Text-world annotation and visualization for crime narrative reconstruction

HO, Yufang, LUGEA, Jane, MCINTYRE, Dan, XU, Zhijie and WANG, Jing <http://orcid.org/0000-0002-5418-0217>

OXFORD UNIVERSITY PRESS | Digital Scholarship in the Humanities

Text-world annotation and visualization for crime narrative reconstruction

Journal:	*Digital Scholarship in the Humanities*
Manuscript ID	DSH-2018-0007.R2
Manuscript Type:	Full Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Ho, Yu-Fang; University of Huddersfield, Linguistics and Modern Languages Lugea, Jane; Queen's University Belfast, School of Arts, English and Languages McIntyre, Dan; University of Huddersfield, Linguistics and Modern Languages Xu, Zhijie; University of Huddersfield, Department of Informatics Wang, Jing; Sheffield Hallam University, Department of Computing
Keywords:	Text World Theory, information visualization, legal discourse, Meredith Kercher case, Contextual Frame Theory, cognitive computational modelling

Text-world annotation and visualization for crime narrative reconstruction

Abstract

In order to assist legal professionals with more effective information processing and evaluation, we aim to develop software to identify and visualize the key information dispersed in the unstructured language data of a criminal case. A preliminary model of the software, *Worldbuilder*, is described in Wang et al. (2016). The present article focuses on explaining the theory and vision behind the computational development of the software, which has involved establishing a means to annotate discourse for visualization purposes. The design of the annotation scheme is based on a cognitive model of discourse processing, Text World Theory, which describes and tracks how language users create a dynamic representation of events (i.e. text-worlds) in their minds as they communicate. As this is the first time Text World Theory has informed the computational analysis of language, the model is augmented with Contextual Frame Theory, amongst other linguistic apparatus, to account for the complexities in the data and its translation from text to visualization. Using a statement from the Meredith Kercher murder trial as a case study, we illustrate the efficacy of the augmented Text World Theory framework in the careful and purposeful preparation of linguistic data for computational visualization. Ultimately, this research bridges Cognitive and Computational Linguistics, improves the TWT model's analytical accuracy, and yields a potentially useful tool for forensic work.

Keywords: Text World Theory; Contextual Frame Theory, cognitive computational modelling; information visualization; legal discourse; Meredith Kercher case.

1. Introduction

The human mind is a powerful and effective processor of written and spoken communication, being able to perform complex operations at various linguistic and metalinguistic levels. Reading and interpreting crime narratives (e.g. witness/suspect statements) in a legal case, however, is perhaps more complex than everyday communication and is one of the major demands in the work of legal practitioners. In cases where the texts contain complex and contradictory information, assessing the probative value of the claims contained therein can be a challenge. Investigators often

have to rely on intuition, experience, and memory of crucial details to make sense of the complex claims, arguments, and propositions contained within statements.

Concerning the professional evaluation of crime narratives, we contend that legal practitioners across the board can be supported in performing their respective inferential tasks more effectively if they are equipped with a visual understanding of (i) the key information components of WHO-WHEN-WHERE-WHAT HAPPENED in a statement, and (ii) the dynamic cognitive process and the overall pattern of epistemic (un)certainty projected in the statement. We thus propose that a solution to these issues can be found by combining cognitive linguistics and computational approaches to language analysis (two fields not often married, but see Olivier and Tsujii, 1994; Narayanan, 1999a, 1999b; Veale and O'Donoghue, 2000; Feldman, 2004; Barnden, 2006; Kristiansen, *et al.*, 2006; Veale, 2006; Manzanares, 2010; Mahlberg, *et al.*, 2017 for notable exceptions). As part of such an endeavor, we aim to develop a technique for systematically identifying and visualizing key communicative components in legal texts. To provide such support, a robust linguistic analytical framework is required. A cognitive model of discourse processing, Text World Theory (TWT), is applied and adapted to underpin our computational visualization methods. We contend that the adapted TWT framework enables the formal identification[1] of specific elements in texts that are important for producing computer-based visualizations and for accounting for all the facts that are put forward in a legal case. Being able to visualize key information and its truth-status will facilitate the analysis of the nature and quality of propositions given by the various parties involved in a case and, ultimately, assist legal practitioners in making informed interpretative decisions about the probative value of the statements in question.

This article focuses on expounding the theoretical and methodological framework and its potential applications (for the prototype of our computational development, *Worldbuilder*, see Wang, *et al.*, 2016a, 2016b). In Sections 2-4, we review and evaluate how TWT could be of use to crime narrative analysis and explain how a linguistic annotation scheme is designed to incorporate the range of information that is needed for computer-assisted visualization processing. After this theoretical exposition, we then demonstrate in Sections 5-6 the potential application of the proposed text-world annotation-visualization model via a case study. We present a text-world analysis of a witness statement drawn from the high-profile Meredith Kercher murder case (2007-15). The statement made by one of the accused, Raffaele Sollecito (see Appendix), is manually annotated, diagrammed, and analyzed, currently using open source software for illustration purposes.[2] We then conclude in Section 7 by discussing critical issues in the future development of *Worldbuilder* and automatic data annotation.

2. Text World Theory

For the last two decades, much research in linguistics and discourse studies has incorporated a strong cognitive component in its analytical models, focusing on cognitive processes that enable human beings to understand and interpret language and, more particularly, to associate text with experiential knowledge and context. Researchers working on language and cognition regard narrative as a means of relocating the reader's deictic center, so that the reader experiences the effect of entering the spatio-temporal world of the narrative and moving around within it. There has been a body of research which looks at the human ability to construct mental representations of 'worlds' or 'frames' which are held in place while stretches of text are processed (Duchan, *et al.*, 1995; Segal, 1995a, 1995b; Werth, 1995a, 1995b; Emmott, 1997; Werth, 1999; Gavins, 2007).

The cognitive framework we adopt for testing crime narrative annotation and visualization is based on Gavins' (2005, 2007) revised version of Text World Theory, developed from Werth's original conceptualization (1994, 1995a, 1995b, 1999). TWT posits that all discourse situations are divisible into three manageable levels of conceptual activity. These are characterized as follows: *discourse-world*, *text-world*, and *world-switch*.[3] ***Discourse-world*** refers to the immediate situational context in which the language event takes place; it is a context co-constructed and shared by the discourse participants (for example, speaker/listener in a conversation or writer/reader in written communication). ***Text-worlds*** are then created in the minds of the discourse participants, being invoked by the language used in the context. Each text-world has its own spatio-temporal parameters, meaning that a person experiencing this world must reset his/her deictic orientation by responding to 'world-builders' (i.e. time, location, enactors [also known as characters]) provided in the text. The variants to schematic frames (time, places, names, objects, *etc.*) thus enrich the text-world (Werth, 1999: 259).

Given the dynamic nature of the discourse process, there are sometimes changes in the initial text-world parameters that lead to a 'world-switch'. ***World-switches*** may be indicated by a deictic shift in time and/or location. When the spatio-temporal information is altered directly by discourse participants (e.g. through a shift in tense or the use of a spatial or temporal adverbial), a deictic world-switch occurs. Sometimes discourse participants may project an inaccessible state of affairs (i.e. entering into one's mind typically by means of expressing wishes, intentions, and beliefs, *etc.*). These types of mental representations usually convey a particular attitude or epistemic distance to the proposition being expressed and, as such, are categorized as ***modal-worlds*** in TWT terms. Thus, throughout the discourse process, in addition to deictic world-switches, it is also possible to detect switches to a modal-world triggered by linguistic indicators of a person's attitude or knowledge/belief with regard to a

particular topic.

The multi-layered world model provides an extremely useful framework for studying legal discourse. One particular value of this framework lies in the principle of ***(in)accessibility*** with respect to domains of the worlds. That is, text-worlds that are created by means of shifts in space or time (e.g. discourse participants' flashbacks) will constitute domains directly accessible to other participants. However, if a world-switch is introduced by means of a projection from one's own mind, the domain in question (i.e. a modal-world) will *not* be directly accessible from the outside world. As modal-worlds are often projected from a certain viewpoint, they stipulate situations which cannot (as yet) be confirmed (Werth, 1995a: 75). It should be noted that the concept of information (in)accessibility has a bearing on the assessment of whether a given proposition is verifiable (admissible) or unverifiable (inadmissible) as evidence to a court of law.

As this is the first time TWT has informed the computational analysis of language, the model is augmented with Emmott's (1997) Contextual Frame Theory (CFT) and Ryan's (1991) Possible World Theory, to account for the complexities in the data and its translation from text to visualization. As we will discuss in Section 3, Emmott's theory provides not only an account of the way in which readers form a mental picture of a fictional context (i.e. 'frame') during the reading process, but also how readers respond cognitively to contextual and textual changes indicated in the text. As Emmott's CFT explores more forcefully the dynamic relationship between what is written in the text and what happens cognitively to a reader, it is particularly useful in addressing the concept of cognitive 'switches' in text processing. Ryan's theory is useful in conceptualizing the discourse structure of non-fictional text type such as legal statements (see Section 4.1).

3. Revisiting the Concept of Cognitive 'Switches' in Text Processing

Whenever a text describes an event, the reader usually learns about the context from a series of linguistic cues: (i) who is described or present in the physical environment? (ii) where is the action/event/state located or being described? (iii) what is the approximate time of the action/event/state? Emmott (1997) adopts the term ***frame***, an equivalent notion to 'text-world', to indicate the reader's mental representation of the context. The term ***contextual monitoring*** is also introduced to refer to the reader's awareness of the context changes in a text. The reader is given information about each new context as it occurs (i.e. who is doing what with whom in a particular place at a particular time) but must hold this information 'in mind' as specific events are described. For the reader of narrative texts, context building and monitoring is complicated by the fact that narratives often switch suddenly from one location and/or one time to another.

In the real world, we cannot suddenly be transported back or forwards several years in time and we cannot be instantaneously located miles away. In text-worlds, however, the reader's focus of attention may be suddenly shifted from context to context in this way. Therefore, to make sense of the story, the reader needs to be able to recognize such 'switches' within a text.

3.1 Inferring Continuity or Change: A World-Switch or Not?

During the reading process, the reader is actively involved in constructing and updating their mental representations of the contexts. Emmott (1997) discusses the extent to which maintaining a context involves the mental effort of making inferences about continuity and change. The narrative may shift to a completely new context. The notion of ***frame switch*** thus indicates the process by which the reader 'ceases to directly monitor one frame and starts monitoring another' (Emmott, 1997: 147). In TWT terms, the reader cognitively makes a world-switch.

However, as people present in a frame (or in a text-world) may move in and out of a particular context, in such cases, the reader must assume some contextual continuity. Emmott (1997) introduces the term ***frame modification*** to indicate that the same contextual frame remains primed but that the frame must be altered to take account of a change to the set of people in a context. Take the following sentences in Sollecito's statement (see appendix), for example:

> Example 1 (sentence 3):
> *On November 1st, I woke up and had breakfast with Amanda then she went out and I went back to bed….*
>
> Example 2 (sentences 4-8):
> *Then around 13:00-14:00 I went to see her at her house. Meredith was there too. Amanda and I had lunch while Meredith did not have lunch with us.*
> *Around 16:00 Meredith left in a hurry without saying where she was going. Amanda and I stayed home until about 17:30-18:00.*

As Emmott (1997: 143) states, a frame may be modified by adding or removing people. After modification, the same frame remains in force even though the set of people has changed. Both examples illustrate the instances of frame modification. In a sense, we remain in the same text-worlds we constructed, even though the set of people populated in these two text-worlds have changed.[4]

To keep track of continuity and change within narrative, the reader needs to monitor contextual information, in which time-space reorientations represent the major boundaries. The difficulty lies in deciding the temporal boundaries of a specific context. Monitoring temporal switches is more complex than the reader's other assumptions

about location change, given that time described in a text may move at different speeds and it is usually moving steadily onwards as events are described (Emmott, 1997: 149). Sometimes time will move in such a way that the reader cannot expect the same grouping of people to be in the same situational context any longer. The narrative may move backwards in time (e.g. flashbacks) or there may be a leap into the future (e.g. the following day, week, month, or year) which suggests that the frame that has been monitored may no longer be primed and that it is necessary to switch to a new frame (Emmott, 1997: 150). Given that TWT, as Lugea (2016: 91) observed, does not provide a principle means of distinguishing between the simple passage of time within a text-world and a switch in the temporal settings to a new text-world, Emmott's distinction between frame modification and frame switch is crucial in conceptualizing the updating process of text-worlds, namely, to decide if a world-switch occurs or not. Fig. 1 illustrates how the spatial and temporal parameters play a role in monitoring world-switches.

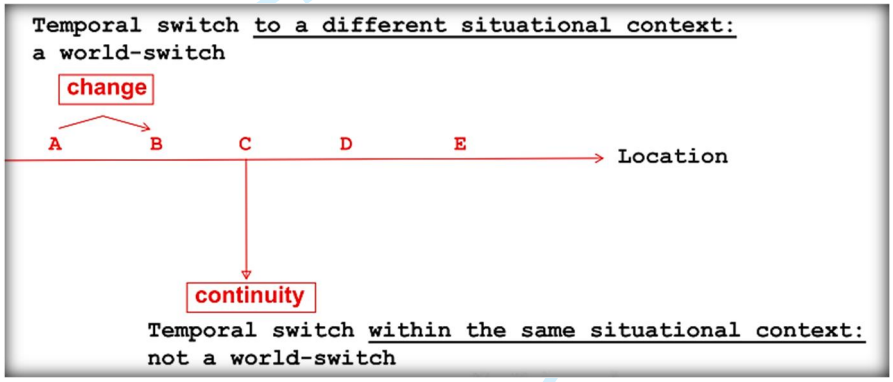

Fig. 1 Inferring context continuity or change

When a change of location (from A to B) is described in texts, a temporal switch necessarily co-occurs despite the absence of overt textual signals. The change of situational context constitutes a world-switch. By contrast, when a sequence of temporal events is described as happenings in the same situational context (C), no world-switch occurs, as illustrated in Example 2 above. The cognitive work involved in assuming continuity or change of a situational context must be taken into account in annotating and diagramming the text-world structure of a legal text.

3.2 Distinguishing Text Types: Framed Text vs. Unframed Text

Narrative is usually defined as a succession of events (Herman, 2002). Very often events are described as they take place within a particular context. As a result, these events are 'brought to life' for the reader, being 'acted out' rather than presented in summary form. This type of presentation, according to Emmott (1997: 236), is called ***framed text***, since the reader needs to monitor the same contextual frame over a stretch of text. Framed

text locates a succession of events in a particular context (see Example 2).

Another type of presentation, ***unframed text***,[5] by contrast is when events are summarized and presented as background information to the main story line. A stretch of unframed text does not describe one specific occasion in context, but rather presents its readers with generalization (e.g. in the introductory paragraphs of a story) (Emmott, 1997: 238). The text is unframed because it does not explicitly place us in any specific context and does not require monitoring by a contextual frame. The sentence, *Amanda always carries a big handbag*, arguably exemplifies the unframed text type.

As Emmott rightly observes, in addition to the frequent switches from one contextual frame to another (i.e. world-switches), it is also common for a narrative to switch from describing a particular context to providing background information that the reader needs in order to make sense of the story. Since these two text types are often interwoven in a narrative, it is important for the reader to be aware of these different types and be able to recognize signals of a switch between text types. The distinction between (framed) events-in-context and (unframed) decontextualized generalizations is important and has a textual realization which needs to be considered in developing a cognitive model of discourse processing, annotation, and visualization.

One focus of our research lies in keeping track of all the salient information presented in crime narratives (i.e. who is involved in an action/event in any particular place and/or at any particular time). The tracking of textual and contextual information about people, events, places, and time is a complex information-processing task and needs to be done by examining the complexity of linguistic patterns in a statement. Our research thus involves identifying these elements in a narrative and attempting to annotate them in a systematic way for visualization purposes. In the next section, we report how we structuralize the TWT framework as well as the rationale behind the linguistic annotation scheme that we have designed.

4. TWT Annotation Scheme

As a cognitive model of discourse processing, Text World Theory was developed initially to account for and diagram how discourse participants create mental representations of the discourse. It focuses on relating linguistic choices to cognitive structure and processes that underlie the production/reception of language. Based on the framework specified in Sections 2 and 3, Fig. 2 presents the structured scheme for a computer-based operational purpose.[6] The following sub-sections then explain the rationale of the annotation scheme design.

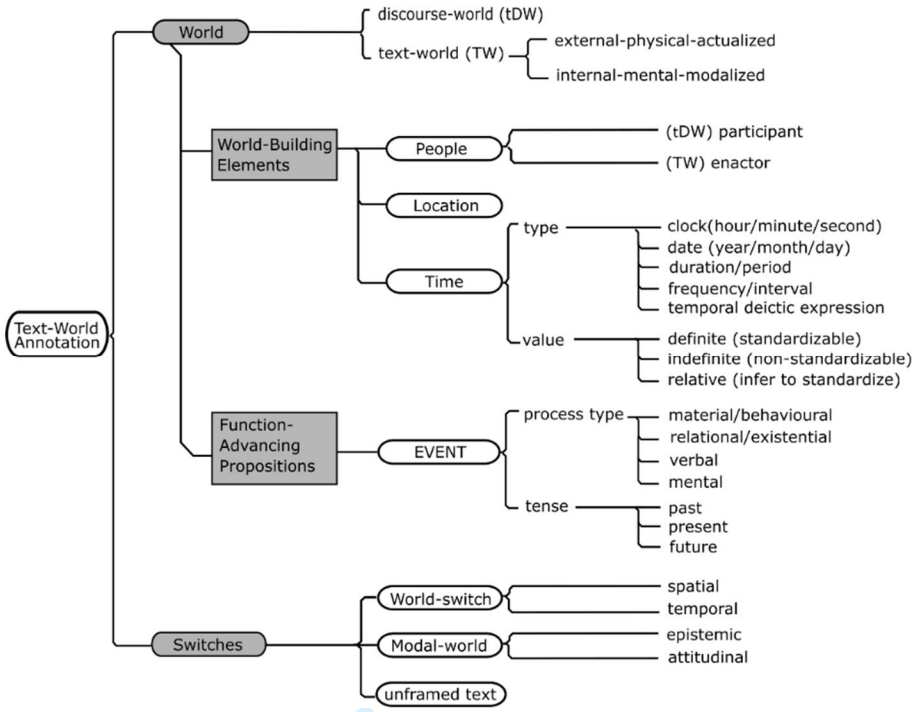

Fig. 2 TWT structured framework for data annotation

4.1 Defining 'World'

Before we explain the design of the annotation scheme, one important 'world'-related concept needs to be clarified at the outset; that is, to draw a crucial distinction between *actual* and *textual* discourse-world while processing legal documents.

In TWT, the discourse-world is the immediate real-life context, whereas text-worlds are the mental representations constructed by the participants from the language used in the particular context in question. While examining written legal documents such as witness/suspect statements, the real-time discourse setting is always that we as readers, or any legal practitioners, situated in different spatial-temporal coordinates in the actual world, are engaged in processing the text produced in a specific legal setting. Our *actual* discourse-world is *split* (Gavins, 2007) in space and time from the real negotiated 'discourse-world' – that is, the legal setting that we mentally constructed by responding to world-builders specified in the official written statement. Taking the first paragraph of the statement studied here (see Appendix) as an example, the discourse structure is illustrated in Fig. 3:

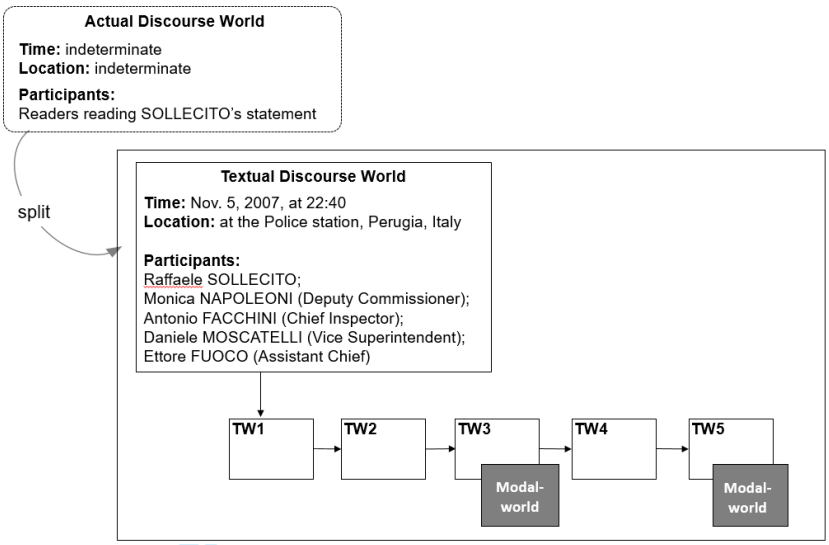

Fig. 3 Actual discourse-world vs. Textual discourse-world

From the theoretical perspective, the legal discourse setting is the initial 'text-world' that we mentally construct while processing the official statement. However, considering that text-world, in essence, is different from the other text-worlds evoked by witness/suspects' (in this case, Sollecito's) personal narratives relating the past events, we think it is necessary to make a clear distinction. To avoid confusion, we thus propose to reinvoke Ryan's (1991) typology of 'Actual World', 'Textual Actual World' and 'Alternative Possible Worlds' in conceptualizing the discourse structure of non-fictional text types such as legal statements. We refer to the legal setting as the *textual* discourse-world (tDW), which is split from our *actual* ontological discourse-world (see Fig. 3). While processing any legal statement given by a witness/suspect, we are in a sense re-centering into a new system of **actuality** and **possibility**, in Ryan's (1991) terms, where multiple text-worlds or modal-worlds are projected from the crime narratives. It should be noted that in our annotation scheme, tDW always refers to the legal discourse setting (e.g. police station or courtroom) specified in the official document.

As the discourse-world of such written statements is split in time and space between the participants, readers are in a way 'communicatively handicapped' (Verdonk, 2008: 375) because we do not physically share the abundance of contextual elements (for instance, the prevailing mood among the tDW participants, the gestures, voices, facial expressions and body language) as much as we do in a face-to-face discourse-world. There is thus no direct way for us to negotiate the meaning of the text with the text producer; there is only the written text from which we can increment our knowledge about the details of a case.

As we know, in criminal proceedings, the truth-value of a proposition has decisive implications for the admissibility of evidence. Regarding the propositions which cannot

be verified as true, TWT has a way of structuring such information. When linguistic indicators of attitudinal or epistemic modality (i.e. degree of certainty) are used, the information introduced by such indicators creates a modal-world and is stored as possible situations, which would need to be retained and verified until further evidence is acquired.

In the annotation of text-worlds, we mark-up TW-*external/physical/actualized* and TW-*internal/mental/modalized* (see Fig. 2). This is to make a distinction between a text-world in which EVENTs[7] (i.e. action/event/state) are presented as actual happenings (something that is going on *out there* in reality), and a modal-world that is projecting *inner* experience (something that is going on *inside* oneself, e.g. cognition, emotion and imagination, etc). Any proposition marked with attitudinal or epistemic modality would be categorized as *internal/mental/modalized* projections, and the EVENTs projected within such modal-worlds will be flagged for special attention, to be examined further for their factuality or credibility. It is important, however, to stress that not all the EVENTs reported in a factual manner will necessarily be deemed to be true (as we will illustrate later in the statement analysis, see Section 6). Likewise, neither the projected EVENTs marked with epistemic uncertainty are necessarily false or unreliable (as we have exemplified in Ho, *et al.*, forthcoming 2018). Given that these two types of information require us to increment our knowledge in a different way, we thus annotate the text in such a way that we can identify the patterns of the propositional content in a statement for ease of further evidential analysis.

4.2 World-Building Elements: WHO-WHEN-WHERE

Text-worlds are mentally constructed by the discourse participants. By responding to the ***world-building elements*** (i.e. deictic, referential and descriptive information provided in a text), they create mental representations of the contexts. These contextual elements include ***time***, ***location***, and ***enactors*** (people populated in a text-world). Take Example 1 again for illustration: *On November 1st, I woke up at around 11, I had breakfast with Amanda then she went out and I went back to bed*. The linguistic reference to time (on *November 1st, at around 11*), enactors (*Amanda* and *I*=Sollecito) and location (Sollecito's house, as inferred from the contextual information), come together to form our mental representation of that text-world.

With regard to people being referred to in a text, they are categorized as tDW participants or TW enactors. The linguistic reference to people is usually realized through nouns, proper nouns or pronouns (e.g. in this case, Chief Inspector *Antonio Facchini, Amanda, Meredith, I*). Spatial location is usually realized through noun phrases and adverbials or prepositional phrases specifying place (e.g. *we left the house*, *we went into town*, *we stayed in the center*) or place deixis (*here/there*).

The annotation of time is far more complex than identifying references to people and location in a text (for a comprehensive literature review on the temporal structure of natural language discourse and temporal annotation, e.g., TimeML, see Mani, *et al.*, 2005). Temporal information is usually conveyed through the tense/aspect of verbs or time adverbials. Temporal information associated with verbs will be addressed later when we discuss the annotation of function-advancing propositions (i.e. the verb groups of EVENTs process types) to capture WHAT-HAPPENED in the discourse. While marking up the time element of world-building, the annotation focuses specifically on adverbials of time in a text, to identify *when* something happened, for *how long* and *how often*.

In our annotation scheme (see Fig. 2), temporal expressions are categorized into five types (adapted from Setzer and Gaizauskas, 2001; Pustejovsky, 2005; Alonso, *et al.*, 2011):

- **clock** (hour/minute/second): *at around 11, until about 17:30-18:00…*
- **date** (year/month/day): *on November 1, 2007…*
- **duration/period**: *for about two weeks, during that time…*
- **frequency/interval**: *always, occasionally, sometimes, daily, every week…*
- **temporal deictic expression**: *that night, the following morning…*

Each annotated temporal information will then be assigned an information value (by human annotators, at the current stage), that is, to decide if the time information is clearly stated, or unspecified and indeterminate, or to be inferred via contextual information. Time value is categorized into three groups in the annotation scheme:

- **Definite:** temporal expression in text that explicitly refers to a specific point in time (e.g. *on November 1*) and can be normalized to some standard format (typically represented as YYYY/MM/DD, e.g. *2007-11-01*) and be directly placed on a timeline.
- **Indefinite:** temporal expression that is unspecified and thus cannot be normalized and placed on a timeline.
  For example, *earlier, recently, always, frequently, generally, never, occasionally, often, rarely, regularly, seldom, sometimes, usually, etc.*
- **Relative:** temporal expression that is not explicitly stated but can be normalized by taking into account some contextual information. For example, time adverbs such as *now, today, tonight, tomorrow, yesterday* can be inferred and normalized if we know the corresponding reference time. The reference time can either be the time of tDW (i.e. the legal document creation time) or another temporal expression as referred to in the crime narrative. For example,
  (a) *I have known Amanda for about two weeks.*
  (b) *On November 1, at around 11… The following morning around 10:00…*

Example (a) indicates a specific duration of time up to the present, with the date of tDW (see Fig. 3) as the reference end point of the duration. In example (b), the temporal deixis *the following morning* denotes the consecutive next day after *November 1*, a date that was referred to earlier in the text.

Crime narratives usually contain plenty of temporal information. It is crucial to annotate and make all this information explicit so as to support subsequent inferential tasks. Compared to definite and indefinite temporal information, the annotation of relative temporal expressions is more challenging. As described above, contextual information has to be identified (by human annotators) to determine the temporal relation between a temporal expression and its reference time.

4.3 Function-Advancing Propositions: WHAT-HAPPENED

In the TWT model, the world-building elements of a text form a kind of static conceptual background (i.e. text-world) against which certain events and activities are played out. ***Function-advancing propositions*** work to develop and advance EVENTs within the text-worlds; they are expressed in some way by the participants to drive the discourse forwards in order to achieve certain communicative purposes. Function-advancing propositions are realized in verb groups and mapped onto Hallidayan ***process types*** described in Systemic Functional Grammar (SFG) (Halliday, 1985, 1994; Halliday and Matthiessen, 2004).

In SFG, one of the functions of language is to represent or construe human experience (through the grammatical unit of the clause): to describe the happenings and states of the real (and unreal) world. It is the means by which we make sense of 'reality'. At the simplest level, language use reflects the users' view of the world as consisting of 'going-on' (verbs) involving people or things (nouns) that may have attributes (adjectives) and which go against background details of place, time, manner, *etc.* (adverbials). If we use functional labels (i.e. labels that indicate the role played by each element of the representation), we can express the 'content' of clauses in terms of: (a) ***processes*** indicating what kind of action/event/state is being described; (b) ***participants*** involved in each process type, e.g. Actor/Recipient, Sayer/Receiver, Senser, *etc.*; (c) ***circumstances*** specifying the when, where, why and how of the process (Thompson, 2014: 92). The detail of *participants* and *circumstances* has been covered when we discussed the annotation of the world-building elements to capture the WHO-WHEN-WHERE information in a text. In this section, with recourse to the Hallidayan process types, we focus on exploring WHAT-HAPPENED as described or projected in crime narratives.

Following SFG, function-advancing propositions may take the form of six different types of process: *material*, *mental*, *relational*, *verbal*, *behavioral* and

existential (for details, see Halliday and Matthiessen, 2004, Ch. 5; Thompson, 2014). ***Material process***, construing outer experience, refers to physical actions or events in the real world. It describes processes of *doing* and *happening*. By contrast, ***mental process*** serves to construe the inner experience of *sensing* (i.e. happenings within one's mind). Verbs of perceiving, thinking, and feeling are in this group. In addition to the outer and inner aspects of our experience, there is a third type, ***relational process***, which indicates states of *being* (including *having*) and serves to identify (identification) or to characterize (attribution). ***Verbal process*** refers to putting thoughts into words. Verbs of *saying* (e.g. *tell, report, explain, ask, speak*) are in this group. These are the four major types of process in that they are the cornerstones of English grammar in its guise as a means of structuring experience.

In addition to the four main types, there are two subsidiary process types. Each of these shares some of the characteristics of the major types. On the borderline between 'material' and 'mental' is the ***behavioral process***: representing the physical manifestation of mental states (e.g. *laugh, cry, smile, stare*). On the borderline between the 'relational' and the 'material' is the category of ***existential process***, representing that something exists or happens. Existential clauses typically have the verb 'be', by which phenomena of all kinds are simply recognized 'to exist' (e.g. *there were drops of blood in the sink*). As Thompson (2014: 109) suggests, it is in fact possible to see these two process types as sub-categories rather than as groups on a par with the four main types. To avoid confusion in annotation, we thus suggest that (a) behavioral process to be subsumed within material process, and (b) existential process to be subsumed within relational process (see Fig. 2).[8] Table 1 illustrates these concepts using examples taken from the statement under study (see Appendix):

Function-advancing propositions: EVENT process types		**Textual examples**
Material / Behavioral	*doing* (actions) or *happening* (events)	I *woke up*; we *stayed* in the house; M *left* in a hurry.
Relational / Existential	states of *being* or *having* or *existing*	*there were* drops of blood in the sink the rest of the bathroom *was* clean
Verbal	*saying*	*she told me* she was going to the pub…; *she said*…; *I answered*…
Mental	*sensing* (perception, cognition, emotion)	I *don't remember* what we did; I *am certain* I had dinner

Table 1 Function-advancing propositions and textual examples

It should be noted that verbal and mental processes usually have a clausal complement, which in Functional Grammar is called ***projection***. If a projected clause is used, the process and participants in the projected clause are annotated and analyzed separately.

In our previous discussion of temporal information annotation, we focused specifically on adverbials of time in text, to identify *when*, *duration* or *frequency* of happenings as well as relative temporal references. It was noted that time information is also realized through verb tense/aspect – the grammatical expression of time reference. Thus, while annotating function-advancing propositions (i.e. EVENTs process types) in text, we will also tag the time information associated with each process type identified.

Time is a universal, non-linguistic concept with three division: *past*, *present*, and *future*. By *tense* we understand the correspondence between the form of the verb and our concept of time. In English grammar, however, the term 'tense' is applied to a series of verb forms that express not only position in time, but also additional *aspectual* or *modal* properties of the verb groups. *Aspect* concerns the nature of the verbal action or state (for example, a complete event, an ongoing or repeated situation, *etc.*). Verbs are also often conjugated for *mood* (i.e. modality), which marks a speaker's attitude or commitment to the truth value of the verbal action or state, relating to such conditions as intention, uncertainty, evidentiality, possibility, *etc.*

In fact, it is usually difficult to untangle these features of the language. Very often any two of tense, aspect, and mood (or all three) may be conveyed by a single grammatical construction (Quirk and Greenbaum, 1973: 40). Since in many cases the three categories are not manifested separately, our identification of the linguistic segment of verb forms (process types) would be using a combined tense–aspect–mood system. In addition to marking up **finite verb phrases** that convey tense and modality, **non-finite forms of verbs** – the infinitive (*to-V*), the *-ing* participle, and the *-ed* participle – will also be identified and annotated accordingly.

4.4 World-Switches: Annotating Cognitive Text Processing

TWT aims to capture the dynamic nature of the discourse process through the notion of *world-switches* (Gavins, 2007). World-switches in a text may be indicated by a deictic world-switch in time and/or location, or a switch to a modal-world based upon a linguistic indication of a speaker/writer's attitude or knowledge/belief concerning a particular topic, as realized in modalized propositions (see the *mood* choices in verb forms). These two types of world-switches in the TWT framework could be considered as 'frame switches' in Emmott's (1997) term. They indicate our cognitive processing of 'framed text' that locates a succession of events in a particular context which needs to be monitored over a stretch of text.

However, as discussed in Section 3.2, in addition to the frequent world/frame-switches, it is also common for a narrative to provide summarized or generalized background information (i.e. 'unframed text') for the reader to make sense of the story. Given that these two text types, (framed) events-in-context and (unframed) decontextualized generalizations, are often interwoven in a narrative, it is also important for us to recognize signals of a switch between text types in our annotation of cognitive discourse processing. Table 2 sums up the types of switches that may occur in text processing and the textual examples observed from the statement under study.

Cognitive 'switches'	Textual examples
world-switch	e.g. *around 13:00-14:00, I went to see her at her house* (deictic world-switch based on changes in temporal/spatial references)
modal-world	e.g. *we went into town,* but *I do not remember what we did* (knowledge, belief or hypothetical expressions)
unframed text	e.g. *Amanda always carries a big handbag.* (un-sequenced, habitual action/event or generalization)

Table 2 Cognitive 'switches' and textual examples

This section briefly outlined the rationale of the TWT annotation scheme being designed specifically for crime narrative data annotation and information visualization. When it comes to annotation, as with any linguistic categories, some cases will fall more neatly into one category, whereas others may be more marginal. It is not always easy to decide, for example, which type of process we are dealing with. There will be indeterminate cases that could equally well be analyzed as different types. For the time being, we propose to use multiple tags to deal with indeterminate cases.

5. Potential Application to Crime Narrative Reconstruction

We argue that the text-world annotation scheme we propose here has a potential application in the narrative reconstruction of crimes. On the one hand, the TWT cognitive framework provides a comprehensive toolkit for accounting for all the facts that are put forward in a legal case. World-building elements encompass WHO-WHEN-WHERE information regarding all the 'worlds' constructed from crime narratives, while function-advancing propositions and world-switches cover WHAT-HAPPENED in the discourse. On the other hand, it enables the formal identification of specific elements in texts that are important for producing computer-based visualizations.

After raw data is annotated based on the structured TWT framework, the data can then be further processed for quantitative analysis and information visualization. The annotated data can be used to generate two different visualization outputs (see Fig. 4 for the tentative schematic view of the outputs): (a) text-world diagrams: graphical

representations of the macro-level cognitive structure reconstructed from a legal text; and (b) timeline visualization of the annotated EVENTs, which are re-sorted from their textual occurring sequence into chronological order and anchored on a timeline, and the EVENTs projected within TW-internal/mental/modalized (as marked in grey) will be flagged for special attention.

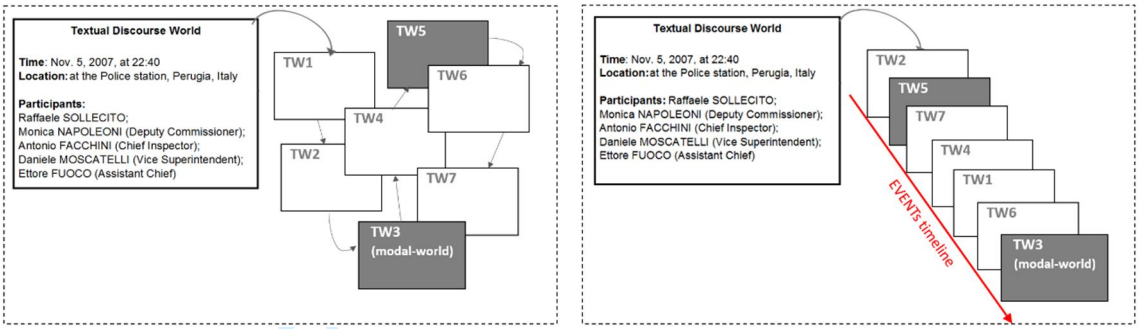

Fig. 4 Proposed visualization outputs

The first type of visualization output is to understand the overall cognitive process and the degree of epistemic (un)certainty underlying each statement for linguistic analysis. As we know, EVENTs (i.e. process types) as reported in a statement do not always follow the chronological sequence (e.g. flashback). The timeline visualization output thus aims to re-sort the temporal sequence of the EVENTs in order to reconstruct the story line in a crime narrative. As a result, it is likely to be of significant value in the reconstruction and comparison of contentious events.

As shown in Fig. 5, the annotated EVENTs in text will be anchored on a timeline. EVENTs that are presented as actual happenings (in TW-*external/physical/actualized*) will be described in the upper part of the timeline. In contrast, any propositions that are embedded within TW-*internal/mental/modalized* (i.e. modalized proposition with projected clause) would then be described beneath the timeline. Given that this type of information is 'inaccessible' to discourse participants, its credibility or factuality thus needs to be examined further.

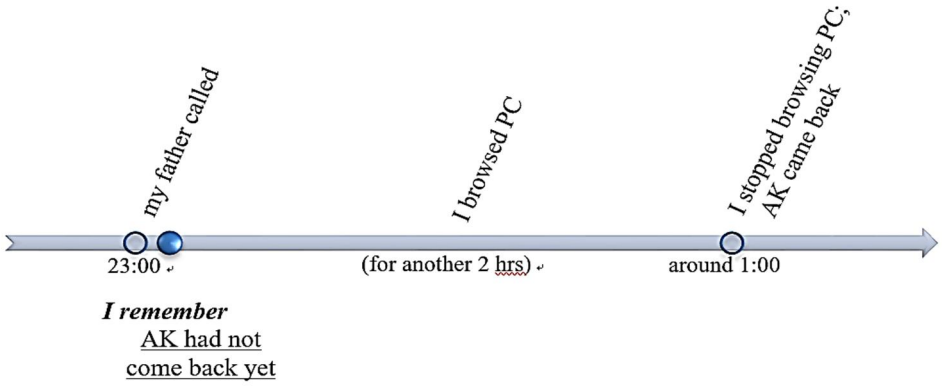

Fig. 5 Material process types in a timeline: *actualized* vs. *modalized*

Take Fig. 5, for example. Sollecito claims that Knox *had not come back yet* from 23:00 (when *[his] father called*) to 1:00 next day, which is approximately the time of the murder. This crucial claim indicates that he is not corroborating Knox's alibi. However, given that this claim is presented as a material event projected in a modal-world, as signaled by the epistemic modal verb *remember*, it thus needs to be marked as a situation which cannot (as yet) be confirmed.

The purpose of this distinction in our annotation is to reflect a speaker/writer's degree of epistemic (un)certainty regarding their propositions. Nevertheless, it is worth pointing out that such distinction *does not* guarantee that events reported as actual happenings in reality should necessarily be deemed as truthful. We will come back to this point when we analyze our case study in Section 6.

In brief, taking advantage of the heuristic value of visual representation, we will be able to draw out interesting features of arguments that were not previously readily apparent. As a result, we propose that this is likely to lead to improved performance in all reasoning tasks, including those routinely carried by police investigators, lawyers, and judges in criminal proceedings. In Section 6, we intend to demonstrate how this proposed model might work in practice.

6. Pilot Study: A Witness Statement from the Meredith Kercher Case

The text used for the purpose of illustration in this article is taken from the Meredith Kercher murder case. In this case, an American student Amanda Knox was charged, along with her Italian boyfriend, Raffaele Sollecito, with the murder of Kercher, her English flatmate in Italy. The case involves 3 countries and has attracted international media attention for more than 7 years. After 2 convictions and 2 acquittals, the dramatic legal battle was brought to an end definitively on March 27, 2015.

After the discovery of Kercher's body on November 2, 2007, Knox and Sollecito went to the police station several times to testify as a witness. In the late evening of November 5, Sollecito was called to the police station to clarify some inconsistencies in his original statement. When confronted by the police, Sollecito rescinded his original claim that Knox has been with him on the night of the murder, thereby removing his support for Knox's alibi. He now said that that she might have gone out on the night of the murder and he had been home alone. This section examines the cognitive structure and credibility of the statement Sollecito made to the police on November 5, 2007 (see Appendix, with sentences numbered for ease of reference. For the text-world analysis of the three controversial statements Knox made to the police on the same day, see Ho *et al.*, 2018).

It should be pointed out that, although the police interview is in a Q&A format, the investigators' questions are not recorded in the official written documents (neither

in the Italian original nor its English translation). This impedes our examination of the discourse from a holistic perspective. The analysis, therefore, focuses on Sollecito's responses to the police questioning.

6.1 Text-world Diagram and Process Types Analysis

In order to manually diagram the text-worlds projected in the statement under study, to illustrate the TWT approach to crime narrative analysis, we currently make use of VUE (Visual Understanding Environment), concept mapping software developed by Tufts University[9] (See Lugea, 2012, 2016; Ho *et al*., 2018 for the application of VUE in text-world diagramming). Based on the theoretical framework specified in Sections 2 and 3, here we outline the diagramming process and the cognitive structure underlying the statement.

In the text for analysis (see Appendix), the page header provides the background description that serves as an official record of this legal document. This information is recorded from a third-person point of view, from which we form our mental representation of the initial police interview context (tDW), a conceptual backdrop against which the main text, Raffaele Sollecito's first-person narrative account (TWs), unfolds. The tDW thus constitutes a starting point for the text-world diagram, which is constructed based on the world-builders specified at the header, that is, at the police station in Perugia, at 22:40 on November 5, 2007 and populated by the discourse participants, Raffaele Sollecito and four police officers.

We use a rectangular box to indicate the textual discourse-world (tDW), and rounded rectangular boxes shaded with different gray scales to indicate different kinds of text-worlds (TWs). Light gray is used to mark TW-*external/physical/actualized*. Dark gray indicates TW-*internal/mental/modalized* (i.e. modal-world). Horizontal curved arrows are used to indicate the occurrence of deictic word-switches. Vertical arrows indicate the continuity of a situational context, namely, the temporal shift that occurred within the same situational context (see 'frame modification' discussed in Section 3.1). Fig. 6 thus presents the overall cognitive structure of text-worlds of the statement under study.

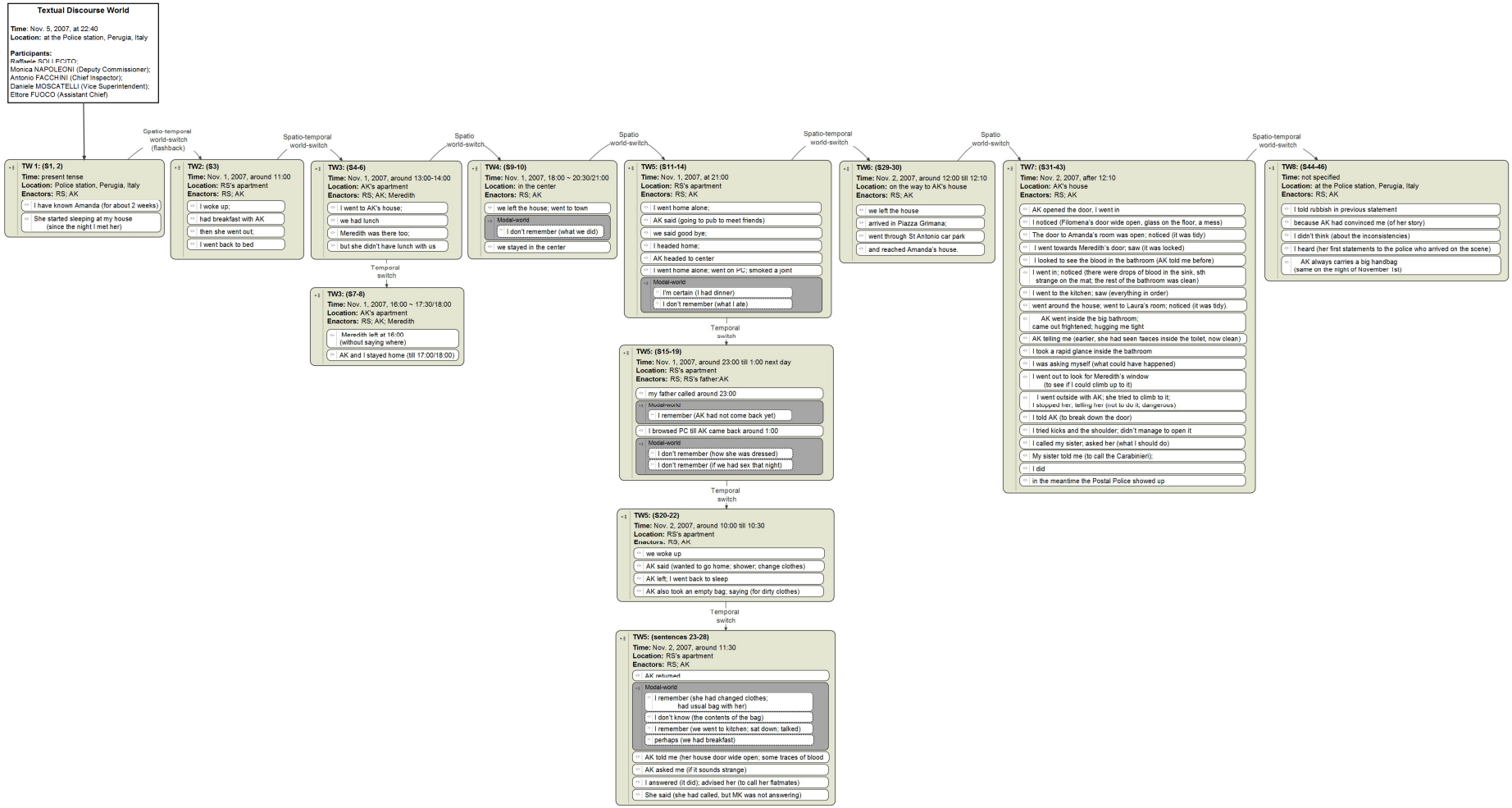

Fig. 6 Text-world diagram of the statement under study

There are total 46 sentences (872 words) in Sollecito's responses to the questioning. Sentences 1-2 (conceptualized as TW1) report his acquaintance with Knox; sentences 3-28 (TWs 2–5) present his flashbacks relating what happened on November 1, the night of the murder. Sentences 29-43 (TWs 6–7) then report what happened on the next day after the crime was revealed. Sentences 44-46 (TW8), arguably, could be seen as unframed texts, namely, de-contextualized, general, static description Sollecito made in the tDW context.

Along with the world-building, we also annotate 'what-happened' as reported or projected in the text-worlds. In our annotation of EVENTs process types, we take into account both **finite** and **non-finite forms of verbs** (Gerund: *V-ing*; Infinitive: *to V*; and *-ed* /*-ing* participles). For example, the clause *my sister told me to call the police* will be counted as having two verb groups and get two tags (verbal + material); whereas the sentence *Meredith left in a hurry without saying where she was going* is counted as having three verb groups and get three tags (material + verbal + material).

Fig. 7 shows the quantitative result of the total verb groups in the whole text and our manual annotation of the process types. As stated earlier, currently we use multiple tags as a means of dealing with indeterminate instances. The 13 verbs that are tagged with multiple process-types are categorized separately and will be discussed later.

Total V-groups (including non-finite V) in the text		**157**
Material / Behavioral	*doing* (actions) or *happening* (events)	89 (57%)
Relational / Existential	states of *being* or *having* or *existing*	20 (13%)
Verbal	*saying*	17 (11%)
Mental	*sensing* (perception, cognition, emotion)	18 (11%)
V with multiple tags		13 (8%)

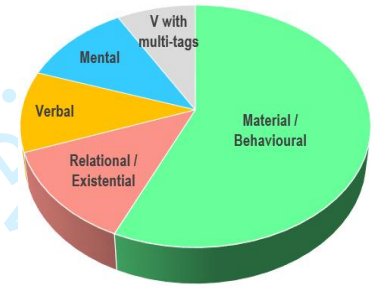

Fig. 7 Quantitative result of manual annotation

As shown in Fig. 7, the total number of verb groups in the text is 157, among which 11% are verbal process types distributed throughout the text (e.g. *Amanda told me…*; *she said…*) and 13% are relational process types (e.g. *there were drops of blood in the sink*; *her room was in a complete mess*), which mostly occur in sentences 31-43 (TW7) where Sollecito describes what they saw at the crime scene on the day Kercher's body was found.

The most salient types of process in this statement are those involving physical actions/events. 57% of the verb groups are material process types that Sollecito uses to report his (and Knox's) whereabouts, doings and happenings on the night of the murder

as well as the day after. In most of the material clauses, Sollecito and Knox are the 'actors', in Systemic Functional terms. For example:

> *On November 1st, I woke up at around 11, I had breakfast with Amanda then she went out and I went back to bed. Then around 13:00-14:00 I went to see her at her house…. Amanda and I stayed home until about 17:30-18:00.*

Given that most of the propositions are presented in a factual manner, this cognitive pattern indicates a high degree of epistemic certainty and confidence on the part of Sollecito with regard to the propositional content of his statement.

By contrast, mental process types that project what is inside Sollecito's mind (i.e. modal-worlds) account for only 11% of the total verb groups in the statement. If we observe Fig. 6 in more detail, we will see that most of the modal-worlds occur in TWs 4 and 5, in which Sollecito recounts what he and Knox did on the night of the murder.

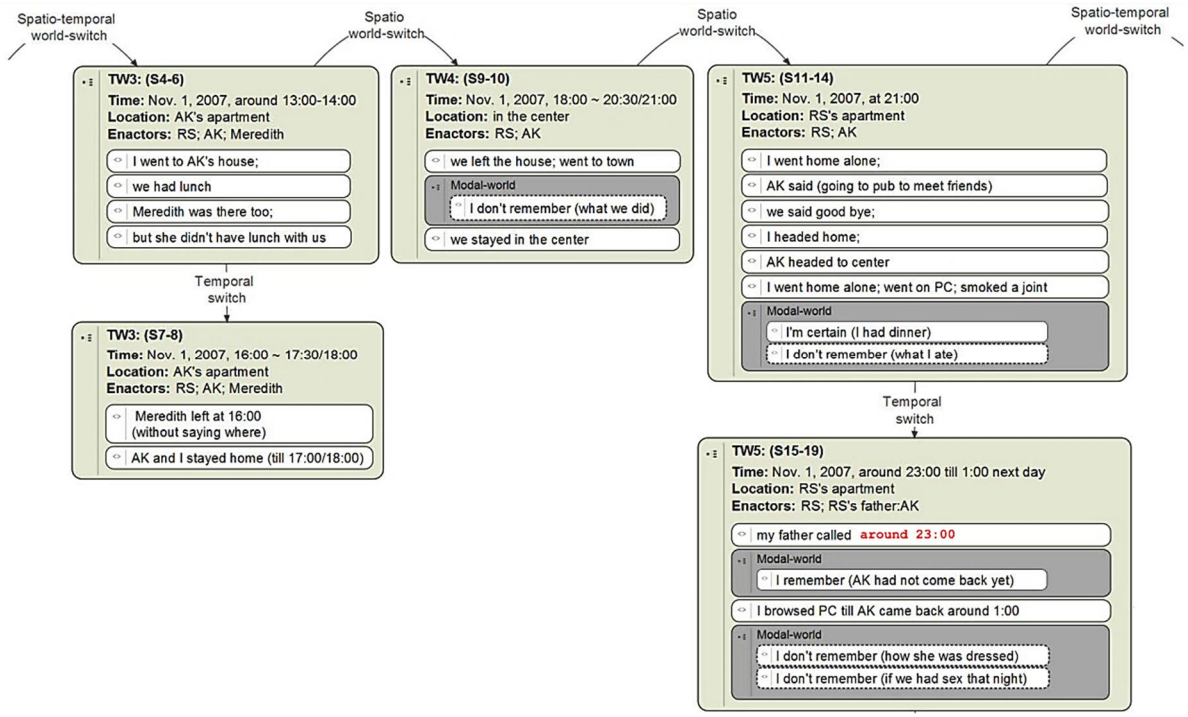

Fig. 8 Modal-worlds projected in the statement

As shown in the close-up view in Fig. 8, in the dark gray area that indicates TW-*internal/mental/modalized*, these mental process clauses convey Sollecito's various degrees of epistemic distance to the crucial claims (i.e. the projected clauses) he made in the statement. For example, in TW4, he claims that he has no memory at all of *what [they] did* during the three hours in the center; whereas in TW5, with regard to Knox's alibi, he *remembers* that at around 23:00 when his father called *[Amanda had not come back yet]*. As mentioned earlier, given that this type of mental projection is inaccessible to discourse participants, such claims thus need to be marked as a situation which cannot (as yet) be confirmed. Their credibility or factuality thus needs to be examined further.

Now, let us turn to examine in detail the 13 verbs that are tagged with more than one process-type. 11 of the 13 instances are tagged as mental/material in that they are mental-state verbs that express a conscious physical act involved in perception (through the five senses, e.g. *seeing*, *hearing*), as listed below:

- *she found* [the entrance door wide open and some traces of blood…]
- *I noticed* [that Filomena's door was wide open with some glass on the floor…]
- *I noticed* [that it was tidy].
- *I… saw* [that it was locked].
- *I noticed* [there were drops of blood in the sink…]
- *I… saw* [that everything was in order]
- *I… noticed* [it was tidy].
- *she had seen* [feces inside the toilet]
- *I just took a rapid glance* inside the bathroom
- *to see* [if I could climb up to it]
- *I heard* [the first statements]

As defined in SFG, a mental process clause always involves a *Senser* (the one that feels, thinks, wants or perceives) and a *Phenomenon* (something that is felt, thought, wanted or perceived), which may be a person, a concrete object, a thing or a 'fact' (Halliday and Matthiessen, 2004: 208). The term 'fact' here is defined as a type of embedded projection where no Senser is involved (see the examples above). Facts can be sensed – perceived or felt – but they cannot do anything or having anything done to them. (NB: it is important to realize that 'facts', as the term is used here, have nothing to do with 'truth'.)

The common pattern shown in the 11 mental/material instances is that they are all *perceptive* mental process involving the perceiving phenomena that are presented as 'facts' (as introduced via attributive relational or existential clauses). They have properties that set them apart from the other subtypes of mental-state verbs. That is, unlike the other cognitive or emotive verbs that project Sollecito's internal 'inaccessible' mental states, the verbs of perception here (e.g. *found*, *noticed*, *saw*, *seen*, *heard*) project specific and definite existing phenomena (e.g. *door wide open/locked*, *drops of blood*, *feces*) that are actually available for outside evaluation or verification. Thus, in the mental-material continuum, these examples could arguably be categorized as near-material, given that the embedded clauses are not expressions of his mental states, but the existing objects or situations in the external world.

In summary, the overall cognitive pattern of Sollecito's statement indicates a high degree of confidence with regard to his knowledge and belief of the claims he made in the statement. More than 80% of information is presented in a factual manner or as factual information. Now we shall proceed to examine the probative value of the information contained in this statement.

6.2 *Temporal EVENTs Cross Comparison*

In Section 5, we proposed to develop a timeline visualization output that will re-sort the annotated EVENTs (i.e. process types) into a chronological sequence and place them on a timeline, so as to reconstruct the story line in a crime narrative. It is envisaged that, if we anchor a witness/suspect's reported actions/events on a timeline and have it in parallel comparison with timelines of events reported by any other witnesses, as well as with a timeline of hard facts (e.g. CCTV, PC or phone call records), we will then be able to readily detect any contentious information, for example, clashes in time and/or location.

As discussed in Section 6.1, the cognitive structure of the statement and the process type analysis demonstrate Sollecito's high degree of confidence regarding his claims. At first sight, it seems that none of these propositions contradict any of the others. They all seem logically consistent and could conceivably be true. In this section, we thus intend to demonstrate how a cross comparison of different timelines in parallel might be of significant value in the reconstruction and comparison of contentious events.

Following Sollecito's statement, Fig. 9 shows his whereabouts and doings on the night of the murder (see TWs 3–5), in comparison with the other two parallel timelines: a timeline of events reported by a witness named Popovic[10] and a timeline of hard facts.

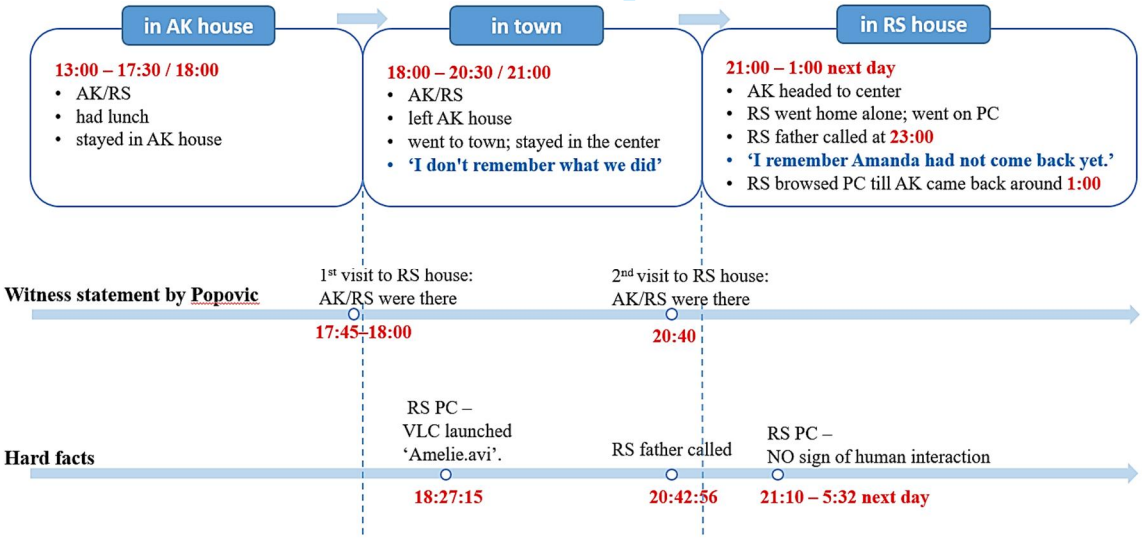

Fig. 9 EVENTs timelines: Sollecito vs. Popovic and hard facts

Sollecito claims that he spent the whole afternoon (from 13:00 until about 17:30/18:00) with Knox at her house. They then went to town and stayed there from 18:00 to 20:30/21:00. Afterward, he went home alone without Knox and browsed his PC from 21:00 till 01:00 next day when Knox came back to his place. After cross comparison of the timelines, four crucial contradictions could be observed:

(1) The witness Popovic (hearing of March 21, 2009) testified that on the evening of November 1 she went to Sollecito's house twice, and on both occasions, both Knox and Sollecito were at home. She testified that her mother was sending her a suitcase via a coach from Milan, which was expected to arrive in Perugia at midnight. She thus stopped by Sollecito's at **around 17:45–18:00** to ask for his help.[11] Later, her mother called her again saying that she did not manage to send the suitcase. So Popovic, after finishing her class at 20:20, walked to Sollecito's house and arrived at around **20:40**, to tell him that she no longer needed the help. Popovic testified that on both occasions it was Knox who answered the door. These two temporal events reported by Popovic clash with Sollecito's claim that they were both in town (i.e. not at Sollecito's house) from 18:00 to 21:00.

(2) As shown in the timeline of hard facts, Sollecito's computer shows a record of VLC media player launching a movie 'Amelie.avi' at **18:27:15**, which suggests that the couple were in Sollecito's house at that particular time.

(3) Sollecito claimed that his father called him 'at [his] home number' at around 23:00. This temporal event is presented by Sollecito as 'actual happening' with great epistemic certainty, and yet, in fact, the phone record shows that his father called him at **20:42:56**. Moreover, it should be noted that the time of Popovic's second visit (**20:40**) seems to coincide with the actual telephone call time of Sollecito's father. Both events seemingly point to a higher credibility that Sollecito and Knox were at his house at that particular time.

(4) Sollecito claims that he browsed his computer alone at home from 21:00 till 1:00 next day. However, the record of his computer shows no signs of human interaction from **21:10 till 5:32 next day**, a hard fact that contradicts his claim.

As illustrated in Fig. 9, multiple timelines for parallel comparison could be of great use in detecting conflicting claims and evaluating whether a statement is reliable enough to be deemed admissible as evidence in a criminal trial. Our analysis has shown that Sollecito's statement, which was admitted as evidence to the court in the criminal prosecution, may not be as reliable as it appears to be. Taking advantage of the heuristic value of visual representation of propositional content, we are able to draw out interesting features of arguments that were not readily apparent. We believe that our proposed approach, in the long term, is likely to be of value to forensic linguists, police officers and legal practitioners in improving criminal investigation and evidential decision-making process and thereby potentially preventing miscarriages of justice.

7. Conclusion and Future Work

This article reports on the work we have done to develop an annotation scheme based on Text World Theory to support the visualization of complex language data. The annotation scheme underpins our computational development of *Worldbuilder*, aiming to annotate and visualize the complex and often competing crime narratives produced in legal discourse. Using a statement from the Meredith Kercher murder trial as a case study, we demonstrate how the revised TWT framework provides a comprehensive toolkit for accounting for all the facts that are put forward in a legal case and enables the identification of specific linguistic elements in texts that are important for computational visualizations. This cognitive discourse model has been greatly reinforced by Contextual Frame Theory (Emmott 1997), an important augmentation that allows us to account for the complexities of cognitive text processing for visualization purposes. Bridging Cognitive and Computational Linguistics, our research improves the TWT model's analytical accuracy and yields a potentially useful tool for forensic work.

There remains work to be done in improving the reliability and applicability of the text-world analytical model and advancing the *Worldbuilder* tool. The challenges include: (i) improving the scale, complexity, and rigor of data annotation, from manual tagging by human researchers at the current stage to the future development of automatic annotation, and (ii) technicality and effectiveness of information visualizations, for example, more intuitive and accessible visual representations which can simplify the communication of complex information. More sophisticated data quantification also needs to be developed for improving the accuracy of linguistic evidential analysis. We suggest that addressing these challenges should be an aim of future research.

Funding

This work was supported by the University of Huddersfield [URF207].

References

Alonso, O., Baeza-yates, R., Strötgen, J., and Gertz, M. (2011). Temporal Information Retrieval: Challenges and Opportunities, *1st Temporal Web Analytics Workshop*, pp. 1-8.

Barnden, J. A. (2006). Artificial Intelligence, Figurative Language and Cognitive Linguistics. In Kristiansen, G., Achard, M., Dirven, R., and Ruiz de Mendoza Ibáñez, F. (eds), *Cognitive Linguistics: Current Applications and Future Perspectives*. Berlin: Mouton de Gruyter, pp. 431-459.

Duchan, J. F., Bruder, G. A., and Hewitt, L. E. (eds) (1995). *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale: Lawrence Erlbaum.

Emmott, C. (1997). *Narrative Comprehension: A Discourse Perspective.* Oxford: Oxford University Press.

Feldman, J. A. (2004). Computational Cognitive Linguistics, *Proceedings of the 20th International Conference on Computational Linguistics*, Geneva, Switzerland, pp. 1114-1118.

Gavins, J. (2005). (Re)Thinking Modality: A Text-World Perspective. *Journal of Literary Semantics,* **34**(2): 79-93.

Gavins, J. (2007). *Text World Theory: An Introduction.* Edinburgh: Edinburgh University Press.

Halliday, M. (1985). *An Introduction to Functional Grammar.* London: Arnold.

Halliday, M. (1994). *An Introduction to Functional Grammar.* London: Arnold.

Halliday, M., and Matthiessen, C. (2004). *An Introduction to Functional Grammar.* London: Routledge.

Herman, D. (2002). *Story Logic: Problems and Possibilities of Narrative.* Lincoln: University of Nebraska Press.

Hidalgo-Downing, L. (2000). *Negation, Text Worlds and Discourse: The Pragmatics of Fiction.* Stamford, CT: Ablex.

Hidalgo-Downing, L. (2002). Creating Things That Are Not: The Role of Negation in the Poetry of Wisława Szymborska. *Journal of Literary Semantics*, **31**(2): 113-132.

Ho, Y., Lugea, J., McIntyre, D., Wang, J., and Xu, Z. (forthcoming 2018). Projecting Uncertainty: Text-Worlds in Three Statements from the Meredith Kercher Murder Case. *English Text Construction*, **11**(2).

Kristiansen, G., Achard, M., Dirven, R., and Ruiz de Mendoza Ibáñez, F. (eds) (2006). *Cognitive Linguistics: Current Applications and Future Perspectives*. Berlin: Mouton de Gruyter.

Lugea, J. (2012). Building a Narrative Text-World through Deixis and Modality: A Contrastive Study of Spanish and English. Ph.D. thesis, Queen's University Belfast.

Lugea, J. (2016). *World-Building in Spanish and English Spoken Narratives.* London: Bloomsbury.

Mahlberg, M., Stockwell, P., de Joode, J., Smith, C., and O'Donnell, M. B. (2017). Clic Dickens: Novel Uses of Concordances for the Integration of Corpus Stylistics and Cognitive Poetics. *Corpora*, **11**(3): 433-463.

Mani, I., Pustejovsky, J., and Gaizauskas, R. (eds) (2005). *The Language of Time: A Reader*. Oxford: Oxford University Press.

Manzanares, J. V. (2010). Cognitive Linguistics and Computational Modeling. *Textus*, **23**(3): 763-794.

Nahajec, L. (2009). Negation and the Creation of Implicit Meaning in Poetry. *Language and Literature*, **18**(2): 109-127.

Narayanan, S. (1999a). Moving Right Along: A Computational Model of Metaphoric Reasoning About Events, *Proceedings of the National Conference on Artificial Intelligence AAAI-99*, Orlando, pp. 121-127.

Narayanan, S. (1999b). Reasoning About Actions in Narrative Understanding, *Proceedings of the International Joint Conference on Artificial Intelligence IJCAI-99*, Stockholm, pp. 350-357.

O'Donnell, M., Zappavigna, M., and Whitelaw, C. (2008). A Survey of Process Type Classification over Difficult Cases. In Jones, C., and Ventola, E. (eds), *From Language to Multimodality - New Developments in the Study of Ideational Meaning*. Equinox Publishing, pp. 47-64

Olivier, P., and Tsujii, J. i. (1994). A Computational View of the Cognitive Semantics of Spatial Prepositions, *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, New Mexico, pp. 303-309.

Pustejovsky, J. (2005). Time and the Semantic Web, *Proceedings of the 12th International Symposium on Temporal Representation and Reasoning*, Burlington, VT, USA, pp. 5-8.

Quirk, R., and Greenbaum, S. (1973). *A University Grammar of English*. London: Longman.

Ryan, M.-L. (1991). *Possible Worlds, Artificial Intelligence and Narrative Theory.* Bloomington: Indiana University Press.

Segal, E. M. (1995a). A Cognitive-Phenomenological Theory of Fictional Narrative. In Duchan, J. F., Bruder, G. A., and Hewitt, L. E. (eds), *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale: Lawrence Erlbaum, pp. 61-78.

Segal, E. M. (1995b). Narrative Comprehension and the Role of Deictic Shift Theory. In Duchan, J. F., Bruder, G. A., and Hewitt, L. E. (eds), *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale: Lawrence Erlbaum, pp. 3-17.

Setzer, A., and Gaizauskas, R. (2001). A Pilot Study on Annotating Temporal Relations in Text, *Proceedings of the ACL 2001 Workshop on Temporal and Spatial Information Processing - Volume 13*, Toulose, France, pp. 88-95.

Thompson, G. (2014). *Introducing Functional Grammar.* London: Routledge.

Veale, T. (2006). Computability as a Test on Linguistic Theories. In Kristiansen, G., Achard, M., Dirven, R., and Ruiz de Mendoza Ibáñez, F. (eds), *Cognitive Linguistics: Current Applications and Future Perspectives*. Berlin: Mouton de Gruyter, pp. 461-483.

Veale, T., and O'Donoghue, D. P. (2000). Computation and Blending. *Cognitive*

Linguistics, **11**(3): 253-281.

Verdonk, P. (2008). Book Review: Text Worlds Theory: An Introduction. *Language and Literature*, **17**(4): 373-376.

Wang, J., Ho, Y., Xu, Z., McIntyre, D., and Lugea, J. (2016a). The Visualisation of Cognitive Structures in Forensic Statements, *Proceedings of IV2016: 20th International Conference Information Visualisation*, Lisbon, Portugal, pp. 106-111.

Wang, J., Ho, Y., Xu, Z., McIntyre, D., and Lugea, J. (2016b). A Visualization Method for Understanding Forensic Statements. *IEEE Information Visualization (InfoVis)*, Baltimore, Maryland, USA, 23-28 October.

Werth, P. (1994). Extended Metaphor - a Text-World Account. *Language and Literature*, **3**(2): 79-104.

Werth, P. (1995a). How to Build a World (in a Lot Less Than Six Days and Using Only What's in Your Head). In Green, K. (ed), *New Essays in Deixis: Discourse, Narrative, Literature*. Amsterdam: Rodopi, pp. 49-80.

Werth, P. (1995b). 'World Enough, and Time': Deictic Space and the Interpretation of Prose. In Verdonk, P., and Weber, J. J. (eds), *Twentieth-Century Fiction: From Text to Context*. London: Routledge, pp. 181-205.

Werth, P. (1999). *Text Worlds: Representing Conceptual Space in Discourse*. London: Longman.

Appendix: R. Sollecito's statement to police (Italian version and English translation)

Questura di Perugia
Squadra Mobile
Area Affari Generali

OGGETTO: Verbale di sommarie informazioni da persona informata sui fatti rese da:
SOLLECITO Raffaele, già generalizzato.

Il 05 novembre 2007 alle ore 22.40 , negli uffici della Squadra Mobile della Questura di Perugia. Innanzi ai sottoscritti Ufficiali di P.G. Sost. Comm. Monica NAPOLEONI, Isp. Capo Antonio FACCHINI V. Sov. Daniele MOSCATELLI, Ass. Capo Ettore FUOCO è presente è presente il nominato in oggetto il quale, ad integrazione delle dichiarazioni rese il novembre presso questi Uffici, in merito ai fatti per cui è indagine, dichiara quanto segue:----------

A.D.R. conosco Amanda da circa due settimane. Dalla sera stessa in cui l'ho conosciuta lei ha iniziato a dormire a casa mia.----------

A.D.R. Il giorno 01 novembre u.s. mi sono svegliato verso le 11.00, ho fatto colazione con Amanda poi lei è uscita e io sono tornato a letto. Poi verso le ore 13-14 l'ho raggiunta a casa sua. Nella circostanza c'era anche Meredih. Abbiamo pranzato insieme io ed Amanda mentre Meredith non ha mangiato con noi.----------

A.D.R. Meredith verso le ore 16.00 è uscita frettolosamente senza dire dove andasse. Io e Amanda siamo restati a casa fino a verso le ore 17.30 – 18.00.----------

A.D.R Siamo usciti da casa, siamo andato in centro, ma non ricordo cosa abbiamo fatto.----------

A.D.R. Siamo rimasti in centro dalle ore 18.00 fino alle ore 20.30/21.00. Io alle ore 21.00 sono andato a casa mia da solo, mentre Amanda mi ha detto che sarebbe andata al pub Le Chic, perché voleva incontrare dei suoi amici.----------

A.D.R A questo punto ci siamo salutati io mi sono diretto verso casa mia mentre lei si è diretta verso il centro.----------

A.D.R. Sono andato a casa da solo, mi sono messo al computer e mi sono fatto una canna. Senz'altro ho cenato ma non ricordo cosa ha mangiato. Verso le ore 23.00 mi ha chiamato, sull' utenza fissa di casa mia 075.9660789, mio padre. Nella circostanza ricordo che Amanda ancora non era tornata.----------

A.D.R Ho navigato al computer per altre due ore circa dopo la telefonata di mio padre ed ho smesso solo quando Amanda è rientrata presumibilmente verso le ore 1.00.----------

A.D.R Non ricordo bene come fosse vestita e se era vestita allo stesso modo di quando ci siamo salutati prima di cena.----------

A.D.R. Non ricordo se quella sera abbiamo consumato un rapporto sessuale.----------

A.D.R. La mattina successiva verso le ore 10,00 ci siamo svegliati, lei mi ha detto che voleva andare a casa a farsi una doccia e cambiarsi gli abiti.----------

A.D.R. Infatti verso le ore 10.30 è uscita ed io mi sono rimesso a dormire.----------

A.D.R. Quando è uscita la mattina per andare a casa sua, Amanda ha preso anche una busta vuota, dicendomi che gli sarebbe servita per metterci i panni sporchi.----------

A.D.R. Verso le ore 11.30 è ritornata a casa e ricordo che si era cambiata i vestiti; aveva con se la a sua solita borsa.----------

A.D.R Non sono a conoscenza del contenuto della sua borsa.----------

Raffaele Sollecito

A.D.R Ricordo che siamo andati subito in cucina, ci siamo seduti, e abbiamo parlato per un po', forse abbiamo fatto colazione. Nella circostanza Amanda mi ha raccontato che quando è arrivata a casa sua ha trovato la porta d'ingresso spalancata, e delle tracce di sangue nel bagno piccolo e mi ha chiesto se la cosa mi sembrava strana. Io gli ho risposto di si e gli ho consigliato anche di telefonare alle sue amiche. Lei mi ha detto che ha telefonato a Filomena, mentre ha detto che Meredith non gli rispondeva.----------

A.D.R.Verso le ore 12,00 siamo usciti da casa, percorrendo corso Garibaldi, siamo arrivati a Piazza Grimana, poi siamo passati per il parcheggio di Sant'Antonio e siamo giunti presso l'abitazione di Amanda. Per fare il tragitto ci abbiamo impiegato circa 10 minuti. ----------

A.D.R Appena siamo arrivati lei ha aperto la porta con le chiavi, sono entrato all'interno ed ho notato che la porta di Filomena era tutta spalancata con dei vetri per terra e la sua stanza era tutta in disordine. La porta della camera di Amanda era aperta ed ho notato che era tutto in ordine. Poi sono andato verso la porta di Meredith ed ho visto che era chiusa a chiave, prima di fare questo avevo guardato se fosse vero quello che mi aveva detto Amanda circa il sangue all'interno del bagno. Infatti sono entrato ed ho notato che c'erano delle gocce di sangue sul lavandino e invece sul tappetino ho notato che vi era qualcosa di strano, una sorta di mista acqua e sangue, il resto del bagno era pulito.----------

A.D.R. Sono andato in cucina ho visto che era tutto in ordine, poi mi sono fatto un giro per il resto della casa, sono andato in camera di Laura ed ho notato che era tutto in ordine. In quel mentre Amanda entrava nel bagno grande, adiacente alla cucina, ed usciva spaventata e mi abbracciava forte, dicendomi che prima, quando aveva fatto la doccia, aveva visto delle feci all'interno del water, mentre adesso il water era pulito.----------

A.D.R Io ho solo dato una rapida occhiata all'interno del bagno, fidandomi di quello che mi diceva Amanda.----------

A.D.R. A quel punto mi sono chiesto cosa stesse succedendo, sono uscito dalla casa per cercare la finestra della camera di Meredith per vedere se riuscivo ad arrampicarmi. Sono uscito fuori insieme ad Amanda e lei ha provato ad arrampicarsi, io l'ho subito bloccata dicendole di non farlo in quanto era pericoloso. Poi ho detto ad Amanda che la soluzione migliore era quella di sfondare la porta, ho provato a dare calci e spallate, ma non sono riuscito ad aprirla. Dopodiché ho chiamato mia sorella al cellulare e mi sono consigliato su cosa potevo fare, essendo lei un Tenente dei Carabinieri. Mia sorella m'ha detto di chiamare il 112, cosa che io ho fatto, ma nel frattempo è arrivata la Polizia Postale.----------

A.D.R. Vi ho riferito nel precedente verbale un sacco di cazzate, perché lei mi aveva convinto della sua versione dei fatti e non ho pensato alle incongruenze. Io ho sentito le prime dichiarazioni che ha fatto alla Polizia Postale, intervenuta sul posto.----------

A.D.R. Lei solitamente porta sempre con se una grande borsa che aveva anche la sera del 1 novembre.----------

I verbalizzanti danno atto che il verbale viene chiuso alle ore 03.30 del 6 novembre 2007.----------

L.C.S.----------

Raffaele Sollecito

SUBJECT: Witness statement of person informed of the facts given by SOLLECITO Raffaele, already identified.

On November 5th 2007 at 22:40 in the offices of the Flying Squad of the Perugia Police Headquarters. Before the undersigned of the Criminal Investigation Dept. Deputy Commissioner MONICA NAPOLEONI, Chief Inspector Antonio FACCHINI Vice Superintendent of Police Daniele MOSCATELLI, Assistant Chief Ettore FUOCO is present the above-mentioned who, to supplement the declarations made [November] in these Offices, in regards to the facts being investigated, declares as follows:

I have known Amanda for about two weeks.
From the night that I met her she started sleeping at my house.
On November 1st, I woke up at around 11, I had breakfast with Amanda then she went out and I went back to bed.
Then around 13:00-14:00 I went to see her at her house.
Meredith was there too.
Amanda and I had lunch while Meredith did not have lunch with us.
Around 16:00 Meredith left in a hurry without saying where she was going.
Amanda and I stayed home until about 17:30-18:00.
We left the house, we went into town, but I don't remember what we did.
We stayed in the centre from 18:00 until 20:30/21:00.
At 21:00 I went home alone because Amanda told me that she was going to go to the pub Le Chic because she wanted to meet some friends.
At this point we said goodbye and I headed home while she headed towards the centre.
I went home alone, went on the computer and smoked a joint.
I'm certain I had dinner but I don't remember what I ate.
Around 23:00 my father called at my home number 0759660789.
During that time I remember Amanda had not come back yet.
I browsed at my computer for another two hours after my father's phone call and only stopped when Amanda came back presumably around 1:00.
I don't remember how she was dressed and if she was dressed the same way as when we said goodbye before dinner.
I don't remember if we had sex that night.
The following morning around 10:00 we woke up, she told me she wanted to go home and take a shower and change clothes.
In fact at around 10:30 she left and I went back to sleep.
When she left that morning to go to her house, Amanda also took an empty bag telling me she needed it for dirty clothes.
At around 11:30 she returned and I remember she had changed clothes; she had her usual bag with her.
I don't know the contents of her bag.
I remember we immediately went to the kitchen, we sat down and talked for a while, perhaps we had breakfast.

[26] In that circumstance Amanda told me that when she got to her house she found the entrance door wide open and some traces of blood in the small bathroom and she asked me if it sounded strange.
[27] I answered that it did and I also advised her to call her flatmates.
[28] She said she had called Filomena but that Meredith was not answering.
[29] At around 12:00 we left the house; passing through Corso Garibaldi we arrived in Piazza Grimana, then we went through the Sant' Antonio car park and reached Amanda's house.
[30] To walk there it took us about 10 minutes.
[31] As soon as we got there she opened the door with her keys, I went in and I noticed that Filomena's door was wide open with some glass on the floor and her room was in a complete mess.
[32] The door to Amanda's room was open and I noticed that it was tidy.
[33] Then I went towards Meredith's door and saw that it was locked.
[34] Before this I looked to see if it was true what Amanda had told me about the blood in the bathroom.
[35] In fact I went in and I noticed there were drops of blood in the sink, while on the mat there was something strange – like a mixture of blood and water, while the rest of the bathroom was clean.
[36] I went to the kitchen and saw that everything was in order, then went around the rest of the house, I went to Laura's room and noticed it was tidy.
[37] In that moment Amanda went inside the big bathroom, next to the kitchen and came out frightened and hugging me tight telling me that earlier, when she took the shower, she had seen feces inside the toilet, while now the toilet was clean.
[38] I just took a rapid glance inside the bathroom trusting what Amanda had told me.
[39] At that point I was asking myself what could have happened and I went out to look for Meredith's window to see if I could climb up to it.
[40] I went outside with Amanda and she tried to climb to it, I immediately stopped her telling her to not do it because it was dangerous.
[41] I then told Amanda that the best solution was to break down the door, I tried kicks and the shoulder but I didn't manage to open it.
[42] Then I called my sister on her cellphone and asked her what I should do since she is a Carabinieri lieutenant.
[43] My sister told me to call the Carabinieri (112, the Italian emergency number), which I did, but in the meantime the Postal Police showed up.
[44] In my previous statement I told a load of rubbish because Amanda had convinced me of her version of the facts and I didn't think about the inconsistencies.
[45] I heard the first statements that she made to the Postal Police who arrived on the scene.
[46] Usually she always carries a big handbag that she also had the night of November 1st.

The investigating officials acknowledge that the deposition ends at 3:30 (AM) of November 6th 2007.

Notes

[1] In cognitive linguistics, meaning is conceptual and embodied, which thus raises a fundamental issue about the 'cognitive' nature of our proposed cognitive computational modelling: 'who' is conceptualizing the text-world structure that is identified and visualized via the software? It should be clarified at the outset that we do not use TWT in its usual sense as a model that attempts to reconstruct the discourse processing of *all* real participants. That is to say, we do not consider the reader's subjectivity, schematic knowledge *etc.* in building the text-world. Our model follows the tradition in Computational Cognition (including the fields of Artificial Intelligence and Computational Modelling), in which the mind is understood as analogous with a computer, where discourse processing can be described through TWT. We are in fact using the descriptors of the TWT cognitive framework to operationalize a formal analysis of the discourse-level structures of the text. We intend to provide a systematic linguistic account of the text-driven conceptual structures, to help legal practitioners reflect on the story and information provided by a witness/suspect under interrogation (i.e. one tDW participant). The text-world diagramming thus mainly focuses on the text-worlds as constructed by this particular tDW participant, via closely following his/her spatial and temporal deixis. Take sentence 3, for example: *On November 1st, I* [Sollecito] *woke up and had breakfast with Amanda then she went out and I went back to bed*. The primed text-world would be *I woke up and had breakfast* and then *I went back to bed*, instead of *she* [Amanda] *went out* (which presumably engenders a mental construct of another text-world).

[2] We aim to develop *Worldbuilder*, a web-based visual inference system, that will enable the systematic annotation of data using categories derived from TWT. This system will also comprise data quantification and visualization features that will render the annotated

propositions contained within the language data easy for readers (and legal practitioners) to identify patterns. Given that the features of sophisticated data quantification and intuitive information visualization are still underdeveloped in *Worldbuilder* at the current stage, we use open source software (VUE, Visual Understanding Environment) for our manual text-world analysis and to illustrate the vision behind our computational development.

[3] In Werth's original framework, text-world may be split into different types of *sub-worlds*. Gavins (2005, 2007) argues that the deictic parameters established in the initial text-world will change constantly throughout the discourse, and that multiple text-worlds may occur and may not necessarily be subordinate to the original text-world. She thus argues for a replacement of Werth's *sub-worlds* with a non-hierarchical alternative term, *world-switches*.

[4] In Example 2, from a TWT perspective, this temporal switch also involves a change in the ontological status of the other world-builders, in this case, the enactors. As the enactors at 2pm are not the same set of enactors as those at 4pm, it could be argued that the situational context is no longer the same. It should be clarified that in our cognitive computational modelling for crime narrative analysis, we aim to understand *who* is doing *what* with whom in a particular *place* at a particular *time* described in a text. To this end, text-world conceptual structures may vary according to the reader's 'focus' of attention. In our model, we focus on tracking the story/information provided by the 'who', that is, diagramming the text-worlds as constructed by the tDW participant (who is being referred to in the text as the main enactor, in this case, *Sollecito*). Therefore, we argue from a CFT perspective that the movement from 2pm to 4pm is a form of 'frame modification' as we assume that the main enactors (i.e. *Sollecito* and *Amanda*) are bound to the frame (the

same text-world) in the interim period.

[5] The distinction between framed and unframed text lies in the difference between (contextualized) story and (decontextualized) information. It is an important distinction that has not been covered in TWT framework. It is also a distinction that can be difficult to draw at times. In this article, the authors do not argue that 'unframed text' is a new type of text-world in the TWT framework. Instead, we contend that it is a type of text that may occur in any narratives and thus needs to be considered in our annotation scheme. It should be noted that 'unframed text' may occur in the main narrative level (i.e. tDW) or contain within a TW.

[6] It should be noted that negation is currently not included in our annotation scheme. The feature of negation (e.g., *Meredith did not have lunch with us* or *Amanda had not come back yet*) is particularly interesting when considered from a text-world perspective. Negated proposition engenders the creation of a separate world in which the negated event or situation is represented. It is generally agreed that negation involves the formation of a complex structure with regard to the corresponding affirmative, and this complexity is evident in the different perspectives from which negation has been studied (e.g., Gavins, 2007; Hidalgo-Downing, 2000, 2002; Nahajec 2009). The authors are aware of the significant implication of negation in legal discourse. However, to avoid unnecessary complexity in data annotation and visualization, which is done manually at the current stage, we opt to keep such information in the matrix text-world.

[7] Hereafter we use EVENTs in capital letters as a cover term for all the process types presented in a text. See the annotation of function-advancing propositions.

[8] Verbal process types may sometimes be a source of confusion in annotation (see O'Donnell *et al*., 2008). Various aspects of physical action (and mental state in some cases) may be encoded in verbs of saying (e.g. *criticize*, *thank*, *scream*, *shout*) and result in indeterminate cases. However, they do display distinctive patterns of their own and play an important role in legal discourse, for example, witness/suspect reporting or attributing information to sources. Thus, despite the potential indeterminacy, we decide that verbal process types should be annotated separately.

[9] VUE (http://vue.tufts.edu/) is free to download and compatible with all operating systems. It provides a flexible visual environment for structuring, presenting and sharing digital information.

[10] Jovana Popovic's statement was taken from The Murder of Meredith Kercher Wiki Site, a website created by a group of volunteer editors (some are professional translators and some with the expertise in certain areas, such as forensics, DNA, IT or criminal law), to inform the English-speaking world about the case by providing a unique collection of translations of original documents and evidence presented at trial (accessed 20 December 2017).

[11] This piece of information can be seen in her testimony: 'Quindi il primo novembre passato a casa di Raffaele tipo era 5 e 45 pomeriggio, comunque prima delle 6, perché alle 6 c'avevo altro appuntamento, ho passato a casa di loro e ho chiesto favore ad accompagnarmi alla stazione.'

http://themurderofmeredithkercher.com/Jovana_Popovic%27s_Testimony (accessed 20 December 2017).

Text-world annotation and visualization for crime narrative reconstruction

List of Tables:

List of Figures:

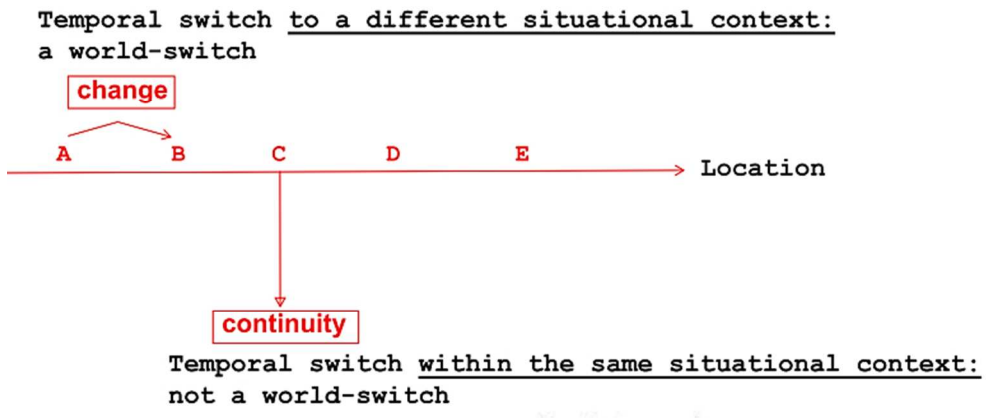

Fig. 1 Inferring context continuity or change

208x87mm (96 x 96 DPI)

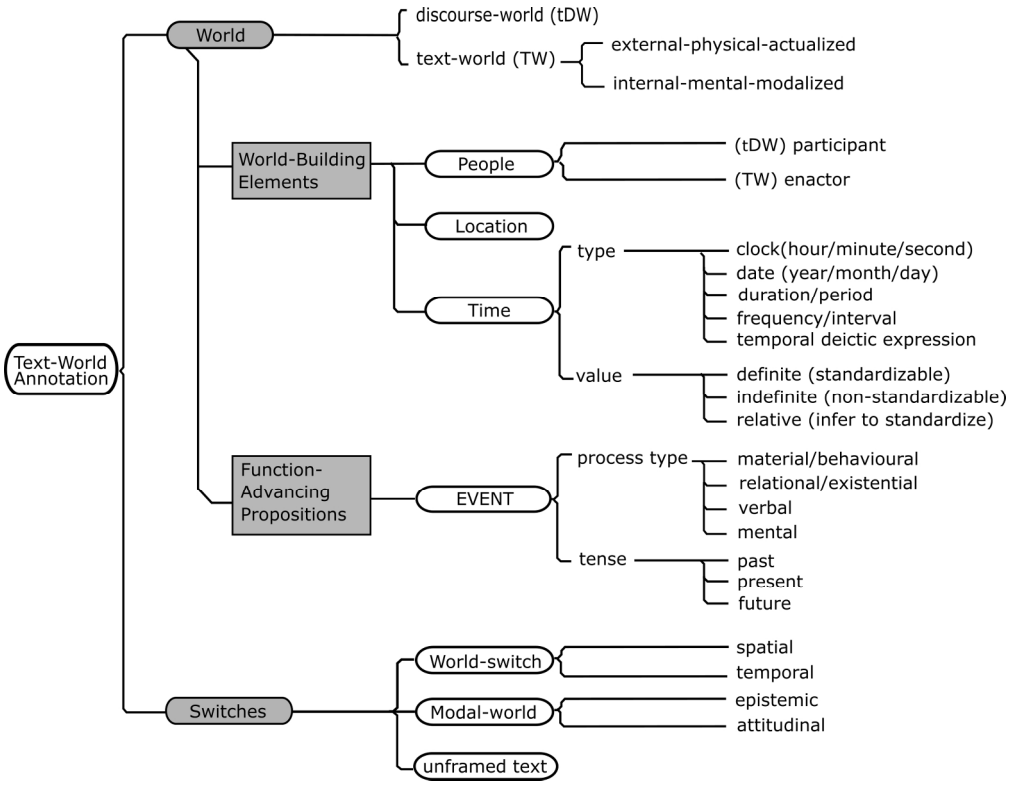

Fig. 2 TWT structured framework for data annotation

380x294mm (144 x 144 DPI)

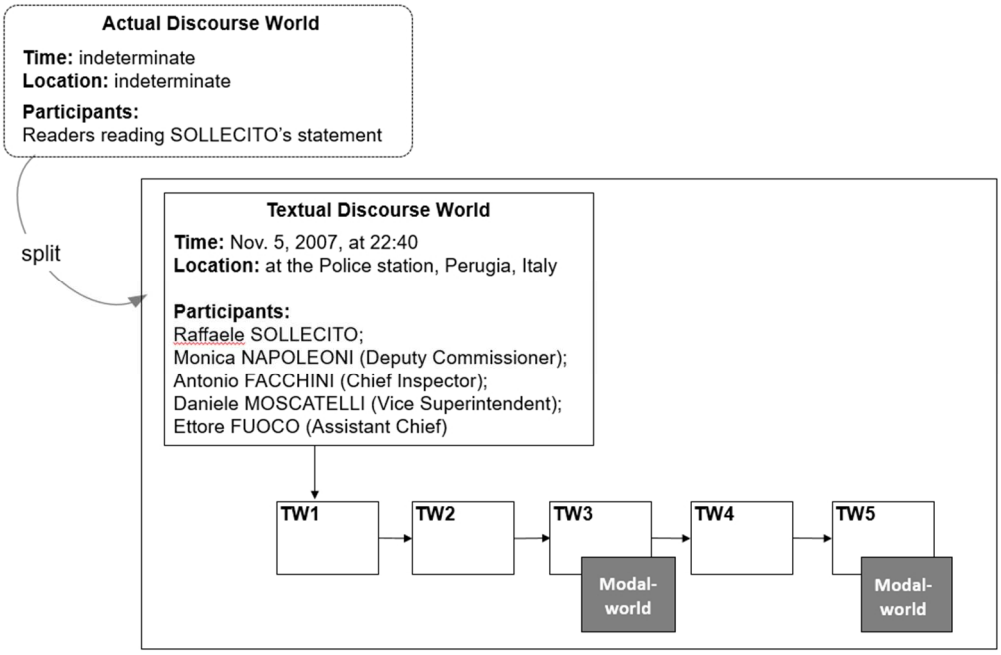

Fig. 3 Actual discourse-world vs. Textual discourse-world

266x173mm (96 x 96 DPI)

Function-advancing propositions: EVENT process types		**Textual examples**
Material / Behavioral	*doing* (actions) or *happening* (events)	I *woke up*; we *stayed* in the house; M *left* in a hurry.
Relational / Existential	states of *being* or *having* or *existing*	*there were* drops of blood in the sink the rest of the bathroom *was* clean
Verbal	*saying*	*she told me* she was going to the pub…; *she said*…; *I answered*…
Mental	*sensing* (perception, cognition, emotion)	I *don't remember* what we did; I *am certain* I had dinner

Table 1 Function-advancing propositions and textual examples

170x85mm (144 x 144 DPI)

Cognitive 'switches'	Textual examples
world-switch	e.g. *around 13:00-14:00, I went to see her at her house* (deictic world-switch based on changes in temporal/spatial references)
modal-world	e.g. *we went into town,* but *I do not remember what we did* (knowledge, belief or hypothetical expressions)
unframed text	e.g. *Amanda always carries a big handbag.* (un-sequenced, habitual action/event or generalization)

Table 2 Cognitive 'switches' and textual examples

279x85mm (96 x 96 DPI)

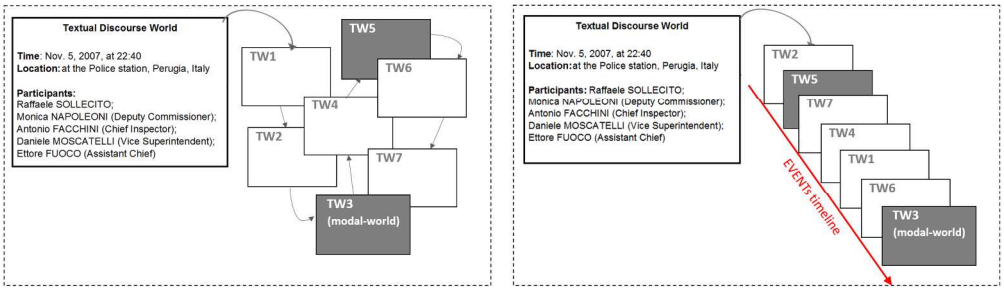

Fig. 4 Proposed visualization outputs

407x115mm (144 x 144 DPI)

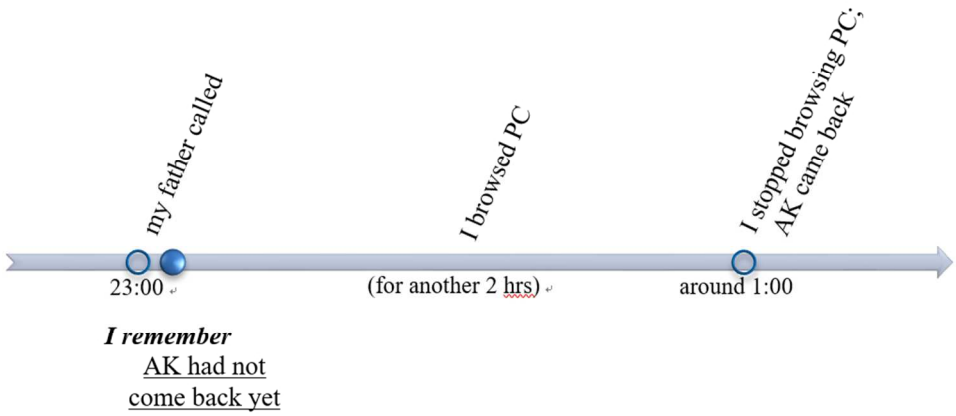

Fig. 5 Material process types in a timeline: actualized vs. modalized

188x81mm (144 x 144 DPI)

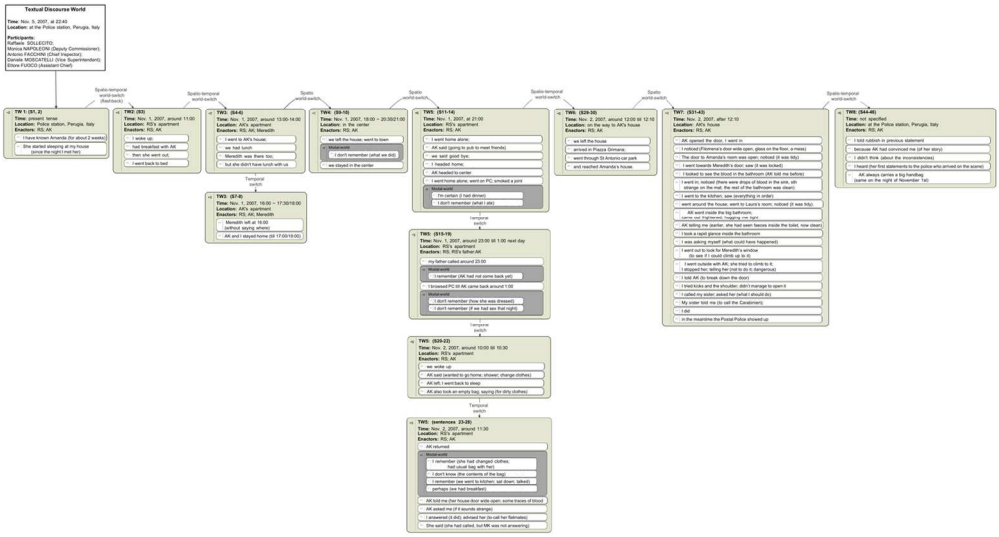

Fig. 6 Text-world diagram of the statement under study

130x69mm (300 x 300 DPI)

Total V-groups (including non-finite V) in the text		157
Material / Behavioral	*doing* (actions) or *happening* (events)	89 (57%)
Relational / Existential	states of *being* or *having* or *existing*	20 (13%)
Verbal	*saying*	17 (11%)
Mental	*sensing* (perception, cognition, emotion)	18 (11%)
V with multiple tags		13 (8%)

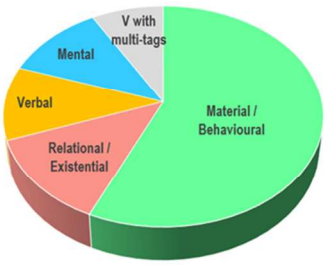

Fig. 7 Quantitative result of manual annotation

260x114mm (96 x 96 DPI)

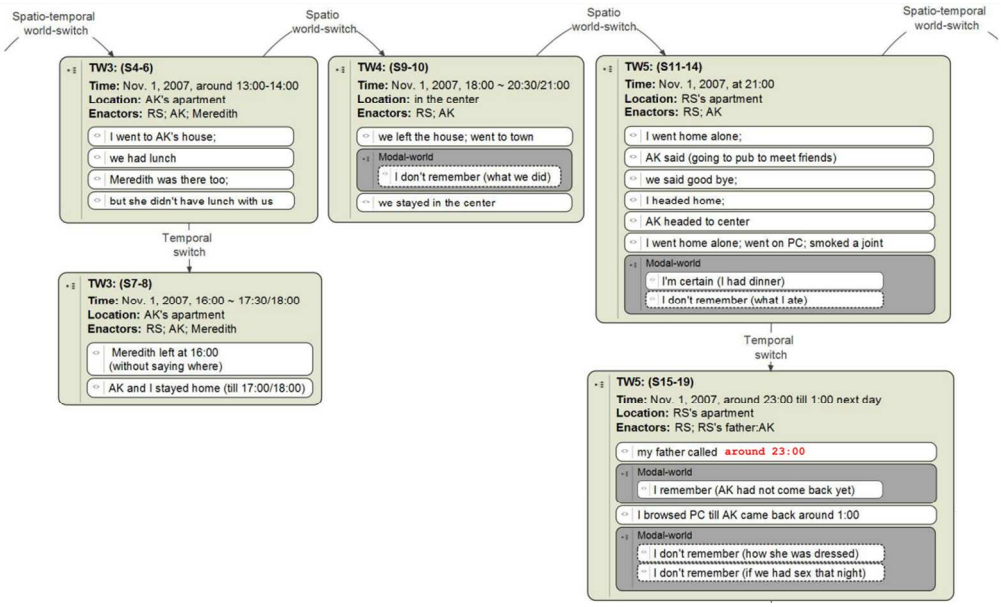

Fig. 8 Modal-worlds projected in the statement

331x198mm (96 x 96 DPI)

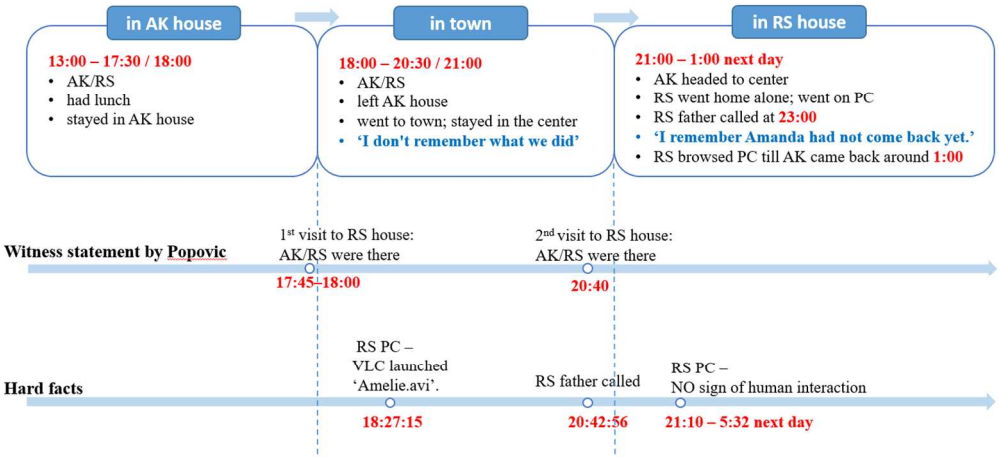

Fig. 9 EVENTs timelines: Sollecito vs. Popovic and hard facts

258x117mm (144 x 144 DPI)