The Effect of Different Writing Tasks on Linguistic Style: A Case Study of the ROC Story Cloze Task

Roy Schwartz¹,², Maarten Sap¹, Ioannis Konstas¹, Li Zilles¹, Yejin Choi¹ and Noah A. Smith¹
¹University of Washington, ²Allen Institute for Artificial Intelligence

Abstract
We show how variants of the same writing tasks can lead to measurable differences in writing style. Our experimenting with the story cloze task (Mostafazadeh et al., 2016a) show that a simple linear classifier informed by stylistic features is able to successfully distinguish among different writing tasks. Combining our stylistic features with language model predictions reaches state of the art performance on the story cloze challenge.

Story Cloze Task as a Style Detection Task
Original Task: Predict Correct Ending given Story prefix
New Task: Predict Correct Ending while ignoring story prefix

John was so excited about his upcoming CoNLL presentation.
His practice proved itself useful, as he got great feedback!
His talk went well.

Linear Classifier

John hates conferences.

Style Features:
- Sentence length
- Word n-grams
- Char n-grams

Language Model Features

\( e^* = \arg\max_{e \in \{e_1, e_2\}} p_{lm}(e|\text{prefix}) \)

The intuition: controlling for the inherent surprisingness of the words in each ending

State-of-the-art Results on the Story Cloze Task

Discussion and Implications
What Causes These Style Differences?
- Writing style can both affect and be affected by mental state (Campbell, 2003; Ireland 2014)
- Does the link between different writing tasks and writing style involves differences in mental state?

Careful Instructions Reduce Style Effects
- Negative results of our style-based model on the COPA dataset (Roemmele et al., 2011)
- Possible reason: annotators were asked to be as brief as possible and avoid proper names

Writing Instructions Affect Writing Style

Binary Classification of Story Endings

<table>
<thead>
<tr>
<th>Same authors, Coherent</th>
<th>Different authors, coherent</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>

Story Cloze Task

<table>
<thead>
<tr>
<th>Coherent:</th>
<th>Different author:</th>
</tr>
</thead>
<tbody>
<tr>
<td>end with past tense verbs</td>
<td>start with proper nouns</td>
</tr>
<tr>
<td>coordinations</td>
<td>end with common nouns</td>
</tr>
<tr>
<td>adjectives</td>
<td>Same author:</td>
</tr>
<tr>
<td>Incoherent:</td>
<td>are (much) longer</td>
</tr>
<tr>
<td>start with proper nouns</td>
<td>are more enthusiastic (&quot;I&quot;)</td>
</tr>
<tr>
<td>end with common nouns</td>
<td>adverbs</td>
</tr>
<tr>
<td>negative words (&quot;hate&quot;)</td>
<td></td>
</tr>
</tbody>
</table>

Most Salient Style Features

Take Home Messages

- The task a writer is given to perform will affect their writing style in detectable ways
- Careful attention must be paid to instructions given to annotators of NLP tasks

* Code available at https://github.com/roys174/writing_style