TOWARDS AN EVENT ANNOTATED CORPUS OF POLISH

Abstract

The paper presents a typology of events built on the basis of TimeML specification adapted to Polish language. Some changes were introduced to the definition of the event categories and a motivation for event categorization was formulated. The event annotation task is presented on two levels — ontology level (language independent) and text mentions (language dependant). The various types of event mentions in Polish text are discussed. A procedure for annotation of event mentions in Polish texts is presented and evaluated. In the evaluation a randomly selected set of documents from the Corpus of Wrocław University of Technology (called KPWr) was annotated by two linguists and the annotator agreement was calculated. The evaluation was done in two iterations. After the first evaluation we revised and improved the annotation procedure. The second evaluation showed a significant improvement of the agreement between annotators. The current work was focused on annotation and categorisation of event mentions in text. The future work will be focused on description of event with a set of attributes, arguments and relations.

Keywords: information extraction; event recognition; corpus annotation

1. Introduction

Event recognition is a subtask of information extraction task. The goal of information extraction is to understand the meaning of a text at some level on which one can catch given type of information and present it in a structured manner. Event recognition focuses on finding in text references to some situations and extracting their descriptions. Event recognition has practical applications in many tasks from the field of natural language processing, like text summarization (Maybury, 1995), discourse analysis, events aggregation and reporting (Vossen et al., 2014; van Erp, Fokkens, & Vossen, 2014; Agerri et al., 2014). Within the Clarin-PL project\(^1\) we

\(^1\)Project web page: http://clarin-pl.eu/.
plan to develop methods and tools for event recognition for Polish. We want to identify event mentions in Polish texts, categorise them on a coarse-grained level and identify event attributes, arguments and relations in order to enable deeper text understanding. In order to create and evaluate such tools we need a practical guideline for event annotation dedicated to Polish and a corpora annotated with events. According to our best knowledge, the most popular and widely used specification for event annotation is TimeML Annotation Guidelines Version 1.2.1 (Saurí, Littman, Knippen, Gaizauskas, Setzer, & Pustejovsky, 2006) (henceforth, TimeML). The specification has been already adopted to several languages, including Spanish (Saurí, Batiukova, & Pustejovsky, n.d.), Catalan (Saurí, Batiukova, & Pustejovsky, n.d.), French (Bittar, 2010) and Italian (Caselli, Bartalesi Lenzi, Sprugnoli, Pianta, & Prodanof, 2011). Applying an existing guideline for another language requires a careful study of that language phenomena and might need some adjustments concerning language-specific issues. In the following sections we present results of our work on adaptation the TimeML specification to Polish language and evaluation of the specification on Polish texts. In Section 2 we present a definition of event concept and what we understand as an event on the ontology level. Section 3 contains a typology of event categories and motivation for event categorization. In Section 4 we define the event mentions for Polish as a text-level representation of events. In Section 5 we present results of two evaluations of the guidelines for Polish performed on the KPWr corpora (Broda, Marcińczuk, Maziarz, Radziszewski, & Wardyński, 2012). In Section 6 we present a detailed procedure for event mention annotation which was created as a result of first evaluation. Section 7 presents a summary of our current work and future plans including event description with attributes, arguments and relations.

2. What is an event?

Event is one of the primary concepts in almost any upper-level ontology. According to the Oxford Dictionary event "is a thing that happens or takes place".\(^2\) In the Suggested Upper Merged Ontology ontology (Pease, 2011) (henceforth, SUMO) an event is represented as a concept called Process,\(^3\) which is defined as following:

> The class of things that happen and have temporal parts or stages. Examples include extended events like a football match or a race, actions like pursuing and reading, and biological processes. The formal definition is: anything that occurs in time but is not an object. Note that a process may have participants ‘inside’ it which are objects, such as the players in a football match”

In other words event is anything that takes place in time (date, time and/or duration) and space (has a location), may involve agents (executor or participants), may contain or be part of other events and may produce some outcome (object). In our work we will consider as event all situations which can be mapped onto the Process concept or any concept which is a subclass of Process in the SUMO

\(^2\)http://www.oxforddictionaries.com/definition/english/event
\(^3\)http://sigma-01.cim3.net:8080/sigma/Browse.jsp?lang=EnglishLanguage&flang=SUD-KIF&kb=SUMO&term=Process
Towards an Event Annotated Corpus of Polish

ontology. The relations of event with other concepts from the SUMO ontology are illustrated on the Figure 1.

Figure 1: Event relations with other concepts from SUMO ontology.

The states are also treaded as events (Saurí et al., 2006) but they have a specific ontological status. It isn’t simple to map the words denoting a state to Process in the SUMO ontology. As Vendler said suggestively, states are “that puzzling category in which the role of verb melts into that of predicate, and actions fade into qualities and relations” (Vendler, 1957, p. 109). This feature brought us to individual treatment of the mentions of the states.

3. Event Categories

We used seven coarse-grained categories of events, i.e. action, state, reporting, perception, aspectual, intensional action and intensional state. The categorisation was based on the TimeML guideline with some modifications. Instead of the occurrence term we used action. The occurrence category from TimeML refers only to specific temporarily located events. Generics — actions which refer to some general rules (for example, a boil event in sentence “Water boils in 100 °C”) are not tagged. We noticed that the distinction between specific and generic events can be applied to any category of events what indicates that the event generality should be defined as an event attribute rather than its category. Taking into account Polish terminological tradition (e.g. Laskowski, 1998), we’ve decided to use the term action instead occurrence,4 to accent its generality and to make visible the key opposition between the two core categories: action and state. In addition this change emphasizes the distinction between the state/action and intensional state/action. The remaining categories can be treated as auxiliary categories, as they refer to another events and introduce some additional information about the event.

4Although they are sometimes treated as synonyms (e.g. Mourelatos, 1978).
TimeML specification doesn’t introduce the higher level classification. Still the categories of events can be divided into four groups in respect to two factors: *dynamicity (course in time)* and *event argument* (see Table 1). The *course in time* factor divides events into static and dynamic events. The static events endure or persists over some period of time and though they may provide the potential change, they do not constitute a change (Mourelatos, 1978, p. 192). The *event argument* factor indicates if the event have (or might have) an argument that is an event. For example the *start* event indicates the beginning of some other event.

<table>
<thead>
<tr>
<th>Without an event argument</th>
<th>With an event argument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td>state</td>
</tr>
<tr>
<td>Dynamic</td>
<td>action</td>
</tr>
<tr>
<td></td>
<td>perception</td>
</tr>
<tr>
<td></td>
<td>reporting</td>
</tr>
<tr>
<td></td>
<td>aspectual</td>
</tr>
<tr>
<td></td>
<td>intensional action</td>
</tr>
</tbody>
</table>

In addition we’ve noticed that actions or states which connect with an event argument could be divided into two groups. Occurring some of the events in the text (i.e., *perception*, *reporting* and *aspectual*) signals that an event which is an argument occurred (or should have occurred) in the real world and occurring the other (i.e. *intensional state* and *intensional action*) doesn’t gave such certainty. It wasn’t the statement of TimeML specification authors but we treat it as the important remark for the future processing of extracted events.

Furthermore, we’ve decided to introduce a separate category for synsemantic verbs that occur with nominalizations (*light predicates*). Since they have specific grammatical function, they are described in the section on event mentions.

Figure 2 shows the final classification of events.

### 3.1. Action

*Action* represents a dynamic situation which occurs in time and space. The event could have some type of outcome that can be a product, achievement or change from one state to another.

Examples: build, dance, jump, running

### 3.2. State

*State* represents a static situation. It refers to object attributes (Apresjan, 2000, pp. 47–48) or situations which are stable and does not change over given period of time (Laskowski, 1998, p. 153).

Examples: sleep, stand

### 3.3. Reporting

*Reporting* refers to a dynamic situation where an agent inform about an event or narrate an event. If the *reporting* refers to an *action* or a *state* then it is a strong indication that the *action* or *state* took place or was true.

Examples: report, explain, relation, telling
3.4. Perception

Perception refers to a physical perception of an event by an agent. This class indicates that the agent was an observer of the event. The perception event is a strong indicator that the observed event took place or was true.

Examples: see, hear, observation, hearing

3.5. Aspectual

Aspectual refers to a dynamic situation which indicates a change of a phase of another event. The change can be (following TimeML):

1. Initiation — an event was started,
2. Reinitiation — an event was stopped and started again,
3. Termination — an event was stopped before it was completed,
4. Culmination — an event was completed,
5. Continuation — an event is continued.

If the aspectual event refers to an action or a state and it is not referred by any intensional action or state, then it is a strong indication that the action or state was true for some period of time.

Examples: start, stop, continuation, interruption

3.6. Intensional action

Intensional action is a situation where an agent declare his or her will to perform an action or give a command to another agent to perform an action. We cannot infer if the action was or will be performed in the future. The possible groups of intensional actions (following TimeML) are:

1. Attempt — the agent tried to do X but failed to accomplish it.
2. Delay — the agent postpone some action in time.
3. Avoid — the agent prevent same action which may happen.
4. Ask — the agent asks somebody to do something.
5. Promise — the agent promises to do something.
6. Propose — the agent propose to do something.

Comparing to TimeML we removed two groups of events from this category, i.e.: investigation (investigate, delve) and naming (name, nominate, appoint, etc.). Those two groups does not require any other event as an argument, thus they can be treated as an action.
Examples: try, delay, promise, ordering

3.7. Intensional state

Intensional state is a state which refers to some possible actions or states. It indicates, than an agent refers to some possible event, which may or may not occur in the future. Most of the intensional states are connected with mental activities, emotions and needs. The possible groups of intensional states (following TimeML) are:

1. Thinking — agent A thinks about doing X.
2. Will — agent A want to do X.
3. Expectation — agent A expect X to happen.
4. Emotions — agent A is afraid of doing X.
5. Needs — agent A need to do X.
6. Be ready — agent A jest ready to do X.
7. Ability — agent A jest able to do X.

Examples: believe, fear, wish, desire

4. Event Mentions in Text

Authors of TimeML state that “events may be expressed by means of tensed or untensed verbs, nominalizations, adjectives, predicative clauses, or prepositional phrases” (Saurí et al., 2006, p. 3). In our approach the list was limited to three types of mentions: tensed or untensed verbs (including participles and gerundial forms), nominalizations (in the wider lexicalistic sense) (Jędrzejko, 1993, pp. 53–56) and adjectives. The complete list of mention types is presented at Figure 3.

4.1. Verbs

Statement that events may be expressed by verbs seems almost axiomatic. Philosophers studying processes and their typology, e.g., Ryle (1949), Vendler (1957) or Kenny (1963) developed differentiated classificatory schemes of ‘action verbs’ or verb denotations (Seibt, 2004). Despite some difficulties, their approach had
a great impact on the future classifications including Polish tradition. For Laskowski (1998, p. 152–153), situation is a denotate of sentence constituted by verb.

It was important to introduce a method of annotating periphrastic predication. Ewa Jędrzejko points out several types of complex predicates (Jędrzejko, 2011, pp. 34–37):

- Standard nominal predicates \([\text{VCOP} + \text{NKONKR} // \text{Nabstr} // \text{Adj} // \text{Adv}]\),
- The so-called modal predicates \([\text{VMOD} + \text{VINF}] + \ldots\),
- The phase-aspectual complex predicates \([\text{VFAZ} + \text{V} // \text{NA}] + \ldots\),

---

**Figure 3:** Categories of event mentions.

<table>
<thead>
<tr>
<th>Predicate</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>in simple predicate</td>
<td>yes</td>
</tr>
<tr>
<td>copula</td>
<td>no</td>
</tr>
<tr>
<td>modal</td>
<td>yes</td>
</tr>
<tr>
<td>aspectual</td>
<td>yes</td>
</tr>
<tr>
<td>generic</td>
<td>yes</td>
</tr>
<tr>
<td>in periphrastic metaphorical expressions, idioms</td>
<td>yes</td>
</tr>
<tr>
<td>impersonal form</td>
<td>yes</td>
</tr>
<tr>
<td>predicative</td>
<td>yes</td>
</tr>
<tr>
<td>infinitive</td>
<td>yes</td>
</tr>
<tr>
<td>subject complement (ellipsis of copula)</td>
<td>no</td>
</tr>
<tr>
<td><em>wnienie</em>-like verb</td>
<td>no</td>
</tr>
<tr>
<td><em>to</em></td>
<td>no</td>
</tr>
<tr>
<td><em>żal</em>...</td>
<td>yes</td>
</tr>
<tr>
<td>adverbs</td>
<td>no</td>
</tr>
<tr>
<td>context</td>
<td>no</td>
</tr>
<tr>
<td>accent, intonation</td>
<td>no</td>
</tr>
<tr>
<td>pause</td>
<td>no</td>
</tr>
<tr>
<td>gerund/nominalization</td>
<td>yes</td>
</tr>
<tr>
<td>participle</td>
<td>yes</td>
</tr>
<tr>
<td>adjective</td>
<td>yes/no</td>
</tr>
</tbody>
</table>
• The most common type of the VNA with basic ‘generic’ verbs [VGENER +
NA//NE//Nabstr] + ....
• Periphrastic predicates in the strict sense of the term [VMETAF// METAPRED
+Nabstr//NA//NE ] + ....\(^6\)
• So-called phraseological predicates, i.e. ‘typical’ idioms functioning as verbs
[Vmetafor + N + ◦ + ◦ + ] IDIOM

We’ve decided to exclude some verbs from annotation. Tagged elements should
introduce enough information to classify the situation. Our assumption was that
copulae and other auxiliary verbs (e.g. components of analytic future tense) are
semantically (referentially) empty so they are not very useful for event extraction.
Verbal part of modal predicates, phase-aspectual predicates and predicates with
generic verbs may be called light predicate or light verb (Jespersen, 1965). Ac-
cording to Zolotova, Onipenko, and Sidorova (1999) they are modifiers (phase and
modal) and compensators (accompanied by deverbal noun) There is no agreement
concerning the definition and the semantics of light verbs (Kotsyba, 2014) but it
was valid to include them to annotation as they carry a grammatical and very
general but sufficient lexical meaning. We tag both elements of such predicates
because they are relevant to different kind of event information — after that two
tagged events will be linked as identical.

4.2. Nominalizations
Nominalizations in which the original verb appears as a gerund or deverbative
noun should be treated as equivalent of verb predications. According to Topolińska
the result of nominalization is a nominal phrase correlated on syntactic level with
a sentence or verbal realization of the same propositional model (Topolińska, 1984,
p. 355). Still, we don’t consider nominalization as a simple transposition. Even
the transformations, that aren’t structurally motivated, base on the principle of
semantic transposition and suppletivism (“miłość” — “kochać”, “uroda” — “ładny”,
“klęska” — “pokonać”) and they could be secondarily verbalized by means of synse-
mantic verbs\(^7\) ("czuć radość", “mieć urodę”, “ponieść klęskę” etc.) (Jędrzejko, 1993,
pp. 46–47), so we decided to include them.

We have made an exception for one category. At the first stage of annotation
we don’t mark the nominalizations for the states due to the specific ontological
status of this situations and the features of their nominalizations (Mourelatos, 1978,
pp. 204–210). It is an open question if we need to recognize them in the future.

4.3. Other mentions
Taking into account the scope of our task it was important to consider all predi-
cative expressions. Jodłowski (1976, pp. 31–33) introduced one of the first and
basic classification. It includes many types of nonverbal predicates such as ac-
cent, intonation, context, pause or adverbs. Still, identification of these mentions
would require context analysis or some additional data (e.g., conversational), so we
decided to exclude them from annotation.

\(^6\)concrete’ verbs are used in these constructions metaphorically not only in a verbalising
function” (Jędrzejko, 2011, p. 36).
\(^7\)light verbs
As Saurí et al. (2006) stated events may be expressed by adjectives. Some of them are nominalizations so it is valid to annotate them. We have decided that the mentions that introduce other situations (i.e. that have an event argument) are the most important.

5. Annotators Agreement
The inter-annotator agreement was measured on randomly selected documents from the Corpus of Wrocław University of Technology called KPWr (Broda et al., 2012). We used the *positive specific agreement* (psa) (Hripcsak & Rothschild, 2005) as there are no negative decisions to count to measure the agreement between two linguists. The documents were annotated using the Annotator perspective from the Inforex system\(^8\) (see Figure 4) (Marcinczuk, Kocou, & Broda, 2012).

![Figure 4: Document annotator perspective in Inforex.](http://www.nlp.pwr.wroc.pl/inforex/)

In the first iteration we randomly selected 100 documents. The results are presented in Table 2. The agreement for event mentions without categorisation was ca. 85% and with categorisation it drops to 68%. The results show, that the most confusing categories were *state* (36.98%) and *light predicate* (39.60%). The best agreement was achieved for *aspectual* (86.79%) and *action* (77.89%). We have carefully analyzed the discrepancies between annotators and introduce some clarifications in the guideline. The major changes were:

- Categories which require another event as an argument (*aspectual, perception, reporting, intensional action and intensional state*) can be annotated regard-

---

Table 2: Agreement between two linguists (A and B) after first iteration.

<table>
<thead>
<tr>
<th>Mention category</th>
<th>A and B</th>
<th>Only A</th>
<th>Only B</th>
<th>( psa )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentions without categorization</td>
<td>3184</td>
<td>393</td>
<td>664</td>
<td>85.76%</td>
</tr>
<tr>
<td>Mentions with categorization</td>
<td>2561</td>
<td>1016</td>
<td>1287</td>
<td>68.98%</td>
</tr>
<tr>
<td>action</td>
<td>2085</td>
<td>766</td>
<td>418</td>
<td>77.89%</td>
</tr>
<tr>
<td>state</td>
<td>213</td>
<td>92</td>
<td>634</td>
<td>36.98%</td>
</tr>
<tr>
<td>perception</td>
<td>20</td>
<td>2</td>
<td>37</td>
<td>50.63%</td>
</tr>
<tr>
<td>reporting</td>
<td>39</td>
<td>29</td>
<td>28</td>
<td>57.78%</td>
</tr>
<tr>
<td>aspectual</td>
<td>46</td>
<td>4</td>
<td>10</td>
<td>86.79%</td>
</tr>
<tr>
<td>intensional action</td>
<td>23</td>
<td>19</td>
<td>21</td>
<td>53.49%</td>
</tr>
<tr>
<td>intensional state</td>
<td>115</td>
<td>61</td>
<td>70</td>
<td>63.71%</td>
</tr>
<tr>
<td>light predicate</td>
<td>20</td>
<td>41</td>
<td>20</td>
<td>39.60%</td>
</tr>
</tbody>
</table>

less the argument is directly stated in the sentence or not (for example the event argument might be omitted or referenced by a pronoun).

- We have formulated a procedure for recognition state mentions. We have defined the following criteria:
  - If the mention as a lexical unit is present in the plWordNet\(^9\) (Piasecki, Szpakowicz, & Broda, 2009) then the synset containing the lexical unit must be a direct or indirect hyponym of an artificial synset call “state verb”.
  - State has no dynamic.
  - State does not change in time.
  - Verb representing state has imperfect aspect.
  - Verb representing state does not have a perfective form.

- Passive construction does not indicate a state. Event category results from the semantic of the verb, not the grammatical construction. For example sentences “John was killed” and “Tom killed John” represent the same action of killing a person named John.

- We defined a procedure to determine if given mention should be annotated and with what category. The procedure consists of a set of yes-no questions. The procedure is presented in Section 6.

Next, we have performed a second evaluation to check, if the clarifications improved the agreement between annotators. In the second iteration we randomly selected another set of 50 documents. The documents were annotated by the same two linguists. The results for second evaluation are presented in Table 3. The agreement for mention annotation without categorisation raised from 85% to 93%. There were noticeable improvement in annotation of the categories which were annotated with unacceptable agreement (state and light predicate) in the first iteration (form 36.98% to 74.26% for state; from 39.60% to 50% for light predicate). The overall agreement in the second evaluation was significantly higher. However, there are

Table 3: Agreement between two linguists (A and B) after second iteration.

<table>
<thead>
<tr>
<th>Mention category</th>
<th>A and B</th>
<th>Only A</th>
<th>Only B</th>
<th>psa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentions without categorization</td>
<td>2427</td>
<td>140</td>
<td>225</td>
<td>93.01%</td>
</tr>
<tr>
<td>Mentions with categorization</td>
<td>1856</td>
<td>346</td>
<td>430</td>
<td>82.71%</td>
</tr>
<tr>
<td>action</td>
<td>1531</td>
<td>253</td>
<td>198</td>
<td>87.16%</td>
</tr>
<tr>
<td>state</td>
<td>135</td>
<td>45</td>
<td>81</td>
<td>68.18%</td>
</tr>
<tr>
<td>perception</td>
<td>23</td>
<td>1</td>
<td>6</td>
<td>86.79%</td>
</tr>
<tr>
<td>reporting</td>
<td>21</td>
<td>18</td>
<td>14</td>
<td>56.76%</td>
</tr>
<tr>
<td>aspectual</td>
<td>26</td>
<td>2</td>
<td>8</td>
<td>83.87%</td>
</tr>
<tr>
<td>intensional action</td>
<td>20</td>
<td>8</td>
<td>57</td>
<td>38.10%</td>
</tr>
<tr>
<td>intensional state</td>
<td>88</td>
<td>13</td>
<td>48</td>
<td>74.26%</td>
</tr>
<tr>
<td>light predicate</td>
<td>12</td>
<td>6</td>
<td>18</td>
<td>50.00%</td>
</tr>
</tbody>
</table>

still two categories with low agreement which should be verified, i.e. intensional action with agreement of 38.10% and light predicate with agreement of 50%.

6. Mention Annotation Procedure

After the first iteration of measuring the annotation agreement we formulated a procedure for mention annotation and classification. The procedure consists of a series of yes-no questions.

Input: M — Mention

Q1  Does M name an action (activity, accomplishment or achievement) or state (or their collections) in the real world?
    YES ⇒ goto Q2
    NO  ⇒ don’t mark

Q2  Is M a complex predicate?
    YES ⇒ goto Q3
    NO  ⇒ goto Q4

Q3  Is M a metaphor or idiom?
    YES ⇒ mark head only — goto Q7
    NO  ⇒ goto Q4

Q4  Is M a predicative expression?
    YES ⇒ don’t mark
    NO  ⇒ mark separately both elements:
      • verb (copula) — goto Q5;
      • complement — goto Q1.

Q5  Is M the LIGHT PREDICATE?
    YES ⇒ goto Q6
    NO  ⇒ goto Q7
Q6  Is M connected with nominalization?
    YES ⇒ mark as LIGHT PREDICATE
    NO  ⇒ don’t mark

Q7  Does M has an event argument?
    YES ⇒ goto Q9
    NO  ⇒ goto Q8

Q8  Could M have an event argument?
    YES ⇒ goto Q9
    NO  ⇒ goto Q13

Q9  Is M a state?
    YES ⇒ mark as INTENSIONAL STATE
    NO  ⇒ goto Q10

Q10 Does M inform about a phase of the situation?
    YES ⇒ mark as ASPECTUAL
    NO  ⇒ goto Q11

Q11 Does M describe the reporting of other situation?
    YES ⇒ mark as REPORTING
    NO  ⇒ goto Q12

Q12 Does M describe the perception of other situation?
    YES ⇒ mark as PERCEPTION
    NO  ⇒ mark as INTENSIONAL ACTION

Q13 Is S the state?
    YES ⇒ mark as STATE
    NO  ⇒ mark as ACTION

7. Conclusions and future work
The evaluation showed that the annotation of event mentions is relatively simple —
the agreement after second iteration was 93%. The categorisation of event mentions
causes more problems and the agreement drops to 82%. This shows that the task is
not trivial and if we want to obtain a good quality of data with high agreement the
final annotation of whole KPWr corpus will require the “2+1” approach. This means
that each document in the corpus will be annotated separately by two linguists and
the differences will be evaluated by a third linguist — supervisor.

The future plan is to prepare guidelines for event description with attributes,
arguments and relations. The attributes, we are considering, are:

• generality — is the event specific or general,
• polarity — is the form of the mention affirmative or negative,
• modality — is it assertoric, optative, imperative or interrogative,
Towards an Event Annotated Corpus of Polish

• tense and aspect

Most of them (in particular generality) have to be annotated manually. Although there are tools\(^{10}\) that could be used for automatic annotation of tense and aspect.

The events will be linked with their generic arguments, i.e. agent, time and location. In the last step we will mark the relations between the events. The categories of relations include identity and references between reporting, perception, aspectual, intensional action and intensional state and their event arguments.

References


\(^{10}\)Morphological taggers, i.e. WCRFT (Radziszewski, 2013).


van Erp, M., Fokkens, A., & Vossen, P. (2014). Finding stories in 1,784,532 events: Scaling up computational models of narrative. In Workshop on Computational Models of Narrative (CMN’14), Quebec City, Canada, July 31 – August 2.


Acknowledgment

Work financed by the Polish Ministry of Science and Higher Education, a program in support of scientific units involved in the development of a European research infrastructure for the humanities and social sciences in the scope of the consortia CLARIN ERIC and ESS-ERIC, 2015–2016.

The authors declare that they have no competing interests.

The authors’ contribution was as follows: MM, MO and JK are responsible for the ideas in the research and the methodology; MO and MM created annotation guidelines and were responsible for inter-annotator agreement increasing and performed data analysis, TB, MW collected the examples; all authors participated in drafting the manuscript.

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (http://creativecommons.org/licenses/by/3.0/pl/), which permits redistribution, commercial and non-commercial, provided that the article is properly cited.

© The Authors 2015

Publisher: Institute of Slavic Studies, PAS, University of Silesia & The Slavic Foundation