CA

Annotating Narrative Levels: Review of Guideline No. 1.

Meredith A. Martin

10.22.19

I am reviewing the Narrative Boundaries Annotation Guide, I had several concerns.

First, the authors are clear and concise when describing the project purpose: of training and testing a computer's ability to identify the beginnings and endings of narratives (which the authors call *narrative boundaries*). They are also clear and concise when describing what they mean by annotation in this project: "the record of human judgments identifying where a narrative begins and ends." When the authors arrive at the definition of narrative, however, their clarity and concision begin to unravel a bit. Since there are several theories of narrative, I suggest that the authors include some language like "for the purposes of this study" before launching into their definitions, which are necessarily oversimplified.

In the last paragraph of page two, the three forms could be elaborated more helpfully. Contiguous (or, later, un-interrupted), embedded, or interrupted narratives become important later on, so the guide could highlight that importance by bulleting the forms of narrative it deems important, and foreshadowing how deeply and how intricately the authors will attempt to distinguish between embedded or interrupted.

2.1 should be "Uninterrupted Narrative." I suggest they begin with "the simplest kind of narrative is an uninterrupted one. The section 1.2, titled"narrative boundaries," should end with "after the last word in the narration" and then continue

1

with "The next three sections" etc. (I have marked this change in the document using track change).

Embedded narratives were clearly explained and seems like a feasible for annotators, though arduous.

Marking boundaries for interrupted flashbacks and flashforwards seemed incredibly complicated. Interruptive narratives are the most complicated, and it seems that "Time Shifts" need their own heading (3) since they are not the three main kinds of narrative but may contain the three main kinds within them.

The example from 2.4 is too lengthy; I suggest adding a shorter example.

What happens if there are two types of narration happening (a flashback inside a dream)? Would the phrase "like I had done every morning since she left" be a flashback? The charts are useful but without multiple colors (and with only the example of Murakami to go on) I think that there is, at least as the guidelines are currently presented, quite a bit of room for interpretation and error. The human annotators would need to be a very large group, indeed, with a high level of fluency. For this to be a useful exercise, I can only imagine that the study would need hundreds of annotators.

Television scripts might be moved to an entirely different guideline - I feel that this guideline is already too complicated. I would eliminate scripts OR other kinds of narrative, but having both is too much.

Each section would benefit from additional examples drawn from authors other than Murakami. The beach / pirate example could be useful in each section, and since the levels of narratives accumulate, keeping the beach / pirate example active throughout and adding a literary text example alongside the beach / pirate would be a welcome addition and would more easily train the annotators. Or perhaps using a simpler narrative? Though the guidelines are helpful I would feel baffled by multiple levels of narration. I also can't imagine taking on this annotation task voluntarily.

The procedures seem complicated. How many short stories will the annotator receive? What does it mean to say "Each narrative has an ID number"? This makes no sense to me. I also don't understand "narrative name." Examples here, again from the pirate / beach story or a simpler story than IQ84, would be helpful.

"If possible, please use different colors for each narrative"? I think the authors should assign colors. The authors could use a more robust bibliography for narrative theory.