

Peer evaluation and types of revisions in EFL writing

Yamashita, Yuko

Department of Human Science, Center for Applied Perceptual Research, Kyushu University

Nakajima, Yoshitaka

The Faculty of Languages and Cultures, Kyushu University

Ueda, Kazuo

The Faculty of Languages and Cultures, Kyushu University

Gerard B. REMIJN

International Education Center, Kyushu University

他

<https://doi.org/10.15017/7152004>

出版情報：言語科学. 49, pp.67-77, 2014-03-31. The Faculty of Languages and Cultures, Kyushu University

バージョン：

権利関係：



Peer evaluation and types of revisions in EFL writing

Yuko YAMASHITA*, Yoshitaka NAKAJIMA*, Kazuo UEDA*, Gerard B.
REMIJN**, David HIRSH***

**Department of Human Science, Center for Applied Perceptual Research, Kyushu
University, Japan*

***International Education Center, Kyushu University, Japan*

****Department of Education and Social Work, University of Sydney, Australia*

Abstract

The purpose of this study was to explore how the peer evaluation influenced the types of revision changes made in English as Foreign Language (EFL) writing. Participants consisted of 47 first-year students who took an EFL writing course at a national university in Japan. The students' English proficiency was at an intermediate level on average. They were asked to write an essay and to revise it after it was evaluated by three of them. They exchanged their first drafts with peers to evaluate each essay with one of two different categories of evaluation items; Category A focused on meaning aspects, and Category B on surface aspects. Their first and revised drafts and the filled evaluation sheets were collected. Each revision made by the students was classified as either a surface change or a meaning change. The majority of revisions were surface changes, regardless of the evaluation category, but 30-40% of the changes were meaning changes, which brought new information to or discarded some information from the texts. It seems that the evaluation category did not influence the type of revisions the students made.

1. Introduction

Peer review in first language (L1) and second language (L2) learning is an effective tool to assist educators. Peer review has several benefits, such as encouraging collaborative learning (e.g., Liu & Sadler, 2003; Tsui & Ng, 2000), developing student autonomy (Miao, Badger & Zhen, 2006), and promoting social interaction (Hirose, 2009). Some studies compared peer-, self-, and teacher-assessment, and showed that peer assessment was internally consistent in that peer-ratings were less biased than self-assessment (see Matsuno, 2009; Patri, 2002). For example, Matsuno (2009) explored self-, peer-, and teacher-assessments of English essays written by Japanese university students,

by using an evaluation sheet with 16 items related to the content, organization, grammar, vocabulary and punctuation. The results showed that peer-assessors were internally consistent, and that their rating did not depend on their own writing performance.

One of the main issues is how students incorporate peer review into their work. A number of studies on peer review of L2 writing have explored the type of revisions made by students (e.g., Connor & Asenavage, 1994; Miao et al., 2006; Paulus, 1999). Faigley and Witte (1981) developed a taxonomy of revision changes, which has been widely used to clarify changes in essay writing. Their taxonomy puts emphasis on the distinction of revision whether or not a revision change influences the meaning of a text. They tested their system with groups of writers at different language proficiency levels. The results showed that most revisions of inexperienced writers were surface changes, and only 12% of revisions were meaning changes, whereas 34% of expert adults' changes were meaning changes. Hall (1990) and Paulus (1999) explored how peer review influenced the types of revision, using the same taxonomy, and found that surface changes were the majority of revisions in English as a Second Language (ESL) classrooms. As Faigley and Witte (1981) observed, inexperienced writers tended to limit their revisions to correcting mistakes, and rarely dared to substitute a phrase, add an example, or rework a structure, which would have influenced the meaning of a text.

Although the aforementioned studies shed light on the type of revision changes in L2 writing, there has been a lack of studies and corresponding teaching methods which could elicit revisions to influence the meaning of a text. In the present study, we surmised that peer review with an evaluation sheet including items in the category of meaning aspects might facilitate students to make meaning changes, whereas an evaluation items in the category of surface aspects might facilitate them to make surface changes. We took up the peer evaluation method of Matsuno's (2009) study, which explored self-, peer-, and teacher-assessments of English essays written by Japanese university students, by using an evaluation sheet with 16 items. We divided these items into two categories: Category A, which included overall impression, content and organization, and Category B, which included vocabulary, sentence structure, grammar and mechanics. Category A and Category B corresponded to the taxonomy of revisions proposed by Faigley and Witte (1981) whose analysis method thus was suitable for our present study. We were interested in whether the evaluation category would influence the types of revision.

2. Method

2.1 Participants

The participants were first-year students in an EFL writing course at a national university in Japan. The students' English proficiency was at an intermediate level on average, based on their TOEFL-ITP scores, which they obtained at the beginning of their first semester. The class sizes were ranged from 24 to 26 students. Three students were not included in further analyses, as they did not write the required number of two essays and/or did not receive written feedback twice. In total, 25 students from Class I and 22 students from Class II were included in the current analysis. The students were not randomly assigned to a control group and to one or more experimental groups for two reasons. First, the present study was conducted in an attempt to increase instruction effectiveness during an L2 writing course. Second, it was technically difficult to assign the students randomly to different groups in the education curriculum.

2.2 Procedure

The total number of L2 writing classes was 15 in a semester. During class, the students received essay writing instructions, concerning text organization, content, vocabulary and mechanics. During the sixth and ninth week, the students performed two essay writing tasks, each on a different topic. The first topic was "Favorite places to visit" and the second "The necessity of the smart phone/mobile phone". The topic of the writing task was announced a week before the writing class, so that the students had sufficient time to gather supporting information. The students were given an essay sheet, and were instructed to write a composition on the given topic for 35 minutes. After that, the students exchanged their first drafts with three peers. They were instructed to read each other's first draft carefully and to evaluate each essay with evaluation sheet A or B (see Table 2) for 15 minutes. Matsuno's (2009) assessment criteria were divided into two categories: Category A including overall impression, content and organization, and Category B including vocabulary, sentence structure, grammar and mechanics. We used a five-point Likert scale for each item (see Appendix). As shown in Table 1, the students in class I used Category B in the first writing task, and Category A in the second. Those in Class II used Category A in the first writing task and Category B in the second. The first draft and the three filled evaluation sheets for this draft were handed to each writer. The students were then asked to write a second draft for 20 minutes.

Table 1. Settings

	Class I (n = 25)	Class II (n = 22)
Writing task 1:	Evaluation sheet:	Evaluation sheet:
Favorite places to visit	Category B	Category A
Writing task 2:	Evaluation sheet:	Evaluation sheet:
The necessity of the smart phone/ mobile phone	Category A	Category B

Table 2. Assessment criteria of the evaluation sheet

Evaluation sheet:	Overall	1. Overall Impression
Category A	Content	2. Amount
		3. Thorough development of thesis
		4. Relevance to an assigned topic
	Organization	5. Introduction and thesis statement
		6. Body paragraph and topic sentence
		7. Conclusion
		8. Logical sequencing
Evaluation sheet:	Vocabulary	9. Range
Category B		10. Word/idiom choice
		11. Word form
	Sentence structure/	12. Use of variety of sentence structures
	Grammar	13. Overall grammar
		14. Spelling
	Mechanics	15. Essay format
		16. Punctuation/ capitalization

(Adapted from Matsuno, 2009)

2.3 Data analysis

Each revision made by the students in the second draft was categorized as either a surface change or a meaning change, using Faigley and Witte's (1981) taxonomy (see Table 3).

Table 3. Type of revisions

Surface changes	Formal changes	Spelling/capitalization
		Tense/number/modality
		Abbreviations/contractions
		Punctuations
		Formatting
		Morphological changes
Meaning changes	Meaning-preserving changes	Additions
		Deletions
		Substitutions
		Permutations
		Distributions
		Consolidations
	(do not affect overall summary)	Additions
		Deletions
		Substitutions
		Permutations
Meaning changes	Microstructure changes	Distributions
		Consolidations
	Macrostructure changes	Additions
		Deletions
		Substitutions
		Permutations
		Distributions
		Consolidations

(Adapted from Faigley & Witte, 1981)

Surface changes are changes which do not influence meaning and do not add or discard information. Surface changes are further divided into two subcategories: formal changes (e.g., spelling, tense, and punctuation), and meaning-preserving changes, which introduce new wordings keeping the original concepts in the text

implicitly or explicitly. Meaning changes are changes which affect the meaning by adding or discarding information. Meaning changes are also divided into two subcategories: microstructure changes and macrostructure changes. Microstructure changes make sentence sequences to be understood as consistent and do not affect the overall summary. Macrostructure changes, on the other hand, affect the overall summary.

3. Results

The 25 students in Class I and the 22 students in Class II completed the evaluation for three peers, and the number of evaluation sheets was 75 and 66, respectively. Each evaluation sheet consisted of eight items. The means of all item scores in Category A (e.g., content and organization) from Class I and Class II were 4.27 (SD = 0.77) and 3.78 (SD = 0.89), respectively, and the means of all item scores in Category B (e.g., vocabulary, grammar and mechanics) were 4.17 (SD = 0.8) and 4.20 (SD = 0.63), respectively.

We checked the first and the second draft of each essay, and counted the revisions which the students made in the second draft. Table 4 and Table 5 show the classification results in Class I and Class II. Table 6 shows the type of revisions and the examples of revisions made by the students.

Table 4. Classification in Class I (n = 25)

	Evaluation sheet: Category A	Evaluation sheet: Category B	Total
Surface changes	171	112	283
Meaning changes	77	75	152
Total revisions	248	187	435

Table 5. Classification in Class II (n = 22)

	Evaluation sheet: Category A	Evaluation sheet: Category B	Total
Surface changes	104	128	232
Meaning changes	65	93	158
Total revisions	169	221	390

Table 6. Type and examples of revisions which the students made

Formal changes	<ul style="list-style-type: none"> • It is dangerous to communicate with people whom we <u>don't meet</u> directly → It is dangerous to communicate with people whom we <u>haven't met</u> directly • Smartphones <u>look</u> cool, stylish and gorgeous, especially iPhone. → Smartphone <u>looks</u> cool, stylish and gorgeous, especially iPhone.
Meaning-preserving changes	<ul style="list-style-type: none"> • <u>Train takes me anywhere in Japan</u> and I can enjoy an change encounter there. → <u>If I transfer one train to another, I can go anywhere in Japan</u>, and enjoy an change encounter there. • I can feel good in the orchard, but it is not easy to help grandfather <u>cutting grapes</u>. → I can feel good in the orchard, but it is not easy to help grandfather.
Microstructure changes	<ul style="list-style-type: none"> • By using it, people can do a lot of things. → By using it, people can <u>not only communicate others</u> but do a lot of things. • They enjoy shopping and chatting. → They enjoy shopping and chatting. <u>I would enjoy doing them with my friends when I was in Kobe.</u>
Macrostructure changes	<ul style="list-style-type: none"> • Mobile phone lovers may say that it is very important and it is vital for our everyday life. → Old fashioned people may say that mobile phone is unnecessary, because it fix people to network and it causes the bad effects for our communication skill.

As shown in Table 4, the 25 students in Class I made a total of 187 revisions when instructed to revise after receiving the Category B assessment. They made 112 (representing 59%) surface changes and 75 (41%) meaning changes. They also made a total of 248 revisions when instructed to revise after receiving the Category A assessment. They made 171 (representing 68%) surface changes and 77 (32%) meaning changes. Table 5 shows the total revisions in Class II. The 22 students made a total of 169 revisions when instructed to revise after receiving the Category A assessment. They made 104 revisions (61%) for surface changes and 65 revisions (39%) for meaning changes. They also made a total of 221 revisions when instructed to revise after receiving the Category B assessment. They made 128 revisions (57%) for surface changes and 93 revisions (43%) for meaning changes.

Figure 1 shows the frequencies of revisions (%) in the four above-mentioned subcategories of revisions. Meaning-preserving changes were in the range of 40-50% in both classes. Formal changes were in the range of 10-20% in both classes. It is also to be noticed that microstructure changes were in the range of 30-40% in both classes, which was higher than the percentage range of formal changes.

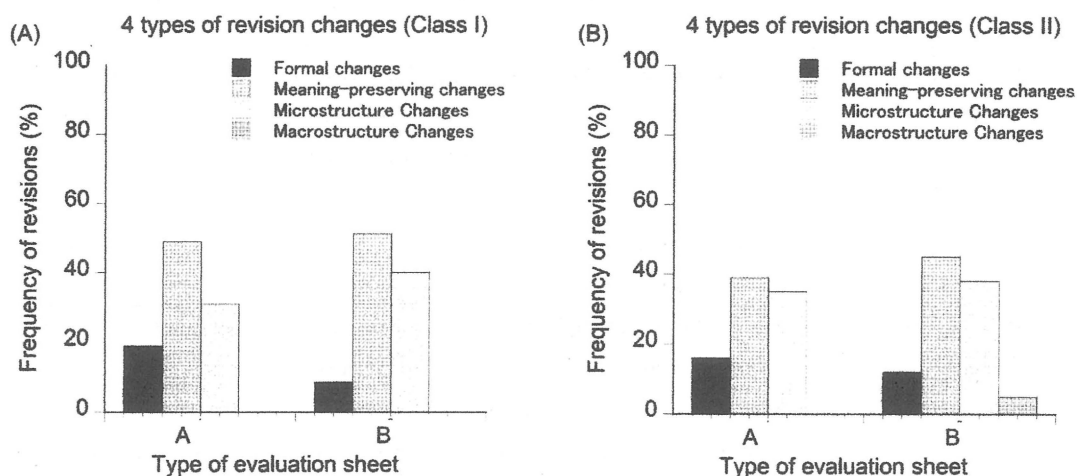


Figure 1. 4 types of revision changes from Class I (left panel) and Class II (right panel).

We read each essay carefully, and found that surface changes included corrections of spelling and tense, and additions of a sentence and a phrase, which did not change meanings. Meaning changes included addition of an example, reworking of the structure, and deletions of a sentence or a phrase, which brought new information to or discarded some information from a text. Below is an example of how a student made meaning revisions.

First draft

Today, most of people have mobile phone and use it anytime and anywhere. Our daily life was changed by the effects of mobile phone. Many people may say that this is a good change, but I don't think so.

Mobile phone effects our communication skills. Mobile phone enable us to talk through the internet, and the changes which we can talk with other people directly is decreasing. Therefore, the number of people who are not good at talking other person may think increasing today.

Second draft

Today most of people have mobile phone and use it anytime and anywhere. Our daily life was changed because of many kinds of advantages of the mobile phone. Now, I think mobile phone is become a vital thing for our everyday life.

Mobile phone enables us to contact with other people as soon as possible wherever we are. This is very useful in the situation that we want to tell the information quickly for other person. For example, when we want to know today's homework and ask it our friends.

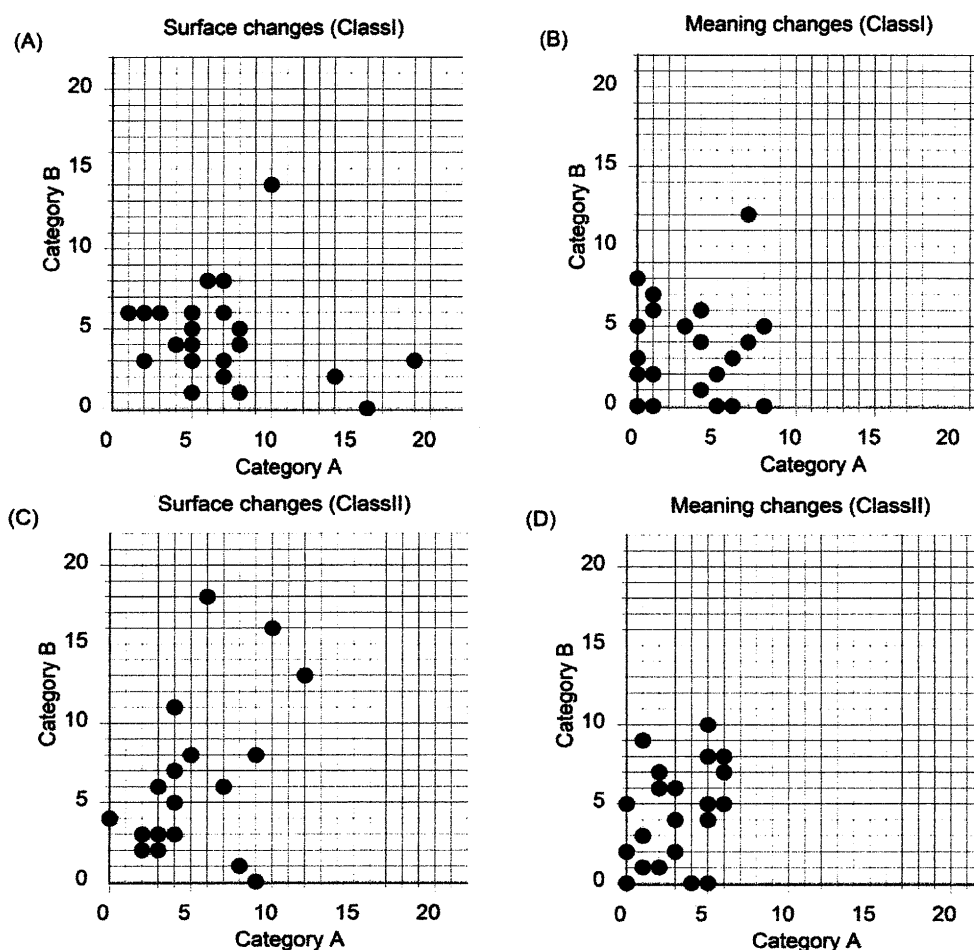


Figure 2. Frequencies of surface/meaning changes which each student made, after receiving assessments in Category A and Category B.

We also explored the tendency of revision making of each individual student. Figure 2 shows the frequencies of surface and meaning changes which each individual student made after receiving assessments in Category A and Category B. Regardless of the evaluation category, most of the students made only less than 10 meaning changes across both classes. With regard to surface changes, the results were more variable; there were a few students who made more than 10 surface changes after assessment in either Category A or Category B.

4. Discussion

The current study investigated how peer evaluation influenced the types of revision changes made in EFL writing. The results showed that the majority of revisions were surface changes. This corresponds to previous studies (e.g., Hall, 1990; Paulus, 1999), in which surface changes dominated in English as Second Language (ESL) classrooms. Students in these studies typically focused on correcting errors. The amount of meaning changes may increase with the student's knowledge of English.




It is also noticed that 30-40% of the changes were meaning changes, including addition of content and reworking of structure. It seems that the category of evaluation sheet did not influence the type of revision changes students made. The procedure in which the students had to read their peers' essays carefully is very likely to have facilitated the students to generate new content for their own essays. This needs to be explored systematically. The present small-scale study provided some pedagogical implications leading to future studies on EFL writing.

References

- Connor, U., & Asenavage, K. (1994). Peer response groups in ESL writing classes: how much impact on revision? *Journal of Second Language Writing*, 3, 257-276.
- Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, 32, 400-414.
- Hall, C. (1990). Managing the complexity of revising across languages. *TESOL Quarterly*, 24, 43-60.
- Hirose, K. (2009). Student-student written interactions during peer feedback in English writing instruction. *Annual Review of English Language Education in Japan*, 20, 91-100.
- Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2(3), 193-227.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 75-100.
- Miao, Y., Badger, R., & Zhen, Y. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, 179-200.
- Patri, M. (2002). The influence of peer feedback on self- and peer assessment of oral skills. *Language Testing*, 19(2), 109-131.
- Paulu, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265-289.
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147-170.

Appendix

Example of five-point Likert scale for items in evaluation sheet

	Too many mistakes Ineffective Very poor		Average	Very few mistakes Effective Very good	
					
	1	2	3	4	5
1. Overall impression	1	2	3	4	5
2. Amount	1	2	3	4	5