Tense use in dialogue and narrative discourse

- In earlier work [2], we used a parallel corpus based on the novel Harry Potter and the Philosopher’s Stone (J.K. Rowling) and its translations.
- In order to investigate tense use in dialogue, [2] separates the narrative parts from the dialogue parts in the text. Looking at a more narrative-oriented chapter (Ch. 1), and a more dialogue-oriented chapter (Ch. 17), the use of the tense categories Perfect and Past is as follows:

<table>
<thead>
<tr>
<th></th>
<th>English (original)</th>
<th>Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perfect</td>
<td>Past</td>
</tr>
<tr>
<td>Narrative</td>
<td>0</td>
<td>668</td>
</tr>
<tr>
<td>Dialogue</td>
<td>41</td>
<td>163</td>
</tr>
</tbody>
</table>

Table 1: Tense use in chapters 1 and 17 (translation of HP)

- Main findings: differences in tense use between narrative and dialogue; no Perfect use in the narrative part of the novel.

Hypothesis

The contrast between dialogue and narrative has to do with temporal orientation.

- The dialogues likely contain more utterances of what is currently going on (relative to the story time), whereas the narrative parts tell a story that happened in the past.
- The traditional view is that the English Perfect conveys current relevance; this would explain the occurrence of Perfects in here-and-now-oriented dialogue, and no occurrences in past-oriented narrative.

This leads to the testable prediction that dialogues with a different temporal orientation have a different tense use.

More data from Harry Potter

- We add data from Chapter 16. Chapters 16 and 17 both contain dialogue, but Chapter 16 is more present-oriented.

<table>
<thead>
<tr>
<th></th>
<th>English (original)</th>
<th>Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
<td>Past</td>
</tr>
<tr>
<td>Ch16 dialogue</td>
<td>182</td>
<td>53</td>
</tr>
<tr>
<td>Ch17 dialogue</td>
<td>126</td>
<td>129</td>
</tr>
</tbody>
</table>

Table 2: Tense use in Chapters 16 and 17 of HP

- The modal construction has got to was removed from the counts.
- The present orientation of Chapter 16 is confirmed by the higher Present ratio.
- However, the number of Perfects is lower in Chapter 16 than in Chapter 17, whereas the hypothesis predicts a higher number in more present-oriented dialogue.

The need for computational tools

- In order to analyze more and larger sets of dialogue data (both written and spoken), we need computational tools that can analyze the temporal structure of a text, but that are also sensitive to dialogue.
- Such tools will allow us to formalize the notion of ‘temporal orientation’.
- They will also benefit a range of other applications that require access to the temporal structure of dialogue.

- PerfectExtractor
- TARSQI
- TIMV
- SitEnt
- TLink Classifier
- Dialog Act recognition systems

Software links in handout/proceedings paper.

Evaluation of computational tools: task 1 / 2

- TARSQI [5] is a system that automatically recognizes events and temporal structure in a text. It was designed for the newswire domain, but we apply it to dialogue data.

   b. Did John read a book yesterday?
   c. Has John gone? 

- Problem 1: English tense constructions are not recognized correctly (has... gone as two events).
  ⇒ other tools, such as PerfectExtractor, perform better on subtask 1
- Problem 2: TARSQI does not annotate for speech act type, so it makes no distinction between assertions and questions ((1a) == (1b)).
  ⇒ the tools need to be sensitive to dialogue structure

Dialogue Acts: task 3

- The structure of dialogue is typically annotated in the form of Dialogue Acts (DAs), covering question-answer contrasts, but several other details in addition.
- Pilot study: we ran our PerfectExtractor to extract Perfects, Present and Past from the Switchboard corpus and cross-tabulate these with DAs.
- We find a significant association between use of the Perfect and question environments (p<0.001).
  - The odds of Perfect use are 1.5 times higher in questions as opposed to statements, although the effect is very small (φ=0.016).
  - The results underline the significance of the above remarks on computational tools having problems with question acts.
- Other DAs show tense preferences as well, see Figure 2. Hedges and agreements appear almost exclusively with a Present, while backchanneling is more Past-oriented.

The hypothesized ‘here-and-now’ orientation of dialogue seems to be confirmed by a larger empirical and computational setting.

Conclusion

- We found differences in tense use between dialogue and narrative. Our results about Perfect use in dialogue are incompatible with the standard view that the Perfect conveys ‘present result’.
- The hypothesized ‘here-and-now’ orientation of dialogue seems to be confirmed by a large use of Present, but not by an increase of Perfects over Past. The use of Perfects relates to the local linguistic environment (negation, lexical aspect, ever, etc.), which requires a closer investigation in subsequent work.
- In future work, we want to further develop computational tools in order to analyze more data from dialogue corpora (both from spoken sources, and written/fictional dialogue), including the BNC corpus and datasets in other languages.

References