Document Cohesion Flow: Striving towards Coherence

Scott A. Crossley (scrossley@gsu.edu)
Department of Applied Linguistics/ESL, Georgia State University
Atlanta, GA, 30303 USA

Mihai Dascalu (mihai.dascalu@cs.pub.ro)
Stefan Trausan-Matu (stefan.trausan@cs.pub.ro)
Department of Computer Science, University Politehnica of Bucharest
Bucharest, 060042 Romania

Laura K. Allen (laurakallen@asu.edu)
Danielle S. McNamara (dsmcnamara1@gmail.com)
Department of Psychology, Arizona State University
Tempe, AZ, 85287-2111 USA

Abstract

Text cohesion is an important element of discourse processing. This paper presents a new approach to modeling, quantifying, and visualizing text cohesion using automated cohesion flow indices that capture semantic links among paragraphs. Cohesion flow is calculated by applying Cohesion Network Analysis, a combination of semantic distances, Latent Semantic Analysis, and Latent Dirichlet Allocation, as well as Social Network Analysis. Experiments performed on 315 timed essays indicated that cohesion flow indices are significantly correlated with human ratings of text coherence and essay quality. Visualizations of the global cohesion indices are also included to support a more facile understanding of how cohesion flow impacts coherence in terms of semantic dependencies between paragraphs.

Keywords: Cohesion Flow; Natural Language Processing, Computational Models; Cohesion Network Analysis; Coherence; Writing Quality

Introduction

Writing is an important aspect of communication because it provides the opportunity to articulate ideas and synthesize perspectives in a persuasive manner that is often independent of time and space constraints (Crowhurst, 1990). Learning how to convey meaning competently in written texts is a crucial skill for academic and professional success. Indeed, the writing skills of college freshmen are among the best predictors of academic success (Geiser & Studley, 2001). Importantly, there are a number of attributes of high-quality writing that are directly related to the linguistic features of the written text.

One important element of writing quality is text *cohesion.* Halliday and Hasan (1976) established this term to characterize the cohesive links or ties between the sentences from the same text or fragment of text. Cohesion is a crucial element for easing the understanding of texts, particularly for challenging texts that present knowledge demands to the reader (McNamara, Kintsch, Songer, & Kintsch, 1996).

In addition, there is a general sense that essay quality is highly related to the cohesion of a text, which is reflected in the literature about writing as well as textbooks that teach students how to write (Collins, 1998). However, this assumption is compounded by differences in cohesion types (i.e. local and global cohesion). Local cohesion refers to text features that link short text segments such as sentences together, while global cohesion refers to text features that link larger segments of texts (i.e., paragraphs). Previous studies have indicated that local cohesion indices are generally negatively associated with essay quality (Crossley & McNamara, 2010, 2011) while global cohesion indices are generally positively correlated with essay quality (Crossley, Kyle, & McNamara, in press; Crossley & McNamara, 2011).

The primary difference between local and global cohesion indices is that global cohesion indices tap into text *coherence*, while local cohesion does not (Crossley et al., in press). The distinction between cohesion and coherence is that cohesion refers to the presence or absence of explicit textual cues that allow the reader to make connections between the ideas in the text. On the other hand, coherence refers to the understanding that the reader derives from the text, which may be more or less coherent depending on a number of factors, such as textual features, prior knowledge, and reading skill (McNamara et al., 1996). Thus, coherence plays a central role in terms of the sense that a text creates while addressing a deeper function of discourse (i.e., illocutionary acts; Widdowson, 1978).

The goal of this study is to further research links between global cohesion devices and both text coherence and essay quality to better understand the cognitive underpinnings of language comprehension. We do this by developing new automated indices of global cohesion (termed *cohesion flow* indices) and demonstrating how these indices can be used to predict human judgments of text coherence and essay quality. These indices are provided in the *ReaderBench* tool (RB; Dascalu, Trausan-Matu, McNamara, & Dessus, 2015).

Cohesion and Cohesion Flow

Cohesion at both the local and global levels helps to develop a text that is unified and connected as a result of individual sentences and paragraphs "hanging" together and relating to one another (Celce-Murcia & Olshtain, 2000). In addition to text segments "hanging" together, a well structured document should have a clear and logical movement of ideas highlighted by the topic flow between sentences and paragraphs (O'Rourke, Calvo, & McNamara, 2011).

Knowing that cohesion is a central element of a text's structure, we are interested in examining how flow and cohesion might be structurally constructed in a text (i.e., derived from the proper sequencing of text elements). Thus, we operationalize *cohesion flow* as a measure of a document's structure derived from the order of different paragraphs and of the manner in which they combine to hold the text together. A text that demonstrates strong cohesion flow by linking ideas between paragraphs will likely be a more coherent text. This would allow ideas to bond together and flow smoothly from one paragraph to another, creating a text that readers can more easily comprehend. In addition, the ability to illustrate and automatically assess cohesion flow would enable researchers to observe how text segments in a document fit together and examine how the sequencing of these segments may affect readers' comprehension.

Cohesion Features and Text Quality

Previous research has placed an emphasis on the relationship between cohesive devices in text and human judgments of that text's overall quality and coherence. In terms of essay quality, McNamara, Crossley, and McCarthy (2010) reported that local cohesion indices did not demonstrate significant differences between low and high scored essays, nor did they correlate with essay scores. In a second study, Crossley, Roscoe, et al. (2011) found that two indices of global cohesion (semantic similarity between initial and middle paragraphs, and semantic similarity between initial and final paragraphs) significantly correlated with essay quality. More recently, Crossley et al. (in press) reported that while local features of cohesion (i.e., sentence overlap indices) were negatively related to essay quality, global indices of cohesion (i.e., paragraph overlap indices) were strongly predictive of essay quality. In another study, Crossley and McNamara (in press) found that modifying the cohesion structure in a text at the global level led to increased essay quality scores (as assigned by human raters). They also reported strong links between global cohesion features and the essay quality scores.

Similar findings have been reported in terms of text coherence. In three recent studies, Crossley and colleagues (Crossley et al., in press; Crossley & McNamara, 2010, 2011) examined the degree to which judgments of text coherence were predicted by automated indices of local and global cohesion reported by a number of computational tools. Crossley and McNamara (2010) reported that cohesion indices calculated at the local level demonstrated significant but negative correlations with human ratings of coherence. A follow up study (Crossley & McNamara, 2011) reported similar results for local cohesion devices. However, this study also examined a number of global cohesion indices that calculated overlap between initial, middle, and final paragraphs. These global cohesion indices were positively correlated with judgments of text coherence. Two recent studies found that while local features of cohesion were negatively related to text coherence, global indices of cohesion were strongly and positively associated with text coherence (Crossley et al., in press; Crossley & McNamara, in press) and that modifications to structure of a text in terms of global cohesion led to increased text coherence scores (Crossley & McNamara, in press). Overall, these studies provide support for the notion that expert ratings of text coherence and essay quality are best explained by global and not local cohesion devices.

Current Study

The current study develops and tests automated measures of global cohesion flow based on the Cohesion Network Analyses available within RB. We also demonstrate how visualizations based on these network analyses can be used to illustrate cohesion trends within a text. Our approach is unlike previous operationalizations of global cohesion in that prior studies (e.g., Crossley et al., in press) focused on assessing lexical and semantic overlap between adjacent paragraphs, and have not assessed the flow of cohesion throughout a text. In this study, we apply automated indices of cohesion flow to a corpus of essays written by college level students and examine how well these measures predict human scores of text coherence and overall writing quality. Thus, our goals are to operationalize new measures of cohesion flow and to use these measures to better understand how these features relate to text coherence and writing quality as judged by human raters. Such an approach gives us insight into the textual properties that drive human perceptions of coherence in language. In addition, we examine how visualizations extracted the cohesion flow indices can be used to illustrated text coherence.

Method

Quantifying cohesion flow

Measuring cohesion. In the current study, we limit the perspective of coherence to *semantic global cohesion* that captures text organization in terms of paragraph links (Lapata & Barzilay, 2005). We focus on paragraph links over sentence links because previous studies have reported stronger associations between global cohesion and text coherence and quality compared to local cohesion. We do this by evaluating cohesion flow using three different semantic models: a) semantic distances in ontologies (i.e., Wu-Palmer; Budanitsky & Hirst, 2006), b) cosine similarity in Latent Semantic Analysis (LSA) vector spaces (Landauer, Foltz, & Laham, 1998) and c) Jensen Shannen dissimilarity of Latent Dirichlet Allocation (LDA) topic distributions (Blei, Ng, & Jordan, 2003).

Building the cohesion flow graph. Our cohesion flow evaluation relies on Cohesion Network Analysis (CNA) (Dascalu et al., 2015), a computational model that uses a *cohesion-based representation of discourse*. CNA combines the semantic models discussed above with Social Network Analysis (SNA) to build a multi-layered cohesion graph (Dascalu et al., 2014). Starting from the general CNA approach, a pruned graph consisting of *cohesion*[*i*, *j*] edges, where *i* and *j* are paragraph indexes, is constructed at inter-paragraph level. This cohesion flow graph uses the chronological indexing of text elements (in contrast to the initial cohesion graph that uses undirected edges) and two building criteria:

- *maximum value*: for each paragraph we determine, out of the following paragraphs, which is the closest one in terms of the selected similarity measure (maximum cohesion to following text elements).
- *above threshold*: out of all possible links between subsequent paragraphs, we select only those graph edges that exceed the imposed threshold of average + stdev of all inter-paragraph similarity measures.

Measuring cohesion flow. Based on the previous cohesion graph, a *topological sort* (Cormen, Leiserson, Rivest, & Stein, 2009) is performed in order to observe the referential sequencing of paragraphs. Afterwards, six cohesion flow indices are calculated in order to evaluate cohesion flow at the document level:

- *Absolute position accuracy*: number of paragraphs that, after performing the topological sort on the cohesion flow graph, are in the correct position (the ordered paragraph index is the same as the initial index).
- *Absolute distance accuracy*: the absolute value of the difference of ordered and initial paragraph indexes. A value closer to 0 characterizes a more cohesive text in terms of adjacency links.
- *Adjacency accuracy* determines how many paragraphs follow the idea of adjacency maximum flow: sum of absolute values of (j-i-1) where cohesion[i, j]>0.
- *Average flow cohesion* is determined as the average cohesion in our cohesion flow graph, i.e., average of all cohesion[i, j];
- *Spearman correlation* between the ordered paragraph index and the initial sequence index.
- *Max order sequence* determines how many ordered paragraph indices follow an increasing trend to determine if flow moves forward in a document (i.e., what is the longest sequence that follows an ascending trend).

All the above indices are normalized in [0; 1] by relating to the overall size of the document. In total we developed 36 indices based on the six feature categories above with each feature computed for topological sort and for LSA, LDA, and Wu-Palmer (36 features in total). Documents having fewer than three paragraphs cannot be not considered in our cohesion flow analysis.

Visualizing cohesion flow. In addition to computing automated indices of cohesion flow, we are able to visualize the process in the *ReaderBench* (RB) tool. Starting from the cohesion flow graph, we apply specific SNA metrics in order to obtain an in-depth perspective of the paragraph connectivity (i.e., the strength of paragraph associations). This flow network visualization and modeling plays an important role in understanding and interpreting the obtained results. We have selected *force-based layouts* for visual representation because they provide a comprehensive view of the social network as a planar graph. In this layout, paragraphs gravitate by having their own mass proportional to their relatedness to other paragraphs. Edges are proportional to the inverse semantic relatedness of paragraphs, while elasticity coefficients are used to obtain a more realistic visualization that minimizes edge crossings and the overall network energy. The size of each paragraph can be proportional to its betweenness centrality (i.e., the number of times it acts as a bridge along all shortest paths between pairs of two other paragraphs from the input text).

Validation Corpus

We used the corpus described in Crossley and McNamara (2011), which comprises 315 timed essays collected from undergraduate students at Mississippi State University. During the collection process, students were given 25 minutes to write an essay and no outside referencing was allowed. Such an environment better represents high stakes testing (i.e., SAT writing tests). Two SAT prompts were used and students were randomly assigned one prompt to which they responded. All students were native speakers of English and were in either Composition One or Composition Two (i.e., freshmen composition courses). Each essay was read and scored by two trained raters using both an analytic and a holistic rubric.

The rubric used to score the essays contained one analytic feature (*organization*: the body paragraphs follow the plan set up in the introduction) that evaluated semantic based, global cohesion (i.e., text coherence). A holistic grading scale based on a standardized rubric commonly used in assessing Scholastic Achievement Test (SAT) essays was also included in the rating rubric. The holistic scale and all of the rubric items had a minimum score of 1 and a maximum score of 6. Details on the rubric used can be found in Crossley et al. (in press).

Eight raters with advanced degrees in English and at least 3 years experience teaching university composition classes rated the 315 essays. The raters were informed that the distance between each score was equal. After the raters were trained, each rater then evaluated a selection of the 315 essays. In all cases, each essay was scored by at least two raters. Once final ratings were collected, differences between the raters were calculated. If the difference in ratings on a survey feature was less than 2, an average score was computed. If the difference was greater than 2, a third expert rater adjudicated the final rating. Correlations between the raters (before adjudication) were acceptable with the average correlations across all essay feature judgments and raters at $r = .72$.

Statistical Analysis

We conducted correlation and regression analyses to examine relations between our document flow indices and human judgments of coherence and essay quality. We first conducted correlational analyses to examine associations between the indices and the human scores. Those indices that demonstrated a statistical ($p < .05$) and meaningful (at least a small effect size, $r > .1$) relation were then used in a regression analysis. Indices that were highly collinear ($r > .899$) were flagged, and the index with the strongest correlation with human scores was retained while the other indices were removed. The remaining indices were included as predictor variables in a stepwise multiple regression to predict both human scores of coherence and overall essay quality. The model from the stepwise regression was then used to assign scores for the essays in the test set using a leave-one-out-cross-validation (LOOCV).

Results

Organization Scores

Each of the 36 document flow indices demonstrated a significant correlation with the organization scores. However, after controlling for multicollinearity, only three indices remained. These three indices demonstrated medium effect sizes with human ratings of text coherence (see Table 1) and were used in the subsequent regression analysis.

Table 1: Correlations between RB cohesion flow indices and raters' organization scores

Variable	*r* value	*p* value
Adjacency accuracy (Maximum value based on LDA)	.399	< .001
Spearman correlation of flow versus initial ordering (Above mean+stdev based on LDA)	.382	< .001
Absolute distance accuracy on topological sort (Maximum value based on Aggregated score)	.381	< .001

A LOOCV linear regression analysis was conducted including the three RB indices. These three variables were regressed onto the raters' coherence evaluations for the 315 writing samples. Of these three variables, two were significant predictors in the regression: Absolute distance accuracy on topological sort (Maximum value based on Aggregated score) and Adjacency accuracy (Maximum value based on LDA). The linear regression using the two variables yielded a model that reported $r = .398$ (MAE = .779). The model accounted for 16% of the variance in the human evaluations of coherence.

Essay Quality Scores

Each of the 36 document flow indices demonstrated a significant correlation with the essay quality scores. However, after controlling for multicollinearity, only three indices remained. These three indices demonstrated medium effect sizes with human ratings of essay quality (see Table 2) and were used in the subsequent regression analysis.

Table 2: Correlations between RB cohesion flow indices and raters' scores of essay quality

Variable	r value	*p* value
Adjacency accuracy (Maximum value based on LSA)	.356	< .001
Average document flow cohesion (Above mean+stdev based on Aggregated score)	.317	< .001
Absolute distance accuracy on topological sort (Maximum value based on LDA)	.310	< .001

A LOOCV linear regression analysis was conducted including the three RB indices. These three variables were regressed onto the raters' essay score for the 315 writing samples. Of these three variables, one was a significant predictor in the regression: Adjacency accuracy (Maximum value based on LSA). The linear regression using the variable yielded a model that reported $r = .334$ (MAE = .763). The model accounted for 11% of the variance in the human evaluations of essay quality.

Discussion and Conclusion

While text coherence can be influenced by a reader's prior knowledge and/or reading skill (McNamara et al., 1996), it also depends on the features of the text. Understanding how elements of the text can increase human judgments of coherence is thus an important area of research in discourse processing and in theories of writing because it can provide us with information about the impact of linguistic features on the cognitive processes involved in text comprehension. Previous research has supported the notion that both text coherence and essay quality are associated with global, but not local cohesion features in the text. The findings from this study further this previous research and extend it by introducing new computational operationalizations of global cohesion that are available in the *ReaderBench* (RB) tool. These operationalizations, which are based on cohesion flow, provide new evidence for how text features can combine and interact to create more coherent text that leads to stronger evaluations of writing quality. The findings provide us with a more in-depth understanding about how text cohesion can lead to text coherence and the effects such cohesion has on essay quality.

The primary goal of this research was to develop and test new indices of global cohesion. The indices developed are based on cohesion flow (i.e., the hierarchical topic progression among paragraphs in a text). The purpose of the new indices was to allow for the examination of the unity and connectedness of paragraphs in a text, but not simply through assessing semantic or lexical overlap among the paragraphs. Rather, our indices examine if paragraphs are sequenced appropriately in reference to one another. Thus, the indices do not measure lexical and semantic overlap between paragraphs, but rather overlap among paragraphs.

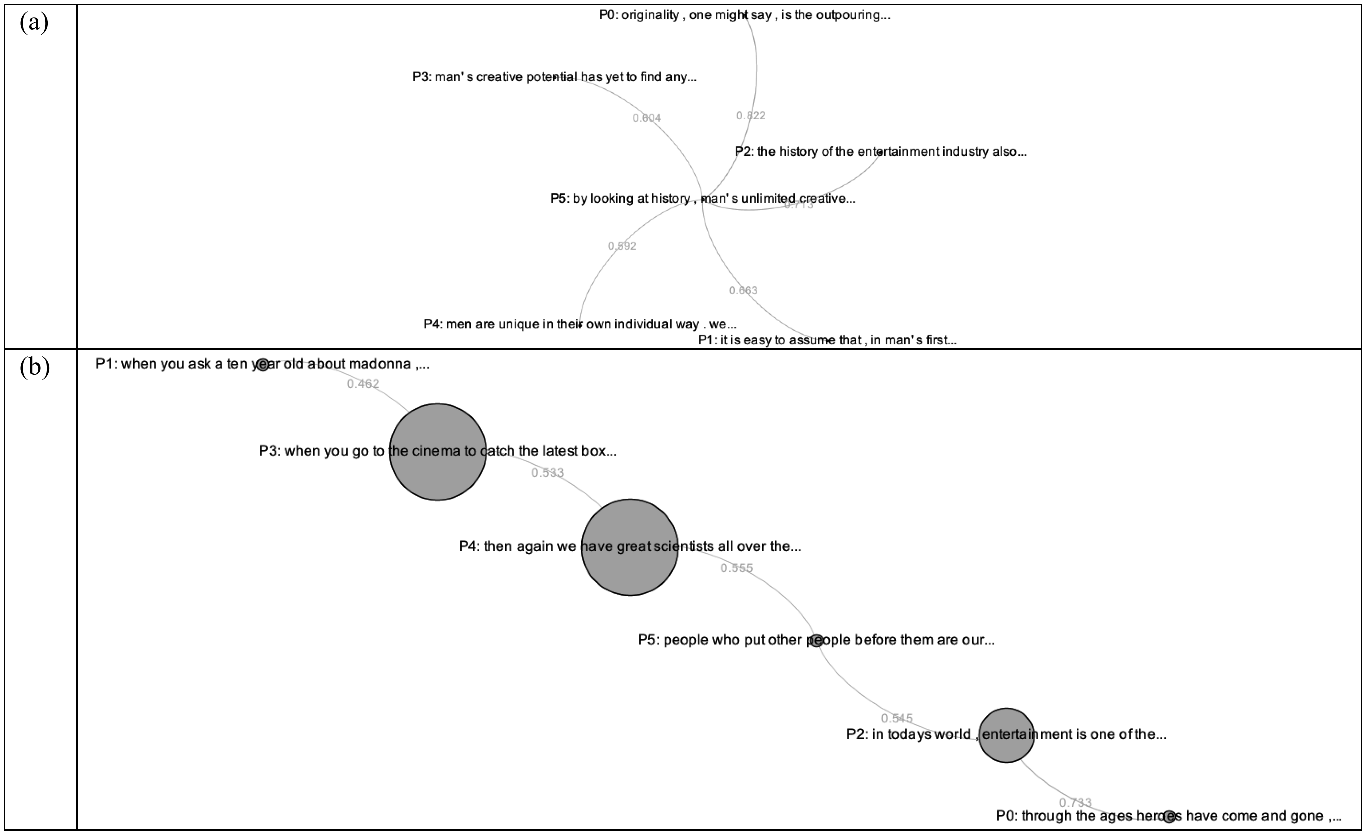

Figure 1: Differences between (a) an essay scored high in organization and (b) an essay scored low in organization.

To do this, we conducted a cohesion flow evaluation derived from a multi-layered cohesion graph that afforded chronological indexing of text elements. Using three different semantic models (LSA, LDA, and Wu-Palmer), we developed 36 indices of cohesion flow that calculated referential sequencing among paragraphs. Our hypothesis was that essays that contained strong cohesion flow among paragraphs would be judged as more coherent and of higher quality. We found this to be true in both cases, providing additional support that automated indices that measure global cohesion through cohesion flow are related to text coherence and essay quality.

We found that each of the 36 indices correlated with human ratings of text coherence and essay quality. Because these indices were measuring a very similar construct, most of the indices were multi-collinear. However, in both test sets, measures based on adjacent accuracy were the strongest predictors of human ratings. This result indicates that texts with the greatest number of paragraphs that flowed together were the most coherent and of the highest quality. These indices were followed by Spearman correlations of flow versus initial ordering, absolute distance accuracy on topological sort, and average document flow cohesion. These indices indicated that a) texts with paragraphs that showed strong correlations or absolute values between the ordered paragraph index and the initial sequence index received higher ratings, and b) texts with high average flow cohesion also received higher scores.

Beyond their predictive power in machine learning models, an important component of the cohesion flow indices is that they are able to be visualized in the RB. As an example, Figure 1 depicts the cohesion flow of two six-paragraph essay (P0…P5) using LSA as semantic similarity function and maximum value as building criteria. The essays presented in this figure were scored 5.5 (essay *a*) and 2.0 (essay *b*) on the organization rating. The figure demonstrates an SNA perspective of the cohesion flow graph in which the paragraphs are the nodes, the edges represent semantic similarity values (LSA) that meet the building criteria, and the size of each node is proportional to its betweenness centrality. A well-organized essay (*a*) summarizes all previous ideas in the conclusions; therefore, the visualized star shape links all paragraphs to the conclusions. Moreover, it is expected that all nodes have a size equal to 0 because we have only in-edges towards the last paragraphs, so no shortest paths exist between pair of nodes excluding the conclusion. In contrast, essay (*b*) highlights an essay that lacks proper sequencing in that the cohesion flow is disrupted by the ordering of paragraphs. In short, the visualization process allows the complex mathematical equations that underlie our cohesion flow indices to illustrate how cohesion flow operates within a text. Such a visualization is helpful in understanding trends in global cohesion patterns within a text and how these patterns are related to text coherence and writing quality.

In conclusion, we have developed new indices of global cohesion that are based on cohesion flow. These indices show significant correlations with human ratings of text coherence and essay quality providing evidence that the coherence of a reader's mental representation is influenced by links among the paragraphs in the text. However, these indices predict only a small amount of the variance in human ratings of text coherence (about 16%) and essay quality ratings (about 11%). These numbers are on par with previous studies examining global cohesion indices and human ratings of text quality (Crossley & McNamara, 2011; Crossley et al., in press), but do indicate that human ratings of text coherence and essay quality are only partially explained by cohesion flow among paragraphs. Future iterations of this research will combine the cohesion flow indices introduced in this study along with other global indices of cohesion and indices related to characteristics of writing such as lexical sophistication, fluency, spelling, and syntactic complexity. Such analyses are beyond the scope of the current study, but we expect them to provide a deeper understanding of how linguistic and semantic elements are collectively related to text coherence and text quality, both crucial elements of discourse processing. In addition, we will explore the degree to which the visualizations are useful to practitioners and writers in order to improve document structure. The overall aim is to provide users with data to maximize cohesion flow between adjacent paragraphs and to develop conclusions with coherent summaries.

Acknowledgment

This research was partially supported by the EC H2020 project RAGE Grant agreement No 644187, as well as by the Institute for Education Sciences (IES R305A080589 and IES R305G20018-02) and NSF (1417997 and 1418378).

References

Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*(4-5), 993–1022.

Budanitsky, A., & Hirst, G. (2006). Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Computational Linguistics, 32*(1), 13–47.

Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and Context in Language Teaching: a Guide for Language Teachers*. New York, NY: Cambridge University Press.

Collins, J. L. (1998). *Strategies for struggling writers*. New York, NY: The Guilford Press.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., & Stein, C. (Eds.). (2009). *Introduction to Algorithms* (3rd ed.). Cambridge, MA: MIT Press.

Crossley, S. A., Kyle, K., & McNamara, D. S. (in press). The Tool for the Automatic Analysis of Text Cohesion (TAACO): Automatic Assessment of Local, Global, and Text Cohesion. *Behavior Research Methods.*

Crossley, S. A. & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of CogSci'10* (pp. 984-989). Austin, TX: Cognitive Science Society.

Crossley, S. A., & McNamara, D. S. (2011). Text coherence and judgments of essay quality: Models of quality and coherence. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of CogSci'11.* (pp. 1236-1241). Austin, TX: Cognitive Science Society.

Crossley, S. A., & McNamara, D. S. (in press). Say more and be more coherent: How text elaboration and cohesion can increase writing quality. *Journal of Writing Research.*

Crossley, S. A., Roscoe, R. D., McNamara, D. S., & Graesser, A. (2011) Predicting human scores of essay quality using computational indices of linguistic and textual features. In G. Biswas et al. (Eds.), *Proceedings of AIED 2011*. (pp. 438-440). New York: Springer.

Crowhurst, M. (1990). Reading/writing relationships: An intervention study. *Canadian Journal of Education, 15*, 155-172.

Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., & Nardy, A. (2014). Mining texts, learners productions and strategies with ReaderBench. In A. Peña-Ayala (Ed.), *Educational Data Mining: Applications and Trends* (pp. 335–377). Switzerland: Springer.

Dascalu, M., Trausan-Matu, S., McNamara, D.S., & Dessus, P. (2015). ReaderBench – Automated Evaluation of Collaboration based on Cohesion and Dialogism. *ijCSCL, 10*(4), 395–423. doi: 10.1007/s11412-015-9226-y

Geiser, S., & Studley, R. (2001). UC and SAT: *Predictive validity and differential impact of the SAT I and SAT II at the University of California*. Oakland: University of California.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion In English*. London: Longman.

Landauer, T.K., Foltz, P.W., & Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes, 25*(2/3), 259–284.

Lapata, M., & Barzilay, R. (2005). Automatic evaluation of text coherence: models and representations. In *19th international joint conference on Artificial intelligence* (pp. 1085–1090). Edinburgh, Scotland: Morgan Kaufmann Publishers Inc.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). The linguistic features of quality writing. *Written Communication, 27* (1), 57-86.

McNamara, D.S., Kintsch, E., Butler Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1–43.

O'Rourke, S., Calvo, R. A., & McNamara, D. S. (2011). Visualizing topic flow in students' essays. *Journal of Educational Technology & Society, 14,* 4-15.

Widdowson, H.G. (1978). *Teaching Language as Communication*. Oxford: Oxford University Press.

Wu, Z., & Palmer, M. (1994). Verb semantics and lexical selection. In *ACL '94* (pp. 133–138). New Mexico, USA: ACL.