Semantic Overlap Computation

Kanishka Misra, Hemanth Devarapalli

1/22/2019

Learner Errors in English

In this example, we consider an error annotated corpus of 1244 short essays, each written by a unique learner of english belonging to one of 16 different first languages (L1).

Since we are covering semantic errors, we extract the errors made in content words (adjectives (J), nouns (N), verbs (V) and adverbs (Y)).

Let's look at an example by a learner whose L1 is Russian:

Incorrect word in English: `r` Replacement as annotated: `r`