
| RESEARCH ARTICLE

From Description to Assessment: An RST-Based Framework for Measuring Discourse Coherence

Rajaa Rhandy

CLAS Laboratory, Faculty of Languages, Arts and Humanities, Ibn Zohr University, Agadir

Corresponding Author: Rajaa Rhandy, **E-mail:** r.rhandy@uiz.ac.ma

| ABSTRACT

Since Thirty-nine years of its publication in 1987, Rhetorical Structure Theory (RST) by Mann and Thompson has established itself as one of the most influential frameworks for describing the rhetorical organization of discourse. However, the practical application of this theory has not gone beyond the qualitative analysis of discourse structure without providing a standardized framework for the evaluation of discourse organization quality. To address this gap and advance the theory, the present study proposes a practical RST- based framework for measuring and assessing discourse organization in written texts. This framework draws upon both the four principles of rhetorical structure theory, namely completedness, connectedness, uniqueness and adjacency along with the discourse coherence principles of Elementary Discourse Unit segmentation. Accordingly, the present article aims at bridging the gap between discourse description and discourse assessment, showing how RST cannot serve only as a descriptive theory of discourse organization but also as a foundation for evaluating its quality.

| KEYWORDS

Rhetorical Structure Theory, discourse coherence, writing assessment, discourse organization

| ARTICLE INFORMATION

ACCEPTED: 01 July 2026

PUBLISHED: 26 July 2026

DOI: 10.32996/jpda.2026.5.3.4

1. Introduction

coherence is the distinctive feature of discourse structure which plays a key role in discourse connectedness and comprehensibility. Different theoretical approaches have been concerned with the description of discourse coherence and how it is established and organized. These theories are subsumed in literature under two basic lines of research, namely the cognitive approach which includes exclusively Sanders et al's psychological approach (1992) and the computational approach which includes Hobbs' theory (1985), Mann & Thompson's Rhetorical Structure Theory (1987), Asher & Lascarides' Segmented Discourse Representation Theory (2003) and Polyani et al's Unified Linguistic Discourse Model (2004). These theories converge and diverge in many respects; while they all share the underlying assumption of the hierarchy of discourse segments into one unified segment which is the elementary unit of analysis; they disagree on the nature of discourse segments. To illustrate, they are conceived either as sentences (Hobbs, 1985); intentional units (Grosz & Sidner, 1986), clauses (Mann & Thompson, 1987; Polyani et al., 2004), or propositions (Asher & Lascarides, 2003; Sanders et al., 1992). Despite the variety of all these discourse theories, they remain all descriptive and fail to provide a practical framework for assessing discourse coherence. In this context, the ultimate objective of this paper is to fill this gap by suggesting a framework that draws upon the basic principles of RST, which is the most influential and widely used descriptive theories of discourse coherence, along with the discourse coherence principles of Elementary Discourse Unit segmentation (Carlson & Marcu, 2001). Accordingly, the present study aims at answering the following questions:

- A) How can Rhetorical Structure Theory be applied to assess coherence?

- B) What assessment framework can be developed from the principles of Rhetorical Structure Theory to assess discourse coherence in writing?

2. Rhetorical Structure Theory as a descriptive theory of discourse coherence

Rhetorical Structure Theory is a descriptive theory which studies the organization of a text in terms of the hierarchically rhetorical relations that functionally exist between its parts and hence provides, a treelike structural analysis of texts where each leaf node represents an Elementary Discourse Unit (EDU) that is incrementally connected to other EDUs. According to RST, each text includes four types of components referred to as “defined objects”, namely relations, schemas, schema applications and structures. discourse relations bind text segments together. Based on this, the schemas establish patterns that help a specific segment of text to be analyzed in terms of other segments. The schema application determines how a schema can be implemented and based on which the structure of a whole text is formed. Thus, coherence is achieved through the hierarchically structured relations which exist between the four components of the RST theory. The smallest parts, portions or spans in a text are referred to as “the nucleus” and “the satellite.” While the nucleus is the most important information in a relation, the satellite is the complementary information.

Mann & Thompson distinguished between two macro types of rhetorical relations, namely mononuclear (hypotactic) relations which relate a nucleus and a satellite and multinuclear (paratactic) relations which hold between two nuclei discourse segments. Another classification, based on the intended effect criterion of rhetorical relations, is between presentational and subject matter relations. Subject matter relations are relations whose intended effect on the reader is to identify a rhetorical relation, while the effect of presentational relations is to increase the reader’s desire to take action. This can take various forms such as belief in, acceptance of, or a favorable appreciation of the action present in the nucleus. Accordingly, subject matter relations include elaboration, circumstance, solutionhood, volitional cause, volitional result, non-volitional cause, non-volitional result, purpose, condition, otherwise, interpretation, evaluation, restatement, summary, sequence and contrast. Conversely, presentational relations include motivation (increases desire), antithesis (increases positive regard), background (increases ability), enablement (increases ability), evidence (increases belief), justify (increases acceptance) and concession (increases positive regard).

3. An RST-Based Framework for Coherence Assessment

3.1 The Measures of Coherence According to RST analysis

Measures of coherence in the present study were developed based on the RST-diagram formation principles namely completeness, connectedness, uniqueness and adjacency. Mann & Thompson (1987) defined these four key principles as follows,

- A. **Completedness:** refers to one schema application that contains a group of text spans that make up the complete text.
- B. **Connectedness:** each text span except the one which includes the entire text is either a minimal unit or a component in another schema application of the analysis.
- C. **Uniqueness:** each schema application is a multi-relational schema where each relation applies to a different set of text spans.
- D. **Adjacency:** the text spans of each schema application form one text span.

Thus, the coherence of each composition was both quantitatively and qualitatively analyzed in terms of different discourse features related to the four basic coherence parameters described above. these features are the total number of elementary discourse units (EDUs), the total number of error free EDUs, the total number of complete EDUs, the total number of rhetorical relations and the total number of error free rhetorical relations and the frequency of rhetorical relations types.

The following table describes and operationally defines each coherence parameter and metric in details

Table1: The Measures of coherence quality metrics with operationalized definitions

Coherence quality metrics	Coherence quality measures	Definitions
Completedness	- The total number of EDUs - The total number of error free EDUs - The total number of complete EDUs	- The total number of discourse units which range from clauses to paragraphs - The overall number of EDUs that maintain discourse coherence without causing any disruptions - The total number of self-contained discourse units
Connectedness	- The total number of rhetorical relations - The total number of error free rhetorical relations	- The total number of rhetorical relations that recursively bind discourse units together - The total number of logical rhetorical relations that recursively bind discourse units together
Uniqueness	- The frequency of rhetorical relations types	- The ratio of presentational and subject matter relations types
Adjacency	- The mean of discourse distance	- The distance between discourse units and discourse relations

The final discourse coherence measure is the mean of dependency distance or discourse distance. It should be noted that both the length of EDUs and dependency length were not taken into consideration as a coherence parameter simply because EDUs, according to our study, range from sentences as minimal units to paragraphs as maximal units. Also, dependency length is concerned with the sentential level, not the discourse level because this metric calculates the distance between the linguistic head and the dependent by summing up all the types of linear distances within a sentence. (Sun & Xiong, 2019) That is why, we relied instead on the size/ length of the text as a whole (dependency distance) since we are interested in coherence at the discourse level and not at the sentence level. dependency distance metric which is a measure of adjacency parameter refers to the interaction between discourse relations and discourse units, i.e., how "discourse units combine with discourse relations to form recursively larger units up to a global unit." (Sun & Xiong, 2019, p.3).

The dependency distance metric is basically rooted in dependency grammar which is concerned with the dependency relation between the linguistic units (words) through the connection between a governor (head) and a dependent and where the distance between these two variables constitutes the dependency distance. Similarly, according to Sun & Xiong (2019), discourse, as a sentence, is produced and received linereally.

Due to the linearity of discourse, discourse structure can be analyzed hierarchically and relationally (Sanders & Vijk, 1996). The hierarchical aspect focuses on the dominance of one discourse unit over another as well as on the distance between connected discourse units, while the latter concerns the semantic or logical meaning of connections between discourse units (i.e., discourse relations). (p.2)

This implies that the relational level includes discourse relations, while the hierarchical level includes the connection of discourse units. Accordingly, the head in RST theory is analogous to a nucleus, while the dependent is analogous to a satellite, so the dependency distance for a RST relation is the distance between the nucleus node and the satellite node and which can be calculated by the following formula:

$$MDD(\text{the sentence}) = \frac{1}{n-1} \sum_{i=1}^n |DD_i|$$

(Sun & Xiong, 2019, p.7)

In this equation, MDD stands for the 'mean dependency distance', n for the number of words in the sentence, and dd_i for the dependency distance of the i th syntactic link of the sentence. Based on the calculation of the mean discourse distance of the sentence, the mean discourse distance of a text can be calculated in the above formula. This is justified by the reason that a rhetorical relation in RST is equal to a dependency relation in dependency grammar. That's why, "a text is treated as a sentence in dependency grammar and it will then add up all values of mean discourse distance and divide this by the number of RST text. The final value is the mean discourse distance." (Sun & Xiong, 2019, p.7) When the node consists of more than one EDU, the node is represented by the nearest EDU, which is the head of a node.

Sun & Xiong (2019) gave an example of how to calculate the mean discourse distance for the text (WSJ_0660) whose root consists of 13 EDUs. The node (1-2) is linked to the node (3-4) by a circumstance relation. In this case, node (1-2) becomes the nucleus, and node (3-4) becomes the satellite. Thus, the mean discourse distance for this text can be calculated in the following way:

$$(2-1) + (4-3) + (4-1) + (8-7) + (7-5) + (12-11) + (13-11) + (10-9) + (10-1) / (13-1) = 21/12 = 1.75$$

In this way, according to Sun & Xiong (2019), the mean discourse distance for a text should not be beyond 3.18 as the average mean discourse distance because "when the mean discourse distance for a text is large, people might find it more difficult to process and understand the structure of the discourse." (p.18) In the present study, we also consider the average 3.18 as a reliable metric of the mean discourse distance.

3.2 Discourse Annotation guidelines

The segmentation of written texts is built on the guidelines of *The RST Discourse Treebank Annotation Manual* developed by Carlson and Marcu (2001) because it is mainly based on the RST. Two to three independent annotators are required to participate in this process. Prior to the segmentation process, all annotators should be briefed about the underlying assumptions and principles of the RST and provided with the discourse tagging manual. Annotators should read each composition several times and manually segment it into elementary discourse units (EDUs) corresponding to clauses and delimit them by square brackets. In this context, it is worth noting that Carlson and Marcu (2001) did not consider clauses which are subjects or objects of a main verb and clauses which are complements of a main verb as EDUs; while they considered relative clauses, nominal post-modifiers and clauses that break up other EDUs as embedded discourse units.

The next step was to determine the rhetorical relations types that hold between two adjacent spans based on the conformity of the cases being examined to the definitions of relations in the RST theory. Following the same protocol of Carlson et al. (1999), when the annotators are unable to determine the type of rhetorical relation existing between two spans either because it does not belong to the taxonomy of rhetorical relations types or because the type of the relations is ambiguous for them, they should write other relation. After the identification of relations, the major chunks of discourse structure should be linked together. The conventions of interpreting the discourse structure of texts under examination following the RST-DT annotation protocol are described as follows,

The text within square brackets denotes a span. Each pair of square brackets is followed by either the uppercase character N, referring to the nucleus span, or the uppercase character S, referring to the satellite span. A pair of two spans (N and S) is respectively followed by a dash and the name of the relation that holds between the spans. The relation name is further followed by parentheses containing the file number (of the source document), and the span numbers (the location of the relation in the document), respectively. In addition, the file number and the span numbers within the parentheses are separated by a colon, and each span number is separated from the other span number by a forward slash. For highlighting a particular signal used, the relevant parts (referring to the relevant textual features) in a span are underlined.

(Das & Taboada, 2014, p.90)

After the annotation process, the annotators should compare the results of their segmentations, especially that the annotation of the same discourse can yield different structures because of disagreement related to the number of relations, the type of relations as well as their signals (Mann & Thompson, 1987; Taboada & Mann, 2006). For instance, in some cases, more than one rhetorical relation type can hold between adjacent spans. However, differences can be resolved by referring to the tagging discourse manual to decide on the most suitable relation. When the annotators cannot agree on any type of rhetorical relation, they can write other relations. To check the inter-annotator agreement, we opt for the same method used by Marcu et al. (1999) to compute Cohen's Kappa coefficient value which "measures pairwise agreement among a set of coders who make category judgments, correcting for chance expected agreement, where $P(A)$ is the proportion of times a set of coders agree and $P(E)$ is the proportion of times a set of coders are expected to agree by chance." (Marcu et, 1999, p.51)

The kappa coefficient can be calculated as follows,

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

According to Landis & Kock's (1977) interpretation of Kappa coefficient values, values which are less than 0.00 indicate a poor agreement, values between 0.00 and 0.20 indicate a slight agreement, between 0.21 and 0.40 indicate a fair agreement, between 0.41 and 0.60 indicate a moderate agreement, between 0.61 and 0.80 indicate a good agreement and between 0.81 and 0.90 indicate almost a perfect agreement. Following this arithmetical measure, the Kappa statistics should demonstrate the degree of agreement and reliability between the independent annotators at various annotation phases for the elementary discourse units, error free elementary discourse units, complete elementary discourse units, the total number of rhetorical relations, the total number of error-free rhetorical relations, the frequency of rhetorical relations types and the mean discordsdistance.

4. Conclusion and recommendation

Building on the basic principles of rhetorical structure theory, the present paper proposes an assessment framework for discourse coherence and hence bridges the gap between discourse description and discourse assessment. However, this framework remains conceptual and needs further empirical investigations. Future research should examine its validity and applicability across different writing genres and proficiency levels. Also, future research is required to explore the different ways of integrating this framework into writing assessment rubrics to enhance EFL learners' writing coherence.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- [2] Grosz, B. Sidner, C. (1986) Attention, intentions and the structure of discourse. *In Computational Linguistics*, 12(3), 175–204.
- [3] Carlson, L., & Marcu, D. (2001). Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 54(2001), 56.
- [4] Hobbs. J. R. (1985). On the coherence and structure of discourse. Technical Report CSLI-85-37, Center for the Study of Language and Information, Stanford, 1985.
- [5] Mann, W. C., & Thompson, S. A. (1987). *Rhetorical structure theory: A theory of text organization*. Los Angeles: University of Southern California, Information Sciences Institute.
- [6] Polanyi, L., Culy, C., van den Berg, M., Thione, G. L., & Ahn, D. (2004). A rule based approach to discourse parsing. Paper presented at the 5th SIGdial Workshop on Discourse and , ACL, Cambridge, MA.
- [7] Sanders, T., Spooren, W., & Noordman, L. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15, 1-35.