

“The clocks had ceased their chiming, and the deep river ran on”: encoding temporal narrative features in modernist and hypertext fiction

Edward Kearns

National University of Ireland, Galway

This paper will describe work that is in progress in the second year of a PhD which seeks to determine whether a valid and useful system of encoding narrative characteristics related to time in fiction can be developed. It then asks how the narratives thus encoded can be quantitatively analysed, using modernist and hypertext fiction as case studies. The project integrates specific concepts of time from classical narrative theory into XML encoding, creating a schema from narrative features that have not been translated into this format before.

As such the paper will first describe the aims of the encoding: to annotate instances of analepsis, prolepsis, changes in narrative level, and stream-of-consciousness narration. I will explain these techniques in terms of narrative theory, and the rationale for encoding them. I will then show my key outputs so far of a new XML schema, full encoding guidelines, and some examples of encoded text from my modernism corpus, before projecting into the next phase of the project by describing methods of quantitative analysis using the encoded text.

In terms of related work, TimeML is the most similar system for narrative encoding with a focus on time. However, TimeML is used to encode time expressions, or references to specific moments in time, such as “last week”, “tomorrow”, etc. It does not engage with the larger transitions of the narrative backwards and forwards in time, the movements described by Gérard Genette as analepsis and prolepsis. Another system, NarrativeML, is less similar to this project, as it does not focus on time; neither does ProppML, as it takes Proppian folktale morphology as its focus. TEI allows some general “characterizations of narrative structure (such as ‘theme’, ‘reconciliation’ etc.)” (“17. Simple Analytic Mechanisms - The TEI Guidelines”), mainly using the element, but this is an adaptation of that element, and is not the main use for which TEI was designed. This mostly speaks to the power of TEI to be used for different kinds of documents, even if it remains more suited to documentary encoding of manuscripts, and their associated metadata. In any case, this application is too general for the needs of this project. Similarly in natural language processing, Reiter, Frank and Hellwig have created “alignment algorithms and experiments that evaluate their performance in establishing cross-document event alignments of narratives” (Reiter et al. 603), but these do not focus on the time features identified by Genette. This theorist is significant to this project because he was the initial major structuralist in the 1960s to focus primarily on time in narrative, and the concepts he suggested for classifying narrative time features are applicable to many kinds of fiction narratives, as are narrative levels. This project integrates those ideas into a new XML schema, incorporating inspiration from XML systems that are used to annotate other narrative features, and their guidelines, but addressing concepts which they do not.

Analepsis and prolepsis are transitions backwards and forwards in story time, respectively. The terms come from Genette, who describes the techniques collectively as anachrony, or “discordance between the two orderings of story and narrative” (36). In part this project is an act of translating terminology from narratology into a machine-readable structure. In this case the corresponding XML tags are <analepsis> and <prolepsis>.

Anachrony “is one of the traditional resources of literary narration” (Genette 36), occurring in literature across several centuries, allowing for comparison between many different genres and time periods. Modernist fiction is a particularly relevant case study because modernist writers were influenced by the changing philosophical and scientific theories of time that were in public discourse at the turn of the twentieth century. Encoding narrative time features allows us to ask whether that influence makes modernist fiction use techniques that disrupt narrative time in a different way to other genres of fiction.

Hypertext fiction is used for comparison because it has been compared to modernist fiction for sharing “modernist devices of jump, fragment, split perspectives, multiple threads, uncertain causalities, *écriture labyrinthine*, and so on.” (Aarseth 87) My case study asks whether modernist fiction uses anachrony and narrative levels in an idiosyncratic way, or whether their use follows similar patterns in hypertext fiction. I am currently in the process of encoding these corpora and will show some of this work-in-progress.

Narrative levels describe stories within stories. The tags for annotating this feature follow the narratological definitions of <FirstDegree>, <SecondDegree>, <ThirdDegree> narrative levels and so on. The terminology for degrees of narrative levels comes from Shlomith Rimmon-Kenan via Manfred Jahn, but Genette also discusses narrative levels in his work. In stream of consciousness narration (annotated as <soc>), the prose follows the thought patterns of a character. That particular phrase was coined by May Sinclair writing about the novels of Dorothy Richardson, although William James also discussed consciousness as a stream without using the specific three-word phrase. Stream of consciousness relates to anachrony and narrative levels because those features sometimes occur within instances of stream of consciousness narration (leading to nested XML tags).

The creation of encoding guidelines is standard practice when creating a system of markup like this. The purpose of the document is to allow any user (in this case from either humanities or computer science) to read the guidelines and then be able to encode text using the schema that the guidelines describe. If the tags that make up the schema are explained well enough, with clear enough rules for their application, then there should be some degree of uniformity in how they are applied by different annotators. On a practical level, as I am the only encoder for this project, having the rules set down gives me something to refer back to when encoding.

This paper will explain the schema in detail, and the reasons for the creation of specific elements and their attributes. It will discuss the encoding process itself, the advantages and disadvantages of manual encoding by an individual, and how these can be mitigated. As such

the paper will inform the audience on both the theoretical and technical details of the project and then seek feedback on the process.

Works Cited:

“17. Simple Analytic Mechanisms - The TEI Guidelines.” *Text Encoding Initiative*, 31 Jan. 2018, <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/AI.html>.

Aarseth, Espen J. *Cybertext: Perspectives on Ergodic Literature*. JHU Press, 1997.

Genette, Gérard. *Narrative Discourse: An Essay in Method*. Cornell University Press, 1980.

Reiter, Nils, et al. “An NLP-Based Cross-Document Approach to Narrative Structure Discovery.” *Literary and Linguistic Computing*, vol. 29, no. 4, 2014, pp. 583–605.