

A CRITICAL REVIEW OF THE TEACHER-EDITED ESSAYS OF COLLEGE EFL LEARNERS

Maria N. Cusipag^{1a}

¹*De La Salle Araneta University, Malabon City, Philippines*
mariacusipag@yahoo.com^a

Abstract

This study reviewed the teacher-edited essays of thirty college EFL learners who were asked to describe two of their friends or classmates. They took an English proficiency test for grouping purposes and were allowed to choose the feedback pattern for them to use to improve their drafts. From their four drafts, their first three drafts were given feedback using a definite pattern. One group used the F-C-FC pattern (draft 1 is reviewed on form, F; draft 2 is reviewed on content, C; and draft 3 is reviewed on form and content, FC). The succeeding groups followed the C-F-FC and FC-FC-FC patterns. Listing the word count, T-units, and ratings was done per draft. These were analyzed using ANOVA, Pearson correlation, stepwise regression, and intra-class correlation. It was revealed that (1) there was no significant difference among the ratings of the different groups of learners who used three different review patterns in their drafts; (2) there were significant differences among the ratings given per draft; (3) draft 2 was the best predictor for the final draft rating; (4) no interaction effect existed between the review patterns and the drafts used; and (5) a significant relationship existed between (a) proficiency and T-units/rating/word count; (b) rating and T-units/word count; and (c) word count and T-units.

Keywords: feedback pattern, interaction effect, proficiency rating, teacher-edited essays

Issues on whether teachers should or should not correct students' compositions have been subjects of concern for researchers as studies on self-edited and peer-edited essays have begun to show merit in the improvement of students' writing performance. At certain stages of students' development in their writing ability, teachers are now aware as to when to help students in improving their essays, when to let their peers help them, or when to let them do the improvement on their own. They should not overlook, however, the preference of students for teacher's correction on their work, instead of them or their peers.

The role of teachers in reviewing students' compositions such as giving feedback on form like grammar, structure of sentences, and mechanics, or feedback on content like organization of sentences/paragraphs, coherence, and providing missing/additional information needed to make the paragraphs clear, is quite an important task. It is a matter of concern not just for research purposes but for monitoring how well their students perform in their classes. Cook (2013) stressed that marking students' papers is one of the most difficult jobs of a teacher. The purpose of in-text marks and comments is to identify the strengths and weaknesses of the students.

Ferris (2004), in her review of past researches (Ashwell, 2000; Ferris &

Roberts, 2001), reported evidence of the importance of teachers' error feedback. She noted that "studies of student opinions about error feedback are very consistent in reporting that L2 student writers value error feedback from their teachers and consider it extremely important to their success" (p. 55).

Guenette's (2007) idea runs parallel to that of Ferris (2004) who also mentioned:

The success or failure of corrective feedback will depend on the classroom context, the types of errors students make, their proficiency level, the type of writing they are asked to do, and a collection of other variables that are as of yet unknown. Teachers must not lose sight of the fact that second language acquisition is slow, gradual, and often arduous, and that corrective feedback is only one of the many factors that contribute to that process (pp.51-52).

Indeed, teachers are expected to review their students' work and give suggestions for improvement not only in writing but in oral form as well. As NCTE (1998-2011) suggests, "writing conferences ... are also very helpful uses of talk in the writing process" (p. 8). Following the suggestions and getting confirmation that what they did was right gives them a feeling of success and may motivate them to do better (Armstrong, 2009).

Ellegaard et al. (2017) analyzed a written data-set of 174 segments of teacher feedback and student response. Teacher feedback is categorized as the *questioning group of feedback* while student response is referred to as *reflection group of responses*. The former consists of open question, wondering question, and leading question, while the latter consists of reflective response, explanation, and students' investigation of their own thinking.

When teachers give written feedback, therefore, meaning is negotiated as students try to ask questions on how to implement teachers' written suggestions. Thus, in this study, both written and oral feedback were considered. Moreover, corrective feedback on form, as well as on content, were used.

Feedback on Form and Content

Ashwell (2000) agreed that "grammar correction is seen as one way of helping writers to improve the accuracy of a piece of writing and in turn, therefore, to improve its communicative effectiveness" (p.229). His study on feedback patterns recommended the use of content feedback in the first draft, followed by form feedback in the second draft and a combination of form and content feedback on the third draft which is the final draft. It was revealed that no significant difference existed when compared against the reverse pattern; neither did it show any significant difference against the pattern of mixed form and content. The sequence in giving the feedback did not affect the students at all, even if both form and content feedback were done in one draft or in separate drafts. The control group where teacher's feedback was given improved better in the formal accuracy of their writing compared

to the group which did not receive any teacher feedback. Regarding content quality, error feedback did not show any marked effect.

Four types of feedback -- zero, content, form, and content and form -- were used by Fathman and Whalley (1990, as cited in Ashwell, 2000, p. 230) in their study. Their subjects, grouped according to type of feedback, included seventy-two ESL students who wrote a story about a sequence of eight pictures. Form and content feedback were composed of underlined grammar errors and general comments which were not text specific. Their original and rewritten compositions were rated by two scorers based on the number of grammar errors per composition, while holistic ratings for content were given for organization, description, coherence, and creativity. It was found that correcting drafts separately on form and content was as effective as giving corrections simultaneously on both form and content. They believed that multiple drafts were not necessary, since revision and editing could be done at the same time on one draft.

Other studies that made use of both form and content were those done by (Ashwell, 2000; Ellegaard et al., 2017; Guenette, 2007). As reported by Ashwell (2000), no difference existed between the groups compared to the *comments only* group where significant progress on general language proficiency was observed. In the interview data of Fazio (2001, as cited in Guenette, 2007, p. 47) it was revealed that the participants paid very little attention to corrective feedback in contrast to comments.

It is the study of Ashwell (2000) that inspired the researcher to use multiple drafts and different feedback patterns as she reviewed students' compositions on form and content.

Proficiency and Linguistic Features (Word Count and T- Units)

The term *proficiency* in second language refers to the “levels of proficiency, the different degrees of actual or required mastery of the second language, or the progression from a basic to a near-native level” (Stern, 2001, as cited in Fatt, 2007, p. 16). Fatt describes it as a fourfold interpretation which deals with the four major skills of listening, speaking, reading, and writing.

A study where proficiency in ESL/EFL writing was revealed to be significantly related to essay writing performance is that of Fatt (2007). In his study on the effects of time constraints and proficiency on ESL essay writing performance, he found out that proficient students fared significantly better than the less proficient students on the composite and component scores of essays.

Similarly, Ferris and Hedgcock (2005) expressed their opinion that “ lower proficiency students may be unable to identify and correct errors even when they have been marked for them ” (p. 85). Thus, the effectiveness of feedback varies with the ability of student writers.

In another study, Gustilo (2011) proved that proficiency is related to linguistic features in essay writing. She reviewed 150 essays written by Filipino college freshmen. Drawing on Karasawa (2003), their assigned average scores were the bases for classifying the essays into different levels of proficiency. Essays rated 1-2 were placed in level 1 proficiency, 3-4 in level 2, and 5-6 in level 3, the highest. She found that most of the *general text characteristics* (fluency/word count, unique words, word per sentence, number of sentence, number of paragraphs), *lexical features* (intensifiers such as really, very), demonstratives (this, that), additive adverbs (both, neither, too), exclusive adverbs (merely, only, just) except conjuncts (however, for example), *clause-level*, and *grammatical features* all have a steady increase across proficiency levels. Her study corroborates previous studies (Becker, 2010; Grant & Ginther, 2000; Karasawa, 2003) that linguistic features have a significant relationship with levels of proficiency in ESL writing.

Often, but not always, a T-unit is a sentence. It is a linguistic feature which is often used in the analysis of written and spoken discourse such as in studies on errors in second language writing. Myllari (2020) defined it as a production unit within a sentence consisting of one main clause and all the subordinate clauses connected to it directly or via another subordinate clause. According to Nordquist (2020), children connect short main clauses with *and*, and they prefer to use few words/T-unit. As they grow, they use appositives, prepositional phrases, and dependent clauses that help increase the number of words/T-unit.

Accounting for student differences in their study, Robb et al. (1986, as cited in Guenette, 2007) calculated the number of error-free T-units and the number of words in error-free T-units per total words in the compositions of their students. They used this as a covariate in their subsequent analysis.

The writing excerpts of more than 2,500 high school students in the Boston area were investigated by Bloomer (2004, as cited in Mattison, 2004). The students were asked to write a response to an excerpt from an article. After breaking down the writing samples into T-units, he counted the number of prepositional phrases in each writing sample, calculated the ratio between phrases and T-units, and then classified the positions of the phrases: introductory, interruptive, or conclusive. He found that the students who scored highest on the IQ test had the highest T-unit to prepositional phrase ratio, and also had the greatest variability in the positioning of their phrases.

Murphy (2006) examined the total number of words and total number of ideas at the discourse level of children diagnosed with reading disabilities. The total number of T-units, MLT-units (mean length of T-unit), clause density, grammatical errors, total number of sentences, and type of sentences were analyzed at the sentence level, while the total number of different words, spelling errors, writing conventions, and lexical errors were analyzed at the word level.

Similar to previous studies, the researcher made use of *word count* and *clauses*, particularly the main clauses (T-units), in relation to the proficiency scores of the participants in her study.

The various issues in the field of writing that the present researcher has come across prompted her to come up with the following research questions: (1) Do the ratings per draft, regardless of feedback pattern, significantly differ from one another?; (2) Do the ratings received by each group of students who were exposed to different feedback patterns significantly differ?; (3) Is there an interaction effect between the drafts and the feedback patterns used?; (4) Is there a relationship between and among proficiency scores and rating/word count/T-units in the teacher-reviewed compositions? (5) Which variables best predict final draft rating?

Methods

Participants

Thirty college EFL students took part in the study. They were asked to describe two of their friends or classmates. They were nineteen years old whose level of English proficiency was *low intermediate*. Their past English grades from their foreign and local English teachers ranged from 60% to 85%.

The participants met twice a week for 100 minutes per meeting. Their course lasted for 18 weeks. Their textbook included reading and writing exercises, as well as topics for essay writing.

Testing and Grouping of Participants

All the thirty respondents took the proficiency test online at the start of the semester. This proficiency test, a low- intermediate Level A2 Reading Test, was conducted by a private testing center. The reading test was given since reading and writing are related. According to the National Council of Teachers of English (2008) of Illinois, “people who read a lot have a much easier time getting better at writing” (p.7). They believe that in addition to making students stronger readers, the practice of motivating them to read books that they enjoy makes them stronger writers. Likewise, Cotton (2008, p. 6) stressed that “increased reading experiences also enhance writing skill development.”

Three groups of participants were formed with each group having high, medium, and low proficiency scores. Scores from 60 to 64.5 were classified as low (L), 65 to 69.5 as medium (M), and 70 to 80 as high (H), with high as Level 3 proficiency (H-3). As in Fatt’s (2007, p. 17) study, their test scores indicated their level of proficiency. The mean scores of the three groups were taken, and since the scores did not differ significantly, $F(2,29) = .061$, $p = .941$, the grouping was the basis for applying the treatment for the feedback patterns and for measuring their fluency/word count and T-units in their essays.

Study Proper

Prewriting Phase

The thirty respondents were asked to read, copy, and use the following prompt as their guide in submitting their essays:

“Compare and contrast any two of your favourite friends or classmates. Use the block-by-block or point-by-point format in outlining before you develop your essay. Send your composition through e-mail to your teacher. For specifications, use Times New Roman, 12 points, double-space, and proper indentation.”

The teacher checked the students' outlines to ensure that they would write better essays. They used either the point-by-point or block-by-block format in preparing their topic outline. As the study of Kellog (1988, as cited in Fatt, 2007, p. 8) revealed, using a written outline, compared to not using one, yielded high scores among his subjects, implying better written texts. The rubric, adopted from DLSU-Manila academic writing scale, was explained so they would know how their essays would be rated. Students were shown a model essay and were asked if it contained the items addressed in the rubric.

Writing Phase

All the drafts had to be written in the computer lab and submitted to the teacher through e-mail. This eliminated biases on poor penmanship in the reading of their essays. In fact, Li (2000, as cited in O'Brien, 2004, p. 20) found that students who submitted e-mailed texts had longer sentences with a consistently higher level of syntactic and lexical complexity. For students who could not finish their work within the limited class period, they were allowed to bring them home and submit them during the weekend. In Fatt's study, it was revealed that the total essay scores and three of the five component scores of essays written in 45 minutes were significantly better than those written in 30 minutes.

No limit as to the total number of words or the number of pages per draft was given to the students. In a study conducted by Stewart and Cheung (1989, as cited in O'Brien, 2004) they recommended that teachers should “simplify writing tasks by removing limitations on the number of words and the required language forms, and ensure the familiarity of the subject matter” (p.7). Oral conferencing with the teacher before writing their second and third drafts was encouraged. Their fourth draft was not reviewed and was considered final.

Choosing of Review Patterns

The three groups chose by lottery one review pattern for their teacher to use in reviewing their essays. These patterns have already been used by them in their previous essays (Cusipag, 2018). They are (1) F-C-FC Pattern: D1 reviewed and revised on form, D2 reviewed and revised on content, and D3 reviewed and revised on both form and content; (2) C-F-FC Pattern: D1 reviewed and revised on content, D2 reviewed and revised on form, and D3 reviewed and revised on both form and content; (3) FC-FC-FC Pattern: D1, D2, and D3 reviewed and revised on both form and content. Group 1 chose the F-C-FC pattern, group 2 chose the C-F-FC pattern, while group 3 chose the FC-FC-FC pattern.

Reviewing of Drafts by the Teacher

The compositions of the three groups of students were given feedback by the teacher according to the pattern that they chose. Only the first three drafts were reviewed but not the fourth/final draft. The feedback on form and content were either underlined or written along the margins of the composition. Feedback on form included errors in grammar, spelling, capitalization, indentation, punctuation marks, and choice of words. Feedback on content included comments and suggestions on thesis statement, organization of ideas, deletion or addition of information, clarity of meaning/comprehension in sentences, originality, and coherence/cohesion.

Revising of Drafts

The writers were advised to revise each draft only after the oral feedback or conferencing with the teacher, but if they had no queries, they were allowed to revise their work right away. Conferencing between the teacher and the writer was done even after class hours during their free time.

Rating the Compositions and Interrater Reliability

Twelve outside raters marked one hundred twenty e-mailed compositions (thirty each for D1, D2, D3, and D4). Such raters used the academic writing scale or scoring guide which was prepared and validated by De La Salle University (DLSU)-Manila English faculty. This guide was modified by the researcher to suit the grading system of the school.

The raters were all college English teachers, and they were given proper instructions on how to rate the papers. Each of the four drafts was marked by three raters.

Based on Cronbach's intraclass correlation coefficient, the ratings provided by the raters were reliable. The agreement among the raters was quite high from D1 to D4 ($r = .913, p < .001$ for D1; $r = .866, p < .001$ for D2; $r = .916, p < .001$ for D3; and $r = .886, p < .001$ for D4).

Listing the Word Count and T-Units

Linguistic features such as word count and T-units were included as variables. The word count was taken from the computer, while the number of T-units per essay were counted by four trained teachers.

Statistical Analysis

The researcher gathered all the raw data regarding the students' proficiency scores, ratings for drafts 1, 2, 3, and 4 with the corresponding feedback patterns used, word count per draft, and T-units per draft. ANOVA, Pearson correlation, stepwise regression, and intraclass correlation were used by an expert to answer the questions raised by the researcher.

Results and Discussion

Based on the analyses done, answers to the research questions raised were revealed, the first of which is shown below:

1. Do the ratings per draft, regardless of feedback pattern, significantly differ from one another?

It revealed that the average draft ratings received by the students were “below average” in the rating scale of the rubric used. This indicates that they need more improvement in both form and content. Among all the draft ratings, the final draft got the highest (76.39), and the first draft, the lowest (73.87). An obvious reason is that the final draft has been the result of rigid revisions in the previous drafts. Moreover, there is a significant difference among the ratings, from D1 to D4, regardless of the pattern used: $F(3,81) = 49.891, p < .001$. The revisions done by the students from D1 to D4 following the feedback given by their teacher were good enough to merit a significant difference among their draft ratings.

The participants were informed during the prewriting phase that their average scores in the four drafts would form part of their final exam grade. This might have motivated them to do their best in implementing the feedback or suggestions given in each of their drafts. As Guenette (2007) said, “... if the students are not committed to improving their skills, they will not improve, no matter what type of corrective feedback is provided.” In this case, they were candidates for graduation; hence, they were committed to improve their drafts to get passing grades.

2. Do the ratings received by each group of students who were exposed to different feedback patterns significantly differ?

As for the feedback patterns used, the F-C-FC pattern got the lowest mean rating (74.17) while the C-F-FC pattern got the highest (76.37). It was observed that the C-F-FC group was the most motivated group in writing their drafts. They often came for conferencing with the teacher while the F-C-FC group was not as aggressive as the C-F-FC group. They were the shy group, and one of them had a part-time job which prevented her from submitting her drafts on time. For the FC-FC-FC group (75.28), not all the members were as serious as the others in revising their drafts, compared to the C-F-FC group. In fact, two had emotional problems as boyfriend and girlfriend, so there were times when they would submit their drafts late.

It revealed that there is no significant difference among the ratings received by each group of students who were exposed to the three different feedback patterns: $F(2,27) = .293, p = .749$.

3. Is there an interaction effect between the drafts and the feedback patterns used?

Between the two independent variables – the different drafts and the feedback patterns – where $F(6, 81) = .312, p = .929$, no interaction effect exists. This implies that the three groups of students did not mind at all the type of feedback given to

them, for as long as they revised their paper based on their teacher's feedback. There were times when smart students asked to change their sentences that were corrected on form, which would prevent them from revising on form. Since the pattern had to be followed, they were advised to do it in the next draft where revision on content was needed.

Any draft can, therefore, be reviewed/corrected using any of the feedback patterns presented to the participants. The teacher can correct on form first, on content first, or on both form and content, whichever is more appropriate or easier for him/her to use on a particular draft. This finding corroborates Ashwell's (2000) finding that "the recommended pattern of content feedback followed by form feedback is not superior to the reverse pattern or to a pattern of mixed form and content feedback" (p. 243).

4. Is there a relationship between and among proficiency and rating/word count/T-units in the teacher-reviewed compositions?

Using Pearson Correlation, investigations were done if there were correlations between and among proficiency scores and rating/word count/T-units in the students' compositions. Findings showed that the proficiency level of the students as determined in their high, medium, and low proficiency scores, has a significant relationship with the essay ratings, word count, and T-units in all the four drafts. The correlation is significant at .01 level (2-tailed).

The fact that the proficiency exam was on reading, this finding might lend support to the reading-writing connection (Cotton, 2008; NCTE, 2008). Thus, the students' ability to revise, both on form and content, correlated with their proficiency scores in reading. Moreover, this finding corroborated that of Watt (2007) where he found significant relationship between proficiency and essay performance of Malaysian college students, and that of Gustilo (2011) where proficiency of Filipino college students is significantly related to word count (fluency), clauses, and other linguistic features in the essays of the students.

Similarly, at .01 level of significance, there is a significant relationship between rating and word count/T-units, as well as between word count and T-units in D1, D2, D3, and D4. Students who obtained high ratings had longer essays (more word count) with more sentences/independent clauses (T-units). Those who wrote longer compositions had more T-units, too, indicating that the more the vocabulary words were, the more were the T-units/sentences and the higher were their essay scores. Again, these findings corroborate Gustilo's (2011) findings.

5. Which variable best predicts final draft rating?

Based on the regression analysis, the best predictor of D4 rating is D2 rating which contributes about 98.2% in the variation in the final draft rating of the students (R Squared = .982, $F = 1500.582$, $p = .003$). The combined effect of D2 and D3 ratings contributes about 98.5% in the variation in the final draft rating (R -squared = .985, $F = 914.31$, $p < .001$).

Conclusions and Recommendations

The motivation shown by the students in revising their multiple drafts must have contributed a lot in the improvement of their essays, such that their ratings differed significantly. The fact that they were candidates for graduation, they had to do their best in order to get a passing grade from one draft to another. Besides, their average score in their drafts formed part of their final exam grade. Thus, writing teachers may give incentives to students performing a difficult task such as writing.

The ratings did not show any significant difference among the three groups of students exposed to the three feedback patterns. The C-F-FC group got the highest rating, which must have been the result of their aggressiveness in having frequent conferencing with their teacher, leadership, and team learning. The three groups, however, were not affected at all by the kind of pattern that their teacher used in reviewing their drafts. No interaction effect existed between the drafts and the feedback patterns used. It did not matter whether the writers revised their papers on content first, on form first, or on both form and content first. Teachers can therefore give feedback to students' essays as the need arises.

As revealed in the study, students with higher proficiency had longer essays, having more word count and T- units. This implies that teachers should not be contented in just giving word meanings or grammar exercises in their classes. Since T - units are basically composed of independent clauses, students should be taught how to write complete sentences starting from simple to complex and apply them in writing paragraphs, gradually moving into longer compositions. As Gustilo (2011) stressed, "Varied sentence structures, for example, those that employ subordinations and introductory participial phrases, among other complex structures, must be explored by students in their writing to avoid monotonous style of writing" (p. 62).

Only ten subjects per pattern were used by the researcher and no control group was used due to limited school population. Other researchers may consider having both experimental and control groups in their future studies.

The linguistic features investigated in the study were limited to the word count and T-units. Had corpus linguistics been used, more features affecting the students' writing performance would have been examined.

EFL writing teachers should continue with their goal of helping their students identify and correct errors on form and content. While more words, sentences, or paragraphs may be necessary to express one's thoughts and ideas, students should not overlook the good features of an essay which are usually included in the rubrics for assessing writing performance.

References

- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method?. *Journal of Second Language Writing*, 9(3), 227-257.
[https://doi.org/10.1016/S1060-3743\(00\)00027-8](https://doi.org/10.1016/S1060-3743(00)00027-8)
- Becker, A. (2010). Distinguishing linguistic and discourse features in ESL students' written performance. *Modern Journal of Applied Linguistics*, 2(6), 406-424.
- Bloomer, D. (2003). Before, on, after: Prepositional phrases as a marker of the intelligence quotient. *Studies in Psychology and Linguistics*, 84, 61-100.
- Cook, S. (2013). *Providing feedback on student writing*. San Jose State University.
<https://www.sjsu.edu/aanapisi/docs/ProvidingFeedbackonStudentWritingbySaraCook.pdf>
- Cotton, K. (2008). *Teaching composition: Research on effective practices*. Northwest Regional Educational Laboratory. <https://educationnorthwest.org/sites/default/files/TeachingComposition.pdf>
- Cusipag, M. N. (2018). The self-reviewed compositions of Taiwanese EFL learners exposed to different feedback patterns. *Harvest*, 14(1), 1-1.
<https://www.ejournals.ph/article.php?id=12970>
- Ellegaard, M., Damsgaard, L., Bruun, J., & Johannsen, B. F. (2017). Patterns in the form of formative feedback and student response. *Assessment & Evaluation in Higher Education*, 43(5), 727-744.
<https://doi.org/10.1080/02602938.2017.1403564>
- Fatt, C. W. (2007). *The effects of time constraints and proficiency on ESL writing performance* [Doctoral dissertation, Universiti Sains Malaysia]. Open Access Repository of USM Research & Publication.
<http://eprints.usm.my/id/eprint/9546>
- Fazio, L. L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority-and majority-language students. *Journal of Second Language Writing*, 10(4), 235-249.
[https://doi.org/10.1016/S1060-3743\(01\)00042-X](https://doi.org/10.1016/S1060-3743(01)00042-X)
- Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here?(and what do we do in the mean-time...?). *Journal of Second Language Writing*, 13(1), 49-62.
<https://doi.org/10.1016/j.jslw.2004.04.005>
- Ferris, D., & Hedgcock, J. (2005). *Teaching ESL compositions: Purpose, process, and practice*. Mahwah, NJ. Lawrence Erlbaum.
<https://doi.org/10.4324/9781410611505>

- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be?. *Journal of Second Language Writing*, 10(3), 161-184. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)
- Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9(2), 123-145. [https://doi.org/10.1016/S1060-3743\(00\)00019-9](https://doi.org/10.1016/S1060-3743(00)00019-9)
- Guenette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16(1), 40-53. <https://doi.org/10.1016/j.jslw.2007.01.001>
- Gustilo, L. E. (2011). Linguistic features that impact essay scores: A corpus linguistic analysis of ESL writing in three proficiency levels. *3L: The Southeast Asian Journal of English Language Studies*, 17(1), 55-64.
- Karasawa, S. (2003). *Patterns of elaboration and inter-language development: An exploratory corpus analysis of college student essays*. <http://Ucrl.lancs.ac.uk/publications/CL2003/papers/karasawa.pdf>
- Mattison, M. (2004). *Word works (126)*. Boise State University: Boise State Writing Center.
- Murphy, C. (2006). Written language of children with a diagnosis of reading disabilities from grades 3 through 6. *Journal of Undergraduate Research*, 8(2).
- Mylläri, T. (2020). Words, clauses, sentences, and T-units in learner language: Precise and objective units of measure? *Journal of the European Second Language Association*, 4(1), 13–23. <https://doi.org/10.22599/jesla.63>
- National Council of Teachers of English Writing Study Group. (2008). *NCTE beliefs about the teaching of writing*. Urbana, Illinois.
- Nordquist, R. (2020). *T unit and linguistics*. ThoughtCo. <https://www.thoughtco.com/t-unit-definition-1692454>
- O'Brien, T. (2004). Writing in a foreign language: Teaching and learning. *Language Teaching*, 37(1), 1-28. <https://doi.org/10.1017/S0261444804002113>