INVESTIGATING THE REVISION PROCESS IN A WEB-BASED PEER REVIEW SYSTEM USING MACHINE LEARNING

Djuddah A.J. Leijen

Centre for Academic Writing and Communication, College of World Languages and Cultures, University of Tartu (ESTONIA)

Abstract

This paper aims to provide an overview about the challenges of investigating the revision process of second language writers using a web-based peer review system. Research has indicated that writing and revising is a difficult task for L2 writers. In order to gain a better understanding of the peer-feedback processes and the effects these have on revision, this study classified different features of peer feedback to determine how these different features can be used to predict whether students revise their text according to the features presented in the feedback comment. This study proposes a method (machine learning) that can be used to investigate the effectiveness of peer review on L2 writing that will make the investigation of large datasets, collected by web-based peer review systems, easier in the future.

Keywords: Writing, revision, web-based peer review system, machine learning.

1 INTRODUCTION

Web-based peer feedback to support the development of student writing is commonly used to support the writing and revision process. However, certain features of peer feedback remain ambiguous. For example, in the context of second language writing, how peer feedback should be organized and the effect this organization has on the writing process remains unclear. (e.g. [1], [2], [3], [4], [5], [6]). According to Goldin, Ashley and Schunn [7], the multidimensionality of peer feedback makes for an infinite set of different configurations. These configurations are reflected by the multitude of research conducted on these specific educational settings. Within this scope, however, there is a need to further investigate feedback in second language writing settings using web-based peer review systems to support the revision process.

Second language (L2) writing research is increasingly influenced by L1 writing research as a result of non-English speaking higher educational institutes having to integrate English as a language for teaching in learning, besides their native language of instruction, resulting in more second language learners being exposed to second language writing as a tool for assessment. Furthermore, technological advances, specifically in global web-based learning opportunities, have provided more accessible data for second language writing research, which in turn have provided students and teachers a more dynamic platform to practice, learn, and share writing tasks. One of these platforms, web-based peer review systems, has received a great deal of attention in L1 writing research in the past few years. Web-based peer-review systems (e.g. SWoRD, CPR, Eli, PeerMark, MyReviewers) are web-based tools that offer students the possibility to experience writing as a process; specifically, as a series of drafts that receive feedback for improvement before the final submission deadline. These systems can be integrated in any classroom setting, of any size, and for any discipline. Feedback is most often given by multiple peers, with or without the intervention of a teacher. From a research perspective, these systems have opened up a whole new chapter for writing research, primarily as these systems collect and store huge amounts of data which would previously have had to be collected manually. More significantly, these systems have enabled researchers to ask different types of questions and to question existing knowledge.

In order to gain a better understanding of the peer-feedback processes and the effects these processes have on the students text for second language writing, it is important to systematically classify different types of feedback students produce and how different classifications act in the feedback and influence students writing. These classifications are compared with results obtained in L1 writing studies and the limited amount of research in L2 writing studies to determine whether web based peer-review systems require a different approach for L2 writers, or whether L1 research is transferrable. The results of this study will provide a classification more specific to L2 writers using peer feedback in a web-based peer-review system and clarify whether the linguistic choices students
make in their formulation of their feedback to their peer influences implementation. In turn, these results could, depending on the outcome, support the development of, either current web-based platforms to support L2 writing, or new web-based platforms specifically targeted to L2 writing. The effects of the third point will depend on the type of results the second point produces, but will be of great value for second language writing educators, and higher education institutes, who have a large student body writing in an second language, and are exploring or using web-based peer review systems in their teaching, and for L2 students to produce better guided peer feedback instructions before they engage in web-based peer reviewing activities.

To gain a better understanding of the features and processes this study proposes to use machine learning as a method for studying larger sets of data reliably: datasets which are collected from web-based peer review systems. This may be achieved by developing robust feature descriptions that can be used to construct (semi-) automatic coding tools, or by identifying crucial features that seem to influence the success of peer reviews in a web-based peer review system across different sets of data.

1.1 Background

Although peer feedback in second language writing is a relatively young field of research, it has already seen dramatic changes since its inception. Most important among these, perhaps, is the general shift in attitudes towards peer feedback as a pedagogic strategy. Whereas early studies (e.g. [8]) found a general resentment among second language towards the idea of learning from peers and preferred teacher feedback, peer feedback is now part of the mainstream of second language writing pedagogy and research. This change became more obvious at the beginning of the 21st century when peer feedback on second language writing emerged as field to be further investigated, being briefly mentioned in the article On the Future of Second Language Writing: A Colloquium [9]. The introduction of computer mediated tools caused a whole new dimension to open up for further exploration as highlighted by another colloquium Changing Currents in Second Language Writing Research [10]. According to Matsuda et al., computer mediated classroom discussion (CACD) might be noteworthy variable worth exploring in second language writing research as the “nature of language use and interaction in computer-mediated communication depends both on the particular interface involved … and on the beliefs and approaches of the teachers involved” (p. 163 [10]). 10 years later and computer mediated classroom discussions (CACD) is not bound by a classroom anymore. Computer mediated interaction is part of nearly all students’ lives. And current developments, such as those presented by MOOCs, are changing the way students interact and give peer-feedback to each other even more profoundly. This is, however, not yet reflected in current research on the use of web-based systems and its effect on writing, specifically for second language writing, which in the context of MOOCs and world Lingua Franca higher education may well represent the majority of students.

L1 and L2 research on peer feedback on writing rarely crossed paths. Peer feedback in L1 writing research were gaining support from studies of Bruffee [11], who explored the collaborative learning potential of students without the interference of a teacher, supported by evidence of studies investigating the benefits students gain by reviewing each other’s work [12], [13]. L2 studies have been much more cautious with the uptake of these claims made in L1 writing research, mainly due to the inevitable clash culture, language proficiencies, and the influence of L1 have on the peer feedback process, as highlighted by Zhang’s [8] study. According to Lundstrom and Baker [6], it would be wrong to assume that findings of L1 studies of peer feedback can be applied to L2 studies. A better understanding is needed how L2 writing teachers, or teachers dealing with L2 writing students, can benefit from peer feedback for their students, and if they benefit, to understand what those benefits are. To find answers to those questions, Lundstrom and Baker [6] compared two groups of students; students who received feedback, but did not give feedback, and students who gave feedback, but did not receive feedback. They concluded that results of L1 research are transferrable to L2 writing research as well; student who gave feedback and revised their work were positively affected, supporting the benefit of collaborative learning [11].

Although peer feedback is considered to be a complicated task for second language writers [14], [15]–[18] a growing body of evidence is being compiled by research, stating that peer feedback can have a positive effect, given a set of specified conditions, on second language writing [6], [10], [19]–[23]. Recent studies exploring the possibilities of peer feedback seem to gain an impulse from technological advances, as indicated earlier by Matsuda et al. [10]. These changes are reflected by Hyland and Hyland’s “State of the art article” on feedback on L2 students’ writing [24]. According to Hyland and Hyland [24], “the role of the computer in both delivering and mediating feedback has become more
visible in practice and in research, p.93”, and have, therefore, changed the face of L2 learners and the
role of peer feedback, specifically in the context of higher educational institutions. It is not surprising to
see an increase of testimonies supporting peer feedback on writing from technology related research
fields (computer assisted language learning, web-based learning, computer mediated communication),
all of which see benefits from peers communicating in an web-based environment for language
learning [25]–[29]. These studies report how computer mediated communication itself provides a rich
learning experience for second language learners, but also highlight that different web-based
platforms provide different results. Díez-Bedmar & Pérez-Paredes [30], for example, compared the
type of peer feedback and the impact of English native-speakers’ feedback on Spanish peers on two
different web-based platforms, wikis and forums. The effects they obtained varied based on these two
different platforms. As predicted by Matsuda et al.[10], different web-based platforms changes the
nature of the language used and strongly influences or highlights communicational features.

One facet of technology supported second language writing that remains unexplored is the use of
web-based peer review systems, which are commonly used platforms supporting peer feedback in L1
writing research; e.g. SWoRD™ [31], Calibrated Peer Review™ [32], [33], Annotationtool [34],
MyReviewers [35], [36], and ELI [37]. Web-based peer review systems, in academia, are generally
associated with the peer reviewing process when submitting a conference paper or a journal article. In
an education setting, the principle is essentially the same; however, the focus is different. Peer-review
systems are strongly embedded in theories that promote active learning, collaborative learning, and
writing by revision as well as consistent with domain-specific tasks [38]. In L1 writing instructions, or
composition research, utilizing peer feedback, using a web-based peer reviewing system, is
considered to be an important additional technique to include [39]. Furthermore, peer review systems
have been shown to support other aspects of learning as well, some of which have proven to be
effective for second language writing. For example, second language writers benefit from an increased
awareness of audience needs and at the same time are given opportunities to practice second
language writing in a meaningful context [40]. In addition, web-based peer review systems support the
finding that feedback from multiple peers improves students writing more than feedback from one
teacher [41]–[43]. Peer-review systems stimulate learners to engage in critical communication on the
writing process and these systems provide an option to do so anonymously, another considerable
advantage [44].

From a writing pedagogical perspective, utilizing peer-review systems would seem a logic
implementation for second language writers in a higher educational setting. Specifically in the context
where English academic writing as a discipline is becoming more important for non-English native
speaking institutions and students, but also in the context of English native institutes who are receiving
an influx of foreign students. In the European context, the introduction of the Bologna process, which
has essentially internationalized European higher education, has also given rise to a greater need for
a common language of assessment amongst an ever-growing body of second language learners.
However, in order to investigate whether the effectiveness of peer-review systems are as effective in
an L2 writing environment as they are in an L1 writing environment, it would be necessary to
understand what makes peer feedback in web-based peer review systems effective in L1 settings, and
to test whether these features are comparable in an L2 environment.

Current research in peer feedback literature using web-based peer review systems are generally
divided in studies focusing on the usefulness of peer feedback [18], [45]–[47] and studies focusing on
the effectiveness of peer feedback on the writing process [14], [43], [48]–[50]. The results obtained by
these two divisions have not always provided consistent results. What is considered useful is not
always considered effective and what is considered effective is not always considered useful. For
example, students consider the use of praise in peer feedback comments to be useful for their
learning of writing [18], [46], however, the inclusion of praise in peer feedback comments is found to
only slightly effect the writing performance [43], [48]. In addition, a similar contrast is found between
directive and nondirective comments. From the students’ perspective, directive comments are
considered to be more useful [47], [51], but nondirective comments are found to be more effective in
findings “may be related to the framings of research questions”. Students’ beliefs and values of what
might be considered useful may not always agree with the actual effectiveness. It would therefore be
more useful to investigate peer feedback effectiveness in second language writing, rather than
exploring their beliefs.
1.2 Peer feedback features

In order to better understand how the peer feedback process influences and affects the revision process, this paper proposes data coming from web-based peer review systems for the following features as determined by different studies for the purpose to better understand L2 writing in web-based peer review systems.

1.2.1 The nature of feedback

According to Nelson and Schunn [50], features that influence the revision process in web-based peer review systems can be classified into cognitive features and affective features. Cognitive features specifically relate to the knowledge and understanding students have about writing. Affective features relate to the way students use language to give feedback to others. In both cases, Nelson and Schunn [50] propose a number of features that can be coded in datasets and can therefore be used to predict whether they influence revision or not.

The cognitive features include, (1) summarization – peers give a short summary about the part to be revised; (2) problem identification – a criticism that is explicitly highlighted by the peer; (3) solution – where peers explicitly offer solution to problems; (4) localization – peers include information where something should be changed; (5) explanation – is given by a peer as to why something should be changed; (6) scope – which indicates whether the feedback is on local issues or global issues of writing. Both 5 and 6 might have, according to Nelson and Schunn, a negative influence on revision. In addition to the cognitive features, affective features can also be measured and include features such as (1) criticism, (2) praise, and (3) summary. In all three cases, language plays a specific role. Whether and how students include criticism, praise, and summary may have a positive or negative influence on revision, according to Nelson and Schunn. In this case, features related to affective language, such as mitigation, is used to measure the effect.

1.2.2 Effect and Affect

Liu and Sadler [40], in their study investigating the effect and affect of electronic peer feedback suggest that, specifically among L2 writers, what students comment on and do not comment on influences revision. According to Liu and Sadler, feedback can be measured by looking at three aspects that influence revision: Area, Nature, and Type. Area identifies feedback to be either oriented to local writing problems (mostly related to language) or global problems (related to writing). In addition, distinguishing between these two, additional differences are related to the nature of the feedback. Both local and global issues can be formulated in such a way that they are either revision oriented, or non-revision oriented. And finally, the type of feedback further classifies the feedback to be evaluative (referring to aspects which are either good or bad), clarification (includes explanations concerning the comment), suggestions (indicates a suggestion for change), and alteration (provides a specific change).

1.2.3 Linguistic features

Leijen and Leontjeva [49] in their study, suggest a more linguistic analysis approach to the coding of datasets. Their main goal was to allocate features that are somewhat embedded in the features indicated earlier by Nelson and Schunn [50] and Liu and Sadler [40], but systematically find linguistic features in these features that will make it easier to code large sets of data. The proposed the following linguistic features: (1) Mitigating devices – such as “I think”, (2) Modal verbs – such as “could maybe”; (3) personal pronouns – such as “you”, “the author”, “one”, “I”, (4) location nouns and prepositions – such as “page” and “line”, (5) error nouns – “mistake”, “problem”, (6) idea verbs – such as “look at” and “consider”, and (7) negative words – such as “bad” and “wrong”.

According to Leijen and Leontjeva [49], if feedback types can be broken down to a specific linguistic unity, coding large sets of data might become easier to perform.

1.2.4 Analyzing revision

Faigley and Witte [52] have perhaps developed the most commonly used taxonomy to investigate revision. In order to understand what type of revision is made, breaking these down into measurable parts would make understanding type of feedback more directly to different types of revision. Perhaps specific types of feedback lead to specific types of revision. As a result, this study proposes, as have many other studies, to include and code not only the feedback but also the revisions in the text using Faigley and Witte’s taxonomy of revision.
Once datasets have been prepared and coded for multiple features, machine learning algorithms can be applied to build predictive models. These models should ultimately be able to shed more light on the different features that lead to revision in web-based environments.

1.3 Machine Learning

Machine learning is a branch of computer science. Machine learning is currently used across disciplines and has become an important real data analysis tool [53]. New simple methods to collect large sets of data (such as web-based peer review systems) call for new methods that will allow increasing understanding of the data itself. According to Abu-Mostafa, Magdon-Ismail, & Lin [54], machine learning is a very useful method to extract meaning from large sets of data, and in order to determine if machine learning can be use to solve a data problem, it has to meet three general features: 1) you assume that a pattern exists, 2) but it is not possible to pin down this pattern using statistics, 3) and you have to have a large amount of data to investigate what the pattern is.

In the case of web-based peer feedback and its influence on revision in the writing process, we assume, and studies have confirmed, that specific features of peer feedback influence a peer to make a specific type of revision or not. The problem is, as there are so many features, it is unknown which features have the greatest influence on revision, and, therefore, can we, when we add and/or remove features, build a model that will lead to an improved revision (for a given set of data or students). Given the suggested feedback features mentioned earlier, if machine learning is able to build a model using all these features, it is likely that the model is not going to predict revision very strongly. Likely, given all features, the chance of revision is as good as chance itself. However, once machine learning removes or adds features, the strength of the prediction might increase, in which case, some features will predict revisions much more than others. In other words, when training students how to give feedback, knowing which features have positive influences on revision and knowing which features have negative influences on revision is important to determine. Additionally, as we established earlier in this paper, as web-based peer reviews can contain an infinite set of different settings, what has a positive influence in one setting, might actually have a negative influence in another, so too for L2 writers and more novice and expert writers.

As such, machine learning is about pattern recognition on a set of coded data. Once we analyse the peer feedback and revision data, which machine learning will split in a ‘training’ set and a ‘testing’ set, for that set of ‘trained’ data it will attempt to recognize a pattern leading to a specified outcome: revision in this case. Once a reliable pattern has been recognized, it will use this pattern to test if this also applies on ‘new’ never seen data. In other words, it will test whether the pattern is also the same for another set of data. If the pattern is observed in other data, we can talk about a reliable predictive model (see Fig 1).

![Diagram of machine learning process](image-url)
Another advantage of machine learning is that it supports replication studies which in turn creates a much better comprehension of all the different configurations peer feedback has on the writing and revision process (as was suggested Goldin, Ashley and Schunn, [7]). This application has also been suggested by Crossley [55] in his paper highlighting the advantage machine learning techniques has on advancing research in second language writing. More specifically for web-based peer feedback, Leijen’s [56] paper examining the possibilities of machine learning to better understand peer feedback and provides examples of a number of studies which have done so (see Leijen [56] for a more comprehensive overview).

2 CONCLUSION

This paper provides an overview of peer feedback and second language writing and introduces machine learning as a method to further investigate the effect peer feedback has on the revision process of second language writers. It further elaborates on a number of features, highlighted by peer feedback studies, which can be used to code large sets of data, generated by web-based peer reviews systems. Finally, the paper ends with the notion that using machine learning as a method will allow for replication studies and will, as a result, be able to advance our understanding of these processes. In addition, although this study focuses on peer feedback on L2 writing, as much is still unknown about this process, the method can be used to test any configuration of peer feedback on writing.

REFERENCES


