ELABORATED FEEDBACK FOR BEGINNING JAPANESE LANGUAGE
STUDENTS: THE BENEFITS AND DRAWBACKS OF SCREENCAST DELIVERY
初級日本語学習者に対するフィードバックの仕方：
スクリーンキャストの長所と短所

Nina Langton, University of British Columbia, Okanagan
ナイナ・ラングトン、ブリティッシュ・コロンビア大学、オカナガン

Giving feedback on written assignments is a significant and time-consuming part of the job of instructors of beginning language classes, but there are many questions concerning the best way deliver effective feedback. Should we flag the error, give verification feedback, give the correct response, or provide more specific information? What is the optimal length or complexity of feedback? Do students read, understand, and use it?

Elaborated feedback provides "an explanation about why a specific response was correct or not and may allow the learner to review part of the instruction," and has been shown to have positive effects on learning (Shute, 2008, p. 160). In other words, it is detailed feedback that tells the learner what was done well and what specific things need to be worked on in order for the student to progress. The question is, how best to deliver this kind of feedback?   This pilot study looks at a relatively new way of giving feedback via screencasting. This technology allows the instructor to produce a digital video that can give a detailed audio and visual explanation of the strong and weak points of the learner's work. This study compares the results of feedback given both by conventional means and by screencasting, examining differences in improvements made in accuracy on revisions of short writing assignments, depending on the modality of the feedback the learner received. The resulting revisions are compared to determine if and how the feedback seems to be understood and used by the students. 19 students in three different classes agreed to participate in the study, but the current report will focus on two students of different levels of proficiency. Student and instructor comments and observations regarding their experience with both conventional and screencast feedback will be included in this report. This is a preliminary and subjective examination, and no statistical analysis has been done. Language proficiency and motivation levels are based on the instructor's observations and the final grade in the class.

Truscott (1996) first sparked the debate on the effectiveness of corrective feedback on writing assignments, arguing that correction of grammatical errors was ineffective, possibly harmful and should be abandoned. Many scholars took issue with this conclusion and a great deal of research into the subject has resulted from it. Shute (2008) produced a review article examining the use of formative feedback, which would include but not be limited to error correction, and concluded that elaborated feedback would enhance learning. While feedback that could be construed as critical or controlling, or that makes comparisons to peers or other external norms can have negative effects on learning, in general, elaborated feedback that is given in "small enough pieces so that it is not overwhelming," is "as simple and focused as possible" and is linked "clearly and specifically to goals and performance" (Shute, 2008, p. 177) can benefit the learner. She also makes the following recommendations based on the over 100 studies she examined. In general higher-achieving learners will be better able to handle indirect feedback, such

as error flagging, hints, and cues, whereas lower proficiency students will benefit from direct, elaborated, correct answer feedback (pp. 180-181). Van Beuningen, et. al. (2011) go further and argue that direct, comprehensive correction is the most beneficial type of feedback for *all* levels of learners.

> Only when the grammatical errors were corrected directly did pupils show evidence of a long-term learning advantage...Direct correction is better for suited for grammatical errors and indirect correction is better suited for nongrammatical errors. (pp. 32-33)

The instructor in the current study has been giving feedback using conventional red-pen mark-up on handwritten assignments, or a combination of inserting comments and flagging errors with highlighting on word-processed submissions. However, when students handed in revisions, it was clear that they often had not understood the hints and comments that they received as feedback. Bailey (2009) also comments on the problems students have interpreting their written feedback, reporting that they often find it to be inconsistent, vague, and frustrating. The question becomes, then, how best to deliver this feedback in order to help students reach their learning goals.

Studies on audio-only feedback have suggested that this type of feedback is more detailed (Lunt & Curran, 2010); is perceived by students to be more personal, understandable and clear (Olesovea et. al., 2011); is more effective in conveying nuance (Ice et. al., 2007); and shortens the perceived distance between instructor and student (Morra & Assis, 2009). Screencasting is one step beyond audio feedback, in that it incorporates a visual component. This is the process by which a video capture of the computer screen activity and accompanying audio is created into a digital file that students can access on their own device. Recent studies indicate that screencasting results in a higher number of overall revisions on subsequent drafts compared to conventional feedback (McGarrell & Alvira, 2013); that the linking of both visual and audio cues aids memory and understanding and increases student receptivity of the feedback (Edwards et. al., 2012); and that it provides inspiration and motivation for the student in their future work (Mathisen, 2012).

These positive results prompted the investigator of the current study to try the methodology in the beginning Japanese language classes she instructs. The process is as follows: students submit their digitized assignment via the online learning management system, the instructor downloads it to her desktop, does an initial mark-up of areas that need comment, then launches the software, and reads through the submission making oral comments to the student. The software program then produces the video and a URL link to it, which is posted back to student on the LMS. The program used at the instructor's university is a proprietary software called Mediasite, but other options such as Camtasia or Jing can also be used effectively.

In the fall term of 2014 the instructor/investigator used a third-semester Japanese language class for a pilot study. The students were divided into two groups in a cross-over design. They were required to submit a first draft and then a subsequent revised draft on two different topics over the course of the term. One assignment received conventional feedback, and one screencast feedback. The feedback was intended to focus on grammatical accuracy, as well as style and organization. The students were also

required to submit a feedback response form explaining how and why they fixed their errors or dealt with suggested changes. The students that agreed to participate in the study also submitted a questionnaire. In the current study, the investigator is looking to see if there is any difference between the first and second drafts in accuracy and corrections of errors depending on the method of feedback; in other words, is there any indication that screencast feedback makes it easier both to give and receive feedback that is detailed yet comprehensible.

Wolfe-Quintaro et. al. (1998) recommend using either the error-free T-units per t-unit ratio or the errors per t-unit ratio for judging accuracy. A T-unit is defined as the shortest grammatically allowable sentence. Using the error-free T-units ratio does not indicate the severity of error, so it was decided to use the errors per T-unit ratio. In a study such as this, the issue of what to count as an error is important. Should errors be tallied for failure to use the kanji that students are supposed to know, or slight misspellings, or style or vocabulary suggestions? The investigator decided to focus on significant errors that affect comprehensibility to a greater or lesser extent, and exclude orthographic errors and suggestions for style or organization.

The current study focuses on two students of different levels of proficiency. Star (names are pseudonyms) is a high proficiency, heritage language learner who is motivated to improve her language ability in order to communicate with family members. Totoro is a mid-proficiency, highly motivated student who plans to go to Japan to teach English and works very hard to improve her language ability.

Star's first submission received a conventional mark-up. The essay, on the topic of her favourite *arubaito*, contained 24 T-units and 785 characters. The assignment was for a minimum 350 characters, so this was well over the requirement. The instructor flagged 12 areas of text indirectly with highlighting (mostly for kanji conversion) and inserted five comments totalling 73 words. There were only two grammar errors of concern, so the ratio of significant errors to total T-units is .083. (With this measurement, the lower the number, the higher the accuracy.) One of the errors had been flagged with highlighting indicating there was a problem.

人がたくさん来るて会社からボーナスが…

This error could have been either a significant mistake in verb conjugation or just a simple typographical slip, but it was easily corrected once attention was drawn to it. However, the student attempted to correct the second error but was not successful, despite instructions for correction that were fairly specific.

「あなたはいい人 plain form of verb とおもって、…

The purple highlighting is the code for kanji conversion. We can see the student's thought process in the student feedback response form, where she stated

あなたはいい人と思う(changed to plain form of verb)
Star thought that the instructor meant the plain form of the verb that had been highlighted. She then joined this sentence to another without an appropriate conjunction, so in effect she added a new mistake to her significant error count. Therefore, after fixing

one error and adding a new one, her significant error to T-unit ratio remained the same in her revised draft. In retrospect, the feedback could have been clearer if です was used instead of "verb", but it is clear that for this student, who is fairly advanced, the feedback was not comprehensible.

Star's second submission, for which she received screencast feedback, was on the topic of her favourite place. This time she wrote 910 characters in 39 T-units. Her significant errors per T-unit was 0.18, which is slightly higher than her first assignment, but this could indicate that she is attempting more complex sentences. The instructor responded with a 7 minute 52 second video of 1014 words, 661 of which were comments on the text. Star watched all or part of this video seven times for a total of 24 minutes 21 seconds. In her revised draft, Star's significant errors to T-unit ratio fell to .069, so her accuracy improved considerably. However, not all of the screencast feedback was successfully incorporated into the revised draft. Two of the significant errors in Star's first draft involved relative clauses. In response to にわ鳥をうんだたまごでした, the instructor provided a 95-word oral response, in which she attempted an explanation of how the particles would work when trying to make one sentence modify a noun. ("Ok, here with your next sentence we also have a little bit of a problem with the grammar of the relative clause. *Niwatori ga unda tamago deshita.* So like if you are joining the sentences, the first one would be *niwatori ga tamago o unda,* right? Then you are making that sentence, *Niwatori ga tamago o unda,* you are making the part to modify *tamago*, so you take *tamago* out of the first sentence and it would be *tamago o,* so *o* is attached to *tamago,* but *niwatori* should be followed by the particle *ga.*) Unfortunately, this response was overly detailed and confusing. In fact, even though the student was given the correct sentence in a direct manner, she reverted to the simple sentence にわ鳥がたまごをうんだ in the revision, so obviously she did not understand the feedback. It was not clear that the first sentence given to her was actually the correct way to construct the relative clause. In other words, the quantity of the feedback did not necessarily lead to an improvement in the complexity of the writing.

Totoro received her feedback in the reverse order, with screencast comments on the first submission and a conventional mark-up on the second. Her first submission on her favourite *arubaito* contained 19 T-units, 457 characters, and 29 significant errors, resulting in a significant error to T-unit ratio of 1.52. Essentially, then, every sentence has at least one error. The instructor responded with a 19 minute 45 second video of 2932 words, of which about 2700 are comments. Totoro viewed all or parts of this video five times for a total of 43 minutes and 44 seconds. In her revision, Totoro's significant error to T-unit ratio decreased to 0.42, so the improvement is dramatic. Almost all of the feedback given to the student was direct, elaborated feedback. The instructor gave Totoro the correct form and explained why it should be that way. For example, for the sentence 私はいろいろなアルバイトがしたです。Totoro received a detailed explanation about the direct object marker. ("So when you have a verb like *suru*, to do, you have to do <u>something</u>, right? So the something that you do is what is called the direct object, and that is marked by the particle *wo*. So you need to *arubaito wo suru*. You have to keep the *wo*'s, keep that sort of together in your head as a unit. So depending on what you want to do with *suru*, if you just want to say "I did various *arbaito*", then you would just use *iroiro na arubaito wo shimashita.* But if, I think what you are trying to do is I have had the experience of doing something, so that would be *arubaito wo shita koto ga*

arimasu.") In her response feedback regarding this comment, Totoro wrote: "The particle was changed to を because アルバイト is the object being acted upon. Also を したことがあります was used to properly convey that 'I have had the experience of' doing various part-time jobs." So while in some ways her response may be imitative, it can be hoped that she was reinforcing or consolidating in her mind what the instructor had stated in the screencast.

For Totoro's second submission, she wrote on her favourite movie. In this assignment, she wrote 379 characters in 16 T-units, so it is somewhat shorter than her first assignment. There were 177 words and 69 characters in nine inserted comments, and the instructor flagged 6 additional areas of text for grammatical errors. She had 11 significant errors, so the resulting significant error to T-units ratio is 0.687. In her revision, Totoro corrected five of the 11 significant errors, but made four new errors, incorrectly changed two errors, and did not change one error. As a result, her total decreased to seven significant errors, for a ratio of .437. She did not pay attention to some of the direct corrections she had been given, but she did correct some of the indirect flagging. So it makes it difficult to say how effective overall this feedback was for her, and her responses to the feedback don't always make her thought processes clear. For example, where the particle を had been flagged with 会う、she merely writes "changed を to が," which demonstrates that she did not learn how to correct her mistake from the indirect flagging.

If we compare percentage improvement for both students (Figure 1), we can see that both produced better significant error per T-unit ratios with the screencast feedback.

Star	Draft 1	Draft 2	% Change
Conventional	.083	.083	0.0
Screencast	.18	.069	61.7
Totoro			
Conventional	.687	.437	36.4
Screencast	1.52	.421	76.3

Figure 1. Significant error per T-units ratio

However, it would be far too simplistic to make any conclusions about the efficacy of one method or the other based on these very simple and subjective observations. All kinds of factors come into play, including of course the fact that we expect students are learning from various sources, not just feedback, over the course of the term.

It could be argued that the effectiveness of screencasting for particular students can best be judged by their own preferences. Screencast feedback was introduced to the students in the fall semester of 2014, and those continuing to the spring semester were allowed to choose which way they wanted to receive their feedback on subsequent assignments. Both Star and Totoro chose screencasting, and in fact more students in the class chose screencasting than conventional feedback. Star commented that conventional feedback "can't further explain what the error is or why it is incorrect," but is "easier to refer back to when making edits and changes vs. trying to find the correction in the video." This in fact is a challenge with the videos, and while it is possible to put bookmarks in them, it still is a problem when students want to return to one specific spot. Several other students in the class commented on this as well. Regarding the screencast, Star commented that "waiting for the video to load was annoying," but it is "more in-

depth/comprehensive" and "interesting because I've never used this method of correcting before, so I think I paid attention / learned more." This comment indicates that there is a certain novelty factor in the use of screencasting that can be exploited. However, when the novelty wears off, it may be that the screencast methodology loses some effectiveness.

Totoro noted that conventional feedback "seemed like it wasn't as detailed, and there were less 'for example' type statements to help draw parallels to similar mistakes." On the positive side, however, "it's easy to refer back to and less time-consuming to read over." Screencast feedback, however, "was much more detailed and provided many more examples, which helped me fix similar mistakes in other assignments." This comment addresses the ability to transfer this learning to new contexts. However, screencast feedback "was quite long, and not as easy to refer back to because of this." It is difficult to say what the optimum length for screencast feedback would be, but depending on the number and severity of mistakes, the length of time required to provide elaborated feedback will vary.

Meraid, another study participant, chose conventional feedback, and one of her comments about her choice speaks to the idea of social presence or distance. She reported, "I take criticism poorly, so it's easier for me to see feedback rather than hear it." Much of the literature assumes that increasing instructor presence and reducing the distance between the instructor and the learner are good things, but some students do not seem to want that. Meraid is embarrassed to hear her mistakes, so she would rather receive corrections in a less personal way. She also neatly sums up the potential advantage of screencasting when she compares the two methods:

> Conventional: Easy to see and correct. Best for assignment.
> Screencast: Best for actual knowledge.

If done well, screencasting can potentially be a better way of delivering detailed explanations to help students who really want to put the effort in to improve their skills. For students who just want to complete the assignment and get an acceptable grade, conventional feedback is sufficient. Giving learners the choice of feedback method allows them to tailor it to their learning goals.

Obviously there are limitations to this study in that it is only looking at one example of each type of feedback for two participants. In order to reach any definitive conclusions about the effectiveness of screencasting feedback for beginning students of Japanese, longer term studies involving more participants will be required to see if and how screencast feedback helps students to transfer what they learn from it to new writing samples.

From the instructor's point of view, the technology is reasonably easy to learn and use. Some studies claim that the time for the instructor to produce the feedback is less with screencast feedback than conventional (Edwards et.al., 2012; Mathisen, 2012), but the instructor in the current study did not find this to be the case. The time required to provide elaborated feedback will depend on the type of assignment being marked, the proficiency of the student, and how comprehensive the feedback is. At the same time, the feedback cannot be so long and detailed as to overwhelm the learner, who may be tempted to just "turn it off." A balance between "enough" and "too much" must be

reached, and it is the instructor's job to try to ascertain this balance for each individual student.

The perceived shortening of distance that Meraid commented on may have a down side for the instructor as well. It might feel to the instructor like he or she is speaking directly to the student, that this is an "interactive" experience, but in fact it is a monologue and the instructor is unable to tell when or where the student is not understanding. Therefore, if the instructor is going in to a lot of detail that the learners are not understanding, the feedback may end up being just as confusing and frustrating for them as conventional feedback might be. The instructor in the current study also noticed a tendency, when speaking the feedback, to get distracted by the next problem and neglect to go back and complete or clarify comments on the first one. In addition, the mid-sentence stoppages, repetitions, and slip-of-the-tongue errors that sometimes occur in conversation-style feedback may not be interpreted correctly by the learner, and there is a danger that these mis-speakings then become "carved in stone." Organized, clear, concise, and unambiguous comments are essential in this type of feedback. In some ways, conventional mark-up may have less teacher-induced errors because the instructor has the opportunity to immediately see and correct what has been written, as opposed to having to pause, go back and review what was said. Mediasite does not allow an instructor to record over and correct errors, and that is a drawback with this particular programme. A tablet, on which physical marks and visual cues could be captured on the video, might aid in comprehension of the feedback.

Despite some of the drawbacks associated with the use of this methodology for delivering elaborated feedback, it appears to hold considerable potential for use with motivated students who want detailed feedback on their writing assignments. It also offers another choice for students to fit their own learning styles and preferences. Further study will help to determine its effectiveness.

References

Bailey, R. (2009). Undergraduate students' perceptions of the role and utility of written assessment feedback. *Journal of Learning Development in Higher Education* 1, 1-14.

Edwards, K., Dujardin, A.F. & Williams, N. (2012). Screencast feedback for essays on a distance learning MA in professional communication: An action research project. *Journal of Academic Writing* 2(1), 95-126.

Ice, P., Curtis, R., Phillips, P., & Wells, J. (2007). Using asynchronous audio feedback to enhance teaching presence and students' sense of community. *Journal of Asynchronous Learning Networks* 11(2), 3-25.

Lunt, T. & Curran, J. (2010). 'Are you listening please?' The advantages of electronic audio feedback compared to written feedback. *Assessment & Evaluation in Higher Education* 35(7), 759-769.

Mathisen, P. (2012). Video feedback in higher education: a contribution to improving the quality of written feedback. *Universitetsforlaget, Nordic Journal of Digital Literacy* 7(2) 97-116.

McGarrell, H. & Alvira, R. (2013). Innovation in techniques for teacher commentary on ESL writers' drafts. *Cahiers de l'ILOB* 5, 37-55.

Morra, A.M. & Asis, M.I. (2009). The effect of audio and written teachers responses on EFL student revision. *Journal of College Reading and Learning* 39(2), 68-

Olesova, L.A., Weasenforth, D., Richardson, J.D. & Meloni, C. (2011). Using asynchronous instructional audio feedback in online environments: A mixed methods survey. *MERLOT Journal of Online Learning and Teaching* 7(1), 30-42.

Shute, V.J. (2008). Focus on formative feedback. *Review of Educational Research* 78(1), 153-189.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning* 46, 327-369.

Van Beuningen, C.G., De Jong, N.H. & Kuiken, F. (2011). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning* 62(1), 1-41.

Wolfe-Quintero, K., Inagaki, S. & Kim, H.Y. (1998). *Second language development in writing: Measures of fluency, accuracy & complexity.* Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawai'I at Manoa.