

An Experiment in Student Peer Assessment

An element of peer assessment seems to be increasingly desirable in undergraduate programmes, and what follows is a bare-bones description of my (rather ambitious) attempt to have first year language students correct and mark one piece of each other's work¹. The experiment was time-consuming – it was to last about 6 weeks – but also something of an eye-opener. It involved a group of 15 first-year French specialists (13 female, 2 male) whom I saw for one hour per week in the capacity of language skills tutor, and who had typically been awarded BS and CS at A-level. Learners at this level are used to being evaluated by tutors but still tend to react to the mark awarded rather than the detailed comments written on their work: first years often find it difficult to differentiate between concepts such as «right / wrong», «good / better», «grammatically correct but awkward / genuine French expression», and many individuals, despite constant encouragement, make minimal use of corrected work. Students also rarely get to see their peers' work, from which they can both learn a great deal and gain in self-confidence. By introducing a carefully planned element of peer assessment I hoped to encourage the students to think about evaluation in a more constructive manner, and was able to conclude that first-year students find it extremely difficult but very useful to take a step back from their work and to look at assignments in French from another perspective.

¹ This brief report is not intended as a practical or theoretical guide to language tutors. My intentions are simply to share an interesting experience. The theory behind my experiment is based on D.A. Kolb's «Learning Cycle» (Planning, Doing, Reflecting, Thinking), *Experimental Learning: Experience as the Source of Learning and Development* (New York: Prentice Hall, 1983).

Description of Study

Aim of Study

The aims and objectives of the exercise were quite modest. Given that «even our best students rarely display understanding»,² I wanted to (1) encourage students to think about their work critically in order to bring them to an awareness of their own linguistic strengths and weaknesses, and (2) give them a new way of considering their language assignments.³ As a consequence of approaching a piece of academic work in a new way, the students would, it was hoped, automatically gain new insights into the French language itself. I felt sure that, in theory at least, students could spot and correct errors in exercises and handouts that they would normally miss in their own work. (How many times do students react to corrected written work with the comment: «did I *really* write that?»). The experiment itself was not too complicated in principle: I intended to give the students one of their peers' assignments to correct and mark. The assignment was a 300 word minimum, language essay – *Un héros français* – marked 50% on content, 50% on language. Of course, I was less concerned with the final mark than with the process itself. However, given that students always seem to be (quite naturally) obsessed with marks, I retained a numerical grade for the assignment to foster their interest.

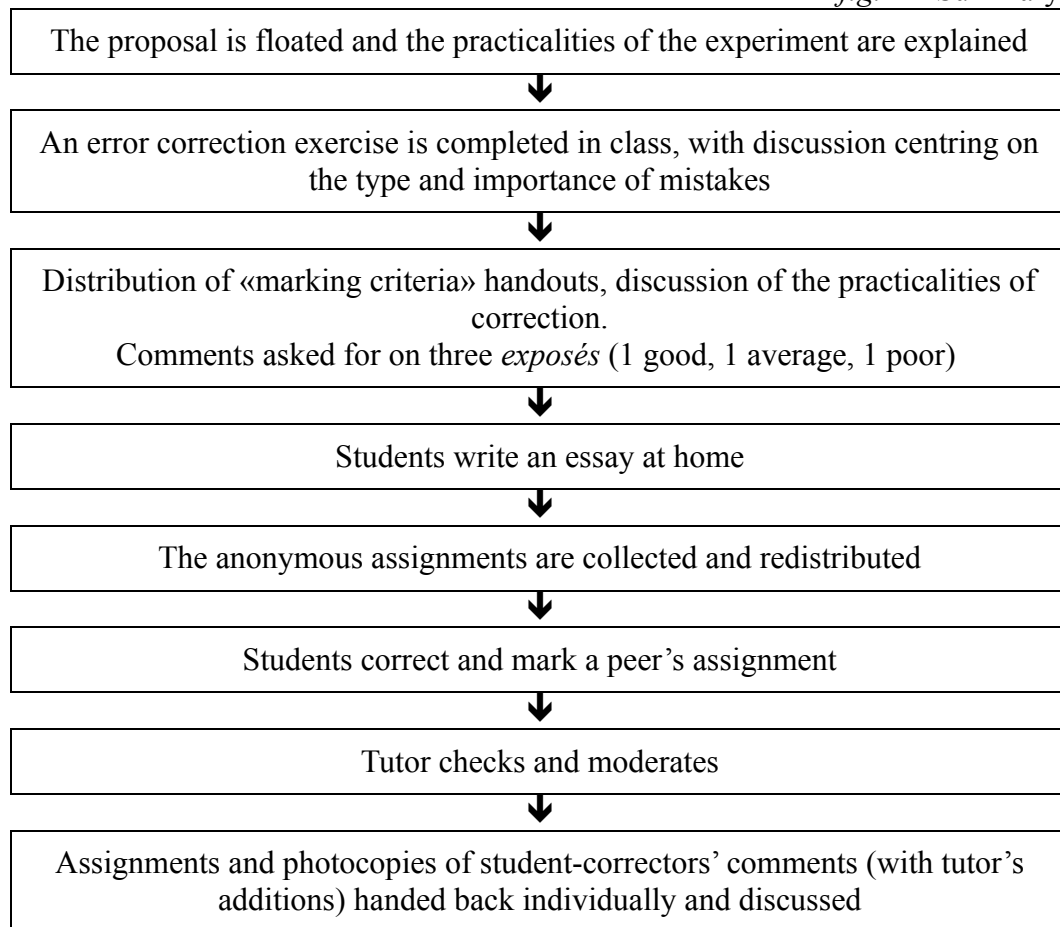
Methods (*see fig. 1*)

Beforehand, I spoke to the class about what I wanted to do and why. I explained:

- that I'd never done this before,
- that the whole class needed to approve of the experiment – individuals had the right of veto,

² H. Gardener, «Opening minds», *Demos* 1 (1993): 2-5.

³ F. Smith believes that it is not possible to separate understanding from thinking as mental processes: to enable a student to think is to enable a student to understand, *To Think: In Language, Learning and Education* (London: Routledge, 1992), in particular 12–37.

fig. 1 Summary:

- that I believed everyone was capable of correcting a good number of the mistakes commonly made at this level,
- that the assignment concerned was an essay,
- that the final mark was much less important than the process of correction, and
- that I would oversee the entire process to exclude the possibility of miscarriages of justice. The students were made aware that my intervention would be once the essays were corrected and marked, but before they were handed back to their authors.

Pre-Experimental Error Correction Exercise

The next step was to do an error correction exercise in class, in which students were asked (1) to correct mistakes and (2) to decide to what degree a particular mistake was important.⁴ At first the students merely corrected perceived errors, but the *mise en commun* was surprisingly successful: the students had heard me comment on what I consider to be «elementary errors», and they freely discussed not only what in *their* minds constituted «bad» mistakes. They highlighted genders [«we have dictionaries...»],⁵ adjectival agreements [«my teacher at A level went on a lot about this...»], irregular verbs [«spelling mistakes are acceptable, but ever since I learned *faire*, *avoir* and *être* I've been warned to get irregular verbs right...»].⁶ They were also able to discuss the circumstances surrounding errors: Is it a typing error ? Was the author trying to use complicated syntax or advanced expressions ? Should credit be given for attempting difficult linguistic constructions and falling into error in the process...?⁷

This was followed by an exercise, prepared at home, in which students were given three essays on the same subject to «correct». They were asked to decide on an order of merit, and then to comment on what they liked and did not like in the essays. Given the varying levels of intellectual and analytical development in the group, giving specific criteria for the completion of this task was delicate. A brainstorming session proved extremely useful as it meant that many of these criteria came from the learners themselves. In this way, the students were already

⁴ Through pure serendipity I had actually done this exercise a couple of weeks previous to the beginning of my experiment. This turned out, I think, to be something of an advantage, since the experiment proposal referred to a common group experience. Also, too many exercises of a similar type and in too short a space of time will inevitably lead to loss of interest.

⁵ A colleague suggests that previously given examples of dictionary errors ought to be transparent, e.g. *Oh puits, je suis foré...*

⁶ My students' comments appear to corroborate Gardener's thesis that «students – especially good students – learn to regurgitate what they have been taught but never actually change their underlying beliefs», *art.cit.*, p.3.

⁷ The student consensus was that it should.

using analytical and multi-purpose processing skills, even before the essays were distributed. With a minimum of guidance, they discussed in small groups the relationship between the «content» and «language» elements of an essay, and then considered the following (real) openings to a first-year essay:

i/ La semaine dernière il y a eu une grève en France. Beaucoup de français ne sont pas allés au travail. Je n'aime pas les grèves, parce que si on veut travailler, on ne peut pas.

ii/ A cause de la grève, la France s'est trouvée la semaine dernière plus ou moins paralysée. Pourtant, malgré les inconvénients évidents le public, comme d'habitude, a soutenu les grévistes.

It was decided by the students that the second was better, because it went further than a simple description of events and it neatly integrated analysis, opinion and fact.⁸ The essays I then distributed were based on real (but anonymous) essays submitted in a previous year. Students were encouraged to consider not only *that* a mistake had been made, but also *why* and *what effect* the mistake had on the sentence / text.⁹ (Of course, they were more concerned to learn what effect the error had on the final *mark*, and in its way, this is not wholly counter-productive: serious mistakes lose more marks...). In order to foster interaction and so that the students took responsibility for the direction of their discussion, I left the room to allow collation of results on the board.¹⁰ Ten minutes later their elected foreman gave a brief presentation of opinions and results. I then distributed handouts explaining how I come to give a piece of work a mark out of 100, highlighting both language and content; this was to give the students solid criteria / a point of reference when they came to mark an assignment. We also spoke a little about the practicalities of marking (making comments, underlining, correction symbols, giving alternatives, pointing out why something is wrong...).

Before the actual marking exercise took place, I anticipated several problems. Absences were not thought to be too important, even though the student(s) would

⁸ This is exactly the sort of exercise for which electronic conferencing could prove invaluable. With contact hours at a premium, brainstorming and preliminary discussion via a package such as *First Class Client* can prepare classes and seminars most effectively.

⁹ This exercise is intended to bring style to the students' attention, and none of the essays contained any grammatical errors. It was quite easy to put together the essays reflecting good and bad practices, but it proved much more difficult to compose an 'average' essay, and any any case the presence of the «average» version only complicated matters. In future I shall only distribute two versions of the essay, the «good» and the «less good».

¹⁰ There was a general consensus, although the students certainly found this exercise more challenging than I had anticipated. W.G. Perry would not put this down to the students' varied linguistic abilities but to different stages of intellectual progression: Duality, Multiplicity and Relativism do not come to every individual at an identical stage in their development. Cf. *Forms of Intellectual and Ethical Development* (New York: Holt, Rinehart, Winston, 1970).

miss out on important follow-up. Shyness and in particular over-marking were, I thought, the serious potential pitfalls: would students be happy to openly criticise their colleagues' work? The assignment would be submitted anonymously so as to guard against unhelpful confrontation.¹¹ Over-generous marks would be moderated, and if a student felt incapable of giving a mark, a comment would suffice. In fact, in the light of the encouraging discussions we had already had, I made the comment obligatory and the mark optional (but highly desirable).

There was, however, a moment of hesitation: a weak student came to see me to say that he did not consider himself capable of marking work, his own marks being only in the 30s and 40s. But, he said that he understood his mistakes when I gave him back a corrected piece of work, and he also said that he found it easier to see problems in other people's work. «I know something's wrong but often I don't know why» – this encouraged me to continue with the experiment, and he did not really need *persuading*, rather *reassuring*. I hoped that this exercise would give him some much-needed confidence. So, we went ahead, the students were given a week to compose short essays at home and I then redistributed them when they were handed in. By the following class the students had all corrected the work and written the comments (mostly in English but some in French) of which I had stressed the importance. They had also suggested marks. It was now up to me to check and moderate before returning the essays to their authors.

The marks were not excessive and the corrections / comments had, for the main part, been made with great diligence. I was quietly impressed – *de* was noted to be missing after *témoigner*; the spelling of *nacquit* was corrected; *une manque* was spotted by several markers; and *Il a vu son futur ennemi. C'était une belle femme. Elle chantait dans un concert* was appropriately amended to *Il a vu son futur ennemi, une belle femme qui chantait dans un concert*. It was necessary to adjust only the odd mark, up or down a little. One student-marker had correctly commented on the multitude of elementary errors in an essay, but this was not reflected in the mark awarded: I adjusted from 58% to 51%, anxious to keep within the same degree class. I also raised several marks – still trying to respect the original degree class – when I suspected that the marker's zeal resulted from being over-delighted with her-/himself in spotting a mistake: *Paul Gauguin, la tombe de Gainsbourg à la Rive Gauche* and *il voyaga beaucoup* hardly merit

¹¹ Essays were to be typed, with dates of birth or symbols replacing names.

triple underlinings! In all, I only changed the class of two essays: 58% to 65%, and 64% to 70%. In both cases, the well-researched and clearly presented content deserved more credit. Otherwise, learners had been remarkably accurate in placing their particular essay within a given class. One problem that I had not anticipated, however, was that one marker – anonymous, of course – had been extremely vicious in her/his comments, writing a page-and-a-half of vitriolic criticism on the work I recognised to be that of the student who had previously come to see me to voice his worries! There was no way that I could give these comments back to this student – «childish argument ... badly written ... mistakes that twelve-year-olds don't make ... stupid, basic errors» – even if it meant falsifying the experiment. So, I dictated comments to a colleague, trying to be loyal to the few positive aspects of the marker's correction, and wondered how I could tell the marker that she/he – I did not recognise the writing – had gone much too far (not only were the comments cruel, but the assignment was not *that* bad). So, another week later, I returned the students' work. I had photocopied the comments made on each essay, including my own comments / additions. I took the final 15 minutes of the class to see students individually in order to return (a) their essays and (b) the comments they had made as markers. I did not say anything to the student whose work was so viciously criticised, simply giving him the «new» comments. As for the student who had lambasted his work, I diplomatically pointed out that she had been insensitive and over-zealous, and admitted that I had not given her final comments to the student; but I also made the point that marking was an art and that the problem was in her tone rather than in her corrections. The general consensus was that the exercise had been useful and interesting, but made considerable demands on the students' time.

Results

Student feedback was fairly positive, although members of the group were more comfortable to have their work marked by their peers than to mark their peers' work. It has been interesting to compare those errors corrected by three members of the group with mistakes they themselves made: Student A was able to correct *la groupe*, *la monde* and *le question*, but in his own essay wrote *le mort* and *la ministre*; Student B corrected verb conjugations quite competently – *il mort en France*, *elle a devient professeur* – but missed many agreements in her own essay; and Student C managed to spot *les personnes âgés*, *cette patriotisme français* and

Article

elle fut une française remarquable, but herself wrote *il fut né* and *ses idées sont bien connu*. Most of the mistakes that students corrected were of the «nuts and bolts» type – agreements, genders, verbs – although more than one corrector commented, on coming across bad syntax or an incorrect sentence structure, that «it sounds wrong» or «I don't understand». It would be extremely interesting to undertake a thorough analysis of the relationship between the mistakes made by the students and the mistakes they manage to spot and correct in the work of others.¹²

Conclusion

On the whole the experiment was a positive if time-consuming experience. The various stages of preparation, based loosely on Kolb's «Learning Cycle» – Planning, Doing, Reflecting, Thinking – were certainly important, and I will consider going through them again, even without the final marking exercise (10-week modules do not make for easy continuity).¹³ I believe that the students *did* manage to gain a new perspective on the whole written assignment exercise, and their outlook seemed more mature afterwards, even if their linguistic abilities had not been miraculously improved. My worries that by exposing students to imperfect French I would be suggesting to them new ways to fall into error appear unfounded. However, the object of the exercise was that students might experience a new pedagogical and evaluative *methodology*. In the absence of a formal evaluation form distributed amongst those students involved – a regrettable oversight – I was left with the impression that some students did try to take a step back from their own written work when the experiment was over. One member of the group told me that she found it useful to come back to an essay or translation a day or two after she initially prepared it; another student told me that she exchanged work with a colleague in another group before handing it in, and that although it remained just as difficult to spot problems with her own work, she managed to pick up on all sorts of mistakes in that of her friend. These are, of course, comments that tutors have been making to students for years; but it is

¹² Presumably, learners at a more advanced stage of intellectual development are more able to analyse and criticize their own work. However, some of the improvements in composition and checking techniques must surely be put down to the reinforcement of good learning practices: coming back to work a day or two after a piece of work was composed; re-reading and checking with a peer; systematically looking up the genders of unfamiliar words and the conjugation of unfamiliar verbs; automatically looking for agreements...

¹³ This type of exercise would also lend itself to a computer-based error-correction course designed to give students valuable practice in revising text, a process that many students are reluctant to do in the case of their own work.

encouraging to hear such simple words of good practice coming, unprompted, from the mouths of the students themselves.

fig.2

Some Examples of Students' Corrections Amended by the Instructor

Student correction	Instructor's correction
<i>Gaugin</i> , triple underline	single underline
<i>à la Rive Gauche</i> , triple underline	single underline
<i>il voyaga beaucoup</i> , triple underline	single underline
<i>e</i> added to p.p., <i>elle a voyagé vers...</i>	<i>e</i> deleted
auxiliary changed to <i>être</i> : <i>elle a monté l'escalier</i>	<i>avoir</i> reinstated
T. of <i>faire</i> changed to perfect: <i>comme ce jour-là il faisait beau, il est parti...</i>	imperfect reinstated
pl. noun, <i>Le week-end elle allait...</i>	In this context, «les week-ends» = anglicism
<i>le</i> replaced by <i>l'</i> : <i>le héros</i>	<i>le</i> reinstated
capital <i>a</i> given to <i>les anglais</i>	lower-case <i>a</i> reinstated

A.P. Tudor

University of Hull