

Beyond proficiency: Linguistic features of exceptional writing
Towns, S.G. and Watson Todd, R.

English Text Construction, Vol 12, No. 2, pp. 265-289. 2019.

The definitive version of this article was published as Towns, S. G., & Watson Todd, R. (2019). Beyond proficiency: Linguistic features of exceptional writing. English Text Construction, 12(2), 265-289.

<https://doi.org/10.1075/etc.00029.tow>

Beyond proficiency: Linguistic features of exceptional writing

Abstract: Many studies have investigated the correlations between linguistic features and human judgements of writing quality. These studies usually investigate either proficient student writing or exceptional literary writing. The current study attempts to bridge these two perspectives by comparing proficient writing to award-winning exceptional writing using movie reviews written by bloggers and Pulitzer Prize winners. A range of linguistic features representing syntactic complexity, lexical complexity, and lexical cohesion were analyzed using both automated and interpretive methods. It is found that some, but not all, of the trends seen in writing development studies continue on to exceptional writing, with lexical sophistication and lexical cohesion through conceptual associations making the largest contributions to the differences between proficient and exceptional writers.

Keywords: movie reviews, lexical cohesion, lexical complexity, syntactic complexity, writing quality

1. Introduction

The evaluation of English writing quality and linking this quality to the presence of certain linguistic features has been a topic of research for decades. However, the meaning of the term “quality” is dependent on the texts being analyzed. In the case of student writing, proficiency is the goal and quality is interpreted as matching a standard model such as a “readable expository style that will serve for courses and, later, for professional or civic writing assignments” (Shaughnessy, 1979: p. 280). Writing research in the field of literary studies, on the other hand, only analyzes the most exceptional quality of writing, such as texts which are a part of the literary canon (van Peer 2008). These texts are praised as being exceptional due to factors such as the texts’

emotional impact on the reader, or the use of metaphors and symbolism, or the texts' insightful commentary on culture and society.

The methods by which linguistic features in the text are identified can also differ depending on the context and the goals of the research. There are two contrasting prototypical approaches to research into the relationship between writing quality and linguistic features. The first approach is purely quantitative research which attempts to correlate linguistic features with writing quality. Most of this research examines student writing (e.g., Ortega 2003; McNamara, Crossley & McCarthy 2010), but a few studies have also been conducted on exceptional journalism (Louis and Nenkova 2011, 2013a, 2013b). The second approach is often seen in literary studies, where research investigates the use of linguistic features in exceptional texts. This literary research focuses on interpretations of how the linguistic features might affect the reader (van Peer 2008).

Our goal in this study is to gain insights into the nature of exceptional writing through the analysis of linguistic features. To accomplish this goal, we will compare proficient writing to exceptional writing using methodologies found in the quantitative research on student writing. These methodologies identify linguistic features that may form the basis for human judgments of writing quality using correlation statistics. To gain further understanding, however, it is also important to interpret how the various linguistic features identified as influencing quality are employed in exceptional writing, as is done in literary research. Therefore, we will draw on both quantitative and interpretive approaches to researching writing quality.

1.1. Quantitative analyses of linguistic features

The first approach to research into writing quality is exemplified by purely quantitative investigations into the development and proficiency of student writing, whether the texts being analyzed were written by students learning to write in their L1 in grade school (e.g., McCutchen & Perfetti 1982; Beers & Nagy 2009) or in college and beyond (e.g., Berninger, Abbott, Nagy & Carlisle 2010; Crossley & McNamara 2010), or by those learning to write in an L2 in any context (e.g., Grant & Ginther 2000; Ortega 2003). In such pedagogical contexts, students are instructed in writing with the goal being to match a standard model of proficient writing, which might include characteristics such as clear organization, strong argumentation, and appropriate register. To this end, it would be useful to know the extent to which certain linguistic features contribute to human judgements of the characteristics of the standard model. Previous research into student writing attempts to uncover these contributions by investigating whether a score for a particular linguistic feature in the texts correlates with a human evaluation of the texts. With the advent of computer technology, this genre of research is now usually conducted using automated computerized tools such as complexity analyzers or methods such as comparisons of student writing to databases or large corpora. One automated tool that is often used is Coh-Metrix (Graesser, McNamara, Louwerse & Cai 2004). As an example of this type of research, a study using Coh-Metrix to

investigate the correlation between linguistic features and human judgements of writing quality showed that lexical complexity (such as lexical rarity and lexical diversity) and syntactic complexity (such as the number of words before the main verb) are strong predictors of the quality of a text (McNamara, et al. 2010).

The extensive research into student writing development and quality (especially in an L2 context) has not only investigated snapshots of writing and quality judgements at a specific time, but has also allowed researchers to see general trends across different levels of proficiency. This research has traditionally focused on the three main areas of syntactic complexity, lexical complexity, and lexical cohesion. Syntactic complexity research has investigated metrics such as the number of words per sentence, the number of words before the main verb in a sentence, or the number of clauses and T-units per sentence. In general, it has been shown that syntactic complexity increases as students become more proficient (Lu 2010; Bulté & Housen 2014). Another set of linguistic features which is often studied is lexical complexity, using quantitative metrics such as the average length of words, the sophistication of words, or the density of word types. As with syntactic complexity, studies have shown that lexical complexity has a positive correlation with proficiency (Engber 1995; McNamara et al. 2010). The third common area that is studied is lexical cohesion. Cohesion metrics include the number of repeated words, or the percentages of connected topics in a text. Interestingly, the type of cohesion studied most extensively, referential cohesion, initially increases as proficiency increases (similar to syntactic and lexical complexity), but then starts to decrease as students reach higher levels of proficiency (Crossley & McNamara 2011).

While various linguistic features in syntax, lexis, and cohesion show clear trends and patterns when comparing low proficiency and high proficiency student writing, it is unclear whether these trends extend beyond proficient writing to exceptional writing or if there is instead a change at the highest levels of writing quality. To our knowledge, the only quantitative studies into the linguistic features of exceptional writing were conducted by Louis and Nenkova (2011, 2013a, 2013b), who were investigating the differences between proficient and exceptional science journalism in the New York Times. In their study, proficient writing was exemplified by “typical” articles published in the New York Times, while exceptional texts were selected by experienced science journalists. In this way, these studies have a similar methodology to the studies of student writing above in that expert human readers are judging the quality of the texts. It is not clear, however, whether or not the science journalists were consulting specific criteria or grading rubrics in their judgements of exceptional quality, as teachers often do in a classroom setting, other than an attempt to classify the articles by the degree to which the texts were “clear, interesting, beautiful, or well-structured” (Louis and Nenkova, 2013b: p. 88). These proficient and exceptional science journalism texts were then evaluated and compared by Louis and Nenkova on several different linguistic features, such as the use of concrete words by comparing the text to psycholinguistic databases, “beautiful language” which included the use of rare words and rare collocations of words, and the specificity of the words and the sentences in the text. It was found that in the articles judged to be exceptional,

concrete words were more likely to congregate at the beginning and end of the articles. Exceptional writing was also found to use more rare words and rare combinations of words (Louis & Nenkova 2013b). For specificity, proficient writing was found to be more specific than exceptional writing, but the specificity levels in exceptional writing were much more variable from sentence to sentence than in proficient writing (Louis & Nenkova 2013a). While valuable, we believe that these results need to be extended in two ways: first, by increasing the number of linguistic features to investigate, and second, by examining a different genre.

1.2. Interpretive analyses of linguistic features

The second approach to research into writing quality is commonly found in interpretive literary studies. In such research, the focus is not on proficiency as with student writing, but on the exceptional writing of high-quality literature. Such writing might have been authored by Pulitzer Prize winners or they might be texts that have been included in a literary canon. This research is highly qualitative with subjective interpretations of the meaning of the text by literary scholars. Most of this work does not focus specifically on linguistic features, but research in the field of stylistics does consider the effect of linguistic features on the quality of literary texts. As an example, Widdowson's (1992) analysis of "The Rime of the Ancient Mariner" discusses the frequent line-by-line verb tense changes found throughout the poem. Another study investigated the differences in linguistic features between William Shakespeare's "Romeo and Juliet" and Arthur Brooke's "The Tragicall Historye of Romeus and Juliet". Both plays were written around the same time and covered the same story, but Shakespeare's version is the one that has been canonized as being exceptionally well written, while Brooke's has been all but forgotten. It has been argued that one of the reasons Shakespeare's version is so celebrated is simply because the quality of the text is superior and is "resting ultimately on specific concrete [linguistic] measures" (van Peer 2008: 25). Other studies of the linguistic features of exceptional literary writing take a corpus linguistics approach (Mahlberg 2007). These features, for example, have been used to uncover the lexical and grammatical patterns spoken by different characters in "Romeo and Juliet" (Culpeper 2002). Word frequencies, keywords and phraseology have also been investigated in Joseph Conrad's "Heart of Darkness" (Stubbs 2005) as well as in several works of Jane Austen (Fischer-Starcke 2010).

Unfortunately, the interpretive approach of the majority of literary studies makes comparisons between texts difficult as the results are not generalizable enough to use for the analyses of other texts. The interpretive approach does, however, create understandings of how a text uses certain linguistic features to influence the reader.

1.3. Comparing quantitative and interpretive approaches

Most of the studies of both student writing and literary stylistics use a common, if not explicitly stated, method, which compares holistic and/or analytic human judgements of quality with various linguistic features found in the text. Holistic analyses can be seen in an overall score for a student essay (as in McNamara, et al., 2010) or a judgement that a literary text should be enshrined in a canon (as in van Peer 2008), while analytic analyses score writing quality on quantitative metrics of specific areas such as content, organization, vocabulary, language use and mechanics (Connor-Linton & Polio 2014). These holistic and analytic human ratings can then be compared to the quantitative analysis of surface features of the texts such as the average sentence length, the number of rare words used, or the amount of referential cohesion in order to uncover any correlations between the human judgements and the surface features of the text.

Despite these similarities, quantitative and interpretive approaches to the study of linguistic features and writing quality can be differentiated on several different dimensions. The first difference is the definition of “writing quality”. In student writing, the focus is on matching the models of proficient writing, while in the interpretive literary studies, “writing quality” is related to the potential impact that the text has on the reader. The second difference is whether the analysis is on a corpus of whole texts or on specific points in one text. Research on student writing investigates a large number of texts from many different students, while literary analysis usually focuses on specific sections of one text by one author. A third difference is in the method of analysis. Since the student writing corpus is large, the texts must be analyzed quantitatively using automated methods with no room for qualitative interpretation, while the literary studies focus on linguistic features as a starting point for an interpretive analysis of the meaning of the text.

In order to shed light on the nature of exceptional writing, this study will compare exceptional writing to proficient writing using multiple methods and tools for analysis based on the strengths of the two research approaches investigating proficient student writing and exceptional literary writing. We will first conduct a quantitative analysis of linguistic features associated with syntax, lexis, and cohesion in two corpora which cover the same topics. The first corpus comprises texts written by proficient writers, while the second corpus is written by exceptional writers. We will then take an interpretive approach to analyze those linguistic features which seem to contribute the most to the notion of writing quality in exceptional texts. Therefore, there are two main research aims for this study:

1. Which linguistic features contribute the most to distinguish exceptional quality writing from proficient writing?
2. How do the linguistic features of exceptional writing contribute to the quality of a text?

2. Methodology

2.1. Data collection

The main objective in our data collection was to find accessible texts that are directly comparable (that is, texts written on the same topics for the same purpose) from both exceptional writers and proficient writers. However, before we can make determinations on writing quality, we must acknowledge the subjectivity of a judgement of writing quality. Different readers with different purposes will make different judgements of the quality of a text. In the few studies which make similar comparisons, such as those by Louis and Nenkova (2011, 2013a, 2013b), the data consists of professional journalism texts with the exceptional texts being selected by expert journalists. In our study, the texts were also gathered from the journalism field, but in order to compare texts written on the same topic, we selected Internet-published movie reviews of the same movies. In this genre of movie reviews, recognition from the Pulitzer Prize Board is a socially accepted indicator of exceptional quality. Therefore, Pulitzer Prize winners and finalists can be considered as being able to produce exceptional texts. Our proficient writing, on the other hand, was movie reviews which were authored by bloggers who posted their texts on their own personal blogs on the Internet.

Our data selection was undertaken with the assumption that the Pulitzer Prize winners and finalists write exceptionally well in every one of their texts, while the bloggers were at least proficient enough to have a large number of readers, as evidenced by high Google search rankings. To find exceptionally well-written movie reviews, we started with the lists of Pulitzer Prize winners and finalists on the Pulitzer Prize website (The Pulitzer Prizes n.d.). Pulitzer Prizes for movie reviews are given under the Journalism category, in the sub-category of Criticism. This sub-category includes critics who cover art, architecture, books, dance, fashion, movies and music. Since 1973, only five movie reviewers have been honored with a Pulitzer Prize. Of these, two were still frequently publishing movie reviews when the data was being gathered: Joe Morgenstern of the Wall Street Journal (2005 Winner) and Wesley Morris of the Boston Globe (2012 Winner). In order to gather additional data for this study, Pulitzer Prize finalists were also considered. These included Ann Hornaday of the Washington Post (2008 Finalist) and Manohla Dargis of the New York Times (2013, 2015 and 2016 Finalist).

To find proficient movie reviews written by bloggers, we started with websites which host a directory of movie reviews. We selected movie review blogs which had only one author, reviewed current Hollywood movies (as opposed to foreign films, classic films or niche genres such as horror), contained multiple recent reviews of the same movies as our four exceptional writers above, and had a relatively high Google Page Rank. These criteria allowed us to find active movie review sites which have a relatively large following and could therefore be assumed to be written proficiently. The top four movie review bloggers based on Google Page Rank who met all these criteria were chosen for our study in order to match the number of exceptional writers.

As we were collecting the movie review data, we observed that Pulitzer Prize winners and finalists were more likely to review smaller, artistic films, while the bloggers were more likely to write about films that were more popular such as Hollywood blockbusters. Nevertheless, we were able to find movie reviews for a few movies that had reviews written by both sets of writers. Because of this difference in preferences for which movies to review, we chose to analyze movie reviews from one artistic film and one blockbuster. Representing smaller, independent films we selected reviews of Wes Anderson’s *Moonrise Kingdom*, while the Hollywood blockbusters were represented by reviews of *Mission: Impossible - Ghost Protocol*, starring Tom Cruise. Our corpus included four reviews for each movie from each of the two sets of writers, for a total of 16 movie reviews. The small size of this corpus allowed for a manual interpretive analysis of lexical cohesion as described below, but we believe it was still large enough to allow for meaningful quantitative results. Details of the size of the corpora can be seen in Table 1.

Only the text bodies of the movie reviews were considered for this study. The titles of the movie reviews were not considered, nor were any other text, image, or media that might have been found on the Internet webpage along with the review. There was no further editing of the movie review texts, such as changes to capitalization, punctuation, spelling, or grammar.

	Number of Texts	Total corpus size	Min Text Length	Max Text Length
Proficient Writing	4 reviews each of two movies	5443 words	336 words	1649 words
Exceptional Writing	4 reviews each of two movies	7408 words	345 words	1451 words

Table 1: Movie review corpora sizes

2.2. Data analysis

When the Pulitzer Prize committee confers writers with the honor of being named a Pulitzer Prize finalist or a winner, the award is a holistic rating of excellence in writing quality. The current study uses this criterion to separate our movie reviews into exceptional writing (prize winners and finalists) and proficient writing (bloggers). The texts were analyzed for linguistic features associated with syntax, lexis, and cohesion using metrics from previous research in writing quality and writing development focusing on student writing. One of the exceptional texts was then selected for an interpretive analysis of the linguistic features which were found to be representative of exceptional writing.

2.2.1. Syntactic complexity

The linguistic construct of syntactic complexity plays an important role in the measurement of language development and learners' proficiency and flexibility in achieving various communicative goals (Ortega 2003, 2015). It has been found that longer clauses, T-units, and sentences, as well as a higher proportion of complex sentences (Lu 2010; Bulté & Housen 2014), the number of words before the main verb in a sentence (McNamara, et al., 2010), and the number of modifiers per noun phrase (Crossley, Weston, McLain Sullivan & McNamara 2011) have shown correlations with writing quality.

However, several meta-analyses of research into syntactic complexity have shown mixed results with respect to writing quality, even when looking at the same metrics such as T-unit and clause lengths (Crowhurst 1983; Wolfe-Quintero, Inagaki & Kim 1998). In addition, while there is some mixed evidence that syntactic complexity increases as learners become more proficient writers, the trend might not continue into exceptional writing, and shorter, less complex sentences might be more common in exceptional writing (Louis & Nenkova 2013).

For this study, the syntactic complexity of the texts was measured across seventeen different metrics which were divided into four sub-categories: length of production unit, sentence complexity, coordination and subordination, and particular structures. Fourteen of the syntactic complexity measurements were analyzed using the L2 Syntactic Complexity Analyzer (Lu 2010), which is a tool based on results from meta-analyses by Wolfe-Quintero, et al. (1998) and Ortega (2003). Coh-Metrix (Graesser, McNamara, Louwerse & Cai 2004) was used to analyze an additional three syntactic complexity measurements: the number of words appearing before the main verb of the main clause in the sentences of a text, the mean number of modifiers per noun phrase, and a score for sentence syntax similarity.

2.2.2. Lexical complexity

In writing development studies, it has been shown that there is a correlation between the lexical complexity of a text and the perceived quality of that text. Numerous studies into the lexis of student writing have focused on what Read (2000) described as lexical richness, which can be subdivided into density, sophistication, and diversity. The first subdivision, lexical density, is a measurement of the number of content words in a text as a percentage of the total number of words in the text. However, this does not seem to correlate with writing quality (Engber 1995; Lu 2012).

The second subdivision, lexical sophistication (sometimes called lexical rarity), is related to the use of unusual, rare, or advanced words in a text. For example, as learners become more proficient in their writing, they tend to use longer words (Ferris 1994; Grant & Ginther 2000). Psycholinguistic databases can also shed light on the issue of lexical sophistication of a text. Decreasing meaningfulness and familiarity (Crossley & McNamara 2012) and increasing concreteness and polysemy (Crossley et al. 2011) have been correlated with increased writing

quality. Word imageability is another psycholinguistic metric that has been investigated, and in Louis and Nenkova's (2013b) study on exceptional writing, it was shown that exceptional writing had high imageability in the beginnings and the ends of the articles and that the visual words in the articles were connected to a smaller number of topics than in typical writing.

The third and final subdivision of lexical richness is lexical diversity (sometimes termed lexical variation). Lexical diversity metrics such as type-token ratios and other related metrics show the range of the different words that are used. These metrics have also been related to writing quality, as higher lexical diversity is more likely to be found in higher quality writing (Engber 1995; McNamara et al. 2010).

The current study investigated a total of 32 lexical complexity measurements. The single lexical density measurement and the 19 lexical diversity metrics were computed using the Lexical Complexity Analyzer (Lu 2012), while the twelve metrics related to lexical sophistication were computed using a combination of the Lexical Complexity Analyzer and Coh-Metrix (Crossley & McNamara 2011).

2.2.3. Lexical cohesion

The third area which has been investigated for its relationship to human judgements of writing quality is lexical cohesion. Lexical cohesion is the connection of concepts in a text and can be created through word repetition or reference (usually called referential cohesion), through pairs of synonyms, hyponyms, or other logical relationships (Cruse 2011), or through semantic or conceptual associations (Hoey 2005). Cohesion is related to writing quality because it allows the writer to link various topics and ideas into a cohesive whole. Most of the recent studies on the correlation between cohesion and writing quality have focused only on referential cohesion. These studies use automated methods, as referential cohesion is the type of cohesion that is most easily identified by computers. Some studies have shown that more advanced L2 students produce more connectives than intermediate students (Connor 1990; Jin 2001). But other studies claim that at the highest levels of proficiency, the amount of cohesion in student texts is actually lower than that found in texts from students with lower levels of proficiency (Crossley & McNamara 2011).

Many of the studies linking referential cohesion to writing quality use Coh-Metrix to conduct the analysis, since it was originally designed to measure the referential cohesion of a text (Graesser, McNamara, Louwerse & Cai 2004). Coh-Metrix has several indices to accomplish this, including noun overlap, argument overlap, stem overlap and content word overlap. Each of these four indices are measured locally on adjacent sentences as well as globally across all sentences. It is important to note, however, that these are all referential cohesion metrics that look for repetition, matching lemmas, and pronoun references. They do not measure cohesion in a text created by logical relations such as synonyms and hypernyms, or the conceptual semantic associations between

words. Coh-Metrix also does not report which words it has identified as being linked by referential cohesion, and therefore does not give insight into how referential cohesion is being used in the text.

To our knowledge, there is no automated tool available to identify and classify all three types of relationships (referential cohesion, logical relations, and conceptual associations) in a text. Therefore, to gain a more complete picture of the cohesion in our texts, we conducted an additional manual analysis of the three types of cohesion. To accomplish this analysis, first, each text was separated into clauses. Threads of cohesive topics and the words in those threads in different clauses were identified. Each pair of cohesive words was classified as one of the three types of cohesion, and a percentage of the total number of cohesive links was calculated for each of the three types.

2.3. Automation and correlations

The analysis of our movie reviews on syntactic complexity, lexical complexity, and cohesion was mostly conducted through automated methods, using the three web-based tools of the L2 Syntactic Complexity Analyzer (Lu 2010), the Lexical Complexity Analyzer (Lu 2012) and Coh-Metrix (Graesser, McNamara, Louwerse & Cai 2004). In addition, a manual analysis of the three types of lexical cohesion (referential cohesion, logical relations, and conceptual associations) was conducted. In the analysis, each of our texts was given a quantitative score for each linguistic feature. A correlation was then computed between the holistic judgement of the quality of the text (the separation of writers selected by the Pulitzer Prize committee and bloggers) with the analytic metrics in syntax, lexis, and cohesion. This correlation was a comparison of one dichotomous variable (the holistic judgement) and many continuous variables (each of the linguistic feature metrics), so therefore, point biserial correlation (r_{pb}) was used. A metric with a positive point biserial correlation indicates higher values for that metric in the texts written by Pulitzer Prize winners while a negative point biserial correlation indicates higher values in the texts written by bloggers. A significance level of $p < 0.01$ was used due to the low number N as well as the high number of metrics being tested. An effect size for each metric based on the r_{pb} was also computed (following McGrath & Meyer 2006), with cutoffs of 0.10 for small effect sizes, .24 for medium effect sizes, and .37 for large effect sizes.

3. Results

To illustrate how the linguistic features can be identified in a text, it might be useful to compare some surface linguistic features from the beginning excerpt of a movie review written by a Pulitzer Prize winner to the beginning excerpt written by a blogger. Both excerpts are the opening few sentences of movie reviews of *Moonrise Kingdom*. The first excerpt was written by Pulitzer Prize finalist Ann Hornaday and published in the Washington Post.

“Moonrise Kingdom” opens with no music - just the sound of raindrops falling on the roof of a preternaturally cozy house, which the camera gently leads the audience through as the family members inside go about their rainy-day business.

Bathed in apple reds, egg-yolk yellows and an air of studied eccentricity, the house is immediately recognizable as yet another habitat created by Wes Anderson, a film director whose obsession with material culture, nostalgia and nursery comforts borders on the fetishistic.

Of course, for viewers who happen to share Anderson’s taste for boldly framed, bespoke productions - in which everything looks (and most probably is) lovingly handmade and artisanal, “Moonrise Kingdom” will simply offer yet another chance to live, at least for a little while, in the kind of universe only Anderson can create. (You can almost smell the damp canvas and wood polish in that opening sequence.) Those who long ago wrote off the writer-director as insufferably mannered and arcane - the usual term of art is “twee” - well, they’re welcome to stay out in the rain.

That opening-scene house has a name, by the way: Summer’s End, which turns out to aptly capture a vaguely autumnal tale of young love that takes place in early September 1965 - a time of Ford Falcons and mothers who smoked. (Hornaday 2012)

An analysis of this short text of 218 words shows that the syntax is quite complex with long sentences and many clauses in each sentence. The average sentence length of this short text is 36.1 words with each sentence containing an average of 2.8 clauses. Lexical complexity is also high. This short text has eight words which are not in the top 10,000 most frequent English words, when compared to the British National Corpus: *arcane*, *bespoke*, *fetishistic*, *handmade*, *insufferably*, *preternaturally*, *raindrops*, and *twee*. The type-token ratio is also high at .71, meaning that there is very little repetition in this text. (Only three content words occur more than once.) This lack of lexical repetition combined with very few pronouns also means that the referential cohesion in this text is low. However, the number of conceptual associations is very high with several lexical chains covering different topics such as the rain (*raindrops – rainy-day – rain*), the cozy family house (*habitat – house – cozy – comfort – nursery – family – mothers – inside*), and handcrafted art (*handmade – artisanal – art – create – bespoke – production*)

The second excerpt is from a movie review blog site called “Cane Toad Warrior” hosted on Blogspot. The author is listed as “Luke”. The opening few sentences of this review of *Moonrise Kingdom* are given here.

I've seen all of Wes Anderson's films but I usually don't review them because I find it hard to translate their cinematic quirkiness into words. He's a very filmish filmmaker... not reliant on just visuals and not reliant on just words, he spins them both into something that really represents the full potential of the medium. He knows the power of film, and he's a true modern auteur.

Moonrise Kingdom is no exception to this rule, it's such a vivid piece of filmmaking and one of the best films of 2012. It features a whole host of techniques representative of Anderson's genius - signature pans, an unusual colour palette, split screens that linger long after phonecalls finish, a binocular-vision motif, the use of unusual music. Most of all I love the fact that the narrative seems to start halfway through the story. The great writer Kurt Vonnegut once said that you should start as late in a story as possible, and Anderson obviously buys into similar sentiments. When we join Moonrise Kingdom's young adolescent protagonists on their journey they've already gotten to know each other, and Anderson only shows us the beginnings of their relationship later in the film. (Luke 2012)

An analysis of this text excerpt of 208 words shows that the syntax is less complex than the first text, with shorter sentences and fewer clauses in each sentence. The average sentence length of this excerpt is 26 words and each sentence contains an average of 2.5 clauses. Lexical complexity is also lower than the first text. There are six words which are not in the top 10,000 most frequent English words, when compared to the British National Corpus: *auteur*, *filmish*, *filmmaker*, *filmmaking*, *halfway*, *phonecalls*, with two of these not found in the 25,000 most common words (*filmish* and *phonecalls*). The type-token ratio is .67, meaning that there is more lexical repetition in this text than in the first one. (Eleven content words occur more than once.) Pronouns also play a larger role, with 7 pronouns used 18 times. The repetition in content words and pronouns means that the referential cohesion in this text is much higher than in the first text. The number of conceptual associations is relatively low in this case, with the main lexical chain being made up of movie terms such as *film*, *cinematic*, *narrative*, and *protagonists*.

From this introductory analysis, we can see that the exceptional writing is more syntactically complex than the proficient writing with longer sentences and more clauses per sentence. The exceptional writing is also more lexically complex with more rare words and a higher type-token ratio. In addition, the cohesive links in exceptional writing are less likely to be classified as referential cohesion and more likely to be classified as conceptual associations, with there being several different conceptually linked threads containing several words each. The remainder of the results section below details a similar analysis comparing all of the texts in the two corpora across the three aspects of syntax, lexis, and cohesion.

3.1. Syntactic complexity

Syntactic complexity was one of the categories of linguistic features which was analyzed in this study in order to find differences between movie review texts written by Pulitzer Prize winners and those written by bloggers. We have further subdivided this category into four sub-categories: length of production unit, sentence complexity, coordination and subordination, and particular structures. The results for the syntactic complexity metric correlations are shown in Table 2.

Sub-Category	Metric	Tool	r_{pb}	Effect Size
Length of Production Unit	Mean Length of Sentences	L2SCA	0.24	Medium
	Mean Length of T-Units	L2SCA	0.34	Medium
	Mean Length of Clauses	L2SCA	0.12	Small
Sentence Complexity	Clauses per Sentence	L2SCA	0.15	Small
	Sentence Syntax Similarity	Coh-Metrix	0.34	Medium
Coordination And Subordination	Clauses per T-Unit	L2SCA	0.35	Medium
	Dependent Clauses per Clause	L2SCA	0.46	Large
	Dependent Clauses per T-Unit	L2SCA	0.43	Large
	Complex T-Unit Ratio	L2SCA	0.34	Medium
	T-Units per Sentences	L2SCA	-0.34	Medium
	Coordinate Phrases per T-Unit	L2SCA	0.39	Large
	Coordinate Phrases per Clause	L2SCA	0.20	Small
Particular Structures	Complex Nominals per T-Unit	L2SCA	0.36	Medium
	Complex Nominals per Clause	L2SCA	0.18	Small
	Verb Phrases per T-Unit	L2SCA	0.17	Small
	Number of Words Before Main Verb	Coh-Metrix	-0.54	Large
	Number of Modifiers per Noun Phrase	Coh-Metrix	-0.12	Small

Table 2: Point biserial correlations for syntactic complexity metrics

As can be seen in Table 2, most of the metrics for syntactic complexity are positive, meaning that the texts written by Pulitzer Prize winners are more syntactically complex than those written by

bloggers. However, none of these metrics are significant at a value of $p < 0.01$. Four of the seventeen metrics have large effect sizes, namely those concerning dependent clauses, coordinate phrases per t-unit, and the number of words before the main verb. This means that, in general, very few syntactic complexity metrics are able to successfully differentiate between exceptional and proficient texts.

3.2. Lexical complexity

The second category of features which was analyzed to compare the reviews of Pulitzer Prize winners and bloggers was lexical complexity, which was further divided into three sub-categories: lexical density, lexical sophistication/rarity, and lexical diversity. The results for the lexical complexity metric correlations are shown in Table 3.

Sub-Category	Metric	Analysis Tool	r_{pb}	Effect Size
Lexical Density	Lexical Density	LCA	-0.19	Small
Lexical Sophistication/ Rarity	Lexical Sophistication I	LCA	0.53	Large
	Lexical Sophistication II	LCA	0.62 *	Large
	Verb Sophistication I	LCA	0.52	Large
	Verb Sophistication II	LCA	0.74 *	Large
	Corr. Verb Sophistication I	LCA	0.72 *	Large
	Polysemy	Coh-Metrix	-0.73 *	Large
	Familiarity	Coh-Metrix	-0.27	Medium
	Concreteness	Coh-Metrix	0.40	Large
	Imageability	Coh-Metrix	0.43	Large
	Meaningfulness	Coh-Metrix	0.06	None
	Word length by letter	Coh-Metrix	-0.02	None
	Word length by syllable	Coh-Metrix	0.05	None
Lexical Diversity	# of different words (NDW)	LCA	0.46	Large
	NDW (first 50)	LCA	0.56	Large
	NDW Expected random 50	LCA	0.29	Medium
	NDW Expected Sequence 50	LCA	0.15	Small
	Type-Token Ratio (TTR)	LCA	-0.14	Small
	Mean Segmental TTR (50)	LCA	0.16	Small
	Corrected TTR	LCA	0.59	Large
	Root TTR	LCA	0.59	Large
Bilogarithmic TTR	LCA	0.09	Small	

Uber Index	LCA	0.37	Large
Verb Variation (VV)	LCA	0.06	None
Squared VV1	LCA	0.61	Large
Corrected VV1	LCA	0.60	Large
Lexical Word Variation	LCA	0.08	None
Verb Variation 2	LCA	0.32	Medium
Noun Variation	LCA	0.11	Small
Adjective Variation	LCA	-0.24	Medium
Adverb Variation	LCA	0.22	Small
Modifier Variation	LCA	-0.09	None

* Significant at $p < 0.01$

Table 3: Point biserial correlations for lexical complexity metrics

From Table 3, we can see that four of the lexical sophistication features show a statistically significant difference between the two sets of reviews. The results for the remaining metrics are not significant. The positive significant correlations for lexical sophistication suggest that the Pulitzer Prize winners tend to use rare words and the negative significant correlation for polysemy suggests that blog reviews tend to use more common words (since more common words are more likely to be polysemous). These significant findings, along with several large effect sizes for several lexical diversity metrics, suggest that Pulitzer Prize winners use rarer words and a greater variety of words.

3.3. Lexical cohesion

The third and final area of investigation into the differences in linguistic features between our two sets of texts was lexical cohesion. The two sub-categories of lexical cohesion for this study are referential cohesion metrics and ratios of different types of connectedness. The results for the lexical cohesion metric correlations are shown in Table 4.

Sub-Category	Metric	Analysis Tool	r_{pb}	Effect Size
Referential Cohesion	Noun Overlap, adjacent sentences	Coh-Metrix	-0.36	Medium
	Argument Overlap, adjacent sentences	Coh-Metrix	-0.38	Large
	Stem Overlap, adjacent sentences	Coh-Metrix	-0.40	Large
	Content Word Overlap, adjacent sentences	Coh-Metrix	-0.47	Large

	Noun Overlap, all sentences	Coh-Metrix	-0.57	Large
	Argument Overlap, all sentences	Coh-Metrix	-0.51	Large
	Stem Overlap, all sentences	Coh-Metrix	-0.62 *	Large
	Content Word Overlap, all sentences	Coh-Metrix	-0.61	Large
Ratios of types of connectedness	% of Referential Cohesion	Manual	-0.68 *	Large
	% of Logical Relations	Manual	-0.15	Small
	% of Conceptual Associations	Manual	0.80 *	Large

* Significant at $p < 0.01$

Table 4: Point biserial correlations for lexical cohesion metrics

Table 4 shows that all of the referential cohesion metrics have negative correlations, meaning that the texts in our study written by bloggers have more referential cohesion than those written by Pulitzer Prize winners. However, only one of these metrics (stem overlap for all sentences) was significant at $p < 0.01$. For the second sub-category of the ratios of different types of connectedness, there is a strong negative correlation for referential cohesion and a strong positive correlation for conceptual associations. This means that in order to write cohesive texts where ideas and topics are connected, bloggers rely more on referential cohesion such as repetition and pronouns while Pulitzer Prize winners are more likely to use conceptual associations. The 0.80 correlation value for conceptual associations shown in Table 3 is the strongest point biserial correlation of any metric investigated for this study. The lexical cohesion category also has the largest percentage of metrics with large effect sizes. Nine of the eleven cohesion metrics have a large effect size suggesting that the type of cohesion that writers use is an important factor in the perceived quality of their texts.

4. Discussion

Many previous studies have investigated the topic of writing quality, either from a writing development and proficiency perspective, or from the perspective of exceptional literary writing. The current study attempted to bridge the gap between these two perspectives in order to determine if the trends seen in writing development studies extend to exceptional writing, and to determine which linguistic features (if any) contribute to the perceived differences between proficient and exceptional writing. This was achieved by comparing linguistic features in the exceptional writing of Pulitzer Prize winners and finalists to those same linguistic features in the proficient writing of bloggers. Three main categories of linguistic features were analyzed: syntactic complexity, lexical complexity, and lexical cohesion.

For the first category of syntactic complexity, studies have shown that as students develop their writing skills, their writing becomes more syntactically complex (Bulté & Housen 2014). Other studies, however, show mixed results about the role of syntactic complexity in relation to writing quality, both for high proficiency writing (Wolfe-Quintero et al. 1998) and exceptional writing (Louis & Nenkova 2013a, 2013b). In our study, none of the point biserial correlations for syntactic complexity metrics were significant, and very few metrics showed large effect sizes. This indicates that the trend of increasing syntactic complexity seen in writing development studies does not carry on to exceptional writing. In other words, even though on average, exceptional writing is more syntactically complex than proficient writing, the lack of statistically significant results appears to show that syntactic complexity does not play a large role in the perception of writing quality at high levels.

Research into the second category, lexical complexity, and its relationship to writing quality has shown that lexical sophistication and lexical diversity both increase as writers become more proficient (Engber 1995; McNamara et al. 2010). Our study suggests that this trend continues into exceptional writing as well, especially for lexical sophistication. Several of our lexical sophistication metrics returned significant results, and almost half of the metrics for lexical sophistication and lexical diversity showed large effect sizes. This would imply that lexical complexity could potentially be used to distinguish between proficient and exceptional writing.

The third category of linguistic features associated with lexical cohesion provided perhaps the most interesting results of this study. Many of the lexical cohesion metrics returned significant results, and almost all of the metrics showed large effect sizes. The results show that referential cohesion decreases as writing quality increases from proficient to exceptional, continuing the trend seen in writing development studies. However, texts require a certain amount of cohesion to be comprehensible, and our results show that writers of exceptional texts use other techniques to create the connections to allow the necessary inferences to be made. As referential cohesion decreases, the number of conceptual associations increases and takes the place of the referential cohesion. To our knowledge, there have not been any studies which have investigated conceptual associations as a linguistic feature related to writing quality, but our strong results show that perhaps this is a promising area to investigate further.

4.1. Interpretative analysis of linguistic features

So far in this paper, we have been reporting and discussing the quantitative results from our study. However, quantitative metrics such as the ones above only tell part of the story. Such statistics do not show how the various metrics actually contribute to the creation of quality in a text. Therefore, this section will discuss how the relative paucity of referential cohesion and the extensive use of lexical sophistication and conceptual associations might contribute to a text being perceived as

exceptional. To do this, we return to Hornaday's review of *Moonrise Kingdom*, from which an excerpt was shown earlier, and analyze it qualitatively.

We have seen that at the level of exceptional writing, lexical cohesion is likely to be manifest through conceptual associations rather than through the references of repetition and anaphora commonly used at lower levels of quality. Given that referential cohesion is less favored at exceptional levels, we will start by examining those few clear instances of it within the review of *Moonrise Kingdom* as these may indicate aspects of writing where referential cohesion cannot be replaced. All of the referential cohesion in the excerpt given above is related to just three concepts: the title and director of the movie (*Moonrise Kingdom* and *Wes Anderson*), and the opening scene *house*. The first case of the repetition of the title *Moonrise Kingdom* helps to fulfill the requirements of the genre of movie reviews. The other two threads of referential cohesion show that the main purpose of this text is to discuss the artistic directing style of *Wes Anderson* as instantiated by the style of the *house*.

As discussed above, the linguistic feature which gave the strongest correlation results was conceptual associations. To evaluate this linguistic feature qualitatively, we first note how the author of this text weaves several threads of conceptual associations through the excerpt, such as the following:

- *home – habitat – house – cozy – comfort – inside*
- *family – mothers – nursery*
- *raindrops – rainy-day – rain*
- *handmade – artisanal – art – create – bespoke – production*
- *reds – yellows – autumnal – September*
- *taste – looks – smell – sound*
- *apple – egg-yolk*

This text is, of course, a movie review, and so there are also several associations fulfilling the purposes of the movie review genre: *camera – audience – viewers – film – framed – sequence – scene – tale – director* as well as some descriptive adjectives describing the well-known director (*Wes Anderson*) and his style: *eccentricity – mannered – arcane – fetishistic – twee*.

These conceptual association threads contribute to the quality of the writing in three ways. First, long threads of related words create a detailed atmosphere in the writing. In the case of the “home” and “family” threads, the reader can easily imagine a feeling of protection afforded by a *nursery* and the comfort that the *cozy* house and the *family* provides. Secondly, each thread contains associated words within the thread, but the words themselves are not always referring to the same situation or the same object. In the “rain” thread, the first mention of *raindrops* is talking about the weather outside the house, while the final *rain* is talking about moviegoers who do not like *Wes*

Anderson films (*they're welcome to stay out in the rain*). In this way, associated words are being used to serve different themes that are found throughout the review. Thirdly, there are cases where some words have associations with more than one thread. The concept of *damp canvas* has a clear link to the “arts and crafts” thread, but it also has a more subtle link to the idea of wetness, which relates to the rainy weather. Another subtle link can be found between the colors and wetness provided by the word *bathed*. In the text, the house is *bathed* in reds and yellows, but it is also being bathed by the rain at the same time. All of these inter-linked conceptual associations contribute to the exceptional quality of this text by creating a highly coherent and atmospheric feel.

Lexical complexity also plays a role in differentiating exceptional texts from proficient texts. Both the use of rare words and a high diversity of words used can provide the reader with a deeper and more complex understanding of the ideas that the author is presenting. For example, both the excerpt of the proficient writing and the exceptional writing discuss the unique style of the director of the movie, but the exceptional text provides much more detail than the proficient text. For example, the concept *visuals* in the proficient text becomes *looks... lovingly handmade and artisanal* in the exceptional text, while *quirkiness* in the proficient text is expanded into *insufferably mannered and arcane* in the exceptional text. If we expand this analysis to the entire texts, even more examples can be found. The movie's *story* or *narrative* in the proficient text is described as *an alternately soothing and unsettling mythology of young love* in the exceptional text, while the *young adolescent protagonists* in the proficient text are described together as *supremely solemn and earnest adolescent protagonists* in the exceptional text.

In other words, the linguistic features of lexical complexity and conceptual associations work together to provide the reader of an exceptional text with the nuanced concepts that the author is presenting. Many of the words highlighted in the conceptual association threads (*cozy, polish, eccentricity, arcane, artisanal, bespoke, etc.*) are relatively rare words. In addition, the exceptional text often expands on concepts found in the proficient text by using specific, rare words. These linguistic features combine in the exceptional text above to not only convey the key idea that the text discusses the artistic style of Wes Anderson, but to also to give an evocative and rich meaning to both the description of the house on a rainy day in autumn in the opening scene, as well as a qualitative assessment of the art and craft of the film and the film director himself. Similar results can be seen in the exceptional movie reviews of the blockbuster *Ghost Protocol*. However, many of the key conceptual association threads and the lexical complexity in those movie reviews were focused on an aging Tom Cruise playing the role of a globe-trotting, athletic super-spy, rather than focusing on artistry and symbolism as seen in the reviews of *Moonrise Kingdom*. Therefore, while the genre and subject of the movie seem to affect the topics of the lexical complexity and the conceptual association threads, exceptional texts appear to employ more of each when compared to proficient texts regardless of movie genre.

The remaining linguistic features investigated in this study, such as those studied in previous research on syntactic complexity, do not appear to show differences between proficient and exceptional writing. Previous research on student writing has shown that there is a difference in levels of syntactic complexity between less proficient and more proficient writers. However, as writers reach proficiency, there appears to be a plateau in the amount of complexity in their texts since there is no significant difference in these linguistic features between proficient and exceptional writing. This is perhaps unsurprising, as there is a limit to how complex a text can be before it becomes too difficult to read. In the case of lexical complexity and conceptual association, however, increases continued from the level of proficient writing to that of exceptional writing. We would expect that, at some point, there is also a plateau in the amount of these features as well. If all of the lexis in a text is highly infrequent and complex, very few people would understand it. Similarly, increasing the amount of conceptual association might also have limits. In other words, syntactic complexity, lexical complexity, and lexical association all show increases in texts as the quality of the texts increases, but they all appear to reach plateaus at different levels of quality. In our case, syntactic complexity plateaued at the proficient level, while lexical complexity and conceptual association were still relevant for determining the difference between proficient and exceptional writing.

In this study, we started from a human judgement of exceptional writing and sought to determine how exceptional writing might differ from proficient writing based on the linguistic features found in each corpus of texts. The quantitative and interpretive results suggest that both lexical complexity and conceptual associations contribute to the differentiation between proficient writing and exceptional writing. There are, however, other potential causes of the differences in linguistic features which might be investigated further in future research. As noted in the methodology section, the Pulitzer Prize winners and finalists were more likely to review artistic, independent movies, while the bloggers were more likely to review popular blockbusters. We attempted to mitigate genre effects by analyzing reviews of the same two movies, but further research may uncover other differences in linguistic features due to differences in movie genres or differences in the perceived quality of the movie by the reviewers. For example, the polarity of the movie review (that is, whether the review was positive or negative) could potentially influence the linguistic features found in the review. Also, there may be differences in the writers' background and experience. Bloggers may feel a need to legitimize themselves and their expertise, and this may affect the number of personal pronouns and referential cohesion in the text. Additional insights might also be gained by conducting a multifactorial analysis on these two corpora to quantitatively determine how these linguistic features might be working together to influence human judgements of writing quality.

Our study suggests that if writers who are able to follow a standard model of proficient writing would like to improve their writing further, then perhaps the areas of lexical complexity and conceptual association should be taken under consideration. The ability to write a text with high

lexical sophistication requires access to a large mental lexicon, which might be built by extensive reading along with an appreciation of how the judicious use of sophisticated words might allow the text to be precise and nuanced. Writers should also be aware of the lexical cohesion in their texts, and whether or not this cohesion is created by simple reference and repetition or created by complex conceptual associations. But perhaps more importantly, this study shows that not only do linguistic features such as lexical complexity and lexical cohesion play a role as separate constructs in the judgement of quality as seen in the quantitative results, but that they are related to each other and support each other as seen in the qualitative results. Writers wishing to write exceptionally well should be aware that the sophisticated words in a text should be conceptually associated with both the main themes of the text as well as to multiple conceptual association threads, and through these threads, they should connect to other sophisticated lexis found in the text. In this way, writers can create coherent texts containing detailed, rich content which can help to achieve the goal of making an impact on the reader.

References

- Beers, Scott F. & William E. Nagy. 2009. Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing* 22 (2): 185-200.
- Berninger, Virginia W., Robert D. Abbott, William Nagy & Joanne Carlisle. 2010. Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research* 39 (2): 141-163.
- Bulté, Bram & Alex Housen. 2014. Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing* 26: 42-65.
- Connor, Ulla. 1990. Linguistic/rhetorical measures for international persuasive student writing. *Research in the Teaching of English*: 67-87.
- Connor-Linton, Jeff & Charlene Polio. 2014. Comparing perspectives on L2 writing: Multiple analyses of a common corpus. *Journal of Second Language Writing* 26: 1-9.
- Crowhurst, Marion. 1983. Syntactic complexity and writing quality: A review. *Canadian Journal of Education/Revue Canadienne de L'éducation*: 1-16.
- Crossley, Scott A. & Danielle S. McNamara. 2010. Cohesion, coherence, and expert evaluations of writing proficiency. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, Stellan Ohlsson & Richard Catrambone (eds). Cognitive Science Society, 984-989.

- Crossley, Scott A. & Danielle S. McNamara. 2011. Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life-Long Learning* 21 (2): 170-191.
- Crossley, Scott A. & Danielle S. McNamara. 2012. Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading* 35 (2): 115-135.
- Crossley, Scott A., Jennifer L. Weston, Susan T. McLain & Danielle S. McNamara. 2011. The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication* 28: 282-311.
- Cruse, Alan. 2011. *Meaning in language: An introduction to semantics and pragmatics*. Oxford UK: Oxford University.
- Culpeper, Jonathan. 2002. Computers, language and characterisation: An Analysis of six characters in Romeo and Juliet. In *Conversation in Life and in Literature: Papers from the ASLA Symposium, Association Suedoise de Linguistique Appliquee (ASLA) 15*, Ulla Melander-Marttala, Carin Ostman & Merja Kyto (eds). Uppsala, Sweden: Universitetstryckeriet, 11-30.
- Engber, Cheryl A. 1995. The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing* 4 (2): 139-155.
- Ferris, Dana R. 1994. Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly* 28 (2): 414-420.
- Fischer-Starcke, Bettina. 2010. *Corpus linguistics in literary analysis: Jane Austen and her contemporaries*. London and New York: Continuum.
- Graesser, Arthur C., Danielle S. McNamara, Max M. Louwerse & Zhiqiang Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers* 36: 193-202.
- Grant, Leslie & April Ginther. 2000. Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing* 9 (2): 123-145.
- Hoey, Michael. 2005. *Lexical priming: A new theory of words and language*. London and New York: Routledge.
- Hornaday, Ann. 2012. Adolescent love among eccentrics [Web blog post]. <<http://www.washingtonpost.com/go/g/movies/moonrise-kingdom.1221101.html>> (Last accessed on 17 November 2013).

- Jin, Wenjun. 2001. A quantitative study of cohesion in Chinese graduate students' writing: Variations across genres and proficiency levels. Paper presented at the Symposium on Second Language Writing, Purdue University, West Lafayette, Indiana, USA, 15-16 September, 2000.
- Louis, Annie & Ani Nenkova. 2011. Automatic identification of general and specific sentences by leveraging discourse annotations. *Proceedings of 5th International Joint Conference on Natural Language Processing*. Thailand: Asian Foundation of Natural Language Processing, 605-613.
- Louis, Annie & Ani Nenkova. 2013a. A corpus of science journalism for analyzing writing quality. *Dialogue and Discourse* 4 (2): 87-117.
- Louis, Annie & Ani Nenkova. 2013b. What makes writing great? First experiments on article quality prediction in the science journalism domain. *Transactions of the Association for Computational Linguistics* 1: 341-352.
- Lu, Xiaofei. 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15: 474-496.
- Lu, Xiaofei. 2012. The relationship of lexical richness to the quality of ESL learners' oral narratives. *Modern Language Journal* 96 (2): 190-208.
- Luke. (2012). Moonrise Kingdom [Web blog post].
<<http://canetoadwarrior.blogspot.com/2012/10/moonrise-kingdom.html>> (Last accessed on 29 September 2013).
- Mahlberg, Michaela. 2007. Corpus stylistics: Bridging the gap between linguistic and literary studies. In *Text, discourse and corpora: Theory and analysis*, Michael Hoey, Michaela Mahlberg, Michael Stubbs & Wolfgang Teubert (eds). London UK: Continuum, 219-246.
- McCutchen, Deborah & Charles A. Perfetti. 1982. Coherence and connectedness in the development of discourse production. *Text* 2 (1-3): 113-140.
- McGrath, R. E., & Meyer, G. J. 2006. When effect sizes disagree: the case of r and d. *Psychological Methods* 11(4), 386-401.
- McNamara, Danielle. S., Scott A. Crossley & Philip M. McCarthy. 2010. Linguistic features of writing quality. *Written Communication* 27 (1): 57-86.
- Ortega, Lourdes. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24 (4): 492-518.

- Ortega, Lourdes. 2015. Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing* 29: 82-94.
- Read, John. 2000. *Assessing vocabulary*. Cambridge UK: Cambridge University Press.
- Shaughnessy, M. P. (1979). *Errors and expectations: A guide for the teacher of basic writing*. Oxford: Oxford University Press.
- Stubbs, Michael. 2005. Conrad in the computer: Examples of quantitative stylistic methods. *Language and Literature* 14 (1): 5-24.
- The Pulitzer Prizes: Honoring Excellence in Journalism and the Arts. (n.d).
<<http://www.pulitzer.org>> (Last accessed on 10 November 2013).
- van Peer, Willie (ed). 2008. *The quality of literature: Linguistic studies in literary evaluation* 4. Amsterdam, Netherlands: John Benjamins Publishing.
- Widdowson, Henry George. 1992. *Practical stylistics: An approach to poetry*. Oxford UK: Oxford University Press.
- Wolfe-Quintero, Kate, Shunji Inagaki & Hae-Young Kim. 1998. *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu: University of Hawai'i at Manoa.