【Original article】

Optimizing Peer Check for the ESL Writing Classroom

Exploring the Effectiveness of Peer Check Strategies in

Uncovering Global and Local Errors

Mark Guthrie[1]

Frid Joanis[2]

[1] [2]Organization for Promotion of Higher Education and Student Support, Gifu University

Abstract

The purpose of this article is to explore the efficiency and effectiveness of peer review in English writing classes in Japanese universities, and in what guise it should be implemented, by reporting the results of a study conducted on the matter. In the study, participants were thrice tasked with checking model paragraphs for Local and Global errors individually, in pairs, or in small groups of four or five over the course of a semester. The results of the study show that participants working in groups consistently found the most errors overall, in Tests 2 and 3 pairs found the most global errors, while the participants working individually consistently found the fewest errors throughout. Nonetheless, participants working individually, in pairs, and in small groups all improved in the number of errors they corrected by the end of the semester.

Key words: L2 writing, groupwork, pair work, independent work, Japanese English learners,
　　　　　peer check/review/feedback

1. Introduction

The use of feedback is defined by Narciss (2008, p.127) as being in its essence any information given to a learner after their response (i.e. output), aimed at enlightening them about their current learning or performance status. This feedback, in learning contexts, is a vital element of second language acquisition (SLA). In traditional English as a Second Language (ESL) classrooms, students receive feedback from both the teacher and their peers, particularly in

speaking exercises where response is immediate. In fact, Lightbown & Spada (2013), argue that the language classroom is typically the sole location where error feedback is commonly found at a high frequency. The reason for this is that language acquisition is a social undertaking as "learning is not something an individual does alone, but is a collaborative endeavor necessary involving other individuals… [because]... [e]ffective error correction and language learning depend crucially on mediation provided by other individuals" (Aljaafreh & Lantolf, 1994, p. 480).

The effects of peer feedback, review, and/or editing in the classroom has been extensively researched, with a broad range of exploratory reports establishing that its usage is advantageous to both L1 and L2 students (Rouhi & Azizian, 2013; Bijami, Kashef & Nejad, 2013; Baker, 2016). Furthermore, as Lundstrom and Baker (2009) point out, due to its innate collaborative essence, peer feedback has gained increasing popularity in the writing classrooms over the last few decades. They argue that this is primarily due to its evident advantages, particularly for students who provide feedback to their peers. However, the authors of this research have, over many years of personal experience in the Japanese education system, observed scant utilization of peer feedback in English writing classes, which has been previously confirmed by Barker's (2021, p115) view that "[m]any students will have little or no experience of either giving or receiving feedback on written assignments." Barker went on to add that only 25 percent of participants reported having undertaken peer editing of Japanese writing at the junior high or the high school level, and as such were unlikely to be skilled in the area. Thus, while peer feedback has been extensively researched, there are still many points of interest and questions that remain.

The motivation of this study comes from the present authors' desire to improve the efficiency and effectiveness of peer review in English writing classes in Japan. This study seeks to determine whether peer review is best implemented individually, in pairs, or in groups by looking at how many errors students can find in short writing compositions and comparing those findings in each context. While other research has proven the effectiveness of peer review, the present study seeks to explore the efficiency of peer checking in groups.

2. Methods

Context and participants

This study was undertaken with four classes during a 15-week English writing university course with 30-40 students in each class in the fall semester of 2022. The students comprised both male and female individuals who were in their first or second years, aged between 18 and 21. Over the course of the semester, students participated in peer editing within rotating groups consisting of 4 to 5 students. They analyzed one another's 100-to-150-word compositions, which explored various personal themes. Prior to beginning their weekly composition assignments for

the course, the students underwent a three-week review of basic grammatical structures, as well as being made aware of a number of   common errors made by Japanese students taken from *An A-Z of Common English Errors for Japanese Learners* (Barker, 2010). As the majority of the students were not proficient in spoken English, they were permitted in these groups to use Japanese as a method of communication. In this way they were better able to collaborate with one another and engage in "collective scaffolding" (Donato,1994) by combining their linguistic repertoires in order to explain concepts to one another.

For this research, one class consisting of 28 students was chosen as Group 1, another class consisting of 28 students was chosen as Group 2, and two classes each consisting of 30-35 students were chosen as Group 3. Group 1 worked individually, Group 2 worked in pairs, and Group 3 worked in small groups of 4 or 5. As attendance during the semester fluctuated, so did the total number of students in each of the groups studied. Therefore, there were 23-28 students in the individual group for each test, 24-28 students in the pair group (with 12-14 total pairs for each test), and 65 students working in small groups (with 11-13 total groups for each test).

The tests

During the course of the module, the students were issued with three 'test' paragraphs to correct: one (T1) prior to the students' first peer check session in week 3, another (T2) in week 8 prior to the mid-term test, and the final one (T3) in week 14 prior to their final test. Each test paragraph was written using language and vocabulary similar to that which the students' themselves would use. Additionally, they contained intentional mistakes that the students themselves would make and were made aware of through the three-week grammar review as well as through the lessons and feedback provided by the teachers throughout the semester. To ensure that each of these paragraphs were of equal difficulty, they were composed by one of the authors of the paper, reviewed by the other author, and then by three colleagues who also taught the same writing class. However, it should be noted that only the writing classes of the authors were used for this study.

In one class (henceforth known as Group 1) each student was given a paragraph to check individually and were tasked with identifying the mistakes and correcting them, simulating the individual peer check of a partner's work. In another class (henceforth known as Group 2), one paper was given to a pair of students to check the test paragraphs together, and in the remaining two classes (henceforth known as Group 3) the class was separated into groups of 4 or 5 students and asked to work together in checking the test paragraphs. In all cases the students were given seven minutes, the approximate amount of time students would be expected to spend on each paragraph in the classroom setting, to uncover as many errors as possible, after which the teacher revealed the answers to ensure that the exercise was educationally beneficial to the students.

The errors

Each paragraph contained a total of 12 errors, made up of 9 Local errors, those that do not cause problems in communication and understanding, and 3 Global errors, those that hinder communication and disrupt meaning (Touchie, 1986). The Local errors - two capitalization mistakes and one each of verb tense, conjunction, pronoun, punctuation, spelling, pluralization, and subject verb agreement - were all chosen as they represented the most common mistakes that students in previous semesters had been observed to make. (It should be noted here that the authors of this paper made an error in creating the first test, and what was labeled as a *subject verb agreement* error was in fact a *verb tense* error. As such, it has been omitted from analysis, though it remains included in the overall error count). The Global errors—sentence combinations, word order, and paragraph order/flow—however, were not strictly covered in the initial three week grammar and error review, though students should have been taught about them at the high school level and reviewed them in lessons throughout the module.

3. The results

Local vs Global - an overview

Overall, all groups, Group 1 (Individuals), Group 2 (Pairs), and Group 3 (Groups of 4 or 5), generally improved over the course at finding the errors presented in the test paragraphs. In the first test (T1), Group 1 found 31.85% of errors, Group 2 found 42.86%, and Group 3 found 51.92%. In the third test (T3), this rose to 50.00% (an 18.51% increase), 59.03% (a 16.17% increase), and 64.39% (a 12.47% increase), respectively, marking out Group 1 as the most improved overall throughout the course of the module. Though this shows relatively significant improvement, it should be noted that there was a drop in error detection in all groups between the second test (T2) and T3: Group 1 decreased 0.36% in the percentage of found errors, Group 2 decreased 3.15%, and Group 3 decreased 4.55%. This leads the authors of this paper to question whether T3 was more difficult, perhaps subconsciously skewed for difficulty as teachers sought to increase complexity due to higher expectations placed on students as they developed through the course.

Table 1. Total errors for Group 1, Group 2, and Group 3

Total Errors			
	Group 1 (Individuals)	Group 2 (Pairs)	Group 3 (Small Groups)
Test 1	31.85%	42.86%	51.92%
Test 2	50.36%	62.18%	68.94%
Test 3	50.00%	59.03%	64.39%
Average errors found (all tests)	44.07%	54.69%	61.75%

While Group 3 found the most errors in all tests, these were mostly Local errors (T1 67.52%, T2 86.87, T3 79.80%). Additionally, Group 3 found the fewest Global errors in T2 (15.15%) and T3 (18.18%). Between T1 and T2, Group 1 more than doubled their ability to find Global errors (7.14% in T1 to 15.94% in T2), and then again in T3 (33.33%). However, Group 2 showed the greatest improvement in this area, more than quadrupling their ability to identify Global errors in the first two tests from 4.76% to 20.51%, and almost doubling that again in T3, where they identified 38.89%, marking them out as the most adept in this area.

Table 2. Total Local and Global errors for Group 1, Group 2, and Group 3

Total Local Errors for All Groups			
	Group 1 (Individuals)	Group 2 (Pairs)	Group 3 (Small Groups)
Test 1 Local	40.08%	55.56%	67.52%
Test 2 Local	61.84%	76.07%	86.87%
Test 3 Local	55.56%	65.74%	79.80%
Average Local errors found (all tests)	52.49%	65.79%	78.06%
Total Global Errors for All Groups			
	Group 1 (Individuals)	Group 2 (Pairs)	Group 3 (Small Groups)
Test 1 Global	7.14%	4.76%	5.13%
Test 2 Global	15.94%	20.51%	15.15%
Test 3 Global	33.33%	38.89%	18.18%
Average Global errors found (all tests)	18.80%	21.39%	12.82%

Local vs Global - a focused view

As mentioned, Group 3 (groups of 4 or 5) found the most difficulty identifying and correcting Global mistakes. Though they improved in correcting *word order* mistakes, finding 15.35% in T1, 45.45% in T2, and 54.55% in T3, they failed to identify any *sentence combinations* or *paragraph order/flow* issues in any of the three tests. This lies in stark contrast to Group 1 (individuals) and Group 2 (Pairs), who in T1 found *word flow* issues (21.42% and 14.29%), in T2 improved on this (26.08% and 46.15%) whilst also locating 21.74% and 15.38% of *combining sentences* mistakes, and then in T3 adding 21.43% and 41.67% of *paragraph order* mistakes to their improved *word flow* (35.71% and 58.33%) and *combining sentences* (42.86% and 16.67) errors.

Graphs 1-3. Global errors for Group 1, Group 2, and Group 3.
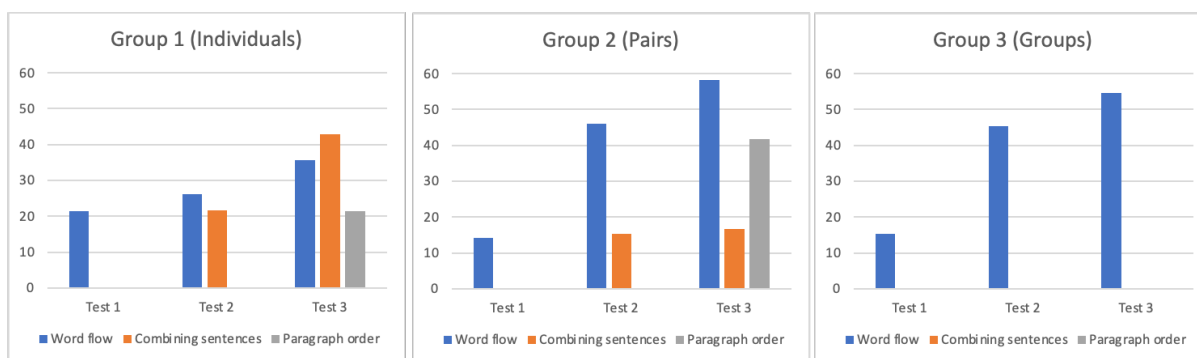
When it came to correcting *conjunctions*, Group 3 was consistently excellent, correcting 92.30% in T1, and 100% in both T2 and T3, while Group 1 improved over the course, correcting 67.86% in T1 and 75% in T3, with pairs seeing similar results (85.71% in T1 and 91.67% in T3). In addition, Group 1 saw some improvement in *punctuation* identification, rising from 61.71% in T1 to 78.57% in T3, though this is a decline from T2, where they located 91.30%. Group 2 and Group 3 fared less well in this aspect, with both Groups showing a steady decline from T1 (64.29% and 100%) to T3 (41.67% and 90.90%), though the latter remained consistently the best at identifying and correcting these mistakes throughout.

Though Group 2 saw consistent improvement in identifying *plural* errors (51.14% in T1, 61.54% in T2, and 66.67% in T3), the only area in which all classes saw continuous improvement throughout was *pronouns*: Group 1 at 42.86% in T1, 43.48% in T2, and 67.86% in T3; Group 2 at 35.71% in T1, 69.23% in T2, and 75.55% in T3; and Group 3 at 30.77% in T1, 90.90% in T2, and 100% in T3, marking out Group 3 as not only the most improved in this area, but also the most consistently accurate from T2 onwards.

One other area that saw a similar pattern throughout Groups 1, 2, and 3 was in spelling. While all improved from T1 to T2 (Group 1 from 17.86% up to 86.96%, Group 2 from 14.29% to 100%, and Group 3 from 53.84% to 100%) every Group dropped dramatically in T3 (10.71%, 33.33%, and 54.55% respectively). This wild fluctuation, particularly the dramatic drop off between T2 and T3, is one of the reasons why the authors suspect that T3 was subconsciously created as more difficult than its predecessor.

Graphs 4-6. Local errors (plurals, pronouns, and spelling) for Group 1, Group 2, and Group 3
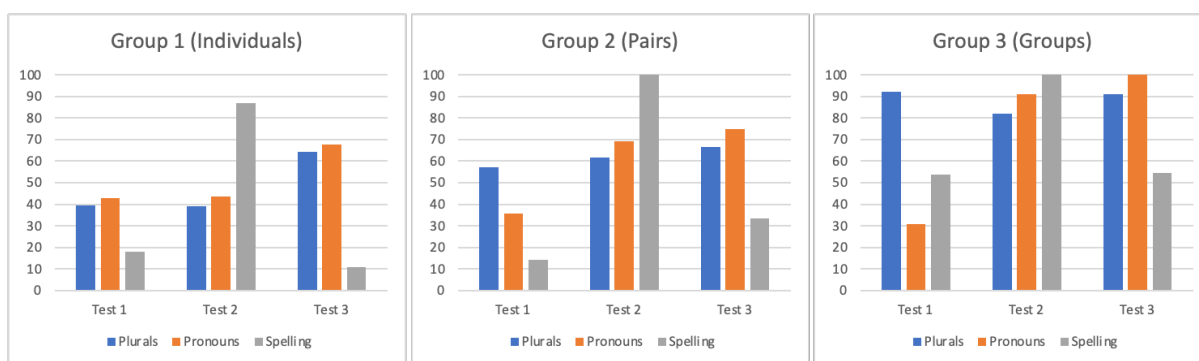
The aforementioned mistake in T1 makes it difficult to gauge progress in *subject verb agreement* error detection, however, between T2 and T3, Groups 1 and 2 showed improvement, with the former rising from 47.83% to 89.29%, and the latter going from 90.90% to 100%. Though Group 2 showed some decline, it was marginal, from 92.31% to 91.67%.

When it came to *verb tense* errors, all Groups showed improvement from T1 to T2 (Group 1 30.35% to 60.87%, Group 2 42.85% to 69.23%, and Group 3 34.62% to 81.82%), however there was a precipitous drop for all in T3 (14.29%, 25% and 45.50% respectively). T2 aside, Group 3 once again showed that they were most adept at uncovering these Local errors, however, due to the dramatic difference between T2 and T3, there is once again the suspicion that T3 was created as more difficult than its precursors.

Graphs 7-9. Local errors (subject verb agreement and verb tense) for Group 1, Group 2, and Group 3
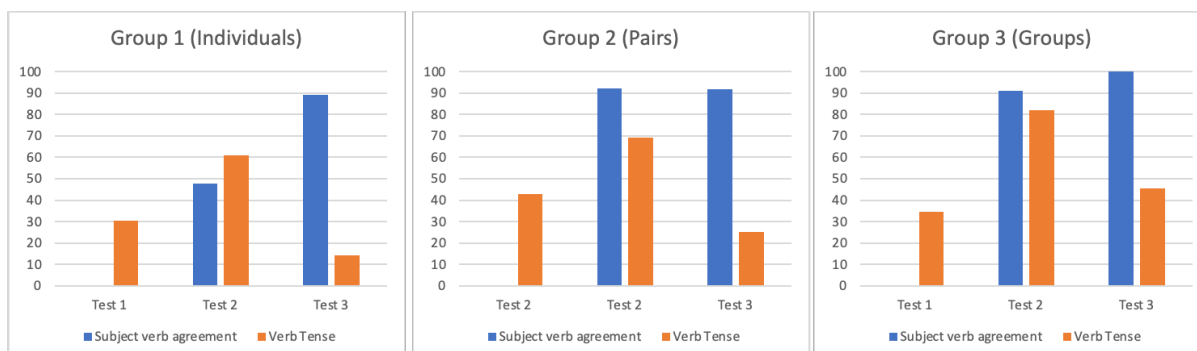
4. The findings

All classes generally improved in both identifying and correcting mistakes within the test paragraphs over the course of the module. This is something that could have been predicted, as students became used to reviewing each other's work, and their writing skill sets developed through both the shared knowledge acquired from their peers and the feedback received from their teachers, which was given both in their paragraph assignments and in the classroom. However, it is clear that throughout the course, Group 3 was consistently more accomplished, something that could perhaps have been expected, as culturally, "Japanese use strategies which maximize the strength of the group" (Yamada, H. 1990, p. 272), and as such would be likely to be predisposed towards proficiency in this environment.

Perhaps the most striking discovery is that, while Group 3 was by some distance more adept at identifying and correcting the more black-and-white Local errors, by stark contrast they consistently lagged behind the other Groups when it came to the Global issues. While by T3, both Group 1 and Group 2 had become proficient at locating and correcting at least some of all three forms of Global errors—word order, sentence combinations, and paragraph order/flow—Group 3 were only able to do so with the former. There is a possibility that this disparity is a result of a differing emphasis on Global errors in the classroom, as the classes from the Group 1 and Group 2 students were taught by one teacher, and the classes of the Group 3 students taught by another. However, this rationale can perhaps be discounted, as following each test, all students were shown the answers, and in every instance the reasoning behind the errors were made clear. As a result, it can be supposed that all students should have been aware of the potential for Global errors in the tests, at least from T2 onwards.

Another supposition that could be made is that the seven minutes allotted for the corrections of the tests, while adequate for Group 1 and Group 2, was insufficient for Group 3 to

identify the more complex concepts of the Global errors. Though proficient at pointing out the more grammar-focused Local issues, problems that have a more concrete conclusion, it perhaps takes more time to reach a consensus amongst four or five people when it comes to the potentially more objective Global issues. One explanation for this could be a reticence to put forth one's opinion in a communal environment, as according to Madeen and Sugimoto (2019), Japan is the world leader in terms of "Strong Uncertainty Avoidance," a situation that leads Japanese students to avoid stating an opinion in a large group on issues about which they may lack confidence.

It could also be theorized that the issue lies on the other side of the uncertainty coin, with the problem stemming from how Japanese work collectively to come to a conclusion, and how the concept of *enryo* affects decision making. Variously translated as 'reserve', 'tact', and 'thoughtfulness', *enryo* culture, "inhibits Japanese speakers from saying directly what they want, and it also makes it culturally inappropriate to ask others directly what they want" (Goddard and Wierzbicka, 1997, pg 237). It could be surmised that this culture permeates into the classroom setting, and though it is straightforward to assess a black-and-white Local error, more thoughtfulness is required for the Global errors, a situation exacerbated by there being numerous others in Group 3 with whom a tactful position was required. This could lead to greater hesitance in putting forward a suggestion and such opinions are withdrawn, or it could be that more time is required for those indirect interactions to come to a successful fruition.

5. Limitations

There are a few limitations to this present study. The most obvious limitation is the small sample size of 128 students. Having limited participants means that our analysis may not accurately represent the broader population of Japanese university students. While this may be true, our analysis suggests that the study of peer checking in this context is an area of interest for further study. Another limitation was the time constraint (of 7 minutes for completion of the peer checking activity) within the study. As mentioned in the previous section, Group 3 likely needed more time to deliberate and build consensus. If given then time, then it is possible that Group 3's results may have been different when looking for Global errors.

6. Conclusions and considerations for future research

From the collected data, it can be concluded that over time and with practice, whether it be individually, in pairs, or in groups, students can improve their ability to identify and correct errors when engaging in a peer checking process. It is also clear that peer-checking in groups is

the most efficient strategy to eradicate most problems, particularly if they are Local errors. However, unless cultural considerations such as social uncertainty and *enryo* are adequately addressed, group-based peer checking will fail to rectify Global errors to a satisfactory degree, and as a result, students will see their writing perplex readers, as "Global errors should almost always be corrected because they cause confusion…" (Freiermuth, 1997, online).

Through data analysis and classroom observation, this paper proposes three strategies that can be undertaken to rectify this problem and thus facilitate students in reaching their potential as writers in English. One would be to have students work in groups to rectify Local errors, before breaking out into pairs to tackle the more complex Global mistakes. Whilst maximizing the potential for error correction, this has the advantage of sidestepping the aforementioned cultural impositions that stall progress. However, this process may be disadvantageous in how it requires an additional stage in the procedure that may lead to confusion or unnecessary time wasting when dividing into pairs. Another approach would be to utilize smaller groups for the peer check process, meaning that there are fewer conversational barriers to gaining a consensus. This potentially combines the most positive elements of both groups and pairs—high levels of both Local and Global error identification— though there is also the contrary possibility, that fewer Local mistakes would be discovered and that the cultural impositions would remain, meaning that Global issue discovery is reduced. Finally, the third strategy would be to include a greater level of peer-check training prior to the start of the process. In this way, teachers could instruct students in how to approach and rectify Global errors, consequently fostering an atmosphere in which those cultural issues can be addressed and overcome, so that students are more adept at raising Global mistakes released from the constrictions of uncertainty and *enryo*.

As the results show, peer checking, particularly in groups and pairs, can be a highly effective tool in writing classes, and students continuously improve with more practice. However, it has become clear that when it comes to ensuring efficiency and accuracy, particularly in the Global issues, further research is required to explore the best of the aforementioned strategies. Yet, should that research crack the efficiency code, peer review can be not only an invaluable teaching tool, it can also help develop students into better and more accomplished writers.

Bibliography

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The modern language journal*, *78*(4), 465-483.

Baker, K. M. (2016). Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education*, *17*(3), 179-192.

Barker, D. (2010). An A-Z of Common English Errors for Japanese Learners (English Edition). BTB Press

Barker, D. (2021). What Students Learn from Peer Editing L1 Writing. 『岐阜大学教育学部研究報告．教育実践研究・教師教育研究』*23*, 115-124.

Bijami, M., Kashef, S. H., & Nejad, M. S. (2013). Peer feedback in learning English writing: Advantages and disadvantages. *Journal of Studies in Education*, *3*(4), 91-97.

Donato, R. (1994). Collective scaffolding in a second language. *Vygotskian approaches to second language research*, 33-56.

Freiermuth, M. R., (1997). L2 Error Correction: Criteria and Techniques, *The Language Teacher,* Issue 21.9, https://jalt-publications.org/tlt/articles/2197-l2-error-correction-criteria-and-techniques

Goddard, Cliff and Wierzbicka, Anna (1997). Discourse and Culture. In Teun A. van Dijk (ed.), Discourse as Social Interaction. London: Sage Publications, pp 231-259.

Lightbrown, P., & Spada, N. (2013). How Languages are Learned: Oxford Handbooks for Language Teachers Oxford: Oxford University Press.

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of second language writing*, *18*(1), 30-43.

Madeen, E., & Sugimoto, H. (2019). THE STATUS OF CREATIVE WRITING AS AN ACADEMIC DISCIPLINE: WHERE IT THRIVES AND DIVES–WITH A SPECIAL FOCUS ON JAPAN AND ITS "KATAIZED" CULTURE. 『東京都市大学共通教育部紀要』*(Journal of liberal arts and sciences at Tokyo City University)*, *12*, 153-163.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In *Handbook of research on educational communications and technology* (pp. 125-143). Routledge.

Rouhi, A., & Azizian, E. (2013). Peer review: Is giving corrective feedback better than receiving it in L2 writing?. *Procedia-Social and Behavioral Sciences*, *93*, 1349-1354.

Touchie, H.Y. (1986). SECOND LANGUAGE LEARNING ERRORS THEIR TYPES, CAUSES, AND TREATMENT. *JALT Journal Volume 8, No. 1*, 75-80

Yamada, H. (1990). Topic management and turn distribution in business meetings: American versus Japanese strategies. *Text - Interdisciplinary Journal for the Study of Discourse*, vol. 10, no. 3, 1990, pp. 271-295.

ライティング教室におけるピアチェックの最適化

グローバルおよびローカルエラーを明らかにする

ピアチェック戦略の効果の探求

ガスリー・マーク [1]

ジョーニス・フリード [2]

[1] [2] 岐阜大学教育推進・学生支援機構

要旨

　本論文の目的は，日本の大学における英語ライティングクラスでのピアチェックの効率と効果，及びどのような方法で実施すべきかを探究することである。このテーマに関する研究の結果を報告することで，そのガイドラインを提供する。研究では，参加者に対して1学期を通じて，段落の例のローカルエラーとグローバルエラーを，個別に，ペアで，または4人から5人の小グループで三度チェックする課題を与えた。研究の結果，グループで作業する参加者が，最も多くの総合的なエラーを見つけることが示され，テスト2とテスト3では，ペアが最も多くのグローバルエラーを発見した。一方，個別に作業する参加者は一貫して最も少ないエラーしか見つけることができなかった。しかしながら，個別に作業する参加者，ペア，小グループで作業する参加者はいずれも，発見するエラーの数を学期末までに改善した。

キーワード：第二言語（L2）ライティング，グループワーク，ペアワーク，独立作業，
　　　　　　日本人英語学習者，ピアチェック／レビュー／フィードバック