Feedback on Feedback: An Analysis of L2 Writers' Evaluations of Proofreaders

Mark Rebuck[1]

Meijo University

Most research into second language (L2) writers' feedback preferences has been in the context of L2 writing classes with students responding to teachers' feedback. The present study, however, sought to ascertain writers' perceptions of corrective feedback provided by proofreaders, both native speakers (NSs) and nonnative speakers (NNSs) of English, recruited by an English language-check service in a Japanese graduate school. The aim of this piece of exploratory research was practical: to improve academic feedback through a deeper understanding of the student writers' perceptions of different feedback styles and strategies. In the study, 15 students ranked the usefulness of the feedback provided by six anonymous proofreaders on the same sample of high-stakes writing. Students were also required to add comments to elucidate the rankings assigned. An analysis of the ranks and the writers' holistic impressions suggested, amongst other things, that most writers preferred direct feedback over indirect, and were more averse to erroneous correcting of content than of language. The study also revealed that feedback provided by NNSs often outranked that given by NSs, although, partly because of the small sample size, no firm conclusions were drawn from this finding.

Key Words: error correction, feedback, proofreading, students' perceptions

1. Introduction

Providing corrective feedback on students' L2 (second language) writing can entail teachers taking up "an enormous amount of time" outside of class (Harmer 2007, 157). It is partly because of its potential for becoming a burden that responding to students' writing is a concern for many writing instructors. It also helps explain the presence of a large body of research on L2 error feedback, research that became particularly active after Truscott (1996) controversially asserted that correcting errors in L2 students' writing is not beneficial, and can even be counterproductive, to students' writing development. Other writers, most notably Ferris (2003), have since provided evidence to support error correction. The debate, however, over the benefits or otherwise of error correction is somewhat academic for many teachers because, as Chandler (2003) reports, numerous studies have consistently shown that learners expect teachers to comment on their written errors and are frustrated if this does not happen.

For the author, an English teacher employed at the time by the Graduate School of International Development (GSID), Nagoya University, to provide an English language-check of students' academic papers, the correction of L2 writing errors was his raison d'etre. Thus rather than seeking to ascertain whether corrective feedback helped students in the long term to become better writers, he considered that a more relevant research focus for his particular context was this: What form of feedback is perceived by students—the users of the English check service—to be most valuable?

Gaining insights into students' reactions to various forms of error correction—in other words, obtaining feedback on feedback—was a necessary first step for improving the quality of feedback provided and, ultimately, raising the standard of writing produced by GSID students. This paper describes an exploratory study conducted to shed light on students' perceptions of different feedback styles and strategies.

---

[1] The research for this paper was completed while the author was employed by the Nagoya University Graduate School of International Development. Since April 2013, he has been in his current position of Associate Professor at the Faculty of Pharmacy, Meijo University where he teaches EMP (English for Medical Purposes).

2. Background

2.1 The research gap

Students' reactions and preferences to written feedback is an area that, as Ferris and Roberts (2001) point out, has been neglected somewhat in many error correction studies. Those studies that have focused on students' perceptions of received feedback have been in the context of EFL and ESL writing classes, usually with the writing teacher providing feedback for a short classroom assignment. These studies have mainly used questionnaires with closed items to ascertain students' preferences (see, for example, Chandler, 2003). This present study differed from previous investigations into writers' perceptions of corrective feedback in several respects. Firstly, feedback was provided for a sample of high-stakes academic writing, specifically an extract from an MA thesis. ("High-stakes" in this context refers to writing whose evaluation could have a significant impact on the writer's life chances). Secondly, those providing the feedback were paid student proofreaders, both NSs (native speakers) and NNSs (non-native speakers of English). Thirdly, rather than responding to closed items on a questionnaire, student writers in this study provided holistic impressions (in writing) of the feedback they received. Finally, while in previous studies students responded to the feedback from a single source, usually their teacher, in this present study they compare, evaluate, and comment on feedback given to the same piece of writing by multiple feedback providers.

A questionnaire survey and semi-structured interviews with student writers and the proofreaders were also conducted concurrently with the error-feedback study focused on in this paper. While space does not permit these surveys to be described in depth, some of their findings will be referred to when they illuminate the results of the error- feedback study. All this research was conducted at a time when the creation of a more attractive academic environment for international students has become a particularly high priority for GSID due to the Global 30 Project, a major Japanese government initiative which aims to attract 300,000 international students to Japan by 2020.

2.2 The setting

GSID is located in Nagoya University, a national university in central Japan. In 2011, the year this study was conducted, 132 of its 273 students were from overseas, with only a handful of these being native English speakers. The remaining students were Japanese. Regardless of their L1 (first language), the majority of students produced their academic papers in English, which is also the language used in many of the graduate school's seminars and lectures. Leki's (1992) description of international graduate students as having "disciplinary knowledge [that] may far exceed their ability to express that knowledge in writing in their second language" (11) is on the whole applicable to GSID students: many possess extensive knowledge of, and practical experience in, their field, but have had little formal instruction in academic writing.   At the time of this study, there was little opportunity for students to remedy this situation because academic writing classes were not provided by the school. Moreover, a questionnaire survey revealed that 12 of the 21 MA student respondents had never taken classes for writing academic English before entering GSID.

2.3 Providers of the English check

While university professors can often request proofreading assistance from native English-speaking colleagues, or can use a research budget to pay for professional proofreading services, their students may not have such access to a feedback provider, whether a NS or NNS. Furthermore, although the number of university writing centers in Japan has increased in recent years, the presence of such a center does not necessarily equate

with the availability of a proofreading service. Nagoya University's writing center, Mei-writing, is a case in point. On its homepage it states that its services "DO NOT cover native check" (block capitals in original)[2] . Like many other writing centers in Japan and around the world, Mei-writing positions itself primarily as a space where writers can discuss their work rather than be the "fix-it" centers that North (1984) lamented some decades ago.

While Mei-writing and a growing number of other writing centers have sought to avoid becoming, in the words of Reid (2006, 106), places where students can "'drop off' their drafts and return to 'pick up' the improved papers," it was partly to provide such a "drop off" service for graduate students that the author was employed by GSID. As far as the author could ascertain, GSID was the only faculty in Nagoya University that employed a fulltime English checker. According to several faculty members, a resident proofreader was, and still is, considered necessary because of the school's relatively large number of overseas students, the fact that much of the teaching is conducted in English and, most importantly, due to the fact that most of the academic writing emanating from the school is in English.

The most demanding part of the author's job was the proofreading of students' MA theses (the focus of this study, theses checking is explained in detail below), but he was also expected to check several PhD dissertations, and papers to be submitted to journals written by faculty and students (as a requirement of the doctoral course, students are required to publish two or more academic papers in peer-reviewed journals). One year after the study reported in this paper, academic writing classes were introduced to GSID, and the author's role expanded to include that of an EAP (English for Academic Purposes) teacher.

Towards the end of the calendar year most students were required to submit MA theses for an English language check. To assist in the checking of around 30 theses over a four-week period, a budget was available to the author to employ students as Teaching Assistants (TAs). At the time of the research reported here, however, university employment regulations made the recruitment of NSs for the TA position difficult. Consequently, the TA team was comprised mostly of NNSs, as can be seen in Table 1, which shows the personal data of the proofreaders (this term will be used interchangeably with *TAs* and *(English) checkers* in this paper) who participated in this study. The suitability of a candidate for the TA position was determined by an informal interview and a trial proofreading. No specific training was provided for the proofreaders, but they did receive a document, prepared by the author, containing a general description of what the job entailed.

---

[2] See <http://meiwriting.ilas.nagoya-u.ac.jp/services/tutorial-2/terms-and-conditions>

Table 1
Profiles of the English language checkers

Code Name	Sex	Country	Education in English--medium environment prior to arrival in Japan	English academic writing experience	Proofreading experience prior to this study
EM	F	Mexico	International school in Mexico	BA and MA theses. Two articles in a departmental in-house journal	Edited university newspaper in Mexico
DB	M	Brazil	No	Two papers published in a departmental in-house journal	Third year of proofreading in GSID
MR	F	Romania	No	MA	No
NB	F	Bulgaria	No	MA	Proofread Japanese to English translations
ZC	F	China	No	MA and one journal article	No
ZE	M	Ethiopia	4 years in a Canadian university	Three journal articles	Some proofreading in Canada for other overseas students. Third year of proofreading in GSID
EG	M	Ghana	Educated in English from primary school	Two journal articles. One book chapter	Checked students' papers when he was a university lecturer in Ghana, but "mostly for content"
OC	M	Cameroon	University in the Philippines and in Holland	Three journal articles	Checked the papers of other students in the Philippines
SA	M	Australia	Majoring in linguistics at a university in Sydney	Reports and essays for university courses	Checked classmates' reports and essays
KA	F	US	Majoring in linguistics at a university in US	Reports and essays for university courses	Checked friends' reports and essays

Note. All but three of the TAs were graduate doctoral students in GSID. Of those who were not, SA and KA were second-year undergraduate exchange students, and OC was a research student in another graduate school of Nagoya University. The age of the TAs ranged from 20 to 38.

2.4 Remit of the English check

At the time of this study, TAs were paid for 10 hours' work, or around ¥10,000, to check a single thesis (averaging around 100 pages). In this limited amount of time (and for this relatively paltry sum) they were advised, but not explicitly instructed, that their main job was to focus on language problems such as grammar errors, awkward or unidiomatic sentence structure, word choice, punctuation, spelling, mechanics, and the like. They were not required to concern themselves with content, which, they were told, was the responsibility of the academic supervisors. It is the case, however, that because the realization of meaning emerges through language, this content-form distinction is to some extent a "false dichotomy" (Hyland & Hyland 2003, 86).

Lillis and Curry (2010) refer to the various interventions that have a direct influence on text production as "literacy brokering" (87). One of these interventions, proofreading, is often considered to be the "last stage" of a piece of writing that has gone through several drafts (Scot and Turner 2008, 1). For many it connotes the idea of polishing a piece of text, but as Harwood, Austin, and Macaulay (2009) point out, the term "may be used differently by individuals to describe a range of interventions" (167). In this paper, the terms *proofreading*

and *checking* and *feedback* will be used interchangeably to refer to all the interventions made by the English checkers in this study.

3. Method

3.1 Participants

The students

A task involving students ranking and commenting on error feedback was conducted during the month-long period for MA theses checking. As stated earlier, the objective of this task was to illuminate students' feedback preferences. Of the 21 students who submitted MA theses, 15 students from ten countries (see Table 2) agreed to participate in the study. All were majoring in areas related to international development studies (GSID positions itself as an academic center preparing people for an "active role in the field of international development, cooperation and communication in the future"[3]).

Before handing these students' theses to the TAs, one short section (around 300 words) from the introduction of each thesis was chosen, and multiple copies of these sections made. The TAs assigned to check any one of the 15 theses from which the short extracts had been taken were instructed to skip the pages that had been allocated for the study. These pages had been clearly marked with pencil and demarcated with post-it notes. The writers of each of these 15 theses were thus returned a thesis containing a few pages that had not been proofread.

The Teaching Assistants

The TAs were divided into two groups, A and B (see Table 2), and the extracts were randomly assigned to either one of these groups. Each TA was handed a clear file containing seven (group A) or eight (group B) extracts and instructed verbally and in writing to proofread the extracts in the same way as they usually proofread complete MA theses. To encourage conscientious engagement with the task, the TAs were told that they would be paid an honorarium of ¥4000 to check the file; it was also stressed that the student writers would likely be using the feedback provided to revise the still unchecked pages in their theses. The TAs were not, however, given any details of the study; informing them that their feedback was to be ranked could have encouraged proofreading that was unusually zealous, or in some other way atypical.

Having two groups made it possible to ensure that the file extracts were not taken from theses that a certain TA had checked as part of his/her main proofreading duties. This guaranteed mutual anonymity between students and TAs and also made the task more manageable for both groups, particularly the students, for whom ranking and commenting on samples by all the checkers would have been too demanding.

3.2 An extra checker: GB

It can be seen in the Table 2 that, although ten TAs were recruited, there is a sixth checker, GB, in both groups. GB is, in fact, the author of this paper. He decided to participate in the study because it provided a unique opportunity to obtain students' views on his own corrective feedback (he placed himself in both groups A and B in order to maintain an equal number of checkers in each group). The author was careful to follow the same instructions he had given to the TAs (i.e., to check the MA extracts no differently from other proofreading jobs).

---

[3] See <http://ocw.nagoya-u.jp/index.php?lang=en&mode=l&page_type=gsid_info>

3.3 The student writers' files

A week was given for the TAs to check the extracts. After all their files had been returned, a code to identify the TA to the author (but not to the students) was written on each of the extracts that the TA had proofread. The extracts were then transferred into students' files. Each of the 15 students could thus be handed a file containing six copies of the same extract from his or her thesis with each extract having been checked by a different proofreader. A ranking table was inserted into this file, along with written instructions (shown in Appendix 1). Students were to rank each checker's feedback according to its usefulness, with 1 being the most useful and 6 the least, and also to write comments explaining the ranks they had assigned. As an example, Appendix 2 contains the table completed by Raahi, a student from India. The comments written by the student writers varied in length, but all wrote at least one sentence to justify each ranking.

4. Results

4.1 The rankings

The ranks given by each student, together with the average ranking (AR), are shown in Table 3 (the lower the AR, the better the overall evaluation for that checker). All names used are pseudonyms. Two main points of interest emerge from the quantitative data below. Firstly, KA and SA, the two NS undergraduate exchange students from the US and Australia respectively, received only mediocre average rankings. Secondly, there was sometimes surprising variability in different students' rankings for the same checker; for example, while five students in Group A considered GB's check to be either the most or second most useful, one student placed him second from bottom.

Table 2
Ranks awarded to each proofreader by students

Group A				Student Writers/Nationality				
Checker's code	**Aliz** Hungary	**Maly** Cambodia	**Dian** Indonesia	**Rupa** Bangladesh	**Miho** Japan	**Eri** Japan	**Maria** Venezuela	**AR**
GB	5	1	1	2	1	1	1	**1.7**
DB	1	4	2	1	2	2	3	**2.1**
ZE	2	3	5	3	4	3	2	**3.1**
KA	3	2	3	5	5	4	4	**3.7**
ZC	4	6	4	4	3	5	5	**4.4**
OC	6	5	6	6	6	6	6	**5.8**

Group B	Student Writers/Nationality								
Checker's code	Kade Indonesia	Ratana Thailand	Sukh Mongolia-	Yujin Japan	Hanna Japan	Raahi India	Kolab Cambodia	Azat Turkmenistan	AR
EM	1	1	1	3	5	4	1	3	**2.3**
GB	3	5	6	1	1	1	2	1	**2.5**
SA	6	3	2	4	3	3	3	4	**3.5**
NB	2	2	3	6	2	5	4	5	**3.6**
MR	5	4	4	2	4	2	6	2	**3.6**
EG	4	6	5	5	6	6	5	6	**5.3**

Note. Students ranked the usefulness of the TA's corrections from most (1) to least (6) useful.

4.2 Analysis and commentary of the qualitative data

The author read the students' comments and, using different coloured highlighter pens, identified key words and phrases in an iterative process until a number of main themes emerged from the data (Hanna's comments, the only comments to be written in Japanese, were first translated into English). The 300 word extracts from the MA theses, totaling 90 in all (six extracts for each of the 15 writers), were also analyzed so as to flesh out, exemplify, and sometimes counter, the comments made by the students. Because of space restrictions it is not possible to report the results of all the proofreaders, but what follows is the author's analysis of the evaluations for 9 of the 11 proofreaders, focusing particularly on the elements (feedback points) of the English check that were positively or negatively rated. A number of excerpts from the MA theses are included to accompany the analysis. All student writers gave permission for their work (i.e., extracts from their thesis and ranking sheets) to be reproduced in this paper, as did the checkers, whose handwritten comments can be seen in the following analysis.

GB

Most of the students who positively evaluated GB commented in some way on the feedback's comprehensiveness. The Japanese students, in particular, seemed to appreciate the relatively detailed checking. Eri commented that "GB corrected articles very carefully…which made me very aware of my bad habits in English." The traditional emphasis in Japanese education on grammatical accuracy may have influenced the Japanese writers' expectations of what constitutes useful feedback. Attention to detail possibly accorded with their "great concern" for error-free work (Hyland & Anan 2006, 515).

When the sentence structure was unclear, GB sometimes wrote a few alternative candidate sentences prefaced by the phrase *Do you mean…?*. This feedback technique was positively evaluated by a number of students, including Hanna, who commented that "the proofreader thought about what writer wanted to say and provided choices." The comments for GB and several other checkers suggest that students were sensitive to whether their writing had been correctly interpreted, and favourable evaluations were received when feedback was perceived to have tallied with the writer's intended meaning. Miho, for example, commented that "GB changed sentences to what I want to say exactly."

Aliz was one of the students whose comments on GB were mostly negative. She ranked him fifth despite his corrections being, on the whole, rather more extensive than DB's, whom Aliz ranked top. A comparison of these two checkers' feedback in the paragraph below (Extract 1a and 1b) reflects the difference in intervention seen throughout the extract. GB's intervention could, however, be regarded as excessive; for example, finding the tone of *merely* too dismissive did probably not justify his changing it to the more neutral *only* without a note of explanation. In addition, although the parentheses on the last line were intended to indicate that the enclosed words could perhaps be deleted, without an accompanying written comment, this uncoded feedback (i.e., feedback which signals an error or potential problem but leaves the writer to diagnose and correct it) probably did little more than confuse.

Excerpt 1a.   GB's feedback on Aliz's extract

Excerpt 1b.   DB's feedback on Aliz's extract

However, Aliz's main reason for awarding GB such a low rank can be found in the comment marked *c* that she made on the ranking table (Excerpt 2a). In an informal follow-up interview, Aliz—who was unaware that the interviewer was also the proofreader GB—elaborated on comment *c*. She expressed annoyance at his comments on a footnote that she believed did not need correcting, because it was a quotation from an authoritative outside source (see Excerpt 2b). She added in the interview that such checking was not only "unnecessary," but was a sign that GB was not "reading my paper carefully."

Looking again at this particular piece of feedback, Aliz's criticism of GB was justified; however, rather than perfunctory reading of her extract leading to this feedback being given, it stemmed, at least in part, from what could be called *fault-finding feedback*. Confronting writing that was unclear and/or error dense—not just in patches but for long stretches (as GB found Aliz's to be)—GB would slip into a mode of proofreading that tended to assume the worst. Writing that was not immediately clear to him at the time, such as the footnote in Excerpt 2b, would be corrected, or at least flagged to the writer. This kind of

negatively-biased proofreading possibly led to other instances of unnecessary or erroneous corrections by GB during his time at GSID. Aliz's feedback drew GB's attention to, and helped him avoid, falling into this fault-finding mode.

Excerpt 2a. Aliz's feedback comments on GB

Excerpt 2b. GB's "unnecessary" correction and comments on a footnote

Sukhr demoted GB to sixth place. A comparative analysis of errors revealed that despite GB offering 16 corrections and comments, he was ranked below EG who made only seven corrections (of which only two were found to be grammatically sound). What could explain this rejection of a relatively thorough proofreading job? Sukhr's criticism of GB's feedback as having "so many questions and recommendations rather than English check" may offer a clue. Completing the MA is for many students in GSID is the culmination of a two-year struggle with English writing. Rather than receiving comments and questions that required them to further grapple with the language, such students may prefer feedback facilitating easy revising; that is, direct feedback with the error identified and the correct form provided.

EM

Those who ranked EM highly commented on the clarity of her corrections. Rather than crossing out whole lines, she generally concentrated on changing individual words and improving clarity with the addition of punctuation. Ratana, one of the students who ranked her top, attested to the way she avoided appropriation while correcting register: "The checker did not just change sentences to be their own style, but more like adapted student's writing to sound more proper in formal, in academic way." EM was also ranked top by Sukh, who commented that she "changed to correct words by him/herself;" praise, most probably, for her unambiguous direct feedback. Excerpt 3 shows how EM interpreted Sukh's intended meaning and corrected with minimum disruption to the sentence structure.

Positive evaluation for EM, however, was not universal. She received negative comments from two of the eight students for correcting with a blue rather than a red pen. One of these, Hanna, ranked EM fifth for using the "wrong colour," despite also positively evaluating her feedback as being "easy to read." The importance Hanna placed on ink colour perhaps reflects the close association in Japanese culture of red with making corrections (the commonly used idiom *aka wo ireru*, literally "to add red," refers to the act of correcting).
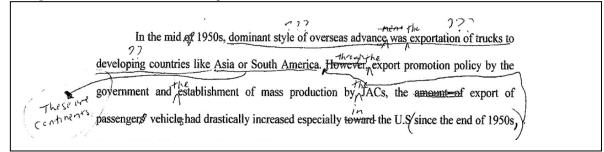
Excerpt 3. EM's correction of Sukh

> In Owen Fiss's article "The right degree of independence", considered the one influential view of judicial independence, as a three strands: independence from the parties to a case, independence from influence by other members of the judiciary, and independence from politics[10].

SA

While SA was positively evaluated by several of the students for providing, in the words of Kolab, "many useful editorial suggestions," his feedback was negatively evaluated by others as lacking in specificity. In Extract 4a, the question marks and underlining indicate that something is not right, but the writer is given minimal explicit feedback as to the nature of the problem. It is perhaps not surprising that Yujin expressed some dissatisfaction with SA's feedback: "Correction without comments. Difficult to know how should I correct." EM's correction (Excerpt 4b), on the other hand, would leave the writer in no doubt as to the revisions required.

Excerpt 4a. SA's correction of Yujin's extract

> In the mid of 1950s, dominant style of overseas advance was exportation of trucks to developing countries like Asia or South America. However, export promotion policy by the government and establishment of mass production by JACs, the amount of export of passengers' vehicles had drastically increased especially toward the U.S since the end of 1950s.

Excerpt 4b. EM's correction of Yujin's' extract

> In the mid of 1950s, dominant style of overseas advance was exportation of trucks to developing countries like Asia or South America. However, export promotion policy by the government and establishment of mass production by JACs, the amount of export of passengers' vehicle had drastically increased especially toward the U.S since the end of 1950s

Raahi's comment indicates another reason for SA's negative evaluation: his occasional insensitivity to the disciplinary context: "At one place, the correction was made where the content didn't remain the same. The context should be well understood before going for any correction (e.g. wage 'costs')." The example to which he is referring is shown in Excerpt 5. The comments for several other checkers also indicated that erroneous feedback tends to invite negative evaluation, probably because it alters the writer's intended meaning. It is also possible that such feedback may lower a writer's perception of the checker's competence, which in turn can, as Burrough-Boenisch (2003, 233) suggests, affect the readiness of an author to accept the feedback as a whole.

Excerpt 5. SA's correction of Raahi's extract

> However, the more or less continuous above average hike in wage ~~costs~~ of software professionals over the years has led to the erosion of the cost advantage that India has enjoyed in the past and put pressure on profit margins of the IT firms. According to a salary survey conducted by Dataquest

Kade, who assigned SA to the lowest rank, commented thus: "The revisions were too little, almost insignificant. There were several grammatical errors in the sample of thesis after the revisions." Indeed, the analysis of Kade's MA sample does show that some errors were left uncorrected, errors that a native speaker, with the advantage of an "intuitive knowledge of English" (Lee 2009, 388) would be expected to spot. This omission, however, may have been deliberate, for in the interview SA commented that, when faced with error-dense text, he would sometimes ignore those errors that could "still be understood."

The second sentence of Kade's comment appears to be referring to erroneous correcting. A few mistakes have indeed been introduced by the checker himself, including the examples in Excerpts 6a and 6b. Wrongly changing the possessive *its* to the contraction *it's* (*with it's higher technology*) is possibly a careless slip, although, as Peck and Coyle (2005, 84) point out, confusing *its* and *it's* is a common error among NS. The erroneous correction of *economic of scale* to *economy of a scale* likely stems, as was the case with Raahi's extract above, from not knowing (and subsequently not checking) the meaning of discipline-specific terminology (in this case *economies of scale*).

Excerpt 6a. SA's correction of Kade's extract

> the domestic market for competition with foreign products. The second reason is that liberalization will attract foreign investment to produce greater total national output with it's higher technology.

Excerpt 6b. SA's correction of Kade's extract

> tariff elimination. Domestic industries claim that they need a period of time to be ready for the competition with imports. They ask for government intervention that protect and allow them ~~achieving~~ to achieve adequate levels of efficiency and ~~economic~~ economy of a scale for international competition.

This proofreader received negative comments from all seven evaluators for the illegibility of her handwriting. She was also negatively assessed by several students for offering few concrete suggestions. Miho, for example, commented that "[KA] asked to rewrite, but doesn't tell what is bad of that sentence." Excerpt 7 is possibly an example of the kind of feedback to which Miho was referring.

Excerpt 7. KA's check of Miho's extract

With rather limited experience of L2 writers, both KA and SA perhaps assumed that alerting students to problems or requesting clarification would be sufficient to enable them to correctly diagnose the problem. However, as Leki (1992, 122) points out, such feedback "may make no sense…if the text already seems clear to the writer." Editing and proofreading can be demanding, particularly when errors render large portions of a text difficult to decipher or incomprehensible, and KA remarked in an interview that she felt "frustrated" at having to tackle writing "full of basic grammar mistakes." This frustration may have found release in terse comments and in the relatively messy handwriting in which they were written.

While, as Belcher (2009) points out, feedback delivered in "anger and frustration" (224) may not be heeded by the writer, KA was one of the few proofreaders to offer any evaluative comments at all, perhaps indicating that she was engaging with the text more than the other checkers (an example of one of these comments is in Excerpt 7 above; difficult to decipher even in the original, it reads "don't start w[ith] *Besides* as it makes a fragment.")

ZC

Eri ranked ZC fifth, commenting that "articles have not really been checked." A look at Eri's writing (Excerpt 8) shows that this is indeed the case: in the first four lines, for example, four definite articles are needed, but ZC offered no feedback at all. All the other checkers, apart from OC, inserted one (KA), two (ZE), three (DB) or four (GB) of the required definite articles (before *YMCA*, *Red Cross*, *food crisis*, and *1980s*). Deckert (2004,106) points out that the frequency with which articles occur makes them particularly difficult for NNSs in general; the fact that Chinese do not have articles would likely have compounded this difficulty for ZC and may partly explain her inability to spot their omission in Eri's writing.

Excerpt 8. ZC's check on Eri's extract

The comments overall suggest that students are generally more sensitive to, and critical of, erroneous feedback related to disciplinary knowledge than to grammar. Dian's response to ZC
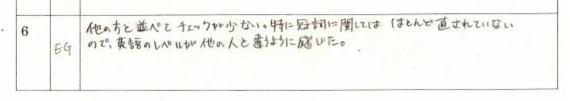
is illustrative of this. Dian writes that she doubts ZC is from "economic field [because] he/she corrected some economic terms which are already correct according to economic definition." In Excerpt 9, which shows the feedback to which Dian is referring, ZC does seem to be hedging her correction with the use of question marks, tentatively offering an alternative or urging the writer to double check; nevertheless, ZC's questioning of what is a key term in Dian's thesis may have undermined somewhat the writer's confidence in the checker.

Excerpt 9. ZC's check on Eri's extract

EG was ranked fifth or sixth by all but one of the students. Comments suggest that his poor ranking was due mainly to insufficient feedback. Hanna commented that "compared to others there are very few corrections…I have a feeling that this TA's level of English is different from that of the others" (see Extract 10 for original Japanese comment). By "different" (translation of the Japanese word *chigau*), Hanna was likely implying that it was lower. EG did, in fact, have native-like proficiency in spoken English and is a published author. Rather than being a reflection of his English level, EG's poor ranking was more likely due to affective factors. In the interview, EG remarked how the work in GSID was different from checking papers in Ghana, where he had lectured: "In Ghana I could understand, but here I had to think for the writer…to break a code." He talked about making "compromises" when faced with writing he could not understand and of feeling "conflict" because "logical process and argument were off-limits," a comment which suggests that making the correcting of content "off limits" adversely affected how EG—and perhaps other checkers— approached their work. All in all, the Ghanaian described the whole experience as "tiring" and "frustrating."

Extract 10. Hanna's comments on EG

OC

Six of the seven students ranked OC bottom. Maria's remark that "there are almost no corrections" represents the sentiment expressed in all the comments. The paucity of corrections is clear when OC's check of Maria's extract is compared with two other checkers: while GB had made around 60 corrections or comments and DB 30, OC made only 4. Moreover, Maria expressed doubts over the veracity of the few corrections made, doubts that were well-founded: each of the four pieces of feedback was erroneous (Excerpt 11 shows one

of these). OC's use of, as Aliz put it, "confusing signs," such as a circle around a word without any explanation, was another reason given for his low rank.

Excerpt 11. OC's check of Maria's paper

In this regard, "donor relationships are central to the lives of most NGOs" (Smillie &Hailey, 2001, 32)[2] in need of funds and assistance to scale up[3], but relationship not always are smooth and fair for both actors. Interference from the donor and clashes in their

Harwood, Austin, and Macaulay (2010) report that proofreaders in a British university sometimes refuse to look at papers by students with a very low standard of English because of "the scope of the changes that would be necessary to make the text readable"(58). For a TA faced with an error-laden piece of writing that he or she feels duty-bound to check, resorting to a strategy of disengagement, whereby a few token corrections are made on each page—what could be called "gesture correcting"— is an understandable, if unsatisfactory, response. OC, however, provided only cursory feedback for all his extracts, regardless of the text's comprehensibility.

Without being confrontational, I sought in the interview with OC to determine if he considered his proofreading to have been satisfactory. In contrast to EG, who stated that he found the job dispiriting, OC's remarks were generally positive. By the interview's end, the author was no closer determining whether OC saw the job primarily as a way to earn some "easy money," or whether he undertook the English check conscientiously, but that his concept of corrective feedback diverged considerably from that of the students whose writing he checked.

5. Discussion

This study found that the amount and form of corrective feedback for a piece of text differed depending on the provider of that feedback. This is to be expected since, as Hyland and Annan (2007) point out, "each of us comes to a text with a certain competence in the language and a particular set of expectations for grammaticality, organization, style and so on" that will influence "judgments about acceptability"(517). The diverse cultural and linguistic backgrounds of the checkers in this study would likely have accentuated such differences in competence and expectations. Moreover, the checkers in this study, unlike student tutors in many university writing centers, received no common preparatory training that could have equipped them with a shared set of strategies for feedback provision. While this lack of training may be a serious shortcoming of the English language-check service, it did provide the study with multiple distinct proofreading styles.

The results of the ranking task suggest that there was sometimes a striking degree of variation in students' evaluations of the same checker. Nevertheless, from the analysis of students' comments a preference for certain kinds of feedback did emerge. Regarding those feedback constituents that were positively evaluated, the majority (9 or more) of the 15 students indicated that feedback was considered useful when:

a)   It was specific, comprehensive, and detailed.
b)   It was clear and easy to understand.
c)   It offered comments and/or suggestions to accompany the corrections.
d)   A choice of different corrections/reformulations was offered.
e)   The corrections did not alter the author's intended meaning.

In contrast, a negative evaluation was generally incurred if the student considered that:
a) Changes had been made that were erroneous, unnecessary, and/or ineffective.
b) Too few corrections had been made.
c) The corrections were difficult or impossible to read.
d) Uncoded feedback such as underlining, circles around words, or question marks had not been accompanied by any clarifying comments.
e) Comments were terse and/or critical.
f) Questions had been asked that did not help the writer determine the nature of the error.

Several of the findings in this study concur with the literature on error feedback. The occurrence of erroneous corrections, for example, was also observed by Lee (2004), who found that half of the errors marked by non-native teachers were not accurate. The negative comments directed at OC and others for unclear feedback is in agreement with surveys that found "both L1 and L2 students resent cryptic codes or symbols that they do not understand" (Ferris 2002, 69). The preference for direct feedback indicated by the participants in this research was shown also by Shin (2007) in her study of Korean ESL students. This finding, Shin asserts, should "not be surprising" since an underlining or a question mark cannot be expected to trigger a corrective revision by the writer when "certain necessary words or structures are simply not there" (363) in the learners' linguistic repertoire.

Student writers in GSID were aware that the proofreading of their theses would likely be done by fellow NNSs, and this knowledge could have negatively influenced their perception of the corrections they received. In light of Gilmore's (2009) comment that one of the disadvantages of L2 peer feedback is that it "may be perceived as less valuable" than that given by the teacher (364), it may be of interest to consider whether the reaction of students who commented negatively on GB's check would have differed had they been informed of his NS status. It is conceivable that this knowledge alone could have positively influenced their perception of his feedback. This also raises the issue of how a writer's expectations of feedback can differ depending on the source of that feedback. Discussions with students in GSID indicated that, although subject (as opposed to language) teachers provide valuable comments on the content of a particular piece of writing, they generally make few, if any, corrections of grammar, spelling and other such errors. Because of its "rarity value," language-related feedback from subject lecturers would perhaps be more appreciated than the same feedback provided by proofreaders from whom it is expected.

5.1 Classroom applications
The ranking and evaluation task could be used by writing teachers who wish to gain insights into how their corrective feedback is perceived by students. Small groups of teachers could collaborate to replicate the task, preferably with students who have never received written feedback from any of the participating teachers. The author has also successfully adapted the task for use as a group peer-editing activity in writing classes.

6. Conclusions

6.1 Changes in GSID influenced by this study
As mentioned earlier in this paper, one reason for the author including himself in the study was to gain insights into how students regarded his feedback. Their comments influenced in several ways his subsequent approach to the English check. He is now, for example, careful not to let any frustration with the writing reveal itself in his feedback. He now uses more "mitigated forms" (Coffin et al. 2002, 118), such as "I think it may be better to…," to soften

the tone of his comments, although it should be pointed out that Hyland and Hyland (2001, 185) found such a softened, hedged tone "carries the very real potential for incomprehension and miscommunication."

The findings of this research helped provide impetus for the introduction of an academic writing course into GSID's curriculum. Explicit focus on form and continuous teacher feedback in these classes is aimed at making students aware of their most common errors. It is hoped that priming students in this way will help them to more efficiently utilize the feedback given in the English check since "expecting a student to meaningfully correct a grammatical error before being introduced to that structure is wishful thinking at best"(Evans et al. 2010, 453).

Moreover, by also instilling an understanding and appreciation of the broader principles of academic writing, the course can serve to counteract one of the possible negative consequences of the English check service; namely, that its presence reinforces the perception that the key to producing a good paper lies in fixing the surface features. This is a view that Swales and Feak (2000, 6) assert does "more harm than good, especially in the long term."

As part of the assessment for the school's academic writing course run by the author, students were required to send a short piece of writing done for one of their content courses to so called "writing consultants," native-speaking exchange students studying at Nagoya University. These consultants not only checked the writing, but also met with each student writer to discuss his or her most common language errors (for which consultants were paid around ¥4,000 a student). At the end of the course, each student gave a short presentation in which they reflect on the feedback they had received throughout the course, and particularly by their writing consultant.

Most students in the writing class expressed great enthusiasm for these face-to-face meetings, considering them extremely useful and illuminative. Particularly for students whose writing contains many global errors that render portions of the text unclear, conferencing between the writer and checker is vital for clarifying the writer's intended meaning. It is also an opportunity for the writer to ask for clarification from the proofreader on any corrections or comments made. Perhaps the ranking of the proofreaders by the students would have been considerably different if it had been based not just on textual feedback, but also on the feedback received during a personal meeting with a proofreader. Future research could compare writers' evaluation of proofreaders based only on written feedback with evaluation of the same proofreaders following a post-proofreading conference.

Although there is much more to academic writing than linguistic accuracy, improving the students' control of the language is important. Too many errors can obscure intended meaning and render any substantial textual improvements difficult at the proofreading stage, at least in the number of hours for which the English checkers are being paid, and without reformulation of the text to an extent that would raise "ethical questions around who is ultimately doing the writing" (Scott and Turner 2008, 3).

While it is expected that writing classes will help students to develop self-editing skills, discussions with students indicate that the English check is regarded as an important step in producing a finished paper that can be shown with pride to prospective employers. The findings of this study, however, raised questions about the usefulness of much of the feedback the students were receiving and prompted the implementation of changes to the English check system at the end of 2012.

Proofreaders now check fewer pages for more money; specifically, rather than checking a complete thesis, which is often well over 100 pages, for ¥10,000, they now look at between 30 to 50 pages for ¥13,500 (as of February 2013). It is hoped that making the workload more manageable will encourage more conscientious checking. While students are instructed to use the corrections and comments made by the proofreader to self-edit the remaining pages of the

thesis, it is likely that a partial check will result in errors remaining in the final paper. Yet, rather than seeking to "correct" all the errors before submission—an aim which was always unrealistic for those students with a lower-level of writing proficiency — proofreaders are now told to place emphasis on making students aware of their main errors and alerting them to text that is difficult or impossible to understand, so that they can work on improving their writing *after* they graduate from the department.

In other words, the English check has become more explicitly didactic. The move to on-screen checking (see below) has helped to accommodate this new emphasis because comment boxes make it easier for the proofreader to give advice on specific aspects of writing by, for example, providing links to relevant URLs. (See Harwood et al. 2012 for a discussion of strategies used by proofreaders in a British university to provide feedback that aimed to be formative.)

A second change in GSID is an administrative one that has made it possible to hire people from outside Nagoya University. Several NSs (all teachers of English as a foreign language) with experience in proofreading were hired for the last English check the author supervised (from December 2012 to February 2013). The situation as it was before the author's departure, however, raised issues of fairness since some students had their papers checked by NSs experienced in checking L2 writing, while others received feedback from less experienced fellow students. In fact, one set of international students proofing the writing of another could be considered as being "peer-editing" rather than "proper" proofreading.

6. 2 Limitations and potential for further research

In this present study multiple student writers passed judgment on the feedback of the same English checker; however, because each student evaluated only his or her own proofread MA extract, the conclusions drawn from the analysis of the feedback need to be treated with caution. Moreover, comparison of feedback by the author was somewhat holistic and instinctive; the analysis could have been enriched had it been made more systematic by, for example, using a checklist on which the occurrence of each error was marked.

Although the number of proofreaders was clearly too few to draw any generalizable conclusions about the relative merits of native and non-native proofreaders, the results of the ranking task suggest that NS status alone is not necessarily a guarantee of superior proofreading skills and serves as a reminder that the general view "that any native speaker can do it" should not be blindly accepted (Scott and Turner 2008, 4). It is perhaps not surprising that inexperienced NSs (both KA and SA were undergraduate students with little experience of academic writing) would be ranked lower than NNSs conversant with the writing problems faced by fellow L2 writers. If a comparison of NS and NNS in terms of the effectiveness of their feedback is to be studied, various factors, such as level of education and editing experience, will need to be considered.

Such a study would also be difficult because, as Gunnel and Shaw (2003) point out, "it is often difficult to tell who is to count as a native speaker" (39). This is exemplified by the response of the Ghanaian TA, EG, to the following interview question: "Do you consider yourself to be a non-native English speaker?" His reply illustrates how difficult the demarcation of NS and NNS can be. He first commented that he considered himself to be a NS only "in certain environments" including in GSID, "where many of the students had not grown up with English"; "in the midst of Americans," however, he felt less sure of his status. Further, he expressed confusion over the term "native speaker" itself: "I actually don't know what they mean by it. If they mean mother tongue, I cannot claim English as my mother tongue, but in terms of learning processes up to a certain level, I can say that I'm a native speaker because I can write and have published papers."

In this study all the proofreaders were required to provide hard-copy feedback. There is often, however, a great difference between feedback written by pen and that given electronically. For example, while EM's feedback on the extracts was rather minimalist, it was considerably more extensive when she corrected entire MA theses electronically. A likely reason for this is that on-screen editing facilitates more radical intervention such as the swapping around of paragraphs. Since this study was conducted the English check has become wholly electronic. Students' papers are exchanged by email and are checked using the track changes function on Microsoft Word. A replication of this study with electronic proofreading should reveal insights into the preferences of students for a mode of feedback provision that is rapidly taking over from the red pen.

Finally, an implicit assumption made in this study is that what students perceive to be useful will actually help to improve the communicative effectiveness of a particular piece of writing. Ferris (2002) reports that the "evidence is fairly conclusive" that "student writers have generally been successful in producing more accurate revisions in response to error feedback" (15). Future research could analyze final versions of MA samples to ascertain the way in which writers act on multiple sources of feedback and discern whether high ranking feedback had actually been the most effective for improving the text. The change to electronic checking could potentially help illuminate the ways in which students pick and choose corrections made by different proofreaders and incorporate this feedback into their writing. This could perhaps be done by having students indicate in dialog boxes the source of the feedback informing each revision. In an increasingly competitive academic environment, proofreader input can be a decisive factor in getting a good grade or being published in a journal. Further research on such input and the way it is perceived could provide valuable pointers for those charged with improving L2 writing.

References

Belcher, Wendy, L. 2008. *Writing your Journal Article in Twelve Weeks: A Guide to Academic Publishing Success*. Thousands Oaks, CA: Sage Publications.

Burrough-Boenisch, Joy. 2003. "Shapers of Published NNS Research Articles." *Journal of Second Language Writing*. 12: 223-243.

Chandler, Jean. 2003. "The Efficacy of Various Kinds of Error Feedback for Improvement in the Accuracy and Fluency of L2 Student Writing." *Journal of Second Language Writing*. 12: 267–296.

Coffin, Caroline, Mary. J. Curry, Sharon Goodman, Ann Hewings, Theresa. M. Lillis, and Joan Swann. 2002. *Teaching Academic Writing: A Toolkit for Higher Education*. London: Routledge.

Deckert, Sharon. K. "A(n)/The/Ø Article About Articles," in *ESL Writers: A Guide for Writing Center Tutors*, ed. Shanti Bruce and Ben Rafoth. 2004. Portsmouth: NH: Heinemann Boynton/Cook, 105-115.

Ferris, Dana. R. and Roberts Barrie. 2001. "Error Feedback in L2 Writing Classes: How Explicit Does it Need to Be?" *Journal of Second Language Writing*. 10: 161-184.

Ferris, Dana. R. 2002. *Treatment of Error in Second Language Student Writing*. University of Michigan Press: Ann Arbor.

Ferris, Dana. R. 2003. "The 'Grammar Correction' Debate in L2 Writing: Where Are We, and Where Do We Go From Here? (And What Do We Do In The Meantime . . .?)." *Journal of Second Language Writing.* 13: 49–62.

Gilmore, Alex. 2009. "Using Online Corpora to Develop Students' Writing Skills." *ELT Journal.* 63: 363-372.

Gunnel Melchers, and Philip Shaw. 2003. *World Englishes: An Introduction*. London: Arnold.

Harmer, Jeremy. 2007. *The Practice of English Language Teaching*. Harlow: Pearson Education.

Harwood, Nigel, Liz Austin, and Rowena Macaulay. 2009. "Proofreading in a UK University: Proofreaders' Beliefs, Practices, and Experiences." *Journal of Second Language Writing.* 18: 166-190.

Harwood, Nigel, Liz Austin, and Rowena Macaulay. 2010. Ethics and Integrity in Proofreading: Findings from an Interview-Based Study. *English for Specific Purposes.* 29: 54-67.

Harwood, Nigel, Liz Austin, and Rowena Macaulay. 2012. Cleaner, Helper, Teacher? The Role of Proofreaders of Student Writing. *Studies in Higher Education* 37: 569-584.

Hyland, Ken, and Eri Anan. 2006. "Teachers' Perceptions of Error: The Effects of First Language and Experience." *System.* 34: 509-519.

Hyland, Ken, and Fiona Hyland. 2001. Sugaring the Pill: Praise and Criticism in Written feedback. *Journal of Second Language Writing.* 10: 185-212.

Hyland, Ken, and Fiona Hyland. 2006. "Feedback on Second Language Students' Writing." *Language teaching.* 39: 83-101.

Lee, Hee-Kyung. 2009. "Native and Nonnative Rater Behavior in Grading Korean Students' English Essays." *Asia Pacific Educational Review.* 10: 387-397.

Leki, Ilona. 1992. *Understanding ESL Writers*. NH: Boynton/Cook: Portsmouth.

Curry, Mary J., and Theresa Lillis. 2010. *Academic Writing in a Global Context*. Abingdon (UK): Routledge.

Norman, W. Evans, James K. Hartshorn, Robb M. McCollum, and Mark Wolfersberger. 2010. "Contextualizing Corrective Feedback in Second Language Writing Pedagogy." *Language Teaching Research.* 14: 445-463.

North, Stephen M. 1984. "The idea of a writing center." *College English.* 46, no.5: 433-446.

Peck, John, and Martin Coyle. 2005 *Write is Right*. Palgrave/Macmillan: Hampshire.

Reid, Joy M. 2006. *Essentials of Teaching Academic Writing*. Houghton Mifflin: Boston.

Scott, Mary, and Joan Turner. 2008. "Problematising Proofreading." *Zeitschrift Schreiben*. http://www.zeitschrift-schreiben.eu/Beitraege/scott_Proofreading.pdf

Shin, Sang-Keun. 2008. "'Fire Your Proofreader!' Grammar Correction in the Writing Classroom." *ELT Journal*. 62: 358-365.

Swales, John.M., and Christine B. Feak. 2000. *English in Today's Research World: A Writing Guide*. Ann Arbor: The University of Michigan Press.

Truscott, John. 1996. "The Case against Grammar Correction in L2 Writing Classes." *Language Learning*. 46: 327-369.

Appendix 1
Instructions to students participating in the ranking task

Dear Raahi

As part of a research study, a few pages of your thesis were given to six proofreaders to be checked. Please use these six different versions to help you correct the relevant pages of your thesis. I would also like to ask you do the following task. The task has 2 steps:

Step 1 **Rank the 6 versions according to how useful they were to you**
Think about which of the versions you found the most useful or least useful for revising your paper. Because we cannot reveal the names of the proofreaders to you, each proofreader has been given a **code,** which you will see written in the **top left hand corner** of each version. Please write the code in the *TA Code* column in the table (see "Ranking and Impression Task" sheet) according to how you useful you found each version. For example, if you found the version checked by the proofreader whose code is *EG* most useful, put *EG* next to 1 in the *Rank* column. On the other hand, you would write the code of the version you found least useful next to 6, and so on.

Step 2 **Write why you ranked the versions in this order**
In the "Impressions" column, write what was good or bad about each version. So, for example, for the version you ranked 1, write why you found it so useful. For the version you ranked as 6, what was wrong with it?

Do not worry so much about correct grammar when you write your impressions (this is not a test!). Just write freely the ideas that come to mind.
日本語で書いてもかまいません。

Please return this folder together with the completed task sheet by＿＿＿.

Thank you for your cooperation

Appendix 2
A completed ranking task table

Ranking and Impression Task

Name   Raahi   Nationality Indian

Note. **Rank 1** will be the English check that you found the most useful and **Rank 6** will be the English check that you found the least useful

The **TA Code** is in the top corner of each English check version

Rank	TA Code	Impressions (Why did you find this English check useful/not useful? What were the good/bad points of this English check?)
1	GB	This English check was the best among available choices. I found it the most useful because — ① The content remained unchanged ② An alternate option was proposed wherever possible. ③ Not only grammar, but the format of the thesis, style etc. were also checked. ④ The corrections were easy to follow.
2	Mf	This English check was also useful for me because — ① The content remained unchanged ② The corrections were easy to follow ③ Some good explanations were made with regard to some mistakes. However, at one place unnecessary correction was made 'of the sector' was marked incorrect, which could change the context.
3	SA	I found this check the third useful. However, some places it was difficult to follow. At one place, the correction was made where the content didn't remain same. The context should be well understood before going for any correction. (e.g. wage 'costs')
4	EM	I found this check useful, but some places the corrections are difficult to follow. And, one correction is made without considering the context (e.g factor of production) At places alternatives are not provided.
5	NB	I found this check less useful because very few corrections are made compared to the other English-checks.
6	EG	I didn't find this useful because one of the corrections could lead to a change in the context and meaning (e.g. 'cost' competitiveness). Such kind of corrections need to be avoided. Very few corrections are found compared to others