on a path in the woods” (example 2), is a process event. The finite verb in this sentence is *are*, which implies that the character in the sentence (*you*) is in motion, because the next sentence in the text is “at the end of the path is a cave,” which suggests that the characters have reached the cave.

A third notable case we observed in our corpus is the use of implied verbs. Despite the missing verb in example 3, “unsure what to make of a scene,” this has still been annotated, as the words, “he was,” are implied in the context of the full text. The inclusion of implied verbs is particularly important for the applicability of a definition of events to multiple languages as not all languages are as verb-focused as English and German. For example, in Bahasa Indonesia, it is possible to form a grammatically correct sentence that does not contain any verbs, as auxiliary verbs do not exist in Bahasa Indonesia.

The fourth example shows the influence of dialogue in fiction, where the first part of the sentence, “You need to make friends, Ryeowook ah,” is spoken. Since this is an opinion stated by the speaker, this is a non-event, as the sentence does not relate to a fact in the narrated world. The second part of the example, “he had said over the dinner table,” is a process event, as the verb focuses on the action of saying the first sentence, which is an action.

The fact that events reported in dialogue are labeled as non-events is quite limiting, because it is common that readers get to know about happenings in the story world through the voices of different characters. The four different categories of events proposed by Gius and Vauth (2022) are not enough for a complete account of all events in a

story. To fill this gap, we make two changes: (1) We introduce complementary labels for speech and thought, and also (2) annotate events within reported speech and thought. Since the goal is to give a fine-grained representation of *how* events are presented, we decided to work with *four* additional secondary labels that distinguish between direct and indirect reports:

1. Direct speech. Example: "Man, am I tired!"
2. Indirect speech. Example: Man, was he tired!
3. Direct thought. Example: "I'm tired!" he thought.
4. Indirect thought. Example: He thought he was tired.

The aim of adding these four extra labels is to be able to analyze in what way speech and thought are used to present events and narrate a story. For example, the phrase: " – You looked through my phone!" would get two labels: *process event* and *direct speech*. The usage of speech and thought in a narrative also influences the certainty and uncertainty of the occurrence of an event. In the phrase: "But they don't want to be friends with me, Appa," the reader will perceive it as the speaker's opinion that they don't want to be friends with them. If this information would be stated by the narrator, this would be perceived as a fact. These complementary labels could therefore be interesting in the analysis of framing and the presentation of information in fiction, but also in news.

Additionally, these labels can be combined with other narrative features for a more nuanced analysis, for example, with labels for the type of narrator (first-, second-, third-person narrator) or focalization, the different points of view from which the action is looked at (Jahn 2021). The presentation mode of events can influence the reader's epistemic stance towards their occurrence. For instance, when events are conveyed through speech, thoughts, or dreams, the reader's confidence that they actually took place may be diminished. Having distinctive labels for thoughts is useful as thought presentation occurs in two contexts: First, it can show that the narrator had direct access to relevant thoughts (Semino and Short 2004), either as a third-person omniscient narrator expressing the thoughts and mental states of the characters in a text, or as a first-person narrator presenting their own thoughts and mental states. In the second context, the narrator does not have such access, but infers the character's thoughts based on external evidence, such as a person's speech, facial expressions, and actions (Semino and Short 2004). Thought presentation, in particular indirect thought, is also associated with the creation of feelings of closeness and empathy for the characters by the reader. Thus, adding these four extra labels to the event categories enables a more thorough analysis of the *syuzhet* of a text. The perception of events in the *syuzhet* is influenced not only by their narrativity but also by presentation modes and focalization. Operationalizing the annotation and classification of events in literary texts taking into account all these variables would be the best-case scenario for a computational narratology of events. However, this has not been done yet by research on literary event detection.

6. Comparison of Computational Narratology in NLP and Literary Studies

To better illustrate the differences between approaches, we compare our narratological model to a narratology-inspired approach for NLP event extraction (Vossen et al. 2021) (see subsection 4.1), which proposes three data structures (or sequences): timelines, causelines, and storylines. For the comparison, we annotate a sample of news (originally used in Vossen et al. (2021)) and a sample of fiction from our corpus. The goal is to show how the domain-specific interests of computational literary studies and NLP for news analysis can lead to different operationalizations of narratological concepts.

6.1 Timeline, Causeline, and Storyline

Figure 1 shows the news sample from Vossen et al. (2021). The temporal relation between all events is expressed in the timeline, whereas only the loose causal relations are included in the causeline, and only explicit explanatory relations that may lead to the climax event are included in the storyline. Figure 2 shows the fiction sample, annotated according to Vossen et al. (2021). Figure 2 shows that timelines, causelines, and storylines do not fully reflect the story presented in fictional texts. First, fiction contains more description (of, for example, surroundings) than news. The timeline of the news sample shows a clear temporal order of events in the text, whereas the temporal order for the description of the grove and the way in which the wolf is stretched out are not explicitly expressed. It can be assumed that the splitting of the grove was created before the stone was placed there, however, it is also possible that the stone was first placed there and the trees grew around it. In genres such as science fiction and fantasy, the environment is not necessarily static, thus complicating expressing all events in a timeline.

Second, the causeline does not contain the description of the grove, the stone, and the way in which the wolf is stretched out. Therefore, this description is not included in the storyline, as the storyline is based on the causeline. However, despite not being part of the causal relations between events, the description of the grove, the stone, and the wolf does contribute to the narrative, since it helps the reader to imagine the scene and contributes to the build-up of suspense, the tension leading to the climax.

Lastly, the storyline that can be derived from the causeline stops at the event **froze**_{e7}, which is the climax of the storyline. Half of the events occurring in the sample, namely those related to the description of the grove and the wolf, are not included in the storyline. However, due to the emphasis on the description of the grove, the stone, and the wolf, the wolf dying appears to be crucial to the narrative. The description of the scene also contributes to the build-up of suspense, thus the event **froze**_{e7} is not actually a climax (according to Vossen et al. (2021)), as there is no falling of action or resolution afterward. Additionally, readers could conclude from this excerpt that the death of the wolf is more important to the narrative than Wilson walking towards and discovering the dead wolf, whereas the storyline only portrays the movements of Wilson.

Police **say**_{e1} that on Saturday around 11:30 p.m. Kimani Gray was **standing**_{e2} outside his home with five other young men before **splitting off**_{e3} when he **noticed**_{e4} two plainclothes officers in an unmarked car. After he “**adjusted**_{e5} his waistband and continued to **act**_{e6} in a suspicious manner,” officials **say**_{e7} the cops got **out**_{e8} of their car and **approached**_{e9} Gray – who allegedly **turned**_{e10} toward them with a loaded .38-caliber revolver in hand. The 30-year-old sergeant and 26-year-old **fired**_{e11} shots.

- *timeline*: [NOW] → includes → **say**_{e1}; **say**_{e1} → before → **say**_{e7}; **say**_{e1} → after → [Saturday around 11:30 p.m.]; [Saturday around 11:30 p.m.] → includes → **standing**_{e2}; **standing**_{e2} → before → **splitting off**_{e3}; **splitting off**_{e3} → simultaneous → **noticed**_{e4}; **noticed**_{e4} → before → **got out**_{e8}; **act**_{e6} → before → **got out**_{e8}; **got out**_{e8} → before → **approached**_{e9}; **approached**_{e9} → simultaneous → **turned**_{e10}; **turned**_{e10} → before → **fire**_{e11};
- *causelines*: **act**_{e6} → circumstantial → **approached**_{e9}; **splitting off**_{e3} → circumstantial → **noticed**_{e4}; **turned**_{e10} → circumstantial → **fire**_{e11};
- *storyline*: **noticed**_{e4} → rising_action → **splitting off**_{e3} → rising_action → **adjusted**_{e5} → rising_action → **act**_{e6} → rising_action → **approached**_{e9} → rising_action → **turned**_{e10} → rising_action → **fired**_{e11}[climax];

Figure 1: Example of the timeline, causeline, and storyline framework applied on news from Vossen et al. (2021).

When Wilson first **heard**_{e1} the sounds of the dying wolf through the corpse of trees, he had **pulled**_{e2} the hunting rifle off his shoulder and **approached**_{e3} warily, **expecting**_{e4} the scene to include fighting foxes, or a stray dog that had **wandered**_{e5} into a snake nest. When he **saw**_{e6} the huge shape of the wolf, he **froze**_{e7}, **unsure**_{e8} of what to make of the scene. The grove of trees **split**_{e9} into a small clearing, and in the center of the circle of grass was a **stone**_{e10} about as tall as Wilson’s waist. The wolf was **stretched out**_{e11} over the top of the stone, head pointed one direction, feet in the other. The stone was **covered**_{e12} in enough blood that it had **dripped**_{e13} down the side of the stone and **coated**_{e14} the dirt around the rock.

- *timeline*: [NOW] → after → **heard**_{e1}; **heard**_{e1} → before → **saw**_{e6}; **pulled**_{e2} → after → **heard**_{e1}; **pulled**_{e2} → simultaneous → **approached**_{e3}; **heard**_{e1} → simultaneous → **expecting**_{e4}; **expecting**_{e4} → simultaneous → **wondered**_{e5}; **saw**_{e6} → before → **froze**_{e7}; **saw**_{e6} → after → **stretched out**_{e11}; **saw**_{e6} → simultaneous → **unsure**_{e8}; **split**_{e9} → before → **heard**_{e1}; **stone**_{e10} → before → **heard**_{e1}; **stretched out**_{e11} → before → **covered**_{e12}; **covered**_{e12} → before → **dripped**_{e13}; **dripped**_{e13} → before → **coated**_{e14};
- *causelines*: **heard**_{e1} → circumstantial → **pulled**_{e2}; **heard**_{e1} → circumstantial → **expected**_{e4}; **saw**_{e6} → circumstantial → **froze**_{e7}; **saw**_{e6} → circumstantial → **unsure**_{e8}; **covered**_{e12} → circumstantial → **dripped**_{e13}; **dripped**_{e13} → circumstantial → **coated**_{e14};
- *storyline*: **heard**_{e1} → rising_action → **pulled**_{e2} → rising_action → **approached**_{e3} → rising_action → **saw**_{e6} → rising_action → **froze**_{e7}[climax].

Figure 2: Timeline, causeline, and storyline framework by Vossen et al. (2021) applied on fiction.

[Police **say**]_{process} [that on Saturday around 11:30 p.m Kimani Gray **was** standing outside his home with five other young men before splitting off]_{stative & indirect speech} [when he **noticed** two plain officers in an unmarked car.]_{process & indirect speech} [After he “**adjusted** his waistband”]_{process & direct speech} [and **continued** to act in a suspicious manner,”]_{process & direct speech} [officials **say**]_{process} [the cops **got** out of their car]_{process & indirect speech} [and **approached** Gray]_{process & indirect speech} – [who allegedly **turned** toward them with a loaded .38 revolver in hand.]_{process} [The 30-year old sergeant and 26-year-old **fired** shots [...]]_{change of state}

Figure 3: News example of the annotation of narrative events. The bold verbs are the finite verbs per annotation span. Note that the word event is omitted from the annotation labels for abbreviation.

[When Wilson first **heard** the sounds of the dying wolf through the corpse of trees,]_{process} [he **had** pulled the hunting rifle off his shoulder]_{change of state} [and **approached** warily,]_{process} [**expecting** the scene to include fighting foxes,]_{non} [or a stray dog that **had** wandered into a snake nest.]_{non} [When he **saw** the huge shape of the wolf,]_{stative} [he **froze**,]_{change of state} [unsure what to make of the scene.]_{stative} [The grove of trees **split** into a small clearing,]_{stative} [and in the center of the circle of grass **was** a stone about as tall as Wilson’s waist.]_{stative} [The wolf **was** stretched out over the top of the stone,]_{stative} [head pointed one direction,]_{stative} [feet in another.]_{stative} [The stone **was** covered in enough blood]_{stative} [that it **had** dripped down the side of the stone]_{change of state} [and **coated** the dirt around the rock.]_{change of state}

Figure 4: Fiction example of the annotation of narrative events. The bold verbs are the finite verbs per annotation span. Note that some events contain implied finite verbs and that the word event is omitted from the annotation labels for abbreviation.

6.2 Narrative Events

In Figure 3 and Figure 4, the two samples are annotated following our definition of narrative events. When comparing the storylines of the news and fiction sample to the annotation of narrative events, it is evident that the build-up and rise in action to a climax (as defined by Vossen et al. (2021)) can be related to the narrative event model. According to this model, all events are processes, except for Gray standing outside and the firing of the shots. Thus, process events in the text seem to build up to the same climax event, which is annotated as a change of state. In Figure 1 the storyline starts with **noticed**_{e4}, whereas [Police **say**] is annotated as a process event.

The firing of the shots is described as a change of state, which puts the emphasis on the police agents shooting at Gray. It is a change of state as the finite verb of the sentence is *fired*. One of the distinguishing properties between changes of state and process events is irreversibility. If the finite verb expresses an irreversible change, the corresponding phrase is a change of state, as the irreversible change has led to a permanent property change of an entity. Firing shots is such an irreversible change, as one cannot reverse firing a shot. An alternative phrasing of the event reported in the last sentence could have a different event type. For example, the same event could be presented from the perspective of Gray (like in the second sentence “he noticed”): “Gray heard gunshots.” This sentence would be annotated as a process event, as the finite verb is *heard* and emphasizes describing a perception.

This can be related to research in which semantic frames are used to analyze perspective and framing in news (Minnema et al. 2022b). For example, in the headline “Cyclist, 70s, seriously injured following collision in Dublin,” the word *collision* triggers the

frame *impact*, showing that the main event in the sentence describes the impact on the cyclist. The same event has also been described with the following sentence: “Driver hits pedestrian with his car, sending the 70-year old man to hospital with heavy injuries.” In this headline, *hits* is the trigger of the frame *cause_impact*, which shows that the main event in this headline expresses the cause of the impact, namely the driver causing the injuries.

The first headline would be annotated as a stative event according to our framework, as the finite verb *injured* describes the physical state the cyclist is in. The low level of narrativity corresponding with this narrative event also corresponds with the frame *impact*, as the impact is described without naming the agent that has caused the accident. The second headline is a process event, as the finite verb is *hits*, which describes a motion. This corresponds with a higher level of narrativity, which fits with the frame *cause_impact*, as this emphasizes the action that caused the impact.

In the fiction sample, the different event categories fluctuate (see [Figure 4](#)). The text starts with a process event, then the level of narrativity moves up to a change of state, and then goes down again to a process event and two non-events. Next, a stative event is followed by a change of state. Then several stative events and two changes of state conclude the paragraph. This fluctuation in level of narrativity cannot be seen in the storyline in [Figure 2](#), as only the first change of state is shown in the storyline.

7. Discussion

To sum up, our model of narrative events can be applied to fiction as well as non-fiction, such as news, and covers both semantic aspects (event types) as well as rhetorical and narratological aspects (presentation modes) that play a crucial role in how events are perceived by readers. Our goal was to propose a general model for the automatic detection of narrative events, as the overview of related work shows that the lack of consensus on a definition of events in NLP has led to a wide variety of frameworks and applications that are hard to compare and relate to each other, making it difficult to adapt an existing approach for events in news to literary texts. Whereas research on historical events has mainly focused on developing frameworks that enable the application of NLP research and techniques on historical texts, we have focused on developing a broad definition of narrative events that can be used by literary scholars as well as other domains. The current limitation is that we still focus on verbs to select the textual span of an event. We are currently experimenting with using our guidelines for annotations on six more languages (Bahasa Indonesia, Dutch, Italian, Korean, Mandarin Chinese, Spanish), and we will modify the guidelines to be applicable more broadly.

Our comparison between the framework by Vossen et al. (2021) and our model of narrative events shows that the annotation of narrative events can be applied to news and is similar to the rise in action to a climax point, as described in the storyline. On the contrary, Vossen et al. (2021)’s framework has strong limitations when applied to fiction, as the rise in action portrayed in the storyline does not align with the fluctuation in action and level of narrativity seen in fiction.

In the future, it would be interesting to analyze further to what extent our model of narrative events can be applied to various languages and domains. Specifically, we showed that the analysis of narrative events as part of the *syuzhet* can contribute to research on framing in news. This line of research has the potential to show how computational literary studies can make a meaningful contribution to NLP research that goes beyond the semantics of texts.

8. Data Availability

Data can be found here: <https://github.com/GOLEM-lab/event-detection-survey>. It has been archived and is persistently available at: <https://doi.org/10.5281/zenodo.17552902>.

9. Acknowledgements

Funded by the European Union. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

10. Author Contributions

Noa Visser Solissa: Conceptualization, Methodology, Writing – original draft

Andreas van Cranenburgh: Supervision, Writing – review & editing

Federico Pianzola: Supervision, Methodology, Writing – review & editing

References

- Baroni, Raphaël (2007). *La tension narrative: suspense*. Seuil.
- Caracciolo, Marco (2014). *The Experientiality of Narrative*. De Gruyter. [10.1515/9783110365658](https://doi.org/10.1515/9783110365658).
- Caselli, Tommaso and Johan Bos (2023). “Investigating Interoperable Event Corpora: Limitations of Reusability of Resources and Portability of Models”. In: *Language Resources and Evaluation* 57 (3), 1107–1137. [10.1007/s10579-023-09643-6](https://doi.org/10.1007/s10579-023-09643-6).
- Conneau, Alexis, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov (2019). “Unsupervised Cross-lingual Representation Learning at Scale”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Ed. by Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault. Association for Computational Linguistics, 8440–8451. [10.18653/v1/2020.acl-main.747](https://doi.org/10.18653/v1/2020.acl-main.747).

- Cybulska, Agata and Piek Vossen (2014). "Using a Sledgehammer to Crack a Nut? Lexical Diversity and Event Coreference Resolution". In: *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14)*. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. European Language Resources Association (ELRA), 4545–4552. http://www.lrec-conf.org/proceedings/lrec2014/pdf/840_Paper.pdf (visited on 11/11/2025).
- Doddington, George, Alexis Mitchell, Mark A. Przybocki, Lance A. Ramshaw, Stephanie M. Strassel, and Ralph M. Weischedel (2004). "The Automatic Content Extraction (ACE) Program-tasks, Data, and Evaluation". In: *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*. Ed. by Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa, and Raquel Silva. European Language Resources Association (ELRA), 837–840. <https://aclanthology.org/L04-1011/> (visited on 11/11/2025).
- Genette, Gérard (1980). *Narrative Discourse: An Essay in Method*. Cornell University Press.
- Gervás, Pablo and José Luis López Calle (2024). "Representing Complex Relative Chronology across Narrative Levels in Movie Plots". In: *Proceedings of the Text2Story'24 Workshop*. Ed. by Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Marina Litvak, 65–76. <https://ceur-ws.org/Vol-3671/paper6.pdf> (visited on 11/11/2025).
- Gius, Evelyn and Michael Vauth (2022). "Towards an Event Based Plot Model. A Computational Narratology Approach". In: *Journal of Computational Literary Studies* 1 (1). [10.48694/jcls.110](https://doi.org/10.48694/jcls.110).
- Hamborg, Felix (2023). *Revealing Media Bias in News Articles*. Springer International Publishing. [10.1007/978-3-031-17693-7](https://doi.org/10.1007/978-3-031-17693-7).
- Huang, Junbo and Ricardo Usbeck (2024). "Narration as Functions: from Events to Narratives". In: *Proceedings of the 6th Workshop on Narrative Understanding*. Ed. by Yash Kumar Lal, Elizabeth Clark, Mohit Iyyer, Snigdha Chaturvedi, Anneliese Brei, Faeze Brahman, and Khyathi Raghavi Chandu. Association for Computational Linguistics, 1–7. [10.18653/v1/2024.wnu-1.1](https://doi.org/10.18653/v1/2024.wnu-1.1).
- Hühn, Peter (2009). "Event and Eventfulness". In: *Handbook of Narratology*. Ed. by Peter Hühn, Jan Christoph Meister, John Pier, and Wolf Schmid. De Gruyter, 159–178. [10.1515/9783110316469.159](https://doi.org/10.1515/9783110316469.159).
- (2013). "Event and Eventfulness". In: *the living handbook of narratology*. Ed. by Peter Hühn, Jan Christoph Meister, John Pier, and Wolf Schmid. Hamburg University. <https://www-archiv.fdm.uni-hamburg.de/lhn/node/39.html> (visited on 11/11/2025).
- Jahn, Manfred (2021). *Narratology 2.3: A Guide to the Theory of Narrative*. English Department, University of Cologne. <https://www.uni-koeln.de/~ame02/pppn.pdf> (visited on 11/11/2025).
- Kishore, Sindhu and Hangfeng He (2024). "Unveiling Divergent Inductive Biases of LLMs on Temporal Data". In: *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*. Association for Computational Linguistics, 220–228. [10.18653/v1/2024.naacl-short.20](https://doi.org/10.18653/v1/2024.naacl-short.20).
- Lai, Viet Dac (2022). "Event Extraction: A Survey". In: *arXiv preprint*. [10.48550/arXiv.2210.03419](https://arxiv.org/abs/10.48550/arXiv.2210.03419).

- Manjavacas, Enrique and Lauren Fonteyn (2022). “Non-parametric Word Sense Disambiguation for Historical Languages”. In: *Proceedings of the 2nd international workshop on natural language processing for digital humanities*. Ed. by Mika Härmäläinen, Khalid Alnajjar, Niko Partanen, and Jack Rueter. Association for Computational Linguistics, 123–134. [10.18653/v1/2022.nlp4dh-1.16](#).
- Minnema, Gosse, Sara Gemelli, Chiara Zanchi, Tommaso Caselli, and Malvina Nissim (2022a). “Dead or Murdered? Predicting Responsibility Perception in Femicide News Reports”. In: *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Ed. by Yulan He, Heng Ji, Sujian Li, Yang Liu, and Chua-Hui Chang. Association for Computational Linguistics, 1078–1090. [10.18653/v1/2022.aacl-main.79](#).
- (2022b). “SocioFillmore: A Tool for Discovering Perspectives”. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Ed. by Valerio Basile, Zornitsa Kozareva, and Sanja Stajner. Association for Computational Linguistics, 240–250. [10.18653/v1/2022.acl-demo.24](#).
- Miori, Deborah and Constantin Petrov (2024). “Narratives from GPT-derived Networks of News and a Link to Financial Markets Dislocations”. In: *International Journal of Data Science and Analytics* 20, 1105–1129. [10.1007/s41060-024-00516-x](#).
- Norambuena, Brian Keith, Tanushree Mitra, and Chris North (2023). “A Survey on Event-based News Narrative Extraction”. In: *ACM Computing Surveys* 55 (14s), 1–39. [10.1145/3584741](#).
- Pannach, Franziska (2023). “‘Orpheus Came to His End by Being Struck by a Thunderbolt’: Annotating Events in Mythological Sequences”. In: *Proceedings of the 17th Linguistic Annotation Workshop (LAW-XVII)*. Ed. by Jakob Prange and Annemarie Friedrich. Association for Computational Linguistics, 10–18. [10.18653/v1/2023.law-1.2](#).
- Pianzola, Federico (2018). “Looking at Narrative as a Complex System: The Proteus Principle”. In: *Narrating Complexity*. Ed. by Richard Walsh and Susan Stepney. Springer International Publishing, 101–122. [10.1007/978-3-319-64714-2_10](#).
- Pichler, Axel and Nils Reiter (2022). “From Concepts to Texts and Back: Operationalization as a Core Activity of Digital Humanities”. In: *Journal of Cultural Analytics* 7 (4). [10.22148/001c.57195](#).
- Pier, John (2003). “On the Semiotic Parameters of Narrative: A Critique of Story and Discourse”. In: *What Is Narratology? Questions and Answers Regarding the Status of a Theory*. Ed. by Tom Kindt and Hans-Harald Müller. De Gruyter, 73–98. [10.1515/9783110202069.73](#).
- Piper, Andrew and Sunyam Bagga (2022). “Toward a Data-driven Theory of Narrativity”. In: *New Literary History* 54 (1), 879–901. [10.1353/nlh.2022.a898332](#).
- (2024). “Using Large Language Models for Understanding Narrative Discourse”. In: *Proceedings of the 6th Workshop on Narrative Understanding*. Ed. by Yash Kumar Lal, Elizabeth Clark, Mohit Iyyer, Snigdha Chaturvedi, Anneliese Brei, Faeze Brahman, and Khyathi Raghavi Chandu. Association for Computational Linguistics, 37–46. [10.18653/v1/2024.wnu-1.4](#).
- Piper, Andrew and Olivier Toubia (2023). “A Quantitative Study of Non-linearity in Storytelling”. In: *Poetics* 98. [10.1016/j.poetic.2023.101793](#).

- Pustejovsky, James, José Castaño, Robert Ingria, Roser Saurí, Rob Gaizauskas, Andrea Setzer, Graham Katz, and Dragomir Radev (2003). "TimeML: A Specification Language for Temporal and Event Expressions". In: *New Directions in Question Answering, Papers from 2003 AAAI Spring Symposium*. <https://aaai.org/papers/0005-ss03-07-005-timeml-robust-specification-of-event-and-temporal-expressions-in-text/> (visited on 11/17/2025).
- Ruppenhofer, Josef, Michael Ellsworth, Myriam Schwarzer-Petruck, Christopher R Johnson, and Jan Scheffczyk (2010). *FrameNet II: Extended Theory and Practice*. Tech. rep. International Computer Science Institute. https://my.eng.utah.edu/~cs6961/papers/FrameNet_book.pdf (visited on 11/11/2025).
- Santana, Brenda, Ricardo Campos, Evelin Amorim, Alípio Jorge, Purificação Silvano, and Sérgio Nunes (2023). "A Survey on Narrative Extraction from Textual Data". In: *Artificial Intelligence Review* 56 (8), 8393–8435. [10.1007/s10462-022-10338-7](https://doi.org/10.1007/s10462-022-10338-7).
- Scheffel, Michael (2013). "Narrative Constitution". In: *the living handbook of narratology*. Ed. by Peter Hühn, Jan Christoph Meister, John Pier, and Wolf Schmid. Hamburg University. <https://www-archiv.fdm.uni-hamburg.de/lhn/node/57.html#> (visited on 11/11/2025).
- Segers, Roxane, Tommaso Caselli, and Piek Vossen (2017). "The Circumstantial Event Ontology (CEO)". In: *Proceedings of the Events and Stories in the News Workshop*. Ed. by Tommaso Caselli, Ben Miller, Marieke van Erp, Piek Vossen, Martha Palmer, Eduard Hovy, Teruko Mitamura, and David Caswell. Association for Computational Linguistics, 37–41. [10.18653/v1/W17-2706](https://doi.org/10.18653/v1/W17-2706).
- Semino, Elena and Mick Short (2004). "Thought Presentation in the Corpus: A Quantitative and Qualitative Analysis". In: *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. Routledge, 114–153.
- Sims, Matthew, Jong Ho Park, and David Bamman (2019). "Literary Event Detection". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Ed. by Anna Korhonen, David Traum, and Lluís Màrquez. Association for Computational Linguistics, 3623–3634. [10.18653/v1/P19-1353](https://doi.org/10.18653/v1/P19-1353).
- Sprugnoli, Rachele and Sara Tonelli (2017). "One, No One and One Hundred Thousand Events: Defining and Processing Events in an Inter-disciplinary Perspective". In: *Natural Language Engineering* 23 (4), 485–506. [10.1017/S1351324916000292](https://doi.org/10.1017/S1351324916000292).
- (2019). "Novel Event Detection and Classification for Historical Texts". In: *Computational Linguistics* 45 (2), 229–265. [10.1162/coli_a_00347](https://doi.org/10.1162/coli_a_00347).
- Sternberg, Meir (1992). "Telling in Time (II): Chronology, Teleology, Narrativity". In: *Poetics Today* 13 (3), 463–541. [10.2307/1772872](https://doi.org/10.2307/1772872).
- van Cranenburgh, Andreas, Karina van Dalen-Oskam, and Joris van Zundert (2019). "Vector Space Explorations of Literary Language". In: *Language Resources and Evaluation* 53 (4), 625–650. [10.1007/s10579-018-09442-4](https://doi.org/10.1007/s10579-018-09442-4).
- Vauth, Michael and Evelyn Gius (2021). *Richtlinien für die Annotation narratologischer Ereigniskonzepte*. Zenodo. [10.5281/zenodo.5078174](https://doi.org/10.5281/zenodo.5078174).
- Vauth, Michael, Hans Ole Hatzel, Evelyn Gius, and Chris Biemann (2021). "Automated Event Annotation in Literary Texts". In: *CHR 2021: Computational Humanities Research Conference*. Ed. by Maud Ehrmann, Folgert Karsdorp, Melvin Wevers, Tara Lee Andrews, Manuel Burghardt, Mike Kestemont, Enrique Manjavacas, Michael Piotrowski, and Joris van Zundert, 333–345. https://ceur-ws.org/Vol-2989/shor_t_paper18.pdf (visited on 11/11/2025).

- Vendler, Zeno (1957). “Verbs and Times”. In: *The Philosophical Review* 66 (2), 143–160. <https://semantics.uchicago.edu/scalarchange/vendler57.pdf> (visited on 11/11/2025).
- Verkijk, Stella, Pia Sommerauer, and Piek Vossen (2024). “Studying Language Variation Considering the Re-usability of Modern Theories, Tools and Resources for Annotating Explicit and Implicit Events in Centuries Old Text”. In: *Proceedings of the 11th Workshop on NLP for Similar Languages, Varieties, and Dialects (VarDial 2024)*. Ed. by Yves Scherrer, Tommi Jauhiainen, Nikola Ljubešić, Marcos Zampieri, Preslav Nakov, and Jörg Tiedemann. Association for Computational Linguistics, 174–187. [10.18653/v1/2024.vardial-1.15](https://doi.org/10.18653/v1/2024.vardial-1.15).
- Verkijk, Stella and Piek Vossen (2023). “Sunken Ships Shan’t Sail: Ontology Design for Reconstructing Events in the Dutch East India Company Archives”. In: *CHR 2023: Computational Humanities Research Conference*. Ed. by Artjoms Šeļa, Fotis Jannidis, and Iza Romanowska, 320–332. <https://ceur-ws.org/Vol-3558/paper3332.pdf> (visited on 11/11/2025).
- Vossen, Piek, Tommaso Caselli, and Roxane Segers (2021). “A Narratology-based Framework for Storyline Extraction”. In: *Computational Analysis of Storylines: Making Sense of Events* 125, 125–140. [10.1017/9781108854221.008](https://doi.org/10.1017/9781108854221.008).
- Wadhwa, Somin, Oktie Hassanzadeh, Debarun Bhattacharjya, Ken Barker, and Jian Ni (2024). “Distilling Event Sequence Knowledge from Large Language Models”. In: *The Semantic Web – ISWC 2024*. Ed. by Gianluca Demartini, Katja Hose, Maribel Acosta, Matteo Palmonari, Gong Cheng, Hala Skaf-Molli, Nicolas Ferranti, Daniel Hernández, and Aidan Hogan. Springer Nature, 237–255. [10.1007/978-3-031-77844-5_13](https://doi.org/10.1007/978-3-031-77844-5_13).
- Xin, Wendy Veronica (2022). “Plot”. In: *Fictionality and Literature: Core Concepts Revisited*. Ed. by Lasse R. Gammelgaard, Stefan Iversen, Louise Brix Jacobsen, James Phelan, Richard Walsh, Henrik Zetterberg-Nielsen, and Simona Zetterberg-Nielsen. Ohio State University Press, 63–81.
- Yan, Zhihua and Xijin Tang (2023). “Narrative Graph: Telling Evolving Stories Based on Event-centric Temporal Knowledge Graph”. In: *Journal of Systems Science and Systems Engineering* 32 (2), 206–221. [10.1007/s11518-023-5561-0](https://doi.org/10.1007/s11518-023-5561-0).
- Zgoll, Christian (2020). “Myths as Polymorphous and Polystratic Erzählstoffe”. In: *Mythische Sphärenwechsel: Methodisch Neue Zugänge zu antiken Mythen in Orient und Okzident*. De Gruyter, 9–82. [10.1515/9783110652543-002](https://doi.org/10.1515/9783110652543-002).