# Using Rhetorical Structure Theory and Entity Grids to Automatically Evaluate Local Coherence in Texts

Márcio de S. Dias[1,2], Valéria D. Feltrim[1,3], Thiago A. S. Pardo[1,2]

[1] Interinstitutional Center for Computational Linguistics (NILC)
[2] University of São Paulo, São Carlos/SP, Brazil
[3] State University of Maringá, Maringá/PR, Brazil
marciosd@icmc.usp.br, valeria.feltrim@din.uem.br, taspardo@icmc.usp.br

**Abstract.** This paper presents a joint model designed to measure local text coherence that uses Rhetorical Structure Theory (RST) and entity grids. The purpose is to learn patterns of entity distribution in texts by considering entity transition sequences and organizational/discourse information using RST relations in order to create a predictive model that is able to distinguish coherent from incoherent texts. In an evaluation with newspaper texts, the proposed model outperformed other methods in the area.

**Keywords:** Local Coherence, Rhetorical Structure Theory, Entity Grids

## 1 Introduction

In text generation systems (as summarizers, question/answering systems, etc.), coherence is an essential characteristic in order to produce comprehensible texts. As such, studies and theories on coherence ([21], [12]) have supported applications that involve text generation ([29], [4], [16]).

By coherence, we mean the possibility of establishing a meaning for the text [17]. Coherence supposes that there are relationships among the elements of the text for it to make sense. It also involves aspects that are out of the text, for example, the shared knowledge between the producer (writer) and the receiver (reader/listener) of the text, inferences, intertextuality, intentionality and acceptability, among others [17].

According to Dijk and Kintsch [8], textual coherence occurs in local and global levels. Local level coherence is present by the local relationship among the parts of a text, for instance, sentences and shorter segments. On the other hand, a text presents global coherence when this text links all its elements as a whole. Psycholinguistics considers that local coherence is essential in order to achieve global coherence [25]. Thus, many researches in computational linguistics have been developed for dealing with local coherence ([1], [2], [5], [9], [10], [11], [13], [15], [18], [20]).

Examples of coherent and incoherent texts are given in Figure 1. Text A is an original text and it is considered coherent. Text B is formed by randomly permuted sentences (a change in the order of sentences) of the original text (Text A). One may see that its coherence is seriously harmed. In this text, not only the reference chain is broken (e.g., in the first sentence, "event" has no previous antecedent), but the discourse organization is also awkward, making it difficult to grasp the main idea.

| **Text A (original, coherent text)** | **Text B (permuted sentences, incoherent text)** |
|---|---|
| The gymnast Jade Barbosa, who got three medals at the Pan American Games in Rio in July, won the election on the Internet and will be the Brazilian representative in the Olympic Torch Relay for Beijing 2008. The torch will travel to twenty different countries, but Brazil will not be at the Olympic route. Therefore, Jade will participate in the event in Buenos Aires, Argentina, the only city in South America to receive the symbol of the Games. The relay will be over on August 8, the first day of the Beijing Olympics. | Therefore, Jade will participate in the event in Buenos Aires, Argentina, the only city in South America to receive the symbol of the Games. The relay will be over on August 8, the first day of the Beijing Olympics. The gymnast Jade Barbosa, who got three medals at the Pan American Games in Rio in July, won the election on the Internet and will be the Brazilian representative in the Olympic Torch Relay for Beijing 2008. The torch will travel to twenty different countries, but Brazil will not be at the Olympic route. |

**Fig. 1.** Example of coherent (Text A) and incoherent (Text B) text

In this paper, we propose a joint model for tackling local coherence in order to be able to automatically differentiate coherent from incoherent (less coherent) texts. In particular, this work is based on principles from other researches, such as that the distribution of entities in locally coherent texts presents certain regularities, which may be evidenced on the entity grids proposed by Barzilay and Lapata [2]. Another assumption is that coherent texts show certain distinct intra- and inter-discourse relation organization [20]. Combining such information in a joint model allows better dealing with the local coherence phenomenon.

For dealing with discourse, we make use of Rhetorical Structure Theory (RST) [21], whose relations are incorporated in the entity grids. A previous work [20] has already considered using discourse in such cases, but in a different way (as we present in the next section). We evaluate our proposal with newspaper texts and show that it outperforms other works in the area, showing the potential of our approach. For the evaluation, we follow the text ordering task proposed by Barzilay and Lapata [2], in which the methods must rank texts (original and sentence-permuted texts) according to their coherence.

Section 2 presents an overview of the most relevant researches related to local coherence. In Section 3, the Rhetorical Structure Theory is briefly introduced. In Section 4, the corpus used in this work is described. Section 5 details the proposed approach in this paper and other methods that were tested. Section 6 shows the experimental setup and the obtained results. Finally, Section 7 concludes this paper.

## 2 Related Work

In a statistical approach, Foltz et al. [10] used Latent Semantic Analysis (LSA) [19] to compute a coherence value for texts. LSA is used to produce a vector for each word or sentence so that the similarity between two words or two sentences may be measured by the cosine measure [28]. The coherence value of a text may be obtained by measuring the cosines for all pairs of adjacent sentences. Foltz et al. obtained 81%

and 87.3% of accuracy applied respectively to the set of texts related to earthquakes and accidents, in English.

Based on Centering Theory [12], Barzilay and Lapata's [2] assumption is that locally coherent texts present some regularities in entity distribution. These regularities are computed by means of an entity grid, i.e., a matrix in which the rows represent the sentences of the text and the columns the entities. For example, Figure 3 shows part of a entity grid for the text passage in Figure 2, both reproduced from [2].

1 **[**The Justice Department**]**$_S$ is conducting an **[**anti-trust trial**]**$_O$ against **[**Microsoft Corp.**]**$_X$ with [evidence]$_X$ that [the company]$_S$ is increasingly attempting to crush [competitors]$_O$.
2 **[**Microsoft**]**$_O$ is accused of trying to forcefully buy into **[**markets**]**$_X$ where **[**its own products**]**$_S$ are not competitive enough to unseat **[**established brands**]**$_O$.
…
6 **[**Microsoft**]**$_S$ continues to show **[**increased earnings**]**$_O$ despite **[**the trial**]**$_X$.

**Fig. 2.** Text with syntactic tags [2]

| | Department | Trial | Microsoft | Evidence | Competitors | Markets | Products | Brands | Case | Netscape | Software | Tactics | Government | Suit | Earnings | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | S | O | S | X | O | – | – | – | – | – | – | – | – | – | – | 1 |
| 2 | – | – | O | – | – | X | S | O | – | – | – | – | – | – | – | 2 |
| 3 | – | – | S | O | – | – | – | – | S | O | O | – | – | – | – | 3 |
| 4 | – | – | S | – | – | – | – | – | – | – | – | S | – | – | – | 4 |
| 5 | – | – | – | – | – | – | – | – | – | – | – | – | S | O | – | 5 |
| 6 | – | X | S | – | – | – | – | – | – | – | – | – | – | – | O | 6 |

**Fig. 3.** Entity Grid [2]

For instance, the "Department" column in the matrix shows that the "Department" entity happens only in the first sentence in the subject (S) position. In turn, in the "Trial" column, it is shown that the "Trial" entity happens in the first sentence in the object (O) position and in the sixth sentence in some other syntactical function that is nor subject or object (indicated by X). The hyphen ('-') indicates that the entity did not happen in the corresponding sentence.

With such a matrix, it is possible to obtain the probabilities of entity transitions in texts. For example, the probability of transition [O -] (i.e., the entity happened in the object position in one sentence and did not happen in the following sentence) in the grid in Figure 3 is 0.09, computed as the ratio between its frequency of occurrence in the grid (7 occurrences) and the total number of transitions of length 2 (75 transitions). From this, a feature/characteristic vector is formed by the probabilities of all the transition types. Such vectors are used to learn the properties of coherent texts in a corpus. Figure 4 shows the feature vector representation of the grid in Figure 3.

| S S | S O | S X | S – | O S | O O | O X | O – | X S | X O | X X | X – | – S | – O | – X | – – |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| .01 | .01 | 0 | .08 | .01 | 0 | 0 | .09 | 0 | 0 | 0 | .03 | .05 | .07 | .03 | .59 |

**Fig. 4.** Feature-vector text representation using [2]

Barzilay and Lapata defined 3 applications to test the model prediction power: text-ordering, automatic evaluation of coherence in summaries and readability assessment. The first two are ranking problems and, according to the authors, present an efficiently learnable model that ranks the texts based on their level of local coherence.

Barzilay and Lapata's approach produced 8 models according to the use (+) or not (-) of syntactical, coreference and salience information. The syntactical information was given by a parser [23] that recognizes the grammatical function of the entities.

Coreference occurs when words refer to the same entity and, therefore, these words may be represented by a single column in the grid. For example, when the text mentions "Microsoft Corp.", "Microsoft", and "the company", such references are mapped to a single column ("Microsoft") in the grid. Salience is related to the frequency of entities in texts, allowing to build grids with the least and/or the most frequent entities in the text.

In the text-ordering task (which is the one that interests to us in this paper), for each original text considered "coherent", a set of randomly permutated versions were produced and considered "incoherent" texts. Ranking values for coherent and incoherent texts were produced by means of the predictive model trained in the SVM$^{light}$ [14] package, using a set of pairs of texts (coherent text, incoherent text). It is supposed that the ranking values of coherent texts are higher than the ones for incoherent texts. Barzilay and Lapata obtained 87.2% and 90.4% of accuracy (fraction of correct pairwise rankings in the test set) applied respectively to the set of texts related to earthquakes and accidents, in English. Such results were achieved by the model considering the three types of information (Coreference+Syntax+Salience+).

Lin et al. [20] assumed that local coherence implicitly favors certain types of discursive relation transitions. The authors used four discursive relations, based on Discourse Lexicalized Tree Adjoining Grammar (D-LTAG) [30], to develop the Discourse Role Matrix, which is composed of sentences (rows) and terms (columns), with discursive relations used over their signaling arguments. Terms were the stemmed forms of the open class words: nouns, verbs, adjectives and adverbs. For example, see the discursive grid (b) for the text (a) in Figure 5.

(S1) Japan normally depends heavily on the Highland Valley and Cananea mines as well as the Bougainville mine in Papua New Guinea. (S2) Recently, Japan has been bying copper elsewhere.

(a)

| S# | Terms | | | |
|---|---|---|---|---|
| | *copper* | *cananea* | *depend* | … |
| S1 | nil | Comp.Arg1 | Comp.Arg1 | |
| S2 | Comp.Arg2<br>Comp.Arg1 | nil | nil | |

(b)

**Fig. 5.** Part of a text and its discursive grid [20]

Figure 5 shows a fragment of the matrix representation (b) of the text (a). Columns correspond to the extracted terms; rows, the contiguous sentences. A cell $C_{Ti,Sj}$ contains the set of the discourse roles of the term $T_i$ that appears in sentence $S_j$. For example, the term "depend" from S1 takes part of the Comparison (Comp) relation as argument 1 (Arg1), so the cell $C_{depend,S1}$ contains the Comp.Arg1 role. A cell may be empty (nil, as in $C_{depend,S2}$) or contain multiple discursive roles (as in $C_{copper,S2}$).

Lin et al. applied their model to the same text-ordering task proposed by Barzilay and Lapata, but now the sentence-to-sentence transitions are D-LTAG relations. They obtained 89.25% and 91.64% of accuracy applied to the set of English texts related to earthquakes and accidents, respectively, improving the previous results.

Another model that used Barzilay and Lapata's approach is the one of Filippova and Strube [9], which implemented the entity model for German and conducted an entity grouping by the use of semantic relations. The Coreference+Syntax-Salience+ model developed by the authors obtained 75% of accuracy as the best result. Iida and Tokunaga [13] used the concepts of entity and coreference to evaluate the coherence of texts written in Japanese and this research obtained 76.1% of accuracy. Freitas and Feltrim [11] applied Barzilay and Lapata's entity model to evaluate coherence in newspaper texts written in Brazilian Portuguese, obtaining 74.4% of accuracy by means of the use of syntactic and salience information applied to the CSTNews corpus [6]. Besides, the authors considered the lemmas of noun phrases (NP) to minimize the lack of a coreference resolution system and used additional Type/Token information [5] to measure the lexical variety of entities in each syntactical function.

## 3 Rhetorical Structure Theory

The Rhetorical Structure Theory (RST) proposed by Mann and Thompson [22] considers that each text presents an underlying rhetorical structure that allows the recovery of the communicative intention of the writer.

In this model, Elementary Discourse Units (EDUs) are connected by rhetorical relations, aiming at coherently organizing discourses. The role of the nucleus (N) or satellite (S) is assigned to each EDU. The nuclei or nuclear EDUs contain the most important pieces of information in the relations and are considered more relevant than the satellites. The satellites, on the other hand, present additional information that helps the reader in the interpretation of the nuclei.

The RST relations are divided into two classes: mononuclear and multinuclear relations. The mononuclear relations are composed of pairs of EDUs that present different levels of importance: one nuclear and one satellite. On the other hand, multinuclear relations link equally important EDUs, which are classified as nuclei.

Figure 6 presents part of a text segmented in EDUs, reproduced from [27]. It is used to exemplify an RST analysis, shown in Figure 7.

| |
|---|
| (1) Many of Almir's – the Pernambuquinho – "courageous" attitudes were dictated by fear. (2) Few people know this, (3) but it is true. |

**Fig. 6.** Part of a text segmented in EDUs [27]

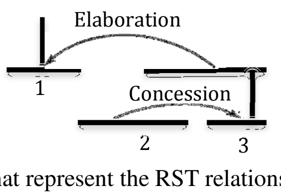


**Fig. 7.** Diagram that represent the RST relations for the text in Figure 6

EDU (1) illustrates the central idea of the discourse, which is the fear that influenced the way Almir (the character) acted. However, EDUs (2) and (3) indicate that few people know about the character feature and the fact that this feature is real. The relations among EDUs (1), (2) and (3) occur as they are recognized in the discourse.

EDUs (2) and (3) are identified as constituents of an ELABORATION relation of the assertion in (1). In RST, this is expressed by the rhetorical relation ELABORATION. EDU (1) corresponds to the nucleus of the ELABORATION relation (the nucleus is indicated by a vertical line and is pointed by the arrow), while EDUs (2) and (3), constitute the satellite of the relation. EDU (3) is the nucleus and EDU (2) is the satellite of the CONCESSION relation.

# 4 The Corpus

The CSTNews corpus [6] has been created for multi-document summarization. It is composed of 140 texts distributed in 50 sets of news texts written in Brazilian Portuguese from various domains. Each set has 2 or 3 texts from different sources that address the same topic. Besides this, the corpus has, in average, 14.9 sentences per text and a total of 5,216 RST relations. Besides the original texts and their RST analyses, the corpus counts with several annotation layers. One may also find single and multi-document manually produced summaries, the identification of temporal expressions, Cross-document Structure Theory (CST) annotation [26], automatic syntactical analyses, nouns and verb senses, text-summary alignments, and, more recently, the semantic annotation of informative aspects in summaries, among other annotations. For this work, we are especially interested in the RST annotation.

For the RST annotation, Table 1 shows the obtained agreement (computed by RSTeval [24]) for the simple textual segments (for the segmentation process, therefore), complex textual segments (given by the internal nodes in the RST structure), nuclearity and relations among segments. Results are shown in F-measure values.

**Table 1.** Agreement results for the RST annotation in the CSTNews corpus

| Criteria | F-Measure (%) |
|---|---|
| Simple textual segments | 0,91 |
| Complex textual segments | 0,78 |
| Nuclearity of each segment | 0,78 |
| RST relations among segments | 0,66 |

According to the results, the agreement among annotators was satisfactory, and is similar to the obtained for other works ([23], [7]) for other languages. Therefore, the annotation is considered reliable and may be used for the purposes of this work.

# 5 Our Approach

Our approach follows Barzilay and Lapata's [2] work, but excludes the use of coreference information (as there is no widely available system for Portuguese, we used the nuclei of noun phrases as entities) and includes the use of RST relation distribution. It captures a pattern of RST discursive relations in coherent texts by using a machine learning technique, creating a predictive model that enables the evaluation of local coherence.

Our grid is formed by sentences (rows) and entities (columns), in which each cell is filled with the RST relations that the entity under focus is part of, also specifying the corresponding nuclearity. Figure 8 shows part of the grid for the text in Figure 6 along with the RST information given by the diagram in Figure 7. Relations are shown in abbreviated forms.

| | **attitudes** | **Almir** | **Pernambuquinho** | **fear** | **people** | **true** |
|---|---|---|---|---|---|---|
| **S1** | elab.Nuc | elab.Nuc | elab.Nuc | elab.Nuc | - | - |
| **S2** | - | - | - | - | conces.Sat<br>elab.Sat | conces.Nuc<br>elab.Sat |

**Fig. 8.** RST relation grid for the text in Figure 6

The Palavras parser [3] has been used to identify the text entities, which are all nouns and proper nouns. Therefore, our approach is prone to parsing errors. However, if such errors are systematic, useful patterns may still be learnt.

As the entity grid was created, the entity distribution was computed sentence by sentence and not by EDUs. The EDUs were not used due to the sparsity of the entity grid, making it difficult the creation of an efficient prediction model.

The sentence by sentence entity distribution was performed for all the possible RST relations in the text. RST relations transitions had length two. For example, the entity "fear" in Figure 8 is present in one relation in sentence 1 (S1) and the possible transitions are computed for all relations. In the case of the transition [**elab.Nuc, -],** there are 4 occurrences (all transitions occur from sentence 1 to 2) out of 6 length-two transitions, considering the grid in Figure 8. Thus, the transition probability [**elab.Nuc, -]** is 4/6 = 0.6666. This way, each text may be seen as a distribution defined over transition types. Now, each text may be represented as a set of transition sequences by using a standard feature vector notation, in other words, each grid *j* of a document $d_i$ corresponds to a feature vector $\Phi(x_{ij}) = (P_1(x_{ij}), P_2(x_{ij}), \ldots, P_n(x_{ij}))$ [2], where *n* is the total number of possible transitions and $P_r(x_{ij})$ is the probability of transition *r* in grid $x_{ij}$. Table 2 shows part of the feature vector for the grid in Figure 8.

**Table 2.** The feature vector representation that uses possible length-two transitions given RST relations and nuclearity information from the grid in Figure 8

| **[elab.Nuc, -]** | **[-, conces.Nuc]** | **[-, elab.Sat]** | **[-, -]** |
|---|---|---|---|
| 0.6666 | 0.1666 | 0.1666 | 0.0 |

It is important to notice that our proposal is different from the one of Lin et al. [20]. We do not use only 4 D-LTAG relations, but the full relation set of RST. We also include nuclearity in the grid, because we believe that coherent texts may follow patterns of nuclei and satellites distributions, which are not considered in Lin et al. approach. Besides, Lin. et al. [20] used the stemmed open-class words and we used the nuclei of noun phrases as entities.

The feature vectors we built were used to create the coherence prediction model with the use of a machine-learning algorithm. In the next section, the evaluation of this proposed model is reported.

# 6 Evaluation

## 6.1 Experimental Setup

The text-ordering task proposed by Barzilay and Lapata [2] has been used to check and to evaluate the performance of our approach compared to other methods. In the text-ordering task, a document is given as a set of sentences and the algorithm investigates the ordering that maximizes coherence. Because of this, random permutations of the original text were generated in order to measure how often a permuted version is ranked higher than the original document. A good model should prefer an original text more often than its possible permutations.

We used 137 out of the 140 texts of the CSTNews corpus and 20 random permutations for each original text of the corpus. Three texts were not used because they did not reach the 20 different permutations defined. We decided to use 20 permutations for each text because this is also the number used by Barzilay and Lapata in their experiment. Thus, the database of this experiment is composed of 2,740 pairs of texts. The SVM$^{light}$ [14] package has been used in the experiment with the ranking option for training and testing our coherence model. Apart from that, the 10-fold cross-validation method was used for achieving a more confident result.

The evaluation metric used was the accuracy measure, which, for a given set of pairwise rankings (an original document and one of its permutations), the accuracy is the ratio of the correct predictions made by the method over the size of the testing set for each fold. The final accuracy for each experiment is the average of the accuracies for each one of the 10 folds.

## 6.2 Other Methods

Besides our proposal, we implemented some other methods from literature, in order to compare our results to the current state of the art. The following methods were chosen based on their importance and used techniques to evaluate local coherence: the LSA method by Foltz et al. [10], the traditional entity grid method by Barzilay and Lapata [2] and the discourse-based method by Lin et al. [20]. All of them were adapted to Brazilian Portuguese, using the appropriate available tools and resources for this language.

The implementation of these methods carefully followed each step of the original ones. The resources used to develop the baselines were: Python 2.7[1] for all the methods, the NLTK package[2] for the stemming required by Lin et al., the Scikit-

---

[1] http://www.python.org/
[2] http://www.nltk.org/

Learn package[3] as in Foltz et al., and Palavras parser [3], mainly used in the implementation of Barzilay and Lapata and Lin et al approaches. We used the RST relations as the necessary discourse information in Lin et al.'s approach.

Barzilay and Lapata's method has been implemented without referential information, since, to the best of our knowledge, there is no available robust coreference resolution system for Brazilian Portuguese and the CSTNews corpus still does not have referential information in its annotation layers.

### 6.3 Results

The LSA method generates a coherence value for each original text and for its permutations. Therefore, the accuracy measure for this method was calculated by the number of times that the coherence value of the original text was greater than its permutation over the total number of text pairs (an original document and one of its permutations). Therefore, for this method, it is not necessary to perform 10-fold cross-validation.

The other methods and our approach were submitted to the text-ordering task using the CSTNews corpus. Furthermore, the implementation of Barzilay and Lapata's approach produced 4 models: (Syntactic+Salience+), (Syntactic+Salience-), (Syntactic-Salience+) and (Syntactic-Salience-), in which Salience is related to the frequency of entities, considering entities with frequency higher or equal to 2.

In the approach proposed here, two variations were created in order to check if the accuracy would improve: the RST relations were grouped according to the relation groups of Mann and Thompson [22] (Variation 1), ignoring nuclearity; the RST relations were not grouped, but the nuclearity information has been removed from the RST relations (Variation 2). Table 3 shows the accuracy of our approaches compared to the other methods, ordered by accuracy.

**Table 2.** Evaluation results, where diacritics ** ($p < .01$) and * ($p < .05$) indicate whether there is a significant statistical difference in accuracy compared to our approach (using t-test)

| **Methods** | **Accuracy (%)** |
|---|---|
| Our Approach | **79.45** |
| Syntactic-Salience- from Barzilay and Lapata | 78.97 |
| Syntactic+Salience- from Barzilay and Lapata | 74.10** |
| Discourse grids from Lin et al. | 70.80* |
| Syntactic+Salience+ from Barzilay and Lapata | 70.73** |
| Syntactic-Salience+ from Barzilay and Lapata | 67.87** |
| Variation 1 of our approach | 66.18** |
| Variation 2 of our approach | 63.99** |
| LSA from Foltz et al. | 58.40** |

The t-test has been used for pointing out whether differences in accuracy are statistically significant. Comparing our approach with the other methods, one may observe that the use of all the RST relations with nuclearity information obtained better results for evaluating local coherence.

[3] http://scikit-learn.org/stable/

In particular, the results showed that the use of nuclearity information of RST relations significantly increased the accuracy (comparing our approach with Variation 2). For Variation 1, the grouped RST relations improved the accuracy in comparison with Variation 2. We believe that this happened because a less sparse grid (which comes from grouping the relations) results in a more efficient learning.

We believe that the use of open-class words in Lin et al.'s model may have been the cause of its significant lower accuracy compared to our approach. Regarding the open-class words, since there are more of them than the number of entities in the corpus, the generated grid was very sparse. This makes it difficult to generate a good predictive model.

The LSA model by Foltz et al. was the approach that obtained the lowest accuracy in comparison with the other analyzed models. This result may be explained by the lack of linguistic information that might have improved its accuracy.

An interesting result is given by the Syntactic-Salience- model, i.e., this approach used neither syntactic nor salience information, but it presented greater accuracy than other more complete models. This was due to few transition types obtained in the grid that only contained information of presence or not of entities in the sentences. As consequence, the grid was less sparse and helped in the generation of a good predictive model. In fact, although it produces worse results compared to our approach, the differences among these models are not statistically significant. Therefore, it is interesting to see that models of different nature may behave quite similarly.

At this point, it is important to notice that we could not directly compare our results to the ones obtained by Freitas and Feltrim [11]. Although they have used the same corpus that us, they used different text permutations (incoherent texts) and some other additional information, which makes direct comparisons unfair.

## 7 Final Remarks

According to the results obtained from the text-ordering task, the use of RST relations and nuclearity was the best among its variation and the other methods from literature. This approach showed to be promising and it may be used for other languages, such as English, as long as there is a corpus annotated with RST relations, a syntactic parser and specific resources that focus on the target language.

As future work, the same methodology employed in this work will be used to develop methods to evaluate local coherence for multi-document summaries with other discourse theories. For this, the focus will be on structuring the discourse to connect sentences from different documents and to establish one or more types of relations among sentences, as Cross-Document Structure Theory (CST) [26] does.

## Acknowledgements

The authors are grateful to FAPESP and University of Goiás for supporting this work.

## References


1. Althaus, E., Karamanis, N., Koller, A.: Computing locally coherent discourse. In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, article 399, Stroudsburg, PA, USA, (2004)
2. Barzilay, R., Lapata, M.: Modeling local coherence: An entity-based approach. Computational Linguistics, vol. 34. pp. 1-34 (2008)
3. Bick, E.: The Parsing System Palavras, Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework, Aarhus University Press. (2000)
4. Bosma, W.: Query-Based Summarization using Rhetorical StructureTheory. In Proceedings of the15th Meetings of CLIN, LOT, Utrecht, pp. 29-44 (2004)
5. Burstein, J., Tetreault, J., Andreyev, S.: Using entity-based features to model coherence in student essays. In Human Language Technologies: In Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 681-684 (2010)
6. Cardoso, P., Maziero, E., Jorge, M., Seno, E., di Felippo, A., Rino, L., Nunes, M., Pardo, T.: Cstnews - a discourse-annotated corpus for single and multi-document summarization of news texts in brazilian portuguese. In Proceedings of the 3rd RST Brazilian Meeting. pp. 88-105 (2011)
7. Cunha, I.; Torres-Moreno, J-M.; Sierra, G.: On the Development of the RST Spanish Treebank. In Proceedings of the 5th Linguistic Annotation Workshop, pp. 1-10. Portland-Oregon (2011)
8. Dijk, T.V., Kintsch, W.: Strategics in discourse comprehension. Academic Press. New York (1983)
9. Filippova, K., Strube, M.: Extending the entity-grid coherence model to semantically related entities. In Proceedings of the Eleventh European Workshop on Natural Language Generations, 139-142 (2007)
10. Foltz, P. W., Kintsch, W., Landauer, T. K.: The Measurement of textual coherence using latent semantic analysis. Discourse Processes, vol. 25, n. 2-3, pp. 285-307 (1998)
11. Freitas, A. P., Feltrim, V. D.: Análise Automática de Coerência Usando o Modelo Grade de Entidades para o Português. In Proceedings of the IX Brazilian Symposium in Information and Human Language Technology, pp. 69-78. Fortaleza, CE, Brazil (2013)
12. Grosz, B., Aravind, K. J., Scott, W.: Centering: A framework for modeling the local coherence of discourse. Computational Linguistics, vol. 21, pp. 203-225 (1995)
13. Iida, R., Tokunaga, T.: A metric for evaluating discourse coherence based on coreference resolution. In Proceedings of the COLING 2012: Posters, pp. 483-494, Mumbai, India (2012)
14. Joachims, T.: Optimizing search engines using clickthrough data. In Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 133–142. New York, NY, USA (2002)
15. Karamanis, N., Poesio, M., Mellish, C., Oberlander, J.: Evaluating centering-based metrics of coherence for text structuring using a reliably annotated corpus. In Proceedings of the 42nd Annual Meetings of the Association for Computational Linguistics, article 391 (2004)
16. Kibble, R., Power, R,: Optimising referential coherence in text generation. Computational Linguistic, vol. 30 n. 4, pp. 401-416 (2004)
17. Koch, I.V., Travaglia, L.C.: A Coerência Textual. 14rd edn. Contexto, São Paulo (2002)
18. Lapata, M.: Probabilistic texts structuring: Experiments with sentence ordering. In Proceeding of the 2nd Human Language Technology Conference and Annual Meeting of

the North American Chapter of the Association for Computational Linguistics, pp. 545-552 (2003)

19. Landauer, T. K., Dumais, S. T.: A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation to coreference resolution. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pp. 104–111, Philadelphia, PA (1997)
20. Lin, Z., Ng, H. T., Kan, M. Y.: Automatically evaluating text coherence using discourse relations. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. vol 1. pp. 997-1006. Stroudsburg, PA, USA (2011)
21. Mann, W. C., Thompson, S. A.: Rhetorical Structure Theory: Toward a functional theory of text organization. Text, 8 (3), pp. 243-281 (1988)
22. Mann, W. C., Thompson, S. A.: Rhetorical Structure Theory: A Theory of Text Organization. Technical Report from Information Sciences Institute (ISI), ISI/RS-87-190, pp. 1-91. University of Southern California, USA (1987)
23. Marcu, D.: The Rhetorical Parsing of Unrestricted Texts: A Surface-based Approach. Computational Linguistics, vol. 26, pp. 396-448 (2000)
24. Maziero, E., Pardo, T. A. S.: Automatização de um método de avaliação de estruturas retóricas. In Proceedings of the RST Brazilian Meeting (2009)
25. Mckoon, G., Ratcliff, R.: Inference during reading. Psychological Review, pp. 440-446 (1992)
26. Radev, D.: A common theory of information fusion from multiple text sources, step one: Cross-document structure. In Proceedings of the 1st ACL SIGDIAL Workshop on Discourse and Dialogue, pp. 74-83. Hong Kong (2000)
27. Ribeiro, G.F., Rino, L.H.M.: A Sumarização Automática com Base em Estruturas RST. Technical Reports from Interinstitutional Center for Computational Linguistics, University of São Paulo, NILC-TR-02-05. São Carlos, Brazil (2002)
28. Salton, G.: Term-Weighting Approaches in Automatic Text Retrieval. Information Processing and Management, pp. 513-23 (1988)
29. Seno, E. R. M.: Rhesumarst: Um sumarizador automático de estruturas rst. Master Thesis. University of São Carlos. São Carlos/SP (2005)
30. Webber, B.: D-ltag: extending lexicalized tag to discourse. Cognitive Science, vol. 28, n. 5, pp. 751-779 (2004)