

Classifying Syntactic Errors in Learner Language – Supplementary Material

1 Classifying SEs across Learner Levels

The W&I dataset includes both corrections of texts by native speakers and by different levels of non-native speakers, we use that to analyze changes across proficiency levels. We extract the distributions of SE types from the train set for each proficiency level A–C, where A is the lowest, and for the native speakers’ text. We compare across levels what percentage of words of a given POS tag are SEs (i.e., changed to a different POS in the correction). Results are presented in Table 1; full confusion matrices are given in Appendix §3.3.

As one might expect, the more proficient the learners, the fewer SEs they make. This holds for learners across all POS types. Native speakers generally make fewer errors than advanced learners, but the trend is mixed. Surprisingly, native speakers rarely make SEs in POS tags that are the most error-prone for learners (top of Table 1). Conversely, native speakers generally do not do better than advanced learners in the least error-prone POS tags (bottom of the table). It is important to note that as the non natives have less data, the borderline changes positive and negative changes might be considered comparable rather than significant changes to one way or another. Further study of this phenomenon is deferred to future work.

Finally, we compare the syntactic proficiency of the learner language in TLE and the A–C proficiency levels of W&I. We compute the percentage of POS tags of each type that were altered in the correction. We find the syntactic proficiency of TLE to be not too high, in most cases between level A and B. Those results agree with the prior knowledge that TLE texts are taken from B1-B2 CEFR (Little, 2006) levels, despite the difference in the native tongues of the learners, which adds noise to the comparison. This result thus provides further validation for SERCL.

| | A | B | C | Native |
|-------|-------|-------|-------|--------|
| SCONJ | 0.804 | 0.864 | 0.923 | 0.942 |
| DET | 0.857 | 0.907 | 0.960 | 0.971 |
| ADV | 0.844 | 0.893 | 0.945 | 0.950 |
| ADJ | 0.875 | 0.923 | 0.962 | 0.972 |
| ADP | 0.891 | 0.935 | 0.969 | 0.976 |
| PART | 0.887 | 0.924 | 0.963 | 0.985 |
| AUX | 0.901 | 0.943 | 0.973 | 0.987 |
| PROPN | 0.902 | 0.930 | 0.966 | 0.968 |
| NUM | 0.897 | 0.929 | 0.960 | 0.950 |
| PRON | 0.908 | 0.930 | 0.963 | 0.953 |
| NOUN | 0.934 | 0.963 | 0.983 | 0.983 |
| CCONJ | 0.922 | 0.944 | 0.968 | 0.971 |
| VERB | 0.945 | 0.964 | 0.983 | 0.980 |
| PUNCT | 0.978 | 0.980 | 0.990 | 0.981 |

Table 1: Percentage of unchanged POS tags per type (rows) and proficiency level (columns) in the W&I dataset. Proficiency levels are A-C where C is the most proficient, the last column is for native speakers. Sorted by the average of columns A-C.

2 Examples of Prevalent Types of SEs Involving Replacements

2.1 POS Changes¹

NOUN → VERB

- When we ~~entrance~~ entered the place our problems began.
- I mean it’s ~~conflict~~ conflicts with your plan, what a pity!

NOUN → ADJ

- These things are very ~~convenience~~ convenient; on the other hand, there are lots of disadvantage points.
- In this way the ~~environment~~ environmental pollution will increase more and more.

¹When sentences contained other errors the corrected tokens were used for clarity of presentation.

ADJ → NOUN

- So while I am at the Camp I like to take ~~many~~ a lot of photographs and climb a mountain.
- He has known research in the ~~deeper~~ depths of his mind.

VERB → NOUN

- I think that I am a very fast ~~swimming~~ swimmer but I don't have good style and technique.
- Obviously, there is no point in saying "Famous people have a right to their own ~~live~~ lives."

VERB → ADJ

- There was only a small sign ~~stamped~~ stuck on the door, saying that it was "closed for repairs".
- Sometimes you ~~surprise~~ are surprised when you check the balance of your bank account.

DET → PRON

- I have a camera, ~~which~~ it is called Minolta X300.
- To finish, I wonder if I need to bring some money or if ~~all~~ everything has been already paid for.

PRON → DET

- I was pleased to receive ~~you~~ your letter recently.
- I can say that the greatest ever invention is the invention of computers, ~~that~~ which has affected both individuals and a society as a whole.

2.1.1 Edge-type Changes²

compound → amod

- Shopping is relaxing, above all on Saturdays when you have finished a ~~work~~ working week and you are expecting a wonderful Sunday.
- Secondly, turning to the end-of-conference party, it is planned to be in Wimbledon Common – nice piece of ~~England~~ English nature.

²In cases of clause-linkage errors, words participating in the relations are underlined; words inducing the relations are in boldface.

cop → aux

- From my point of view I think he was such a very bad actor that it ~~was~~ would be impossible to find another with the same characteristics.
- It should ~~be~~ have been a perfect evening out, but it was the opposite.

xcomp → ccomp

- It ~~made~~ people only spend pocket-money → It ~~meant~~ people only spent pocket-money.
- I went in London for a week of "pure holiday" thinking ~~to have~~ some fun coming to see your musical show → I went to London for a week of "pure holiday", thinking I would have some fun going to see your musical show.

obl → obj

- I'm not so good in tennis but I would like to play it ~~for~~ practice → I'm not so good at tennis but I would like to play it to ~~get~~ some practice.
- However, we have to consider ~~about~~ our human nature.

obl → advmod

- Then all the girls became jealous and stopped ~~to laugh~~ laughing.
- I never had a chance ~~for staying~~ to stay in a tent.

det → nmod:poss

- It is simple to find the way from ~~this~~ your hotel to the conference.
- It allowed me to worry about other things in ~~the~~ my life.

advmod → obl

- And it is even more difficult to predict ~~how~~ what the clothes of the future will look like.
- My work was for the singers, so I could meet singers and talk ~~together~~ with them.

3 Confusion Matrices

3.1 POS and edge-labels changes in TLE

Distribution of edit types in terms of POS tags and edge labels are given in Tables 2 and 3 respectively. Automatic ones are given in 4 and 5.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | VERB |
|-------|-----|-----|-----|-----|-------|-----|------|-----|------|------|-------|-------|-------|------|
| ADJ | 444 | 3 | 27 | 1 | 0 | 16 | 49 | 2 | 0 | 2 | 0 | 0 | 0 | 21 |
| ADP | 4 | 727 | 11 | 0 | 2 | 10 | 5 | 0 | 1 | 0 | 0 | 0 | 6 | 23 |
| ADV | 23 | 6 | 270 | 14 | 4 | 2 | 22 | 0 | 1 | 23 | 1 | 8 | 11 | 11 |
| AUX | 4 | 1 | 7 | 372 | 0 | 0 | 8 | 0 | 8 | 15 | 0 | 0 | 0 | 21 |
| CCONJ | 1 | 4 | 3 | 0 | 40 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 5 | 0 |
| DET | 18 | 11 | 7 | 1 | 0 | 315 | 3 | 1 | 6 | 34 | 0 | 2 | 5 | 5 |
| NOUN | 50 | 5 | 22 | 5 | 0 | 3 | 1621 | 3 | 0 | 14 | 11 | 0 | 3 | 51 |
| NUM | 1 | 0 | 0 | 0 | 0 | 6 | 3 | 21 | 0 | 0 | 0 | 0 | 0 | 0 |
| PART | 0 | 0 | 4 | 4 | 0 | 6 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 1 |
| PRON | 1 | 2 | 6 | 10 | 0 | 32 | 15 | 5 | 0 | 240 | 1 | 6 | 7 | 5 |
| PROPN | 3 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 1 | 48 | 0 | 0 | 0 |
| PUNCT | 0 | 1 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 256 | 1 | 1 |
| SCONJ | 1 | 5 | 12 | 2 | 6 | 1 | 0 | 0 | 5 | 3 | 0 | 1 | 64 | 2 |
| VERB | 37 | 7 | 4 | 24 | 0 | 0 | 46 | 0 | 0 | 2 | 2 | 0 | 0 | 1818 |

Table 2: POS confusion matrix for replacement edits in absolute counts for TLE dataset with manual UD.

| | acl | acl:recl | advel | advmod | amod | appos | aux | aux:pass | case | cc | ccomp | compound | conj | cop | csbj | det | expl | mark | nmod | nmod:poss | nsbj | nsbj:pass | nummod | obj | obl | obl:tmod | parataxis | punct | root | xcomp | |
|-----------|-----|----------|-------|--------|------|-------|-----|----------|------|-----|-------|----------|------|-----|------|-----|------|------|------|-----------|------|-----------|--------|-----|-----|----------|-----------|-------|------|-------|---|
| acl | 58 | 6 | 4 | 0 | 1 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | | |
| acl:recl | 1 | 140 | 3 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | |
| advel | 7 | 3 | 321 | 3 | 0 | 0 | 4 | 0 | 3 | 0 | 3 | 0 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 11 | 0 | 4 | 0 | 4 | 17 | |
| advmod | 2 | 0 | 3 | 270 | 6 | 0 | 13 | 1 | 4 | 5 | 3 | 0 | 1 | 0 | 1 | 8 | 1 | 8 | 1 | 0 | 10 | 1 | 0 | 7 | 24 | 0 | 0 | 6 | 3 | 7 | |
| amod | 0 | 0 | 0 | 5 | 214 | 0 | 0 | 0 | 2 | 0 | 1 | 5 | 4 | 0 | 0 | 15 | 0 | 0 | 1 | 1 | 2 | 0 | 2 | 7 | 3 | 0 | 0 | 0 | 1 | | |
| appos | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | |
| aux | 0 | 1 | 1 | 2 | 1 | 0 | 223 | 3 | 0 | 0 | 1 | 0 | 3 | 23 | 0 | 0 | 1 | 3 | 0 | 0 | 14 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | |
| aux:pass | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 10 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| case | 0 | 1 | 4 | 10 | 1 | 0 | 0 | 0 | 738 | 3 | 1 | 0 | 2 | 0 | 2 | 7 | 0 | 8 | 1 | 1 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 3 | |
| cc | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 3 | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | |
| ccomp | 0 | 2 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 193 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 0 | 11 | |
| compound | 0 | 0 | 0 | 0 | 32 | 0 | 0 | 0 | 0 | 0 | 62 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | |
| conj | 0 | 0 | 5 | 2 | 3 | 0 | 1 | 0 | 0 | 0 | 3 | 0 | 440 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 3 | 1 | 0 | 2 | 0 | 0 | 1 | |
| cop | 0 | 2 | 2 | 6 | 0 | 0 | 32 | 0 | 2 | 0 | 1 | 0 | 3 | 66 | 0 | 0 | 1 | 4 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 1 | |
| csbj | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | |
| det | 0 | 0 | 1 | 10 | 16 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 2 | 2 | 0 | 233 | 0 | 4 | 0 | 25 | 3 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 3 | 0 | |
| expl | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mark | 0 | 0 | 1 | 12 | 1 | 0 | 2 | 1 | 5 | 6 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 79 | 0 | 0 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 2 | 2 | 0 | |
| nmod | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 214 | 5 | 3 | 0 | 0 | 3 | 19 | 0 | 1 | 0 | 0 | |
| nmod:poss | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 7 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 34 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | |
| nsbj | 0 | 0 | 1 | 3 | 0 | 0 | 9 | 0 | 0 | 2 | 0 | 0 | 3 | 5 | 0 | 3 | 6 | 3 | 0 | 1 | 357 | 1 | 0 | 0 | 0 | 4 | 1 | 1 | 6 | 2 | 0 |
| nsbj:pass | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 8 | 27 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | |
| nummod | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| obj | 0 | 1 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 2 | 0 | 0 | 0 | 1 | 4 | 9 | 1 | 7 | 0 | 0 | 457 | 22 | 1 | 0 | 0 | 2 | 9 | |
| obl | 1 | 0 | 17 | 25 | 2 | 1 | 0 | 0 | 7 | 0 | 1 | 0 | 4 | 0 | 0 | 1 | 0 | 4 | 17 | 2 | 3 | 0 | 0 | 26 | 426 | 3 | 1 | 0 | 4 | 5 | |
| obl:tmod | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 9 | 0 | 0 | 0 | 0 | |
| parataxis | 3 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 57 | 0 | 0 | 0 |
| punct | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 256 | 1 | 0 |
| root | 0 | 0 | 0 | 1 | 1 | 0 | 3 | 0 | 1 | 0 | 4 | 0 | 0 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 4 | 2 | 0 | 2 | 0 | 702 | 5 | |
| xcomp | 2 | 0 | 8 | 0 | 0 | 0 | 1 | 2 | 0 | 32 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 3 | 0 | 0 | 0 | 7 | 297 | |

Table 3: Edge-type confusion matrix for replacement edits in absolute counts for TLE dataset with manual UD.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X | Other |
|-------|------|------|------|------|-------|------|------|-------|-----|------|------|-------|-------|-------|-----|------|----|-------|
| ADJ | 3907 | 24 | 61 | 2 | 2 | 192 | 1 | 78 | 4 | 0 | 15 | 4 | 11 | 4 | 0 | 48 | 0 | 0 |
| ADP | 24 | 5998 | 69 | 14 | 11 | 24 | 2 | 63 | 5 | 26 | 39 | 2 | 31 | 82 | 0 | 232 | 0 | 0 |
| ADV | 69 | 67 | 3952 | 41 | 10 | 48 | 4 | 44 | 0 | 16 | 58 | 17 | 37 | 29 | 0 | 52 | 0 | 0 |
| AUX | 3 | 10 | 27 | 5046 | 4 | 7 | 0 | 16 | 0 | 4 | 72 | 4 | 12 | 10 | 0 | 143 | 0 | 0 |
| CCONJ | 2 | 8 | 8 | 0 | 2428 | 3 | 0 | 12 | 0 | 0 | 7 | 2 | 14 | 6 | 0 | 4 | 0 | 0 |
| DET | 33 | 172 | 39 | 26 | 18 | 5451 | 0 | 33 | 6 | 10 | 78 | 0 | 19 | 30 | 0 | 76 | 0 | 0 |
| INTJ | 2 | 2 | 2 | 0 | 0 | 1 | 50 | 3 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| NOUN | 117 | 20 | 33 | 9 | 5 | 73 | 0 | 11288 | 8 | 6 | 19 | 37 | 17 | 8 | 0 | 146 | 0 | 0 |
| NUM | 17 | 7 | 2 | 2 | 0 | 11 | 0 | 7 | 573 | 0 | 2 | 1 | 1 | 0 | 0 | 3 | 0 | 0 |
| PART | 0 | 24 | 9 | 10 | 1 | 8 | 0 | 4 | 0 | 1971 | 14 | 0 | 1 | 22 | 0 | 23 | 0 | 0 |
| PRON | 7 | 46 | 28 | 23 | 10 | 54 | 0 | 23 | 6 | 0 | 9647 | 2 | 132 | 51 | 0 | 45 | 0 | 0 |
| PROPN | 5 | 3 | 2 | 0 | 0 | 36 | 2 | 44 | 0 | 0 | 1 | 1284 | 4 | 0 | 0 | 5 | 0 | 0 |
| PUNCT | 9 | 8 | 10 | 8 | 11 | 6 | 0 | 57 | 0 | 1 | 18 | 3 | 1731 | 7 | 1 | 22 | 1 | 0 |
| SCONJ | 7 | 75 | 28 | 3 | 11 | 9 | 0 | 15 | 0 | 13 | 51 | 1 | 24 | 1846 | 0 | 28 | 1 | 0 |
| SYM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 |
| VERB | 80 | 74 | 47 | 153 | 2 | 22 | 2 | 127 | 2 | 21 | 19 | 6 | 18 | 4 | 0 | 9004 | 2 | 0 |
| X | 1 | 1 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 15 | 0 |

Table 4: POS-type confusion matrix for replacement edits in absolute counts for TLE dataset with automatic UD.

3.2 Output Matrices

We supply the reader with matrices of the systems of the BEA2019 shared task that shared their outputs on dev set (avoiding over analysing the test set).

| | UEDIN-MS | KAKAO&BRAIN | SHUYAO | CAMB.CUED | TOHOKU | GOLD | Average |
|-------|----------|-------------|--------|-----------|--------|------|---------|
| INTJ | 0.86 | 0.86 | 0.71 | 1.00 | 0.43 | 472 | 0.77 |
| PRON | 0.63 | 0.74 | 0.70 | 0.75 | 0.44 | 233 | 0.65 |
| PROPN | 0.54 | 0.44 | 0.96 | 0.60 | 0.64 | 405 | 0.64 |
| ADJ | 0.61 | 0.50 | 0.63 | 0.66 | 0.45 | 7 | 0.57 |
| DET | 0.61 | 0.40 | 0.56 | 0.68 | 0.52 | 166 | 0.55 |
| PART | 0.58 | 0.45 | 0.60 | 0.64 | 0.43 | 158 | 0.54 |
| ADV | 0.58 | 0.51 | 0.57 | 0.66 | 0.33 | 114 | 0.53 |
| AUX | 0.49 | 0.37 | 0.53 | 0.60 | 0.46 | 462 | 0.49 |
| NOUN | 0.56 | 0.35 | 0.53 | 0.53 | 0.44 | 310 | 0.48 |
| CCONJ | 0.45 | 0.61 | 0.42 | 0.59 | 0.24 | 568 | 0.46 |
| SCONJ | 0.49 | 0.48 | 0.49 | 0.54 | 0.29 | 48 | 0.46 |
| VERB | 0.50 | 0.36 | 0.50 | 0.53 | 0.37 | 191 | 0.45 |
| ADP | 0.50 | 0.36 | 0.46 | 0.52 | 0.39 | 584 | 0.45 |
| NUM | 0.46 | 0.48 | 0.42 | 0.50 | 0.31 | 618 | 0.43 |
| SYM | 0.67 | 0.00 | 0.42 | 0.42 | 0.67 | 12 | 0.43 |
| PUNCT | 0.61 | 0.21 | 0.40 | 0.40 | 0.38 | 461 | 0.40 |

Table 6: Upper bound on recall per system. The gold amount which is considered the 100% per row and the average per line are reported on the two rightmost columns. The order of the table is descending on the average.

| | UEDIN-MS | KAKAO&BRAIN | SHUYAO | CAMB-CUED | TOHOKU | GOLD |
|-------|----------|-------------|--------|-----------|--------|------|
| ADJ | 283 | 231 | 292 | 305 | 210 | 462 |
| ADP | 232 | 166 | 214 | 239 | 181 | 461 |
| ADV | 273 | 242 | 268 | 313 | 155 | 472 |
| AUX | 153 | 115 | 165 | 187 | 142 | 310 |
| CCONJ | 71 | 96 | 67 | 94 | 38 | 158 |
| DET | 348 | 229 | 316 | 387 | 294 | 568 |
| INTJ | 6 | 6 | 5 | 7 | 3 | 7 |
| NOUN | 346 | 219 | 328 | 330 | 273 | 618 |
| NUM | 22 | 23 | 20 | 24 | 15 | 48 |
| PART | 97 | 74 | 99 | 107 | 71 | 166 |
| PRON | 367 | 433 | 409 | 437 | 258 | 584 |
| PROPN | 62 | 50 | 109 | 68 | 73 | 114 |
| PUNCT | 116 | 40 | 76 | 77 | 73 | 191 |
| SCONJ | 114 | 111 | 115 | 126 | 67 | 233 |
| SYM | 8 | 0 | 5 | 5 | 8 | 12 |
| VERB | 202 | 145 | 201 | 216 | 149 | 405 |
| X | 13 | 9 | 13 | 15 | 16 | 18 |

Table 7: Amount of SEs per source POS for each of the outputs and the gold standard.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X | Other |
|-------|------|------|------|------|-------|------|------|-------|-----|------|------|-------|-------|-------|-----|------|----|-------|
| ADJ | 6281 | 20 | 21 | 4 | 3 | 109 | 0 | 46 | 0 | 0 | 9 | 8 | 28 | 2 | 0 | 33 | 0 | 0 |
| ADP | 17 | 7649 | 10 | 10 | 0 | 3 | 0 | 25 | 1 | 22 | 16 | 1 | 52 | 35 | 0 | 40 | 0 | 0 |
| ADV | 28 | 18 | 4650 | 23 | 5 | 17 | 8 | 17 | 0 | 3 | 25 | 2 | 95 | 14 | 0 | 18 | 0 | 0 |
| AUX | 2 | 1 | 8 | 5680 | 4 | 2 | 0 | 19 | 0 | 10 | 51 | 2 | 11 | 2 | 0 | 41 | 0 | 0 |
| CCONJ | 0 | 2 | 2 | 1 | 2946 | 2 | 0 | 9 | 0 | 0 | 2 | 1 | 51 | 0 | 0 | 1 | 0 | 0 |
| DET | 15 | 112 | 13 | 13 | 8 | 7175 | 1 | 11 | 1 | 8 | 31 | 2 | 61 | 13 | 0 | 58 | 1 | 0 |
| INTJ | 0 | 1 | 0 | 0 | 0 | 0 | 92 | 1 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 0 | 0 |
| NOUN | 67 | 7 | 19 | 10 | 5 | 35 | 0 | 16050 | 2 | 3 | 18 | 33 | 55 | 5 | 0 | 87 | 0 | 0 |
| NUM | 3 | 1 | 2 | 1 | 0 | 3 | 0 | 6 | 628 | 0 | 0 | 0 | 4 | 1 | 0 | 1 | 0 | 0 |
| PART | 2 | 10 | 6 | 9 | 1 | 2 | 0 | 9 | 0 | 2523 | 13 | 1 | 6 | 20 | 0 | 18 | 0 | 0 |
| PRON | 1 | 13 | 16 | 17 | 6 | 12 | 0 | 25 | 0 | 2 | 8016 | 1 | 235 | 14 | 0 | 25 | 0 | 0 |
| PROPN | 14 | 1 | 0 | 0 | 1 | 16 | 0 | 19 | 0 | 1 | 1 | 1926 | 6 | 0 | 0 | 3 | 0 | 0 |
| PUNCT | 7 | 8 | 10 | 8 | 8 | 5 | 1 | 29 | 1 | 6 | 5 | 6 | 8679 | 11 | 0 | 11 | 0 | 0 |
| SCONJ | 0 | 25 | 7 | 0 | 1 | 2 | 0 | 6 | 0 | 17 | 12 | 1 | 31 | 1882 | 0 | 12 | 0 | 0 |
| SYM | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 3 | 0 | 45 | 1 | 0 | 0 |
| VERB | 21 | 9 | 9 | 44 | 3 | 7 | 1 | 67 | 0 | 5 | 7 | 5 | 20 | 3 | 0 | 9973 | 1 | 0 |
| X | 0 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 3 | 65 | 0 |

Table 8: UEDIN-MS POS confusion matrix on W&I development.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X | Other |
|-------|------|------|------|------|-------|------|------|-------|-----|------|------|-------|-------|-------|-----|-------|----|-------|
| ADJ | 6354 | 16 | 18 | 3 | 1 | 92 | 0 | 37 | 0 | 1 | 2 | 6 | 13 | 0 | 0 | 21 | 0 | 0 |
| ADP | 19 | 7700 | 14 | 6 | 0 | 4 | 0 | 30 | 1 | 12 | 8 | 4 | 11 | 25 | 0 | 47 | 0 | 0 |
| ADV | 15 | 12 | 4768 | 19 | 4 | 13 | 4 | 15 | 0 | 2 | 12 | 3 | 25 | 11 | 0 | 20 | 0 | 0 |
| AUX | 1 | 2 | 10 | 5691 | 3 | 1 | 0 | 22 | 0 | 12 | 41 | 0 | 6 | 1 | 0 | 43 | 0 | 0 |
| CCONJ | 2 | 0 | 1 | 1 | 2979 | 2 | 0 | 6 | 0 | 1 | 2 | 1 | 16 | 0 | 1 | 5 | 0 | 0 |
| DET | 13 | 93 | 21 | 10 | 9 | 7229 | 0 | 12 | 0 | 5 | 17 | 1 | 45 | 10 | 0 | 57 | 1 | 0 |
| INTJ | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| NOUN | 46 | 6 | 20 | 2 | 5 | 33 | 0 | 16123 | 3 | 3 | 22 | 21 | 31 | 5 | 0 | 76 | 0 | 0 |
| NUM | 1 | 0 | 0 | 0 | 1 | 3 | 0 | 5 | 635 | 0 | 1 | 0 | 1 | 1 | 0 | 2 | 0 | 0 |
| PART | 2 | 7 | 4 | 9 | 0 | 0 | 0 | 11 | 0 | 2549 | 11 | 1 | 2 | 13 | 0 | 11 | 0 | 0 |
| PRON | 2 | 11 | 13 | 14 | 7 | 10 | 0 | 28 | 0 | 2 | 8125 | 2 | 128 | 13 | 0 | 28 | 0 | 0 |
| PROPN | 19 | 5 | 0 | 1 | 0 | 13 | 0 | 24 | 0 | 0 | 3 | 1915 | 4 | 0 | 0 | 3 | 1 | 0 |
| PUNCT | 11 | 2 | 8 | 1 | 5 | 2 | 1 | 22 | 0 | 0 | 4 | 2 | 8722 | 4 | 1 | 10 | 0 | 0 |
| SCONJ | 0 | 12 | 9 | 4 | 1 | 3 | 0 | 5 | 0 | 9 | 7 | 0 | 8 | 1929 | 0 | 9 | 0 | 0 |
| SYM | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 3 | 0 | 45 | 1 | 0 | 0 |
| VERB | 17 | 7 | 7 | 30 | 3 | 6 | 3 | 44 | 0 | 7 | 9 | 2 | 11 | 2 | 0 | 10026 | 1 | 0 |
| X | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 6 | 0 | 0 | 1 | 2 | 1 | 0 | 0 | 1 | 62 | 0 |

Table 9: AIP-TOHOKU POS confusion matrix on W&I development.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X | Other |
|-------|------|------|------|------|-------|------|------|-------|-----|------|------|-------|-------|-------|-----|-------|----|-------|
| ADJ | 6333 | 9 | 16 | 2 | 0 | 98 | 0 | 41 | 0 | 0 | 5 | 7 | 26 | 0 | 0 | 27 | 0 | 0 |
| ADP | 10 | 7715 | 7 | 6 | 1 | 4 | 0 | 14 | 0 | 10 | 8 | 1 | 39 | 33 | 0 | 33 | 0 | 0 |
| ADV | 20 | 16 | 4681 | 13 | 1 | 12 | 10 | 6 | 0 | 2 | 12 | 4 | 124 | 9 | 0 | 13 | 0 | 0 |
| AUX | 0 | 0 | 2 | 5718 | 1 | 1 | 0 | 11 | 0 | 9 | 32 | 0 | 11 | 4 | 0 | 44 | 0 | 0 |
| CCONJ | 0 | 0 | 2 | 0 | 2921 | 2 | 0 | 3 | 0 | 0 | 0 | 2 | 86 | 0 | 0 | 1 | 0 | 0 |
| DET | 12 | 59 | 4 | 2 | 3 | 7294 | 1 | 6 | 0 | 5 | 21 | 0 | 81 | 10 | 0 | 25 | 0 | 0 |
| INTJ | 0 | 0 | 0 | 0 | 0 | 0 | 92 | 0 | 0 | 0 | 2 | 1 | 2 | 0 | 0 | 1 | 0 | 0 |
| NOUN | 31 | 3 | 11 | 2 | 3 | 23 | 0 | 16177 | 3 | 0 | 11 | 17 | 48 | 4 | 0 | 63 | 0 | 0 |
| NUM | 3 | 1 | 0 | 1 | 0 | 4 | 0 | 5 | 627 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 0 | 0 |
| PART | 1 | 9 | 3 | 13 | 0 | 0 | 0 | 8 | 0 | 2546 | 8 | 0 | 4 | 18 | 0 | 10 | 0 | 0 |
| PRON | 2 | 12 | 6 | 12 | 0 | 8 | 0 | 17 | 0 | 1 | 7950 | 1 | 346 | 11 | 0 | 17 | 0 | 0 |
| PROPN | 2 | 1 | 0 | 0 | 0 | 12 | 0 | 20 | 0 | 0 | 1 | 1938 | 13 | 0 | 0 | 1 | 0 | 0 |
| PUNCT | 1 | 6 | 3 | 2 | 4 | 0 | 0 | 10 | 0 | 1 | 3 | 0 | 8755 | 5 | 0 | 5 | 0 | 0 |
| SCONJ | 0 | 24 | 7 | 0 | 0 | 1 | 0 | 4 | 0 | 18 | 10 | 0 | 37 | 1885 | 0 | 10 | 0 | 0 |
| SYM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 53 | 0 | 0 | 0 |
| VERB | 16 | 4 | 5 | 27 | 1 | 10 | 1 | 47 | 0 | 1 | 5 | 3 | 21 | 4 | 0 | 10030 | 0 | 0 |
| X | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 4 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 69 | 0 |

Table 10: Kakao&Brain POS confusion matrix on W&I development.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X | Other |
|-------|------|------|------|------|-------|------|------|-------|-----|------|------|-------|-------|-------|-----|------|----|-------|
| ADJ | 6272 | 20 | 30 | 1 | 0 | 100 | 0 | 57 | 3 | 0 | 5 | 9 | 33 | 0 | 0 | 34 | 0 | 0 |
| ADP | 13 | 7667 | 11 | 6 | 1 | 3 | 0 | 20 | 0 | 17 | 14 | 2 | 49 | 37 | 0 | 41 | 0 | 0 |
| ADV | 28 | 18 | 4655 | 15 | 6 | 15 | 6 | 16 | 0 | 1 | 16 | 2 | 115 | 12 | 0 | 18 | 0 | 0 |
| AUX | 2 | 2 | 10 | 5668 | 1 | 2 | 0 | 23 | 0 | 9 | 48 | 3 | 16 | 1 | 0 | 48 | 0 | 0 |
| CCONJ | 0 | 1 | 1 | 1 | 2950 | 3 | 0 | 7 | 0 | 0 | 2 | 0 | 52 | 0 | 0 | 0 | 0 | 0 |
| DET | 17 | 98 | 9 | 11 | 6 | 7207 | 2 | 7 | 0 | 6 | 25 | 1 | 81 | 9 | 0 | 44 | 0 | 0 |
| INTJ | 0 | 0 | 0 | 0 | 0 | 0 | 93 | 0 | 0 | 0 | 0 | 3 | 2 | 0 | 0 | 0 | 0 | 0 |
| NOUN | 64 | 9 | 20 | 4 | 6 | 26 | 0 | 16068 | 11 | 1 | 15 | 26 | 59 | 5 | 0 | 82 | 0 | 0 |
| NUM | 1 | 1 | 2 | 0 | 0 | 2 | 0 | 6 | 630 | 0 | 0 | 0 | 7 | 0 | 0 | 1 | 0 | 0 |
| PART | 2 | 11 | 6 | 10 | 1 | 2 | 0 | 12 | 0 | 2521 | 11 | 0 | 4 | 22 | 0 | 18 | 0 | 0 |
| PRON | 3 | 14 | 11 | 13 | 3 | 11 | 0 | 24 | 2 | 2 | 7974 | 1 | 288 | 14 | 0 | 23 | 0 | 0 |
| PROPN | 12 | 3 | 3 | 0 | 0 | 16 | 1 | 50 | 0 | 1 | 6 | 1879 | 10 | 0 | 0 | 6 | 1 | 0 |
| PUNCT | 6 | 6 | 6 | 4 | 1 | 0 | 0 | 35 | 1 | 0 | 1 | 0 | 8719 | 10 | 0 | 6 | 0 | 0 |
| SCONJ | 1 | 21 | 7 | 0 | 1 | 2 | 0 | 7 | 0 | 18 | 8 | 0 | 38 | 1881 | 0 | 12 | 0 | 0 |
| SYM | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 48 | 1 | 0 | 0 |
| VERB | 17 | 11 | 10 | 34 | 1 | 7 | 1 | 77 | 1 | 2 | 7 | 4 | 26 | 3 | 0 | 9974 | 0 | 0 |
| X | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 3 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 1 | 65 | 0 |

Table 11: SHUYAO POS confusion matrix on W&I development.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X | Other |
|-------|------|------|------|------|-------|------|------|-------|-----|------|------|-------|-------|-------|-----|------|----|-------|
| ADJ | 6259 | 23 | 24 | 3 | 4 | 118 | 0 | 55 | 0 | 0 | 8 | 9 | 30 | 0 | 0 | 31 | 0 | 0 |
| ADP | 15 | 7642 | 8 | 12 | 1 | 5 | 0 | 35 | 0 | 21 | 12 | 1 | 61 | 26 | 0 | 42 | 0 | 0 |
| ADV | 23 | 19 | 4610 | 20 | 9 | 17 | 10 | 21 | 0 | 2 | 21 | 4 | 122 | 19 | 0 | 26 | 0 | 0 |
| AUX | 3 | 1 | 14 | 5646 | 3 | 4 | 0 | 24 | 0 | 12 | 56 | 2 | 13 | 2 | 0 | 53 | 0 | 0 |
| CCONJ | 3 | 1 | 4 | 1 | 2923 | 2 | 0 | 7 | 0 | 0 | 5 | 1 | 69 | 0 | 0 | 1 | 0 | 0 |
| DET | 17 | 122 | 16 | 12 | 10 | 7136 | 1 | 17 | 2 | 8 | 30 | 2 | 72 | 15 | 0 | 63 | 0 | 0 |
| INTJ | 0 | 0 | 0 | 0 | 0 | 0 | 91 | 0 | 0 | 0 | 2 | 1 | 3 | 0 | 0 | 1 | 0 | 0 |
| NOUN | 50 | 9 | 15 | 7 | 8 | 33 | 0 | 16066 | 1 | 5 | 21 | 32 | 60 | 3 | 0 | 86 | 0 | 0 |
| NUM | 3 | 1 | 0 | 0 | 1 | 3 | 0 | 6 | 626 | 1 | 1 | 0 | 6 | 1 | 0 | 1 | 0 | 0 |
| PART | 3 | 10 | 12 | 14 | 1 | 0 | 0 | 12 | 0 | 2513 | 9 | 0 | 3 | 24 | 0 | 19 | 0 | 0 |
| PRON | 4 | 21 | 9 | 17 | 10 | 9 | 0 | 27 | 0 | 2 | 7946 | 3 | 293 | 16 | 0 | 26 | 0 | 0 |
| PROPN | 10 | 1 | 0 | 0 | 0 | 16 | 0 | 32 | 1 | 0 | 1 | 1920 | 6 | 0 | 0 | 1 | 0 | 0 |
| PUNCT | 1 | 4 | 7 | 5 | 4 | 2 | 2 | 28 | 0 | 2 | 6 | 3 | 8718 | 7 | 0 | 6 | 0 | 0 |
| SCONJ | 1 | 20 | 8 | 1 | 1 | 2 | 0 | 6 | 0 | 13 | 12 | 0 | 51 | 1870 | 0 | 11 | 0 | 0 |
| SYM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 48 | 1 | 0 | 0 |
| VERB | 25 | 9 | 11 | 48 | 4 | 6 | 1 | 63 | 0 | 6 | 13 | 1 | 25 | 3 | 0 | 9959 | 1 | 0 |
| X | 0 | 0 | 2 | 1 | 1 | 2 | 0 | 6 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 63 | 0 |

Table 12: CAMB-CUED POS confusion matrix on W&I development.

3.3 W&I by levels

We provide a separate matrix for each level of learner proficiency in the W&I corpus and the native proficiency of LOCNESS.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X |
|-------|-------|-------|------|-------|-------|-------|------|-------|------|------|-------|-------|-------|-------|-----|-------|-----|
| ADJ | 11594 | 70 | 190 | 11 | 11 | 442 | 2 | 456 | 7 | 5 | 36 | 65 | 123 | 12 | 0 | 225 | 0 |
| ADP | 101 | 13880 | 167 | 35 | 33 | 70 | 7 | 233 | 12 | 103 | 82 | 14 | 266 | 193 | 1 | 388 | 0 |
| ADV | 181 | 118 | 8493 | 119 | 78 | 107 | 28 | 173 | 5 | 23 | 134 | 18 | 361 | 80 | 1 | 147 | 0 |
| AUX | 23 | 8 | 59 | 10907 | 9 | 16 | 1 | 111 | 1 | 39 | 324 | 14 | 75 | 8 | 1 | 513 | 0 |
| CCONJ | 13 | 17 | 9 | 4 | 5843 | 9 | 0 | 45 | 3 | 2 | 8 | 0 | 365 | 6 | 0 | 11 | 0 |
| DET | 82 | 522 | 105 | 115 | 66 | 12806 | 5 | 147 | 22 | 22 | 186 | 9 | 341 | 112 | 0 | 410 | 0 |
| INTJ | 4 | 0 | 4 | 2 | 0 | 3 | 312 | 8 | 0 | 1 | 1 | 5 | 9 | 1 | 0 | 5 | 0 |
| NOUN | 372 | 71 | 170 | 71 | 25 | 156 | 12 | 31992 | 28 | 10 | 133 | 235 | 300 | 32 | 1 | 658 | 4 |
| NUM | 8 | 17 | 8 | 7 | 1 | 32 | 0 | 48 | 1405 | 1 | 10 | 1 | 14 | 1 | 2 | 11 | 0 |
| PART | 14 | 79 | 51 | 81 | 11 | 11 | 1 | 63 | 1 | 4845 | 38 | 7 | 32 | 104 | 0 | 125 | 0 |
| PRON | 24 | 94 | 68 | 81 | 48 | 78 | 1 | 165 | 6 | 23 | 19290 | 15 | 1110 | 100 | 0 | 137 | 1 |
| PROPN | 32 | 14 | 8 | 2 | 4 | 45 | 6 | 156 | 1 | 1 | 7 | 3297 | 62 | 3 | 0 | 16 | 0 |
| PUNCT | 21 | 22 | 26 | 9 | 24 | 8 | 5 | 140 | 3 | 11 | 44 | 11 | 17928 | 26 | 1 | 58 | 1 |
| SCONJ | 16 | 113 | 70 | 7 | 20 | 15 | 1 | 46 | 1 | 61 | 107 | 9 | 254 | 3270 | 0 | 76 | 0 |
| SYM | 0 | 2 | 0 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 77 | 1 | 1 |
| VERB | 169 | 59 | 65 | 251 | 9 | 40 | 4 | 393 | 1 | 21 | 43 | 28 | 103 | 21 | 3 | 21005 | 1 |
| X | 1 | 0 | 3 | 1 | 1 | 2 | 0 | 5 | 0 | 1 | 0 | 4 | 2 | 1 | 2 | 1 | 111 |

Table 13: Matrix for POS on W&I training set of A (lowest) level learners.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X |
|-------|-------|-------|-------|-------|-------|-------|------|-------|------|------|-------|-------|-------|-------|-----|-------|-----|
| ADJ | 16232 | 66 | 155 | 14 | 9 | 385 | 2 | 319 | 10 | 2 | 37 | 64 | 121 | 17 | 0 | 158 | 1 |
| ADP | 80 | 19910 | 132 | 37 | 31 | 57 | 4 | 166 | 15 | 61 | 58 | 11 | 263 | 165 | 1 | 307 | 1 |
| ADV | 142 | 96 | 11753 | 126 | 40 | 67 | 27 | 158 | 1 | 31 | 119 | 8 | 365 | 89 | 0 | 134 | 1 |
| AUX | 13 | 6 | 61 | 14903 | 4 | 15 | 1 | 92 | 2 | 33 | 213 | 8 | 92 | 14 | 0 | 340 | 0 |
| CCONJ | 13 | 13 | 11 | 1 | 7499 | 5 | 1 | 30 | 0 | 2 | 11 | 3 | 336 | 8 | 0 | 9 | 1 |
| DET | 70 | 536 | 99 | 88 | 63 | 17932 | 6 | 109 | 20 | 14 | 131 | 6 | 281 | 96 | 0 | 316 | 0 |
| INTJ | 3 | 0 | 6 | 0 | 0 | 0 | 269 | 6 | 0 | 0 | 3 | 3 | 8 | 1 | 0 | 2 | 0 |
| NOUN | 288 | 53 | 124 | 36 | 21 | 151 | 8 | 42212 | 31 | 6 | 71 | 181 | 196 | 32 | 2 | 437 | 2 |
| NUM | 8 | 9 | 3 | 2 | 0 | 31 | 0 | 44 | 1756 | 0 | 6 | 6 | 17 | 1 | 0 | 8 | 0 |
| PART | 16 | 60 | 49 | 43 | 4 | 12 | 0 | 55 | 0 | 6880 | 42 | 8 | 35 | 114 | 0 | 130 | 0 |
| PRON | 14 | 65 | 44 | 58 | 45 | 74 | 1 | 139 | 13 | 15 | 23771 | 9 | 1101 | 100 | 0 | 109 | 0 |
| PROPN | 20 | 8 | 15 | 4 | 2 | 58 | 1 | 169 | 0 | 1 | 9 | 4854 | 59 | 5 | 0 | 11 | 1 |
| PUNCT | 48 | 20 | 30 | 18 | 35 | 21 | 2 | 172 | 0 | 8 | 34 | 14 | 23620 | 26 | 1 | 64 | 0 |
| SCONJ | 10 | 115 | 76 | 10 | 17 | 14 | 0 | 46 | 1 | 70 | 63 | 7 | 277 | 4720 | 0 | 34 | 0 |
| SYM | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 79 | 1 | 0 |
| VERB | 165 | 44 | 65 | 182 | 7 | 37 | 2 | 329 | 2 | 19 | 32 | 25 | 96 | 22 | 1 | 27432 | 0 |
| X | 0 | 0 | 3 | 0 | 0 | 2 | 0 | 8 | 0 | 1 | 2 | 4 | 5 | 0 | 2 | 4 | 143 |

Table 14: Matrix for POS on W&I training set of B (mid) level learners.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X |
|-------|-------|-------|-------|-------|-------|-------|------|-------|------|------|-------|-------|-------|-------|-----|-------|-----|
| ADJ | 14938 | 27 | 81 | 8 | 3 | 174 | 0 | 139 | 2 | 3 | 13 | 23 | 53 | 1 | 0 | 67 | 0 |
| ADP | 35 | 18358 | 60 | 13 | 13 | 14 | 2 | 58 | 3 | 19 | 24 | 1 | 161 | 48 | 1 | 133 | 1 |
| ADV | 64 | 50 | 11003 | 34 | 26 | 41 | 7 | 74 | 2 | 8 | 45 | 2 | 217 | 28 | 0 | 47 | 0 |
| AUX | 8 | 8 | 22 | 12706 | 6 | 4 | 0 | 38 | 0 | 13 | 81 | 5 | 49 | 5 | 0 | 109 | 0 |
| CCONJ | 1 | 12 | 4 | 6 | 6767 | 9 | 0 | 15 | 0 | 2 | 6 | 0 | 167 | 0 | 0 | 3 | 0 |
| DET | 22 | 226 | 26 | 24 | 18 | 16736 | 0 | 32 | 2 | 4 | 49 | 5 | 142 | 19 | 0 | 125 | 1 |
| INTJ | 1 | 1 | 4 | 0 | 0 | 0 | 216 | 2 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 0 | 0 |
| NOUN | 116 | 32 | 29 | 14 | 5 | 63 | 2 | 37581 | 8 | 2 | 38 | 52 | 91 | 4 | 1 | 172 | 2 |
| NUM | 10 | 1 | 3 | 1 | 0 | 9 | 0 | 18 | 1501 | 0 | 7 | 0 | 6 | 0 | 1 | 5 | 1 |
| PART | 3 | 18 | 26 | 19 | 4 | 2 | 0 | 31 | 2 | 6167 | 21 | 5 | 18 | 57 | 0 | 29 | 0 |
| PRON | 8 | 30 | 18 | 31 | 4 | 20 | 0 | 32 | 2 | 0 | 20363 | 3 | 574 | 32 | 0 | 37 | 0 |
| PROPN | 14 | 2 | 7 | 1 | 1 | 25 | 0 | 50 | 0 | 1 | 1 | 4511 | 52 | 0 | 0 | 5 | 0 |
| PUNCT | 25 | 9 | 15 | 5 | 11 | 6 | 2 | 88 | 1 | 5 | 20 | 3 | 21344 | 7 | 0 | 14 | 0 |
| SCONJ | 8 | 59 | 34 | 1 | 5 | 10 | 1 | 19 | 0 | 20 | 32 | 5 | 144 | 4239 | 0 | 16 | 0 |
| SYM | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 4 | 0 | 0 | 0 | 0 | 9 | 0 | 84 | 0 | 0 |
| VERB | 58 | 12 | 18 | 66 | 2 | 13 | 1 | 148 | 2 | 3 | 11 | 4 | 75 | 3 | 0 | 24299 | 2 |
| X | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 120 |

Table 15: Matrix for POS on W&I training set of C (highest) level learners.

| | ADJ | ADP | ADV | AUX | CCONJ | DET | INTJ | NOUN | NUM | PART | PRON | PROPN | PUNCT | SCONJ | SYM | VERB | X |
|-------|------|------|------|------|-------|------|------|------|-----|------|------|-------|-------|-------|-----|------|----|
| ADJ | 1753 | 1 | 5 | 0 | 0 | 3 | 0 | 18 | 0 | 0 | 1 | 4 | 10 | 0 | 0 | 8 | 0 |
| ADP | 4 | 2228 | 5 | 0 | 1 | 1 | 0 | 9 | 0 | 3 | 1 | 0 | 20 | 6 | 0 | 5 | 0 |
| ADV | 6 | 6 | 1264 | 2 | 1 | 0 | 0 | 7 | 0 | 0 | 1 | 0 | 41 | 2 | 0 | 0 | 1 |
| AUX | 0 | 0 | 0 | 1720 | 0 | 0 | 0 | 4 | 0 | 1 | 2 | 0 | 11 | 0 | 0 | 5 | 0 |
| CCONJ | 0 | 2 | 0 | 0 | 763 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 18 | 1 | 0 | 0 | 0 |
| DET | 3 | 17 | 2 | 4 | 3 | 2138 | 0 | 0 | 0 | 0 | 2 | 0 | 27 | 2 | 0 | 3 | 0 |
| INTJ | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NOUN | 10 | 2 | 7 | 0 | 0 | 2 | 0 | 4554 | 1 | 0 | 1 | 8 | 24 | 3 | 1 | 19 | 0 |
| NUM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 153 | 0 | 0 | 0 | 4 | 0 | 2 | 0 | 1 |
| PART | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 643 | 2 | 0 | 4 | 0 | 0 | 0 | 0 |
| PRON | 0 | 1 | 2 | 1 | 2 | 0 | 0 | 1 | 2 | 0 | 1422 | 0 | 55 | 2 | 0 | 4 | 0 |
| PROPN | 5 | 0 | 0 | 0 | 0 | 1 | 0 | 11 | 0 | 0 | 0 | 568 | 1 | 0 | 0 | 1 | 0 |
| PUNCT | 4 | 3 | 2 | 1 | 0 | 1 | 0 | 22 | 1 | 1 | 1 | 1 | 2304 | 3 | 0 | 4 | 0 |
| SCONJ | 0 | 3 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 4 | 0 | 16 | 504 | 0 | 3 | 0 |
| SYM | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 2 | 0 | 12 | 1 | 1 |
| VERB | 2 | 1 | 2 | 4 | 0 | 0 | 0 | 23 | 0 | 0 | 1 | 4 | 12 | 0 | 0 | 2459 | 0 |
| X | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 19 |

Table 16: Matrix for POS on LOCNESS training set of N (native) writers.

3.4 Comparison to NUCLE

Confusion matrix between SeRCL and NUCLE types over the NUCLE train set is given in [17](#). SeRCL types with less than 30 occurrences are omitted. Aggregations are found in the rightmost columns (and do not include -rloc)

| | Wci | Pref | Prep | ArtOrd | Vm | Others | Wtone | Smood | Mec | Pform | Trans | Cit | Wa | Sfrag | WOnc | Wform | Sub | Npos | VO | SVA | Srun | Na | Vform | Spur | Um | WAdv | Vt | max | 2ns | 3rd | sum | max/sum | top3/sum |
|--------------|-----|------|------|--------|----|--------|-------|-------|-----|-------|-------|-----|----|-------|------|-------|-----|------|-----|-----|------|----|-------|------|-----|------|------|-----|--------------|--------------|--------------|--------------|--------------|
| None->DET | 27 | 6 | 0 | 3434 | 0 | 33 | 0 | 0 | 10 | 0 | 7 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 12 | 0 | 23 | 1 | 7 | 2 | 1 | 15 | 3434 | 33 | 27 | 3581 | 0.95890014 | 0.975705103 | |
| CCONJ->ADV | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 55 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 55 | 1 | 1 | 58 | 0.9482758621 | 0.9827586207 | |
| DET->None | 6 | 4 | 0 | 1998 | 1 | 1 | 2 | 0 | 7 | 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 8 | 2 | 66 | 1 | 1 | 2 | 1 | 5 | 1998 | 46 | 8 | 2111 | 0.9366708669 | 0.9815223434 | |
| ADP->None | 20 | 0 | 610 | 2 | 0 | 1 | 1 | 1 | 3 | 0 | 17 | 2 | 0 | 1 | 3 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 4 | 2 | 0 | 3 | 610 | 20 | 17 | 674 | 0.9050445104 | 0.9599406528 | |
| PRON->None | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 111 | 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 2 | 0 | 0 | 111 | 8 | 4 | 131 | 0.8473282443 | 0.9389312977 |
| PART->SCONJ | 32 | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 32 | 4 | 1 | 40 | 0.8 | 0.925 | |
| DET->PRON | 3 | 135 | 0 | 15 | 0 | 1 | 0 | 0 | 5 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 135 | 15 | 6 | 172 | 0.7848837209 | 0.9060767442 | |
| NOUN->DET | 4 | 2 | 0 | 40 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 40 | 4 | 2 | 51 | 0.7843137255 | 0.9019607843 | |
| None->None | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 33 | 7 | 1 | 43 | 0.7674418605 | 0.9534883721 | |
| VERB->None | 14 | 1 | 0 | 1 | 0 | 6 | 23 | 3 | 3 | 0 | 3 | 18 | 0 | 6 | 19 | 0 | 0 | 1 | 0 | 9 | 4 | 0 | 2 | 4 | 458 | 22 | 458 | 23 | 22 | 999 | 0.7646076795 | 0.8397238881 | |
| ADJ->DET | 18 | 3 | 0 | 111 | 0 | 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 111 | 18 | 5 | 151 | 0.7350953377 | 0.8851472185 | |
| VERB->DET | 4 | 0 | 0 | 38 | 0 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 38 | 4 | 2 | 52 | 0.7307692308 | 0.8461538462 | |
| None->PUNCT | 6 | 1 | 0 | 2 | 1 | 8 | 1 | 0 | 915 | 0 | 13 | 51 | 2 | 1 | 3 | 1 | 9 | 4 | 0 | 0 | 236 | 1 | 2 | 0 | 0 | 5 | 915 | 236 | 51 | 1262 | 0.7250396197 | 0.9524564184 | |
| ADJ->None | 7 | 0 | 0 | 2 | 0 | 1 | 5 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 1 | 85 | 7 | 5 | 118 | 0.7203389831 | 0.8220330863 | | | |
| SCONJ->None | 1 | 0 | 62 | 9 | 0 | 1 | 0 | 1 | 1 | 0 | 5 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 62 | 9 | 5 | 87 | 0.7126438782 | 0.873562184 | |
| PUNCT->CCONJ | 1 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 10 | 0 | 112 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 0 | 23 | 0 | 0 | 2 | 0 | 0 | 0 | 112 | 23 | 0 | 159 | 0.704025157 | 0.9119496855 | |
| AUX->None | 0 | 1 | 0 | 0 | 18 | 1 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 19 | 3 | 2 | 0 | 109 | 109 | 19 | 18 | 160 | 0.68125 | 0.9125 | | |
| VERB->ADP | 65 | 0 | 1 | 0 | 0 | 3 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 2 | 3 | 1 | 0 | 0 | 4 | 1 | 0 | 4 | 3 | 0 | 0 | 65 | 5 | 4 | 96 | 0.6778031833 | 0.7708333333 | |
| PRON->PUNCT | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 30 | 0 | 0 | 0 | 1 | 0 | 0 | 30 | 7 | 4 | 45 | 0.6666666667 | 0.9111111111 | |
| NOUN->None | 25 | 1 | 0 | 4 | 0 | 7 | 15 | 2 | 12 | 1 | 3 | 90 | 2 | 4 | 19 | 5 | 0 | 1 | 0 | 3 | 1 | 14 | 0 | 4 | 395 | 9 | 6 | 395 | 90 | 25 | 623 | 0.6340288925 | 0.8186195827 |
| None->NOUN | 92 | 2 | 0 | 3 | 0 | 351 | 0 | 2 | 4 | 1 | 7 | 20 | 3 | 8 | 20 | 2 | 4 | 2 | 3 | 4 | 1 | 6 | 0 | 19 | 3 | 2 | 3 | 351 | 92 | 20 | 562 | 0.6245551601 | 0.8238434164 |
| NOUN->PRON | 1 | 1 | 0 | 9 | 0 | 2 | 3 | 0 | 54 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 9 | 0 | 0 | 1 | 0 | 54 | 9 | 9 | 88 | 0.6162663636 | 0.8181818182 | | |
| PUNCT->None | 0 | 1 | 0 | 1 | 0 | 2 | 2 | 1 | 267 | 0 | 2 | 73 | 1 | 4 | 4 | 0 | 1 | 1 | 0 | 0 | 5 | 1 | 0 | 70 | 1 | 2 | 267 | 73 | 70 | 439 | 0.6082004556 | 0.9339407745 | |
| ADJ->ADV | 25 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 2 | 0 | 18 | 0 | 0 | 0 | 1 | 96 | 0 | 0 | 0 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 96 | 25 | 18 | 158 | 0.6079493667 | 0.8797468354 | |
| ADV->ADP | 41 | 0 | 3 | 0 | 0 | 1 | 2 | 0 | 5 | 0 | 11 | 0 | 0 | 1 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 41 | 11 | 5 | 70 | 0.5857142857 | 0.8142857143 | |
| None->ADJ | 13 | 0 | 4 | 0 | 0 | 55 | 0 | 1 | 1 | 0 | 3 | 0 | 0 | 3 | 6 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 55 | 5 | 0 | 95 | 0.5789473684 | 0.7730473684 | |
| NOUN->PUNCT | 2 | 1 | 0 | 1 | 0 | 2 | 2 | 0 | 57 | 1 | 2 | 2 | 1 | 0 | 4 | 1 | 5 | 2 | 0 | 0 | 14 | 5 | 0 | 1 | 0 | 0 | 57 | 14 | 5 | 101 | 0.5643643636 | 0.7524724725 | |
| ADP->VERB | 68 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 11 | 5 | 1 | 0 | 9 | 7 | 0 | 0 | 68 | 11 | 9 | 121 | 0.5619834711 | 0.7272727273 | |
| None->CCONJ | 10 | 0 | 0 | 0 | 0 | 11 | 0 | 0 | 0 | 0 | 91 | 0 | 0 | 0 | 0 | 1 | 6 | 0 | 0 | 40 | 0 | 0 | 0 | 0 | 0 | 0 | 91 | 2 | 0 | 113 | 0.5366202669 | 0.8711653434 | |
| NOUN->PRON | 14 | 70 | 0 | 5 | 0 | 5 | 1 | 0 | 0 | 9 | 1 | 0 | 0 | 1 | 3 | 5 | 1 | 2 | 0 | 0 | 0 | 4 | 0 | 6 | 0 | 0 | 70 | 14 | 9 | 127 | 0.5511811024 | 0.7322834646 | |
| NOUN->ADP | 24 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 5 | 0 | 3 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 2 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 24 | 5 | 3 | 44 | 0.5454545455 | 0.7272727273 | |
| DET->ADJ | 35 | 4 | 0 | 15 | 0 | 1 | 0 | 0 | 1 | 3 | 3 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 35 | 15 | 4 | 69 | 0.5072465768 | 0.7820606957 | |
| ADJ->PUNCT | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 16 | 0 | 2 | 0 | 0 | 2 | 2 | 0 | 1 | 2 | 0 | 23 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 2 | 0 | 47 | 0.4891617021 | 0.872346255 | |
| SCONJ->DET | 10 | 8 | 0 | 43 | 0 | 3 | 16 | 0 | 2 | 4 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 43 | 16 | 10 | 89 | 0.4831466674 | 0.7752808989 | |
| ADJ->ADV | 12 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 17 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 17 | 12 | 2 | 36 | 0.4672222222 | 0.8611111111 | |
| None->ADJ | 21 | 0 | 0 | 14 | 0 | 1 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 2 | 47 | 1 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 5 | 1 | 47 | 21 | 14 | 100 | 0.47 | 0.82 | |
| NOUN->ADJ | 103 | 4 | 7 | 0 | 0 | 17 | 27 | 1 | 34 | 3 | 3 | 2 | 3 | 11 | 243 | 2 | 2 | 0 | 5 | 0 | 26 | 5 | 6 | 0 | 5 | 9 | 243 | 103 | 34 | 519 | 0.468208925 | 0.732172264 | |
| ADP->DET | 13 | 0 | 1 | 18 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 18 | 13 | 2 | 39 | 0.4615384615 | 0.8461538462 | |
| ADP->NOUN | 24 | 0 | 0 | 4 | 0 | 3 | 0 | 0 | 3 | 0 | 5 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 2 | 0 | 5 | 0 | 1 | 0 | 0 | 0 | 24 | 5 | 5 | 52 | 0.4615384615 | 0.6538461538 | |
| SCONJ->ADP | 27 | 1 | 22 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 22 | 4 | 59 | 0.4576271186 | 0.898300847 | |
| ADV->ADJ | 7 | 0 | 0 | 0 | 6 | 1 | 1 | 0 | 4 | 0 | 0 | 0 | 0 | 1 | 4 | 7 | 0 | 0 | 5 | 0 | 0 | 1 | 1 | 0 | 36 | 8 | 36 | 8 | 7 | 82 | 0.439243942 | 0.6219521955 | |
| VERB->ADV | 28 | 0 | 0 | 0 | 0 | 5 | 6 | 0 | 1 | 0 | 8 | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 1 | 1 | 0 | 1 | 3 | 1 | 2 | 28 | 8 | 6 | 65 | 0.4307692308 | 0.6461538462 | |
| None->PART | 31 | 0 | 0 | 1 | 0 | 18 | 0 | 0 | 31 | 1 | 5 | 0 | 0 | 0 | 0 | 7 | 0 | 109 | 7 | 1 | 0 | 1 | 32 | 16 | 0 | 0 | 6 | 109 | 32 | 31 | 266 | 0.4097744361 | 0.6466165414 |
| None->AUX | 4 | 1 | 0 | 0 | 0 | 51 | 7 | 1 | 0 | 1 | 0 | 2 | 0 | 2 | 2 | 0 | 0 | 5 | 182 | 0 | 0 | 70 | 12 | 0 | 0 | 229 | 229 | 182 | 70 | 569 | 0.4024601569 | 0.845327065 | |
| SCONJ->VERB | 12 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 3 | 0 | 0 | 1 | 1 | 0 | 0 | 5 | 3 | 31 | 3870967742 | 0.6451612903 | | |
| PRON->ADJ | 7 | 6 | 0 | 1 | 0 | 3 | 1 | 0 | 1 | 2 | 16 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 7 | 6 | 42 | 0.380952381 | 0.6904761905 | |
| NOUN->ADV | 49 | 0 | 0 | 1 | 0 | 4 | 11 | 0 | 4 | 0 | 33 | 0 | 0 | 0 | 5 | 19 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 3 | 1 | 49 | 33 | 19 | 133 | 0.3684210526 | 0.7593984962 | |
| None->ADV | 27 | 1 | 0 | 0 | 0 | 47 | 0 | 0 | 17 | 0 | 67 | 1 | 0 | 1 | 5 | 0 | 7 | 0 | 1 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 47 | 17 | 7 | 85 | 0.3621621622 | 0.7621621622 | |
| ADP->ADV | 21 | 1 | 3 | 0 | 0 | 2 | 2 | 0 | 0 | 3 | 18 | 0 | 0 | 0 | 2 | 4 | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 18 | 4 | 58 | 0.3620369555 | 0.7413793103 | |
| AUX->VERB | 40 | 0 | 0 | 0 | 1 | 6 | 2 | 0 | 1 | 0 | 2 | 0 | 0 | 2 | 7 | 3 | 0 | 0 | 0 | 9 | 0 | 0 | 19 | 1 | 0 | 3 | 16 | 40 | 11 | 112 | 0.3571428571 | 0.6696428571 | |
| ADJ->SCONJ | 16 | 0 | 1 | 1 | 0 | 3 | 0 | 0 | 4 | 0 | 12 | 0 | 0 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 12 | 4 | 45 | 0.3555555556 | 0.7111111111 | |
| ADP->SCONJ | 18 | 0 | 13 | 0 | 0 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | |

| | Acc | Dat | Gen | Ins | Loc | Nom |
|------------|------------|------------|------------|------------|------------|------------|
| Acc | 0 | 12 | 27 | 9 | 12 | 19 |
| Dat | 2 | 0 | 6 | 2 | 1 | 2 |
| Gen | 19 | 8 | 0 | 13 | 6 | 18 |
| Ins | 4 | 5 | 6 | 0 | 5 | 8 |
| Loc | 22 | 3 | 9 | 8 | 0 | 2 |
| Nom | 32 | 5 | 30 | 48 | 3 | 0 |

Table 18: Case agreement corrections on adjectives

| | Fem | Masc | Neut |
|-------------|------------|-------------|-------------|
| Fem | 0 | 35 | 19 |
| Masc | 39 | 0 | 28 |
| Neut | 28 | 18 | 0 |

Table 19: Gender agreement corrections on adjectives

| | Plur | Sing |
|-------------|-------------|-------------|
| Plur | 0 | 50 |
| Sing | 86 | 0 |

Table 20: Gender agreement corrections on adjectives

References

David Little. 2006. The common european framework of reference for languages: Content, purpose, origin, reception and impact. In *Cambridge Journal*.