Test 1 (Original String Matching)
Document 1 Score: 0.21
Document 2 Score: 0.41
Document 3 Score: 0.25
Document 4 Score: 0.33
Document 5 Score: 0.49
Document 6 Score: 0.28
Document 7 Score: 0.19
Document 8 Score: 0.23
Document 9 Score: 0.39
Document 10 Score: 0.36
Document 11 Score: 0.15
Document 12 Score: 0.16
Document 13 Score: 0.17
Document 14 Score: 0.16
Document 15 Score: 0.15
Document 16 Score: 0.23
Document 17 Score: 0.46
Document 18 Score: 0.27
Document 19 Score: 0.29
Document 20 Score: 0.38
Document 21 Score: 0.21
Document 22 Score: 0.23

Final accuracy score: 0.2724

Test 2 (T1 w/ref == null condition)
Document 1 Score: 0.21
Document 2 Score: 0.41
Document 3 Score: 0.29
Document 4 Score: 0.33
Document 5 Score: 0.49
Document 6 Score: 0.28
Document 7 Score: 0.17
Document 8 Score: 0.21
Document 9 Score: 0.35
Document 10 Score: 0.32
Document 11 Score: 0.27
Document 12 Score: 0.16
Document 13 Score: 0.17
Document 14 Score: 0.24
Document 15 Score: 0.20
Document 16 Score: 0.17
Document 17 Score: 0.38
Document 18 Score: 0.22
Document 19 Score: 0.29
Document 20 Score: 0.38
Document 21 Score: 0.29
Document 22 Score: 0.24

Final accuracy score: 0.2768

Test 3 (T2 w/ np ≠ "a", "the", "of")
Document 1 Score: 0.21
Document 2 Score: 0.41
Document 3 Score: 0.29
Document 4 Score: 0.33
Document 5 Score: 0.49
Document 6 Score: 0.28
Document 7 Score: 0.19
Document 8 Score: 0.27
Document 9 Score: 0.37
Document 10 Score: 0.39
Document 11 Score: 0.31
Document 12 Score: 0.18
Document 13 Score: 0.17
Document 14 Score: 0.29
Document 15 Score: 0.17
Document 16 Score: 0.23
Document 17 Score: 0.32
Document 18 Score: 0.33
Document 19 Score: 0.31
Document 20 Score: 0.37
Document 21 Score: 0.34
Document 22 Score: 0.32

Final accuracy score: 0.3027

= Increase from prev test

Test 4 (T3 w/ np ≠ "to", "for", "our")
Document 1 Score: 0.21
Document 2 Score: 0.41
Document 3 Score: 0.21
Document 4 Score: 0.33
Document 5 Score: 0.49
Document 6 Score: 0.28
Document 7 Score: 0.19
Document 8 Score: 0.27
Document 9 Score: 0.37
Document 10 Score: 0.39
Document 11 Score: 0.31
Document 12 Score: 0.18
Document 13 Score: 0.17
Document 14 Score: 0.29
Document 15 Score: 0.15
Document 16 Score: 0.23
Document 17 Score: 0.32
Document 18 Score: 0.33
Document 19 Score: 0.31
Document 20 Score: 0.37
Document 21 Score: 0.34
Document 22 Score: 0.32

Final accuracy score: 0.2995

= Decrease from prev test

Test 5 (T4 w/o np ≠ "our")
Document 1 Score: 0.21
Document 2 Score: 0.41
Document 3 Score: 0.29
Document 4 Score: 0.33
Document 5 Score: 0.49
Document 6 Score: 0.28
Document 7 Score: 0.19
Document 8 Score: 0.27
Document 9 Score: 0.37
Document 10 Score: 0.39
Document 11 Score: 0.31
Document 12 Score: 0.18
Document 13 Score: 0.17
Document 14 Score: 0.29
Document 15 Score: 0.17
Document 16 Score: 0.23
Document 17 Score: 0.32
Document 18 Score: 0.33
Document 19 Score: 0.31
Document 20 Score: 0.37
Document 21 Score: 0.34
Document 22 Score: 0.32

Final accuracy score: 0.3027

We are up to 30.27% with just string matching preexisting corefs. Next step will be trying to chunk noun phrases and string match them.