Assessing Annotation Consistency in Universal Dependencies Corpora

Marie-Catherine de Marneffe, Matias Grioni

Jenna Kanerva, Filip Ginter
Current Inconsistencies

this is what the thing is about

This store was what Colorado was all about
this is what the thing is about

This store was what Colorado was all about
Annotation Tool

1. Correct  WrongType  WrongDep  WrongTagOrLemma  Other
   Advmod
   ADV  AUX  NUM  NOUN  PUNCT
   Here  's  two  examples

2. Correct  WrongType  WrongDep  WrongTagOrLemma  Other
   Nsubj
   ADV  AUX  DET  ADJ  NOUN  PUNCT
   Here  's  another interesting  example
## Balancing Precision and Recall for Rich Morphology

<table>
<thead>
<tr>
<th>Language</th>
<th>“Non-fringe” Precision (%)</th>
<th>“Dependency context” Precision (%)</th>
<th>Recall (%)</th>
<th>“Non-fringe” Precision (%)</th>
<th>Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>62</td>
<td>76</td>
<td>66</td>
<td>72</td>
<td>79</td>
</tr>
<tr>
<td>French</td>
<td>65</td>
<td>64</td>
<td>47</td>
<td>76</td>
<td>73</td>
</tr>
<tr>
<td>Finnish</td>
<td>19</td>
<td>21</td>
<td>81</td>
<td>72</td>
<td>75</td>
</tr>
</tbody>
</table>
Leveraging Big Data to Ground Correct Annotation

- CoNLL ‘17 Shared Task data (Ginter et al., 2017)
- Several billion words of automatically parsed web data

<table>
<thead>
<tr>
<th>Language</th>
<th># tree errors</th>
<th>“Non-fringe” Precision</th>
<th>% in Boyd</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>54</td>
<td>41%</td>
<td>38%</td>
</tr>
<tr>
<td>French</td>
<td>74</td>
<td>57%</td>
<td>36%</td>
</tr>
<tr>
<td>Finnish</td>
<td>10</td>
<td>16%</td>
<td>30%</td>
</tr>
</tbody>
</table>
universaldependencies.org/fixit

- Introduce approach to greater UD community
- Intuitive web-based tool to flag inconsistencies
- Inconsistency output for 80+ languages already online