ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **19**, Numbers 1-2, 2016, 31–43

Why Time Resembles Rail Yards? A Way to Look at Time in Text Books

Dan CRISTEA, Andreea MACOVEI

"Alexandru Ioan Cuza" University of Iaşi, Faculty of Computer Science
E-mail: {dcristea, andreea.gagea}@info.uaic.ro

Abstract. We present in this paper a proposal for representing temporality phenomena as expressed by specific free writing techniques, such as temporal discontinuities, transitions from one story to another, flashbacks and change of perspective, phenomena naturally understood by a casual human reader of the text. The long time intention is to develop a technique capable to reproduce on a machine the human reader performances with a sufficiently high accuracy. The diagrams we draw, called time yards, and the associated XML notations include: time tracks, themselves being made out of time segments, the atoms of time segments being the events. The paper makes a preliminary evaluation of the objectivity of recognizing these concepts as catched by a group of annotators.

1. Introduction

Recent approaches related to the interpretation of time in texts address the recognition of temporal expressions, of events and their relative ordering, of timelines and storylines. Time annotation, being it manually done, like in "golden" corpora, or automatically, obtained by learning from trained systems, display a kind of lattice structures, where any two events can either be ordered or not. In the best case, a total ordering of events is obtained (including intersection of intervals and inclusion of time moments into intervals), otherwise only a partial one, or even, in the worst case, no order at all.

Existent systems usually are focussed on texts belonging to the publicistic style (see, for instance the recent workshop on Computing News Storylines - CNewsStory[1]), news being the main type of article, where stories very dense in

[1]The First Workshop on Computing News Storylines, Bejing, China, July 2015.

time data and events rather well ordered are described. The narrator in these texts is either impersonal, perhaps the voice of some news agency, or a notorious chronicle person. Stylistically far from this type of text, belle-letters are quite different in perceiving time. In novels, very often, an ongoing story is interrupted by a flashback, the perspective from which the events are told changes, there are discontinuities in narrations, disparate stories develop without any apparent link between them, followed by surprising encounters of characters and merging of destinies, and the other way round is also possible, characters that have followed for a while common or parallel lives split and their stories continue separately. These complex mise-en-scènes cannot be represented by the simple lattice-like representations mentioned above, neither as linear timelines. An evident reason for that is that comparing and ordering events belonging to separate substories shuffles events instead of putting them in order. It seems therefore natural to consider that, very often, novels include more time axis, one for each distinct theme or fabula development.

1.1. Goals of the paper

Our main aim is to find a computational representation of temporal phenomena occurring in literary texts, there where readers perceive changes of the current direction of time, flashbacks and flashforwards, temporal ruptures, switches between different stories or of narrators and others, preparing thus investigation tools that would make possible to tackle these phenomena by automatic analysis.

In (Macovei and Cristea, 2016) we have used the term *time frames* for temporal axis on which sequences of anchored or unanchored events, bordered in time, making stories by themselves, may occur. We have also proposed a classification of time frames (into narrative, remembering, supposition, general knowledge, and fiction) and looked for clues that would indicate transitions between them, in view of preparing the training of a learning process aimed to segment the text and classify the text spans according to these types.

In this study, we try to go a bit further than in that initial proposal, enhancing the model presented there on several dimensions. First, we explain that time can be better visualised as threads that interlace and separate, much like railways do in complex rail yards. We will call these threads *time tracks*. Secondly, we propose an XML notation for these representations and show a few examples of commute, rupture, join and split points, collected from a small corpus. Third, we make a preliminary evaluation of the annotations in that small corpus and we formulate conclusions.

1.2. Related work and motivation

Although terms such as timeline and storyline have been used in historical studies, dramaturgy and narratology for quite a while[2], only recently they

[2]Marriam Webster's Collegiate Dictionary reports 1951 the year of the first recorded use

received attention in the field of natural language processing, mainly in the context of identifying connected sequences of events. In these researches, *timeline* refers to a representation of events which are chronologically ordered, in general specific for an entity (a character or participant in an action, a geographical place or region)[3] (Chambers and Jurafsky, 2008), while *storyline* has slightly different interpretations for different researchers. For instance, in Laparra *et al.* (2015) *StoryLines* refers to groups of interacting *TimeLines*, or mergers of two or more TimeLines where the same characters or entities are taking part in the action. In this research, more TimeLines may be part of the same StoryLine. So timelines may be said to join, but never to split. This is because storylines in these authors' view describe a kind of text structure which is not related to the flow of time. Events are ordered and anchored in timelines, but the overall relation of the storyline with the time axis seems to be lost.

In Vossen *et al.*'s (2015) computational model, the basic elements of a storyline are: the events (defined by actors, locations and time settings), the anchorings of events to time, a basic timeline (or fabula[4]), as a set of events ordered for time and a set of bridging relations, as a set of relations between events with explanatory and predictive value(s) (rising action, climax and falling action). Events are uniquely represented (as instances) through the Simple Event Model (SEM) (Van Hage *et al.*, 2011), in RDF, and each event can be mentioned more than once in the text, mentions being noted following the conventions of Grounded Annotation Framework (GAF) (Fokkens *et al.*, 2014). Out of events, timelines are computed, as sequences of event instances anchored to time expressions or relative to each other. Sets of events, their corresponding time points and relations that order them are the basic ingredients of a storyline, as the timeline made of those events which have the highest connectivity score. Vossen *et al.*'s interests is to detect storylines in massive streams of news and less to decipher the temporal structure of texts. Thus storylines may perceive as adjacent even events which are distantly placed in time if they happen to correlate well. The term *story* is used to refer the textual expression of such an abstract structure.

Timeline extraction envisages tasks such as event classification and detection, coreference resolution of events, extraction of temporal relations that appear between events, event factuality (Sauri and Pustejovsky, 2008), etc. The aim of the Cross-Document Event Ordering task carried out in SemEval 2015[5] (Minard *et al.*, 2015) was to build TimeLines from English news articles, which differ from other narrative structures, as those described in (Chambers and Jurafsky, 2008;

---

of *timeline* (written *time line*: table listing important events for successive years within a particular historical period), and 1941 for *storyline* (actually *story line*: the plot of a story or drama).

[3]Or *sjuzet* (Brunner, 1986) - a term used in narratology.

[4]The timeless, motionless, underlying theme, the backbone of the story, the "virtual text" (Brunner, 1986). Following the narratology framework of Bal (1997), every story is a mention of a fabula, *i.e.*, a sequence of chronologically ordered and logically connected events involving one or more actors.

[5]http://alt.qcri.org/semeval2015/ task4/.

Chambers and Jurafsky, 2009).

Other systems try to extract storylines or to reconstruct maps of connections that explicitly capture story development (Shahaf *et al.*, 2013) or complex representation (Hu *et al.*, 2014).

Our analysis follows closely previous work, but in the same time we try to formalise a number of textual phenomena which seem not having received much attention before. Very often, a text may contain more fabulae, intertwined, the embedded ones used to hint on the deciphering of the main fabula. Also, two stories that had distinct developments, each focussed on different actors, may join at a certain moment, when the actors meet. The other way round is also possible, when a main fabula splits in two or more others. Temporal ruptures, flashbacks, flash forwards, temporal transitions, etc. are frequent linguistic phenomena that interrupt the ordinary flow of a story. Moreover, unlike in most approaches, our interest is focussed more on free texts than on streams of news.

2. The *time yard* model

We sketch in this section a computational representation of our model, having also a visual correspondent that we call *time yards*.

2.1. The model and visual representations

As said already, our *time yards* are made out of *time tracks*, and time tracks - of *time segments*. Time tracks can be bordered by: *start points* - where they start, *end points* - where they finish, *join points* - in which two different time tracks unify, and *split points* where one time track splits into two separate time tracks. Segments, if inner to time tracks, are bordered by *rupture points*, which are points in the text where the story commutes from one time track onto a different one, or by *commute points*, in which, although the story goes no, it is said from a different perspective, by a different storyteller. If placed at the extreme of tracks, segments can also be bordered on the left side by start or split points and on the right side by join or stop points.

Figure 1 shows an eye bird time yard of a modern novel. Above, in Fig. 1a, the sequence of time segments is displayed, ordered as in the text (the horizontal axis aligns with text offsets). Segments lengths are not related to any fix scale.

Figure 1b shows the resemblance of the time tracks with a rail yard. From this - the time yard name of our model. Segments are placed here on their corresponding tracks, in the order given by the story-time, considered to be the horizontal axis of the diagram. There is no rigorous scale in this representation either, lengths of segments having no immediate correspondence with the time of the story. However, parallel time tracks suggest their simultaneity in the story.

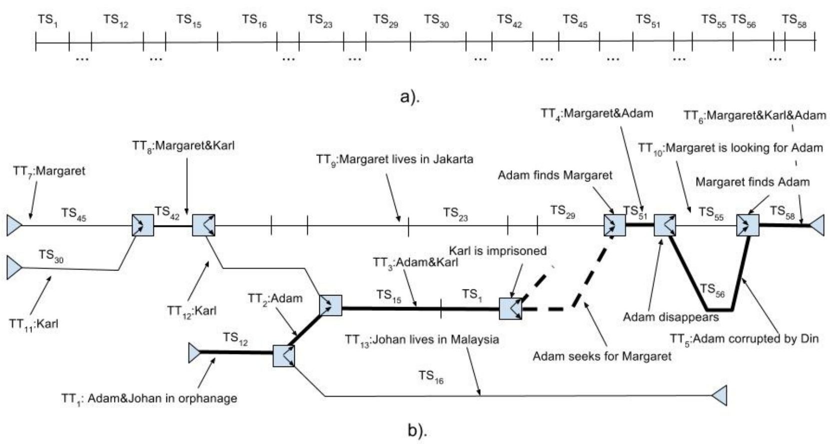

Fig. 1. a) A high granularity chain of time segments in Tash Aw's novel *Map of the Invisible World (MotIW)*. The horizontal axis is the text offset. b) The corresponding time yard: time tracks and segments, with start, stop, rupture, commute, join and split points. The horizontal axis represents the story-time.

Time segments are indexed in Fig. 1 as TS_i, in the order of their occurrence in text, and time tracks (only in Fig. 1b) - as TT_i, in no specific order. Start, stop, join and split points can be easily recognised in the diagram, because they are represented by suggestive symbols, while rupture and commute points are indicated by short vertical lines. For instance, the right border of segment TS_{55}, belonging to TT_{10}, is simultaneously a rupture and a join point. It is a rupture point because the story continues with TS_{56}, which belongs to another time track, TT_5, and it is a join point because it is there where TT_{10} and TT_5 join. In a join point, characters that had separate lives or developments meet and continue from there a common story and in a split point, characters that were involved in a common development separate, each following the story on a separate track.

In bold in the figure we represented Adam's sequence of time-tracks, the main character in the novel. As can be seen, on certain intervals his story coincides with that of other characters: his brother Johan, then Karl, then, after a while, Margaret, a.s.o. Since the horizontal axis in these time yards represent the story-time, it is not compulsory that a text includes a complete description of an interval. As such, segments may not fill contiguously a certain time track. This means that the story may include discontinuities (pauses, misses). In Fig. 1b this is suggested by the continuous lines (segments) and interrupted lines (discontinuities).

The reader is capable to mentally order, following the story-time, the shuffled segments sequently presented in the text (Fig. 1a), as coherent developments narrated by one or more relators, thus reconstructing the time tracks (Fig. 1b).

The segments of a time track are coherently linked together, displaying the involvement of the same character or group of characters. A time track should not necessarily be characterised by a unity of place, but it must be deciphered by the reader as having a unity of action, of intention and development. Often, time tracks narrate long periods of characters' lives.

It should be clear by now that the way we look at texts is at a higher granularity than that in which the events are noted. Our model describes the reader's performance in deciphering temporality in texts, actually in matching text spans with abstract timelines implicitly contained in the story, sometimes from different perspectives, either the narrator's or of characters involved in the story. Thus, it gives a view of the temporal organisation of the text; it is not intended to abstract away from the text level, by extracting a representation matching components of the fabula, as in (Vossen *et al.*, 2015), for instance.

Novels very often include personal interpretations of sequences of events involving characters and these interpretations themselves take story-time. While storying, the author is seen on a timeless track, distant from the time of the plot, which has no significance within the act of interpretation and should be disregarded. But this is not the same when the interpreter is actively acting in the plot. In this case there are two distinct time tracks that the reader perceives: the one of the character telling the story and the one of the story told (interpreted) by this character. In the time yard of Fig. 1b each story-time segment has exactly one interpretation in the text and the interpreter is not shown.

2.2. Examples[6]

The example below shows commute points:

TS_1:[*In the kitchen, waiting for the water to boil, Margaret wondered how was it that*] TS_2:[*she jumped so naturally to embrace Adam.*] TS_3:[*She could not realize what had made her*] TS_4:[*feel so happy to see him again…*]

Here there are 4 segments, all belonging to a time track which could be called Margaret&Adam&Mick separated by 3 commute points. Indeed, TS_2 represents a remember of Margaret, therefore it goes back in time with respect to TS_1 (the *jumping to embrace* event happens before the *water boiling* and *wondering events*). Then, TS_3 follows in time TS_1 (*realizing* is in the same sequence of events as *wondering*). Finally, TS_4 brings the reader back in the same sequence of events as TS_2 (*feeling happy* is simultaneous with *jumping to embrace*).

Following is an example of a rupture point:

TS_1:[*She remembered the words*] TS_2:[*that sometimes, at the time of adolescence, her mother told her. "You are a person apprehended, right?"…*]

[6] The examples are back translations into English (done by the authors) from the Romanian version of *MotIW*.

TS_1 tells about a *remembering event* (placed in the same Margaret&Adam &Mick's time track as the previous example - although not evident here). Then, in TS2, the story has a rupture, bringing the reader back in the remembered times of Margaret&her mother's time track.

The next example displays a split point:

TS_1:[*Up to climb into the truck and disappear after the tarpaulin fall down,*] TS_2:[*he followed them slowly, with his clumsy gait.*]

This is the moment when Adam loses the track of his father, imprisoned by soldiers. Actually TS_2 happens before TS_1, it being the last segment of a track describing Adam&Karl's life together. TS_1 marks the split point, because from now on Adam will be alone for a while and the readers will meet again Karl only to the end of the novel.

Finally we have an example of a join point:

TS_1:[*It was only when she came to the doorway that she realized someone was there, lying on the front steps.*] TS_2:[*He was a boy, a teenager, almost crouched in the womb.*]

Indeed, TS_1 belongs to a time track telling the recent life of Margaret in Jakarta, while in TS_2 she meets Adam. The reader knows about Adam after his separation from his father. So, TS_2 is the first segment of a new time track, which can be called Margaret&Adam.

3. Notations and experimental data

We present in this section a proposal for a set of XML notations capable to describe the features of the model introduced in section 2.

3.1. An XML notation

For reasons of corpus annotation and training, in our experiments we have used the following notations for XML elements and their attributes:

<TT ID NAME LEFT-EP RIGHT-EP/> elements mark **time tracks**. A time track is uniquely identified by an ID. By using the NAME attribute (optional) the annotator may attribute a mnemonic to the TT. The TIME attribute (optional) expresses a precise anchoring in the storytime or a time relationship with respect to another TT. The values of the TIME attribute are IDs of TIME elements (see below). The LEFT-EP attribute describes the left endpoint of the TT. Its values can be: "START" or the ID of an ENDPOINT element of type SPLIT. The RIGHT-EP attribute describes the right endpoint of the TT. Its values can be: "STOP" or the ID of an ENDPOINT element of type JOIN. TT elements are all placed in a TT-SECTION (see below), separate from the body of text.

<TS ID IN-TT NAME TYPE PER>...</TS> mark **time segments**. In an in-line representation, the span of text of a TS element is surrounded by a pair of <TS>...</TS> markings. Same as with TTs, each TS is uniquely identified by an ID. There may be different TSs belonging to the same TT when the text unfolds. This is noted by their identical IN-TT attribute. By using the NAME attribute (optional) the annotator may attribute a mnemonic to the TS. Same as in TTs, values of the TIME attribute are IDs of TIME elements. TSs are characterised by TYPE and PERspective. The classification of types is that from (Macovei and Cristea, 2016), *i.e.*: NAR - for typical narrations in which the story time flows constantly ahead, REM - for flashbacks (remembers) belonging to a character; SUP - for suppositions or speculations (where the time is vaguely attached to a plausible, imaged or supposed, world); GEN - for general knowledge (where there is no time anchor, only statements about generally accepted things); and FIC - for fictions, invented realities, like in movies, plays or novels (most frequent of a NAR type, but whose time limits have no connection with respect to the current story time). Finally, the PER attribute notes to whom the perspective on the current development belongs, a character or the writer.

<TIME ID FROM REL TO/> elements mark **time constraints**, explicitly expressed in text or inferred. Apart from ID, uniquely marking TIME elements, in order to be referred to in TTs and TSs, FROM and TO are IDs of TTs or TSs and REL is a temporal relation. Very shallow RELations between time segments and tracks have been seized by now as necessary, such as BEFORE, IMMEDIATELY-BEFORE, AFTER, IMMEDIATELY-AFTER, SIMULTANEOUS. TIME elements are all placed in a TIME-SECTION (see below).

<ENDPOINT ID TYPE NAME TT1 TT2/> elements mark **endpoints of TTs**. Apart from a unique ID, the TYPE can be “JOIN” or “SPLIT” and TT1 and TT2 are IDs of the two TTs that either join or split in that endpoint. Same as in TTs and TSs, the NAME attribute (optional) can be used by the annotator to assign a mnemonic to the endpoint. ENDPOINT elements are all placed in a EP-SECTION (see below).

Pairs of <TT-SECTION>...</TT-SECTION>, <TIME-SECTION>...</TIME-SECTION> and <EP-SECTION>...</EP-SECTION> elements mark the sections in the XML document where TT elements, TIME elements and EP elements are grouped.

3.2. A small corpus

We proceeded by annotating 2 fragments, of 3 pages each, of the Romanian version of the novel MAP OF THE INVISIBLE WORLD by Tash Aw[7], giving them to a class of 6 and respectively 8 master students in Computational Linguistics. The development period of the time yard model coincided with their

[7]See Acknowledgements.

second term, and they have been informed on the successive revisions. The annotation conventions were practically established interactively, in the class. Although we do not consider them yet as experts in this type of annotation, we were interested to understand whether readers have more or less common perception over time tracks and time segments, even if they may be vaguely signaled or not at all. Students were instructed to delimit (for the time being, just with pencil on paper) the time segments and to note the time tracks the segments belong to. The exercise included only rupture and commute points, which have a much higher frequency than join and split points.

3.3. Evaluation and Discussion

At the moment of writing this paper we have done only a quick inspection and comparison of manual annotations performed by a class of students on a small corpus.

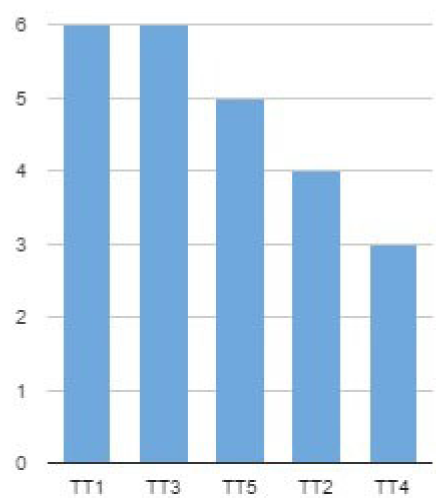

Fig. 2. Recognition of TTs in text 1.
TT1=Adam&Karl now, TT2=Adam now, TT3=Adam&Karl&
a Dutch couple in the past, TT4=Adam in the past, TT5=Margaret now.

The degree of recognition of TTs is described by the histograms in Figures 2 and 3, where on the horizontal axis TTs are placed in the descending order of their degree of recognition by subjects. It can be noticed that in both texts two TTs out of five (TT1 and TT3 in the first text and TT1 and TT2 in the second), *i.e.* 40% of all annotated TTs, have been recognised by all annotators (6 and respectively 8). The following two tracks were recognised by a majority of annotators (75% for the first text and 69% for the second text). The existence of TTs that have been recognised by all or a majority of annotators shows that the notion of time track is an objective, noticeable one. For the tracks on which there is no majority agreement (TT4 in the first text and TT3 in the second) two things can be said: either some annotators perceived tracks there were they are not, taking segments for tracks, or some annotators did not perceive tracks there were they exist, taking tracks for segments included in already signaled tracks. A detailed analysis shows that in both cases the cause of the errors is a wrong understanding of the difference between a segment and a track. For instance, in text 2, a scene including Margaret, Adam and Mick is storied, the author disrupting its sequentiality by a number of commute points that border segments of the same track, but 2 students confused a flashback on the same TT (actually a TS) with a distinct TT. On the contrary, in text 2, three students did not recognise TT4 (Margaret in the past) as a distinct track, instead choosing to assign its TSs to TT1.

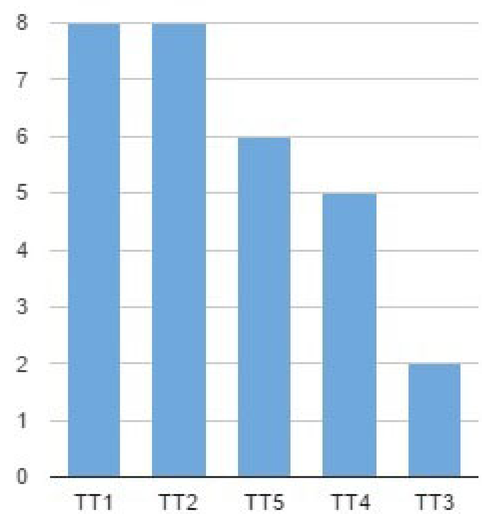

Fig. 3. Recognition of TTs in text 2.
TT1=Margaret&Adam&Mick now, TT2=Margaret's mother in the past, TT3=Adam&Mick now, TT4=Margaret in the past, TT5=Adam&Karl in the past.

4. Conclusions

Our representations resemble ordered graphs, in which nodes are start, stop, commute, rupture, join and split points, edges are time segments and paths are time tracks. Using these graphs, different textual phenomena that involve time can be described: temporal sequences, pauses (a pause is a temporarily stopped fabula by a description or an introspection), time gaps or misses (ellipsis in the narration of the fabula), parallelisms (lack of temporal anchors or explicit placement of events in synchronicity, both leaving space for undetermined ordering, for multiplicity of time axis in the development of the fabula).

The time yard model can be exploited in different ways. First, the realignment of segments in tracks may stay at the origin of a re-storying of the text in the *chronological order* of its developments[8]. As such, the timely shuffled segments of the text, dictated by the authors preference for a certain style, for making more salient certain aspects than others, or simply dictated by the need for digression, can be re-mapped on an imaginative time axis, out of which a summary can be obtained.

Second, a *character's destiny* can be followed on different time tracks, all that include her/his participation, and in the fictitious chronological order. All tracks in which the same character is involved should be in some way sequenced, provided the writer obeyed to the law of nature according to which an entity can be located in exactly one place at a certain moment. Considered for the whole set of characters in the novel, an intermixed set of destinies can be identified, which, like threads/fibers in a tissue, tangle and disentangle, thus making up the whole canvas of the novel.

Third, when deciphered by an artificial agent, the sequentiality of events supposed to characterise each character's destiny, makes up as many sets of constraints as characters, which can be used to check the accuracy of the temporal relationships acquired by the system.

Fourth, from a theoretical perspective, let's notice that theories of discourse, such as Attention State Theory (Grosz and Sidner, 1986), Centering (Grosz *et al.*, 1995), Rhetorical Structure Theory (Mann and Thompson, 1988), all place at their bases the notion of segment, a rather ambiguous one. On the other hand, Veins Theory (Cristea *et al.*, 1998) demonstrates that a generalisation of segment can be obtained if veins are considered. VT claims that veins can be computed beyond segments' borders, this actually making one of its strong points (for instance, the generalisation of Centering is based on this assumption). However, it is natural to think that over a too large zone of a novel, VT cannot be applied anymore, or otherwise artificial referentiality and coherence links can be deduced. The textual domain on which veins can still be computed has never been investigated. Related to this, it is interesting to note that referential links can cross rupture points, provided the corresponding tracks include common entities (persons or things). As such, a cross-investigation of relationships between the two theories would be worth doing.

Our research is at an incipient stage, but we believe the proposed model has a value that can be largely exploited. The immediate step should be a thorough evaluation of inter-annotator agreement. If results, as promised by the preliminary investigation described above, will prove to be sufficiently high, then we may proceed at the construction of a larger corpus, on which to notate also other distinguishing features of time tracks and their constituent segments that might have been missed in this study.

---

[8] As the reader may have noticed, we avoid to say "the order of the events", because the events are only the constituents, the ingredients of the segments and tracks, they are at a too fine granularity for being perceived by our model.

Acknowledgements. Extracts from MAP OF THE INVISIBLE WORLD by Tash Aw. Copyright © Tash Aw, was used by permission of The Wylie Agency (UK) Limited. The Humanitas Fiction Publishing House, Bucharest, has kindly accepted to offer us for research purposes the Romanian version of this novel, translated by Florica Sincu. We are indebted to the class of first year master students in Computational Linguistics from the Computer Science Faculty of "Alexandru Ioan Cuza" University of Iaşi, who, in the Spring term of 2016, have pursued annotations on a few fragments of the Romanian version of this novel.

References

[1] Bal, M. (1997). Narratology: Introduction to the theory of narrative. Trans. Christine van Boheemen. 2nd ed. Toronto: U of Toronto.

[2] Bruner, J. (1986). Actual Minds, Possible Worlds. Cambridge, MASS., Harvard University Press.

[3] Chambers, N. and Jurafsky, D. (2008). Unsupervised Learning of Narrative Event Chains. In ACL (Vol. 94305, pp. 789-797), June.

[4] Chambers, N. and Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 (pp. 602-610). Association for Computational Linguistics, August.

[5] Cristea, D., Ide, N. and Romary, L. (1998). Veins theory: A model of global discourse cohesion and coherence. In Proceedings of the 17th international conference on Computational Linguistics, volume 1 (pp. 281-285). Association for Computational Linguistics.

[6] Fokkens, A., Van Erp, M., Vossen, P., Tonelli, S., Van Hage, W. R., Serafini, L., Sprugnoli, R. and Hoeksema, J. (2013). GAF: A grounded annotation framework for events. In NAACL HLT (Vol. 2013, p. 11), June.

[7] Laparra, E., Aldabe, I. and Rigau, G. (2015). From TimeLines to StoryLines: A preliminary proposal for evaluating narratives. ACL-IJCNLP 2015, 50.

[8] Macovei, A. and Cristea, D. (2016, to appear): Time Frames: Rethinking the Way to Look at Texts, Proceedings of ISA-12, the Twelfth Workshop on Interoperable Semantic Annotation in conjunction with LREC 2016, Portoroz, Slovenia, 28 May 2016.

[9] Minard, A. L., Speranza, M., Agirre, E., Aldabe, I., van Erp, M., Magnini, B., Rigau, G. and Urizar, R. (2015). Semeval-2015 task 4: Timeline: Cross-document event ordering. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015) (pp. 778-786), June.

[10] Grosz, B. J., and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. Computational linguistics, 12(3), 175-204.

[11] Grosz, B. J., Weinstein, S., and Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. Computational linguistics, 21(2), 203-225.

[12] Van Hage, W. R., Malais, V., Segers, R., Hollink, L. and Schreiber, G. (2011). Design and use of the Simple Event Model (SEM). Web Semantics: Science, Services and Agents on the World Wide Web, 9(2), 128-136.

[13] Hu, P., Huang, M. L. and Zhu, X. Y. (2014). Exploring the interactions of storylines from informative news events. Journal of Computer Science and Technology, 29(3), 502-518.

[14] Mann, W. C. and Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. Text-Interdisciplinary Journal for the Study of Discourse, 8(3), 243-281.

[15] Saur, Roser and James Pustejovsky. 2008. From Structure to Interpretation: A Double-layered Annotation for Event Factuality. Proceedings of the 2nd Linguistic Annotation Workshop. The Sixth International Conference on Language Resources and Evaluation, LREC 2008:1-8.

[16] Shahaf, D., Yang, J., Suen, C., Jacobs, J., Wang, H. and Leskovec, J. (2013). Information cartography: creating zoomable, large-scale maps of information. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1097-1105). ACM, August.

[17] Vossen, P., Caselli, T. and Kontzopoulou, Y. (2015). Storylines for structuring massive streams of news. ACL-IJCNLP 2015, 40.