|                             | Dev          | Test |
|-----------------------------|--------------|------|
| English (Wang et. al, 2013) | 86.6         | 80.7 |
| English (Ours)              | 88.9         | 84.8 |
| Spanish (Ours)              | 70.2<br>67.8 | 70.8 |
| Dutch (Ours)                | 67.8         | 69.2 |