What Syntax can Contribute in the Entailment Task

Lucy Vanderwende and William B. Dolan

Microsoft Research
Redmond, WA 98052
{lucyv; billdol}@microsoft.com

Abstract. We describe our submission to the PASCAL Recognizing Textual Entailment Challenge, which attempts to isolate the set of Text-Hypothesis pairs whose categorization can be accurately predicted based solely on syntactic cues. Two human annotators examined each pair, showing that a surprisingly large proportion of the data – 34% of the test items – can be handled with syntax alone, while adding information from a general-purpose thesaurus increases this to 48%.

1 Introduction

The data set made available by the PASCAL Recognizing Textual Entailment Challenge provides a great opportunity to focus on a very difficult task, determining whether one sentence (the hypothesis, H) is entailed by another (the text, T).

Our goal was to isolate the class of T-H pairs whose categorization can be accurately predicted based solely on syntactic cues. This work is part of a larger ablation study aimed at measuring the impact of various NLP components on entailment and paraphrase.

We have chosen to provide a partial submission that addresses the following question: what proportion of the entailments in the PASCAL test set could be solved using a robust parser? We are encouraged that other entrants chose to focus on different baselines, specifically those involving lexical matching and edit distance. Collectively, these baselines should establish what the minimal system requirements might be for addressing the textual entailment task.

2 Details of Microsoft Research Submission

Various parsers providing constituent level analysis are now available to the research community, and state-of-the-art parsers have reported accuracy of between 89% and 90.1% F-measure (Collins and Duffy, 2002, Henderson 2004; see Ringger et al., 2004 for results with a non-treebank parser). There are also efforts to produce parsers that assign argument structure (Gildea and Jurafsky, 2002, and for example, Hacioglu et al., 2004). With these developments, we feel that syntax can be defined broadly to include such phenomena as argument assignment, intra-sentential pronoun anaphora resolution, and a set of alternations to establish equivalence on structural grounds.

Our goal was to establish a baseline for the entailment task that reflects what an idealized parser could accomplish, abstracting away from the analysis errors that any specific parsing system would inevitably introduce. We decided therefore to rely on human annotators to decide whether syntactic information alone is sufficient to make a judgment. Two human annotators evaluated each T-H pair, indicating whether the entailment was:

- True by Syntax,
- False by Syntax,
- Not Syntax,
- Can't Decide

Additionally, we allowed the annotators to indicate whether recourse to information in a general purpose thesaurus entry would allow a pair to be judged True or False. Both annotators were skilled linguists, and could be expected to determine what an idealized syntactic parser could accomplish. We should note at this point that it could prove impossible to automate the judgment process described in this paper; the rules-of-thumb used by the annotators to make True or False judgments could turn out to be incompatible with an operational system.

We found that 34% of the test items can be handled by syntax, broadly defined; 48% of the test items can be handled by syntax plus a general purpose thesaurus. The results of this experiment are summarized in Table 1:

Table 1. Summary of Microsoft Research partial submission; Run1 is without thesaurus, Run2 is with thesaurus

	Without thesaurus	Using thesaurus
True	69 (9%)	147 (18%)
False	197 (25%)	243 (30%)
Not syntax	534 (67%)	410 (51%)

Overall, inter-annotator agreement was 72%. Where there were disagreements, the annotators jointly decided which judgment was most appropriate in order to annotate all test items. Of the disagreements, 60% were between False and Not-Syntax, and 25% between True and Not-Syntax; the remainder of the differences involved either annotation errors or cases where one or both annotators chose Can't Decide. This confirms our anecdotal experience that it is easier to decide when syntax can be expected to return True, and that the annotators were uncertain when to assign False. In some cases, there are good syntactic clues for assigning False, which is why we designed the evaluation to force a choice between True, False, and Not-Syntax. But in many cases, it is simply the absence of syntactic equivalence or parallelism that results in a judgment of False, and most of the disagreements centered on these cases.

3 Results of Partial Submission

Our test results are not comparable to those of other systems, since obviously, our runs were produced by human annotators. In this section, we only want to briefly call

attention to those test items which showed a discrepancy between our adjudicated human annotation and those provided as gold standard. It is worth mentioning that we believe the task is well-defined, at least for the test items we evaluated. For the 295 test items returned in Run1 of our submission, 284 matched the judgment provided as gold standard, so that our inter-annotator agreement on this subset of the test set was 96%.

In Run1 (using an idealized parser, but no thesaurus), there were 11 discrepancies. Of the 3 cases where we judged the test item to be True but the gold standard for the item is False, one is clearly an annotation error (despite having two annotators!) and two are examples of strict inclusion, which we allowed as entailments but the data set does not (test items 1839 and 2077); see (1).

1. (pair id="2077", value="FALSE", task="QA")
   <T> They are made from the dust of four of Jupiter's tiniest moons.
   <H> Jupiter has four moons.

More difficult to characterize as a group are the 8 cases where we judged the test item to be False but the gold standard for the item is True (although 5/8 are from the QA section) The test items in question are: 1335, 1472, 1487, 1553, 1584, 1586, 1634, and 1682. It does appear to us that more knowledge is needed to judge these items than simply what is provided in the Text and Hypothesis. We therefore believe that these items should be removed from the data set, since pairs for which there was disagreement among the judges were discarded. Item 1634 is a representative example.

2. (pair id="1634", value="TRUE", task="IE")
   <T> William Leonard Jennings sobbed loudly as was charged with killing his 3-year-old son, Stephen, who was last seen alive on Dec. 12, 1962.
   <H> William Leonard Jennings killed his 3-year-old son, Stephen.

4 Requirements for a Syntax-Based System

We analyzed our human judgments to establish which syntactic phenomena a robust parser would need to handle in order to complete the entailment task. We can distinguish two categories: the level of syntactic analysis, further described in 4.1, and a set of alternations, described in 4.2. Section 4.3 describes the special handling of syntactic analysis for the purpose of establishing a T-H pair to be False. Most of the examples will be from the subset of judgments that are True and based solely on syntactic cues, because these sentence pairs often isolate the specific phenomena under discussion. We have included a list of syntactic phenomena and alternations for each judgment type in the Appendix, from which the cooccurence of phenomena can also be ascertained.

Additionally, we enumerate in the Appendix those test items, representing only a small fraction (3.5%) of the test set, which can be judged using single word replacement alone. An example of single-word replacement is the following:

3. (pair id="1996", value="TRUE", task="PP")
<T> Iraqi militants abduct 2 Turks in Iraq.
<H> Iraqi militants kidnap 2 Turks in Iraq.

4.1 Syntactic Analysis

The best illustration of the role played by syntactic evidence involves cases where predicate-argument assignment gives clear evidence for the judgment. (4a) and (4b) are good examples:

4. <T> Latvia, for instance, is the lowest-ranked team in the field but defeated World Cup semifinalist Turkey in a playoff to qualify for the final 16 of Euro 2004.

4a. (pair id="1897", value="TRUE", task="IE")
<H> Turkey is defeated by Latvia.

4b. (pair id="1896", value="FALSE", task="IE")
<H> Latvia is defeated by Turkey.

A more straightforward case is for a parser (in most cases, a preprocessing component to the parser) to account for Named Entity Recognition, identifying various expressions of an entity as equivalent, as in (5), where the strings *Reverend Frank Chikane* and *Rev Frank Chikane* refer to the same person.

5. (pair id="847", value="TRUE", task="CD")
<T> On hand to meet him with Mbeki were (…) and director general in the presidency, Reverend Frank Chikane.
<H>On hand to meet him with Mbeki were (….) and director general in the presidency, Rev Frank Chikane.

Other syntactic phenomena frequently observed in the data are T-H pairs that differ only in nominalization, as in (6), coordination, prepositional phrase attachment, and negation.

6. (pair id="1021", value="TRUE", task="RC")
<T> Sunday's election results demonstrated just how far the pendulum of public opinion has swung away from faith in Koizumi's promise to bolster the Japanese economy and make the political system more transparent and responsive to the peoples' needs.
<H> Koizumi promised to bolster the Japanese economy.

We also assume that a parser, broadly defined, will be capable of identifying the inferences invited by the apposition construction and by the predicate-complement constructions. In example (7), if the predicate holds for not only the subject, but also

for the apposition to the subject, then this sentence pair can also be handled straightforwardly:

7. (pair id="1616", value="TRUE", task="IE ")
<T> In 1833, Benjamin Harrison, the 23rd president of the United States, was born in North Bend, Ohio.
<H> The 23rd President of the United States was born in Ohio.

The examples presented above attempt to illustrate each syntactic phenomena in isolation. However, at least half of the T-H pairs require the identification of multiple phenomena simultaneously. The example in (8) requires involves Named Entity Recognition (*Honecker = Erich Honeker*), including identification of spelling variants, two instances of pronominal anaphora (*he = Honecker/Homeker*), and vp-cataphora (*did = build the Berlin Wall*). Nevertheless, if a parser is able to provide this level of syntactic analysis, the system can return a True judgment with confidence.

8. (pair id="621", value="TRUE", task="QA")
<T> Although Honecker led the Communist East German state between 1971 and 1989, he will be remembered most for what he did long before -- building the Berlin Wall.
<H>Erich Honeker built the Berlin Wall.

Finally, the identification of negation naturally plays a significant role in determining entailment, including the identification of morphological variants expressing negation. For Hypotheses that match an embedded clause in the Text, the subordinating conjunction and the semantic type of the main verb is also of importance; this phenomenon was relatively frequent in all but the subset of the test set we judged to be true using syntax alone. Consider examples (9) and (10):

9. (pair id="2025", value="FALSE", task="IR")
<T> There are a lot of farmers in Poland who worry about their future if Poland joins the European Union.
<H> Poland joins the European Union.

10. (pair id="2055", value="FALSE", task="QA")
<T> The fact that Einstein was invited to be the president of Israel is critical to an accurate understanding of one of the greatest individuals in modern history.
<H> Einstein is the president of Israel.

4.2 Syntactic Alternations

By far the most frequent alternation between Text and Hypothesis that a system needs to identify is an appositive construction promoted to main clause in the Hypothesis.

This alternation alone accounted for approximately 24% of the subset of the data we judged could be handled with syntactic analysis[1].

11. (pair id="760", value="TRUE", task="CD")
<T> The Alameda Central, west of the Zocalo, was created in 1592.
<H> The Alameda Central is west of the Zocalo.

Another frequent alternation involves material in a relative clause being promoted to a main clause in the Hypothesis, as in example (12), which includes named entity recognition as well:

12. (pair id="1060", value="TRUE", task="RC")
<T> (…) when Silva was sent in by Rio de Janeiro state Gov. Rosinha Matheus, who also is an Evangelical Christian.
<H> Rosinha Matheus is an Evangelical Christian.

Examples of other frequent alternations that need to be identified are: predicate nominal / premodifier (13), *of*-prepositional phrase / premodifier (14), and *have* / possessive (15).

13. (pair id="1088", value="TRUE", task="RC")
<T> Eight of the 51 Philippine humanitarian troops in Iraq have already left the country, Philippine Foreign Affairs Secretary Delia Albert said early Wednesday.
<H> Delia Albert is the Philippine Foreign Affairs Secretary.

14. (pair id="1096", value="TRUE", task="RC")
<T> A longtime associate of al Qaeda leader Osama bin Laden surrendered to Saudi Arabian officials.
<H> Osama bin Laden is the leader of al Qaeda.

15. (pair id="1010", value="TRUE", task="RC")
<T> (…) photographs of a hazy orange Titan -- the largest of Saturn's 31 moons, about the size of the planet Mercury.
<H> Saturn has 31 moons.

Lastly, there are additional alternations which largely derive from the Information Extraction subset of the test data, where the creators of the test set were requested to select a few targeted types of relations, such as “X was born in Y” and “X is located in Y”, and construct T-H pairs. Such alternations can be found in the appendix.

[1] This distribution is likely to be a result of the instructions given for the creation of the IE subtask in the PASCAL RTE data set, in particular, which focused on several well-known relationship types from IE, such as “born in” and “organizational role”.

4.3 Establishing False Entailment

We found two main categories of T-H pairs that we judged to be False: False, where there was a mismatch in the syntactic structure, and False, where there was no syntactic structure shared by the T-H pair. Although we can annotate this by hand, we are unsure whether it would be possible to create a system to automatically detect the absence of syntactic overlap. Though mismatched main verbs are the primary cue to the absence of overlap, all possible major and minor argument types were found unaligned, each potentially leading to a judgment of False entailment; see the Appendix for the details.

Examples of judging False by mismatch of syntactic structure are those in which the Subject and Verb align (with or without thesaurus), but the Object does not, as in (16):

16. (pair id="103", value="FALSE", task="IR")
<T> The White House ignores Zinni's opposition to the Iraq War.
<H> White House ignores the threat of attack.

The following examples illustrate the absence of shared syntactic structure in the major argument positions. In (17), the entailment is judged False since *baby girl* is not the subject of any verb of *buying*, nor is *ambulance* the object of any verb of *buying*; additionally, there is no mention of *buying* in T at all. In (18), the entailment is judged False because there is no mention of *Douglas Hacking* in the Text, nor any mention of *physician*. While a system using lexical matching might well rule the second example False, there are enough lexical matches in the former that a system using syntax is likely required.

17. (pair id="2179", value="FALSE", task="RC")
<T> An ambulance crew responding to an anonymous call found a 3-week-old baby girl in a rundown house Monday, two days after she was snatched from her mother at a Melbourne shopping mall.
<H> A baby girl bought an ambulance at a Melbourne shopping mall.

18. (pair id="2169", value="FALSE", task="CD")
<T> Scott and Lance Hacking talked with their younger brother at the hospital July 24.
<H>Douglas and Scott Hacking are physicians.

5 Interesting "Not Syntax" Examples

The number of examples that can be handled using syntax, broadly defined, is significant, but more than 50% were judged to be outside the realm of syntax, even allowing for the use of a thesaurus.

Some test items exhibited phrasal-level synonymy, which the annotators did not expect would be available in a general purpose thesaurus. Consider, *X bring together Y* and *Y participate in X* in (19):

19. (pair id="287", value="TRUE", task="IR")
<T> The G8 summit, held June 8-10, brought together leaders of the world's major industrial democracies, including Canada, France, Germany, Italy, Japan, Russia, United Kingdom, European Union and United States.
<H>Canada, France, Germany, Italy, Japan, Russia, United Kingdom and European Union participated in the G8 summit.

There are some examples with apparent alternation, but the alternation cannot easily be supported by syntax. Consider *three-day* and *last three days* in the following example:

20. (pair id="294", value="TRUE", task="IR")
<T> The three-day G8 summit will take place in Scotland.
<H> The G8 summit will last three days.

In other cases, the annotators considered that there were too many alternations and thesaurus replacements necessary to confidently say that syntax could be used. Consider the following example, where *more than half* has to align with *many*, *saying* aligns with *thinking*, and *not worth fighting* aligns with *necessary*.

21. (pair id="306", value="TRUE", task="IR")
<T> The poll, for the first time, has more than half of Americans, 52 percent, saying the war in Iraq was not worth fighting.
<H> Many Americans don't think the war in Iraq was necessary.

6 Discussion and Conclusion

Our goal was to contribute a baseline consisting of a system which uses an idealized parser, broadly defined, that can detect alternations, and optionally has access to a general purpose thesaurus. In order to explore what is possible in the limit, we used two human annotators and resolved their disagreements to produce a partial submission. It is interesting to note that the task is well-defined; of the 295 test items returned in our submission (Run1, without thesaurus), 284 matched the judgment provided as gold standard, so that our inter-annotator agreement on this subset is 96%.

An idealized syntax-based system can account for 34% of the test items, and, with the addition of information from a general purpose thesaurus, 48%. This finding is promising, though we expect the numbers to decrease subject to an implementation with a real-world parser and set of matching rules. An implemented system will also need to take the interaction of various alternations and syntactic phenomena as well. It may well be that there is a limit on the number of interactions an operational system can tolerate before its accuracy declines.

A syntax-based approach appears to be more powerful at deciding when T-H pairs exhibit False entailment: a syntax-only approach categorizes only 9% of the test items as True entailments successfully vs. 24% as False entailments. We have some concern that this imbalance is a consequence of the test creation process, as described in Dagan et al.. The test set authors were instructed for some subtasks to take a given sentence as the Text, and to produce a Hypothesis with either significant word overlap (but False) or no word overlap (but True); the first creation method would favor syntax-based methods. While the current RTE 1 test set is a rich collection of many of the possible types of mappings that can hold between Text and Hypothesis, a test collection that is a representative sampling of the mappings that occur for any subtask will be a valuable resource for system applications that include entailment.

Our baseline results need to be compared with those obtained by the systems using lexical matching and edit distance, as we expect that some of the items that can be handled by syntax alone could also be accounted for by these simpler methods.

We hope that the challenge workshop is well served by offering this study of the capabilities of an idealized parser. While approximately half of the RTE 1 test items are amenable to an approach using syntax augmented with a general purpose thesaurus, it is clear that the remainder of the test set represents an opportunity for work on fundamental entailment and paraphrase problems.

Acknowledgements

The authors wish to thank the organizers of the RTE Challenge and PASCAL for making creating this dataset and making it available. The authors also thank Deborah Coughlin for her linguistic judgments as one of the annotators, and for help with data analysis. We also thank the Butler Hill Group (www.butlerhill.com) for designing and maintaining the annotation environment.

References

1. Michael Collins and Nigel Duffy. 2002. New Ranking Algorithms for Parsing and Tagging: Kernels over Discrete Structures, and the Voted Perceptron. *Proceedings of ACL 2002*, Philadelphia, PA.
2. Ido Dagan, Oren Glickman and Bernardo Magnini. 2005. "The PASCAL Recognising Textual Entailment Challenge."*In the Proceedings of the PASCAL Recognising Textual Entailment Challenge*, April 2005.
3. Daniel Gildea and Daniel Jurafsky. 2002. Automatic Labeling of Semantic Roles. *Computational Linguistics*, 28(3):245-288.
4. Kadri Hacioglu, Sameer Pradhan, Wayne Ward, James H. Martin, and Daniel Jurafsky, 2004. Semantic Role Labeling by Tagging Syntactic Chunks. *Proceedings of the Eighth Conference on Natural Language Learning (CONLL-2004)*, Boston, MA, May 6-7.
5. James Henderson. 2004. Discriminative training of a neural network statistical parser. *Proceedings of ACL 2004*, Barcelona, Spain.

6. Eric Ringger, Robert C. Moore, Eugene Charniak, Lucy Vanderwende, and Hisami Suzuki. 2004. Using the Penn Treebank to Evaluate Non-Treebank Parsers. *Proceedings of the 2004 Language Resources and Evaluation Conference (LREC)*. Lisbon, Portugal.

Appendix

Files of the human judgments described in this paper are available on request from the authors.

- True by Syntax (with/without thesaurus)
- False by Syntax (with/without thesaurus)

Tables 2-5 contain categorization of human judgments regarding the syntactic phenomena and alternations required of a robust parser. Each category is followed by a list of the test items exhibiting the phenomena from the RTE 1 data set. For categories with ten or more examples, the distribution over the subtasks represented in RTE 1 is also given.

Table 2. Syntactic Phenomena Not Involving Alternation

(Counter-)Factive	174, 308, 472, 914, 962, 1076, 1121, 1284, 1300, 1358, 1359, 1421, 1634, 1876, 1907, 1981, 1982, 1983, 2024, 2025, 2041, 2055, 2062, 2121, 2122, 2135, 2167
	CD: 3.7% IE: 14.8% IR: 29.6% MT: 7.4% PP: 22.2% QA: 7.4% RC: 14.8%
Anaphora	298, 621, 936, 987, 1029, 1121, 1312, 1319, 1334, 1618, 1644, 1645, 1646, 1826, 1840, 1848, 1862, 1893, 1907, 2037, 2041
	CD: 4.8% IE: 28.6% IR: 14.3% MT: 14.3% QA: 14.3% RC: 23.8%
Apposition	741, 864, 901, 1144, 1203, 1263, 1387, 1556, 1584, 1616, 1617, 1618, 1648, 1822, 1907, 2036, 2090, 2190
	CD: 22.2% IE: 33.3% MT: 11.1% QA: 16.7% RC: 16.7%
Attachment	724, 893, 897, 1008, 1607, 1667
Coordination	300, 888, 898, 1004, 1042, 1082, 1342, 1370, 1536, 1634, 1640, 1831, 1861, 1895, 1896, 1964, 2014, 2020, 2113, 2129, 2169
	CD: 28.6% IE: 23.8% IR: 14.3% MT: 4.8% PP: 4.8% QA: 4.8% RC: 19.0%
Ellipsis	807, 2112
Existential	310, 840, 875, 1196, 2034
Extraposition	2041

Named Entity Recognition	39, 82, 103, 308, 315, 621, 692, 696, 821, 841, 847, 862, 864, 993, 1060, 1074, 1122, 1123, 1175, 1189, 1196, 1203, 1218, 1300, 1374, 1387, 1507, 1531, 1546, 1549, 1552, 1584, 1590, 1616, 2017, 2032, 2048, 2135, 2144, 2163
	CD: 27.5% IE: 2.5% IR: 15.0% MT: 15.0% PP: 2.5% QA: 25.0% RC: 12.5%
Negation	979, 1004, 1144, 1196, 1301, 1370, 1663, 1826, 1861, 1893, 1981, 1982, 1984
	CD: 15.4% IE: 15.4% MT: 15.4% PP: 23.1% RC: 30.8%
Nominalization	315, 962, 1021, 1092, 1358, 1422, 1625, 2073, 2135, 2176
	CD: 10.0% IE: 10.0% IR: 40.0% PP: 10.0% QA: 10.0% RC: 20.0%
Passive-Active	727, 1053, 1071, 1137, 1263, 1824, 1825, 1862, 1896, 1897, 1901, 1913, 1968, 1984, 1988, 2070
	CD: 12.5% IE: 25.0% MT: 6.3% PP: 18.8% QA: 6.3% RC: 31.3%
Predeterminer	1203, 1445, 2077, 2176
Predicate Complement	692, 938, 1092, 1325, 1468, 2006
Relative/Infinitive Clause	186, 472, 605, 727, 856, 887, 979, 1053, 1060, 1092, 1137, 1196, 1335, 1462, 1480, 1504, 1617, 1618, 1629, 1816, 1900, 1901, 1969
	CD: 13.0% IE: 21.7% IR: 4.3% MT: 8.7% PP: 8.7% QA: 17.4% RC: 26.1%
Spelling variation/error	743, 883, 888, 1556
VP-Cataphora	621

Table 3. Syntactic Phenomena Involving Alternation

“be born” – Appositive “from”	1584, 1621
“be from” – “be born in”	1672, 1687
“be located” – in-PP	1654
“be located” – Appositive	1074, 1123, 1549, 1655, 1944
Adjective/Noun morphology	878, 1203, 1203, 2176
Auxiliary – no-aux	1308
Be – Equi	48, 878, 1144
Be – Appositive	35, 139, 760, 843, 929, 996, 1009, 1011, 1073, 1091, 1093, 1142, 1555, 1613, 1836, 1903, 1905, 1998, 2039
	CD: 10.5% IE: 21.1% PP: 5.3% QA: 10.5% RC: 52.6%

Be – Appositive (flipped)	336, 901, 1032, 1062, 1065, 1134, 1826, 1940, 2037, 2092
	CD: 10.0% IE: 20.0% QA: 10.0% RC: 60.0%
Genitive – Location	35, 828, 1872
Have – Possessive	1010, 2077
Location Alternation	1361, 1820, 1871, 1872, 2082
Non finite – Finite verb construction	1263
Of Prepositional Phrase – Premodifier	864, 1096, 1203, 1325, 1358, 1451, 2037, 2096
Postmodifier – hyphenated	1122
Postmodifier – Premodifier	878, 1627, 2048
PreDeterminer – Restrictive Relative Clause	1662
Predicate Nominative – Postmodifier	1662
Predicate Nominative – Premodifier	166, 739, 739, 1031, 1088, 1096, 1451, 1620
Premodifier – Noun Appositive	39, 1028, 2039
Relative Clause – Main Clause	142, 825, 825, 962, 1007, 1041, 1122, 1609, 1687

Table 4. Single Word Replacement

Single Word	711, 834, 836, 846, 885, 1070, 1282, 1359, 1432, 1445, 1447, 1540, 1611, 1623, 1882, 1952, 1954, 1961, 1962, 1963, 1967, 1979, 1980, 1987, 1994, 1996, 2019, 2049, 2088
	CD: 17.2% IE: 13.8% IR: 3.4% MT: 3.4% PP: 51.7% QA: 6.9% RC: 3.4%

Table 5. Lack of Syntactic Parallelism

Factoid	828, 1820, 1824
Location	1334, 1530, 1531, 1547, 1620, 1624, 1628, 1667, 1672, 1821, 1874, 1999, 2027, 2134
	IE: 50.0% IR: 7.1% MT: 7.1% QA: 21.4% RC: 14.3%
Main verb	727, 887, 938, 965, 967, 993, 997, 1042, 1103, 1133, 1307, 1317, 1335, 1342, 1376, 1380, 1389, 1410, 1424, 1425, 1427, 1441, 1442, 1459, 1472, 1479, 1488, 1504, 1506, 1507, 1516, 1524, 1539, 1546, 1608, 1618, 1622, 1629, 1644, 1647, 1648,

	1680, 1682, 1683, 1686, 1688, 1693, 1840, 1846, 1848, 1849, 1853, 1854, 1860, 1865, 1866, 1869, 1913, 1954, 1979, 1990, 1992, 1998, 2006, 2008, 2014, 2018, 2020, 2021, 2022, 2026, 2029, 2035, 2036, 2038, 2047, 2048, 2053, 2063, 2070, 2090, 2113, 2117, 2125, 2136, 2144, 2152, 2157, 2162, 2163, 2167, 2168, 2176, 2179, 2187, 2190
	CD: 22.9% IE: 21.9% IR: 12.5% MT: 4.2% PP: 9.4% QA: 18.8% RC: 10.4%
Object	103, 1012, 1013, 1071, 1308, 1319, 1321, 1389, 1642, 1816, 1825, 1896, 1901, 2102, 2123, 2124, 2127, 2129, 2145, 2146
	CD: 15.0% IE: 15.0% IR: 25.0% MT: 20.0% RC: 25.0%
Of-Prepositional Phrase	2028, 2092, 2133
Postmodifier	1530
Predicate Adjective	48
Predicate Complement	308, 1174, 1203, 1462, 1552, 1822, 1845
Predicate Nominative	35, 840, 1011, 1032, 1123, 1487, 1498, 1549, 1551, 1553, 1589, 1666, 1675, 2033, 2037, 2039, 2045, 2054, 2056, 2169
	CD: 10.0% IE: 10.0% PP: 5.0% QA: 60.0% RC: 15.0%
Subject	909, 910, 936, 981, 1008, 1016, 1028, 1040, 1055, 1056, 1218, 1301, 1317, 1319, 1325, 1334, 1363, 1367, 1374, 1501, 1536, 1554, 1586, 1590, 1607, 1645, 1646, 1831, 1837, 1848, 1849, 1862, 1863, 1896, 2042, 2043, 2090, 2123, 2127, 2134, 2190
	CD: 12.2% IE: 12.2% IR: 12.2% MT: 14.6% QA: 17.1% RC: 31.7%
Subordinate Clause	1051
Time	1829, 1890