

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/264372709>

TESOL Quartely_What Linguistic Features Are Indicative of Writing Quality? A Case of Argumentative Essays in a College Composition Program

Article in TESOL Quarterly · June 2013

DOI: 10.1002/tesq.91

CITATIONS

83

READS

1,130

3 authors:



Naoko Taguchi

Northern Arizona University

128 PUBLICATIONS 3,335 CITATIONS

[SEE PROFILE](#)



William J. Crawford

Northern Arizona University

20 PUBLICATIONS 516 CITATIONS

[SEE PROFILE](#)



Danielle Wetzel

Carnegie Mellon University

4 PUBLICATIONS 99 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Routledge Handbook of SLA and Pragmatics [View project](#)



Encyclopedia of Applied Linguistics _Pragmatics [View project](#)

Taguchi, N., Crowford, B., Wetzel, D. (2013b). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, 47, 420-430.

What Linguistic Features are Indicative of Writing Quality?

A Case of Argumentative Essays in a College Composition Program

Introduction

This preliminary study investigates linguistic features that distinguish more proficient from less proficient essays written by nonnative speakers of English. Previous studies have shown that certain linguistic features of writing help distinguish the performance of L2 writers with varying levels of proficiency (e.g., Becker, 2010; Cumming, Kantor, Baba, Erdosy, Eouanzoui, & James, 2005; Grant & Ginther, 2000; Hinkel, 2003; Lu, 2011). For example, Grant and Ginther (2000) analyzed general features (e.g., essay length, type/token ratio), lexical features (e.g., conjunctions), grammar (nominalizations, modals), and clause-level complexity (e.g., subordination, use of relative clauses) in 90 Test of Written English (TWE) essays at three proficiency levels (scores 3, 4, or 5). In this study, essays receiving a score of 5 had higher frequency for most of the features than essays with a score of 4 or 3. In another study, Becker (2010) analyzed 43 ESL student essays from three proficiency levels and found that two complexity measures, the number of clauses per t-unit and words per t-unit, distinguished performance between groups. However, a third syntactic complexity measure, number of t-units, did not distinguish written performance.

Among the features that are considered in development of L2 English writing, syntactic complexity, especially t-unit-based measures, are often used as an indicator of more proficient writing and have been accepted widely as a measure of complexity (Ortega, 2003; Wolfe-Quintero, Inagaki, & Kim, 1998). However, some studies have found that higher proficiency

learners do not produce longer t-units (e.g., Bardovi-Harlig, 1992; Smart & Crawford, 2009), leading to a conclusion that t-unit analysis is not reflective of complexity. Rimmer (2006) supports this position and suggests expanding complexity to include phrasal features such as post-noun modification. More recently, Biber, Gray, and Poonpon (2011) have argued that complexity assessed by the dependent clauses measures (t-unit analyses) is not a characteristic of academic writing. In fact, phrase-level complexity (e.g., attributive adjective + noun, prepositional phrase modifying noun) is much more common in academic texts than in conversations.

Several recent studies lend support to this claim. Lu (2011) analyzed a corpus of English essays written by L1 Chinese students and found that complexity measures which best discriminated between proficiency levels were the number of complex nominals per clause (e.g., adjective + noun, possessives, prepositional phrases), as well as the mean length of clause and sentence. McNamara, Crosseley, & McCarthy (2010) found that the best indicators of writing quality included features such as verb phrase complexity (# of words before main verb), lexical diversity, and use of less frequent words, again lending support to the claim that complexity measures beyond the clause-level are relevant indicators of writing proficiency in English.

While there has been some effort to develop a better understanding of the linguistic features that characterize proficient writing, results of previous studies are somewhat inconsistent, suggesting the need to investigate the linguistic-textual features that distinguish lower-level writers from more advanced writers with reference to a larger number of potentially relevant features. Such studies are important because they shed light on the development of L2 writing, which in turn will help to improve writing instruction and assessment practice. The present study is a preliminary effort in this direction and asks: Are there identifiable linguistic features

that are indicative of more proficient writing?

Methodology

This study used placement exam essays written by incoming international students in a private university in the mid-western United States. The participants were around 18 years old. Their native languages were Korean, Hindi, Chinese, Thai, Spanish, Russian, German, and French. About 80% of the students graduated from an English-medium high school. The exam was used to determine whether the students were eligible to take a regular, native-speaker freshman composition class or a nonnative speaker freshman composition class. All nonnative English-speaking students were required to take the exam on-line during the summer of 2011. They were first asked to choose one of two topics and read two texts presenting contrasting views on the topic. Then, they were asked to write an 850-1300 word essay that compared the two views and argued for one perspective over another.

Essays were rated by three native English speakers who were teaching composition classes at the university. Their teaching experiences ranged from three to five years. Scoring rubrics, adapted from Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey (1981), had five categories: language use (25 points), content (30 points), organization (20 points), vocabulary (20 points), and mechanics (5 points). Students with a score of 90 or above were placed into the native speaker version of the composition class; those with a score below 80 were automatically placed into the nonnative speaker composition class. Student essays with a score range of 80-90 were able to appeal for the placement decision if they wished and thus were excluded from the analysis. Inter-rater reliability was .93. Average scores were assigned as a final score.

A total of 116 essays were compiled. Because the purpose of the study is to investigate the relationship between placement decision and linguistics characteristics of the essays, we compared two groups of essays: essays with a score of above 90 ($n=30$) and essays with a score

of below 80 ($n=24$). The former was called high-rated essays and the latter was called low-rated essays.

High-rated essays had a total of 32,918 words, with a mean of 1097 words per essay, while low-rated essays had a total of 21,775 words with a mean of 837 words per essay. Students' L1 backgrounds of the high-rated essay group included Korean, Hindi, Chinese, and Spanish, while those of the low-rated essay group were Korean, Chinese, Spanish, and German. Students in the high-rated group started their formal English study at the average age of 4.57, while the average age of the low-rated group was 6.09. More than 85% of the high-rated group went to English-medium high schools, while the percentage was 77% for the low-rated group.

Data analysis procedures

To select measures of comparison between high- and low-rated essays, we focused on two categories of the scoring rubrics that were worth the most points: “language use” and “content.” We examined descriptions of these two categories to operationalize our measures. In the category of “language use,” an essay was judged excellent if it demonstrated facility in the use of effective, complex constructions, and few or no grammatical errors. Based on Biber et al.'s (2011) findings, we analyzed complexity at two levels: clause- and phrase-level. Tables 1 and 2 display individual complexity features analyzed with examples. The corpora were “tagged” using the Biber automatic grammatical tagger (see Biber et al. 1999: 35–36) to extract the features in Tables 1 and 2.

Table 1. Clause-level complexity measures

Feature	Example
Subordinating conjunctions	Here they express the fact that technology is worthless if people use it for no real purpose
Verb complements	I cannot wholeheartedly say that I agree with Peter Singer's reasoning in regards to the problem at hand.
Noun complements	The fact that humans are moral beings makes it much more important.
Adjective complements	It is true that animals also have a sense of feeling
That-relative clause	This is an environment that best suits their natural desire.
WH-relative clause	They have knowledge from books which came from libraries or bookstores.

Table 2. Phrase-level complexity measures: Pre- and post-noun materials

Feature	Example
Pre-qualifiers	rather, such
Pre-quantifiers	all, half, both
Post-determiners	many, more, only, other, same
Demonstrative determiners	this, these, that, those
Singular definite and indefinite articles	a, an, the
Singular or plural determiners	any, enough, some
Double conjunctions	either
Attributive adjectives	a huge impact
Post-noun modifying prepositional phrase	a man's quest for truth

Turning to the category of “content,” we focused on the use of the source text in the placement essays. According to the rubric, an essay was judged excellent if it demonstrated 1) accurate understanding of and clear responses to the source text and 2) effective use of the source text. There were two ways in which the students responded to source texts: direct reference to the authors (i.e., use of authornames), and use of specific nouns or verbs (i.e., “attributive words”) making reference to the authors or texts (e.g., *advocate*; *argue*; *argument*; *base on*; *believe*; *claim*; *give*; *offer*; *maintain*; *note*; *put it*; *suggest*). Unlike the tagged examples used for phrasal and clausal complexity, source text references were manually coded and counted in the texts.

Results

Table 3 displays analyses of clausal-level complexity.

Table 3. Frequency counts of clause-level complexity features in high- and low-rated essays

Feature	High-rated essays	Low-rated essays
Subordinating conjunctions*	117	140
That clause verb complement**	89.2	74.1
That clause noun complement	20.6	19.5
That clause adjective complements	5.1	7.3
That-relative clause*	44	70.4
WH-relative clause	28.3	32.1

Notes: Frequency counts are normed to 10,000 words. ** indicates higher essays contained more of a given feature. * indicates lower essays contained more of a given feature.

As seen in Table 3, the frequency of certain clausal features (i.e., that-clause noun/adjective complements, and wh-relative clause) are similar in the higher and lower essays. The lower-rated essays had more subordinating conjunctions and *that*-relative clauses, while higher-rated essays had more *that* clause verb complements. Taking all clause complexity features into account, it looks as though the lower-rated essays are slightly more “complex” than the higher-rated essays. That is to say, the lower essays (with the exception of *that* clause verb complements) have either higher or similar frequency counts as the higher essays.

Table 4 displays normed frequency of phrasal features for high- and low-rated essays.

Table 4. Frequency counts of phrase-level complexity features in high- and low-rated essays

Feature	High-rated essays	Low-rated essays
Pre-qualifiers	10.9	9.6
Pre-quantifiers	23.3	22.5
Post-determiners	89.6	81.7
Demonstrative determiners	72.6	76.2
Singular definite & indefinite articles	984.2	963
Singular or plural determiners	19.7	33
Double conjunctions	14.8	12.3
Attributive adjectives**	466	343.9
Post-noun modifying prepositional phrase**	448.1	405.2

Notes: Frequency counts are normed to 10,000 words. ** indicates higher essays contained more of a given feature. * indicates lower essays contained more of a given feature.

In Table 4, pre-qualifiers, pre-quantifiers, post-determiners, demonstrative determiners and double conjunctions had a fairly equal distribution between the two groups. Given the large number of singular definite and indefinite articles in both groups, the difference was also considered to be fairly equal between groups. Singular and plural determiners were used somewhat more in the lower essays. However, the most notable difference was found in attributive adjectives and post-noun modifying prepositional phrases which were both more frequent in the high- than low-rated essays.

Finally, as shown in Table 5, the higher-rated essays referred to the name of the author and used attributive nouns and verbs much more frequently than the lower-rated essays.

Table 5. Frequency counts of source text reference in high- and low-rated essays

Type	High-rated essays	Low-rated essays
Author name**	227.5	161.2
Attributive words**	86.9	63.8

Notes: Frequency counts are normed to 10,000 words. ** indicates higher essays contained more of a given feature. * indicates lower essays contained more of a given feature.

Discussion

The results of this study suggest that there are identifiable linguistic features between the higher- and lower-rated essays. At the clausal level the main areas of difference are with subordinating conjunctions, *that* clause verb complements and *that* relative clauses. Of these three, the two with the largest differences –subordinating conjunctions and *that* relative clauses are more frequent in the lower essays. These findings lend support to three claims: 1) higher proficiency writers do not necessarily produce more complex language at the clausal level, (2) measuring subordination as the sole indicator of complexity is an oversimplification of complexity, and (3) complexity measured by the use of dependent clauses does not necessarily characterize academic writing.

In fact, excessive subordination can be problematic. For example, the text excerpt from one lower-rated essay below contains seven subordinating conjunctions in the 151-word sample. In this example, the writer uses a chain of *if* subordinate clauses (*if one were to...; if frogs did not eat...; if you could eat it...; if a company were to be; ...if not humans...*) followed by two *so* subordinate clauses (*so it is all part...; so instead of focusing on...*). This example shows how a feature of clausal complexity results in an overly complex sentence that is difficult to follow (and likely results in a lower score in language and content areas).

In the case of science, if one were to ask a ecosystems expert what would happen if frogs did not eat the flies, the world would be overpopulated with flies. Now Singer might respond that why don't humans prey on other humans if you could eat it, but this is again going against everything in society. Which brings Pollan's point that it is part of their place. If a company were to be all executives, what would get done? Nothing who will make the TV, this is just how the organization of nature works. In religions, in Christianity for example they state that God put animals in the world for human's use, if not humans would not need food to survive. So it is all part of how the world works, an offset would maybe even destroy the world, so instead on focusing on animals, focus on how humans destroy themselves. <lower-rated essay, file 019>

In addition to subordination, the number of *that* relative clauses is also more frequent in the lower-rated essays. Considering the example below, we hypothesize that a more proficient writer may have used an attributive adjective (*we find effective uses for it/them*) in place of the second relative clause.

It is true that some of the technology is not being used effectively to teach material, but it is a problem that will go away as we find uses for it that are effective.<lower- rated essay, file 024>

The only area where higher-rated essays show a greater tendency toward clausal complexity is the use of *that* clause verb complement. Figure 6 shows the distribution of verb complement types where *that* clauses are relatively more prevalent in mental (e.g., *believe*) or communication verbs (e.g., *argue*, *claim*). In the cases of the verbs *believe*, *argue* and *state*,

higher-rated essays have a greater frequency of verb complement constructions than lower-rated essays. This partially explains why higher-rated essays have a greater number of verb complement clauses. The writing task used in this study asked students to read two source texts of contrasting views on a controversial issue (e.g., ethics of eating meat), compare the two views, and argue for one perspective over another. They were asked to explicitly take one view over another and support why one position was superior to the other. Hence, higher scores in part reflect how logically and explicitly students could argue for one position over another by drawing on examples from source texts to support their ideas. It is possible that essays were judged more effective because (in part) they used a greater number of attributive verbs to make an argument. In the example below, the writer uses two *that* complement clauses headed by *believe* to contrast the different positions taken in the source texts.

As a response of Singer's work "Animal Liberation", Michael Pollan refutes the claim of Singer. While Singer believes that the act of eating flesh is discriminating against the animal, Pollan believes that the act of eating is part of the human-animal relationship, and it is the way some animals are raised that should be frowned upon.
<higher-rated essay, file 046>

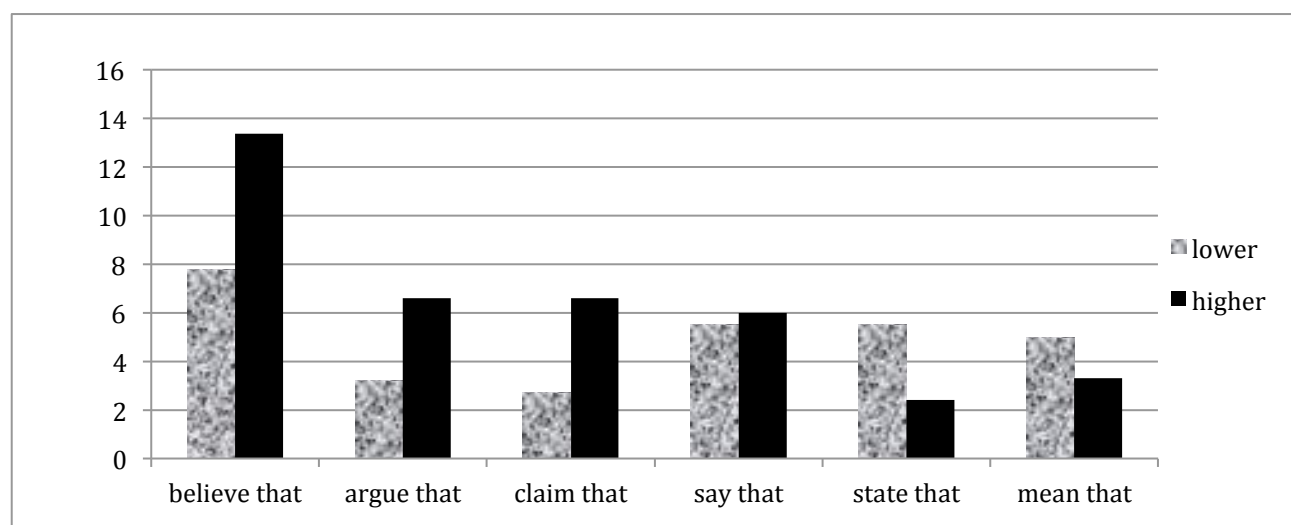


Figure 1. *That* clause verb complements by higher- and lower-rated essays. (normed to 10,000 words)

On the other hand, at the phrasal level, we found that higher-rated essays contain more attributive adjectives and post noun modifying (as opposed to adverbial) prepositional phrases. These findings are in line with the previous literature (Biber et al, 201 and Lu, 2011). As shown in the example below, the higher essays generally used both of these features in a single phrase (attributive adjective are underlined and post-nominal prepositional phrases are in bold).

The Internet is definitely a stimulating tool **for learning** and can play a crucial role **in shaping today's education systems**. < higher-rated essay, file 059>

Post noun-modifying prepositional phrases are also more frequent in the higher-rated essays. The example below shows multiple prepositional complements in a single phrase.

In short there is something akin to **sybiosis** between people in a society. Humans treat each other humanely and take the **interests of others** into account because that is the social norm/ideal. < higher-rated essay, file 003>

Finally, among the number of linguistic and textual features analyzed in this study, one of the clearest differences between the higher- and lower-rated essays is found in frequency of source text reference. Higher-rated essays use both specific author reference as well as attributive language more frequently than the lower essays. In the example below, the writer directly mentions the author seven times in a 91-word sample.

Nowadays, these cases are commonplace and it is definitely a time when animal rights are at the center of many disputes. Among various opinions, Singer and Pollan present two opposing views: Singer against and Pollan for eating animals. I must say, nevertheless, choosing a more sound argument between the two was not completely fair: Pollan's article was written much more recently, and being the critique usually sympathizes better with the readers than does the subject of the critique. In any case, I find Pollan's argument more reasonable than that of Singer. <higher-rated essay, file 094>

Higher-rated essays also contain more specific reference to the author, or to the article itself. In the example below, the writer makes six attributive references in a 77-word sample.

Michael Pollan refutes the claim of Singer. While Singer believes that the act of eating flesh is discriminating against the animal, Pollan believes that the act of eating is part of the human-animal relationship, and it is the way some animals are raised that should be frowned upon. Even though Singer's logic offers significant thoughts and fairness toward animals; however, Pollan's argument is supported with more scientific evidence and more importantly, acceptable to more people than Singer's. < higher essay, file 046>

Conclusion and limitations of the study

This study found identifiable features indicative of higher essay scores. The clearest indicator from our analysis was reference to source text and author but we also found that noun phrase modification (by attributive adjectives and post-noun modifying prepositional phrases) had a tendency to contribute to essay quality. This finding is in line with the more recent work of Biber et. al (2011). Differences were less apparent in subordination measures where the picture was not as clear. Some subordination measures were fairly equal; others showed higher instances in the lower essays (subordinating conjunctions, *that/wh-* relative clauses) and still others showed more cases in the higher essays (verb complement clauses).

This descriptive study is not without its limitation. Increasing the size of the corpus to include more examples of higher and lower-rated essays may show some additional trends. Additionally, including a wider range of linguistic features (e.g., rhetorical features, discourse features, linguistics resources used to distinguish own position from other's) would also be useful.

Finally, the findings reported in this study suggest that writing teachers may want to specifically focus on reference to source texts in argumentative essays. Source text reference not only illustrates comprehension but it also allows the reader to distinguish between the ideas of the reference texts and the ideas of the writer. Additionally, the results of this study show that "complex" writing (at least in the more traditional sense of the term) does not necessarily involve

subordination. Phrasal complexity is prevalent in academic writing because it allows the writer to package information in a concise manner. Writing instruction may benefit from more direct instruction and practice with phrasal complexity.

References

- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26, 390-395.
- Becker, A. (2010). Distinguishing linguistic and discourse features in ESL students' written performance. *MJAL*, 2, 406-424.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & E. and E. Finegan. (1999). Longman *grammar of spoken and written English*. London: Longman.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45, 5-35.
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10, 5-43.
- Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37, 275-301.
- Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45, 36-42.

- McNamara, D., Crossley, S., & McCarthy, P. (2010). Linguistic features of writing quality, *Written Communication*, 27, 57-86.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492-518.
- Rimmer, W. (2006). Measuring grammatical complexity. The Gordian Knot. *Language Testing*, 23, 497-519.
- Smart & Crawford (2009), Complexity in lower-level L2 writing: Reconsidering the T-unit. AAAL, March 2009, Denver, CO.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawai'i at Manoa, Honolulu, Hawai'i.