

On formative assessment

Simon Andrewes uses a case study to show how formative assessment can be a genuine support for learning.

In this article I will look at formative assessment as a PROCESS by means of which progress is traced in order to provide feedback to support and guide learning. I will be writing about WRITING in the context of teaching English for ACADEMIC purposes on pre-session courses¹. Key elements of this process are:

- identifying target language use
- giving appropriate feedback
- establishing clear assessment criteria.

The article builds on a particular case study, with the awareness that this can be no more than a static, cross-sectional view of what was a 10-week process.

Formative assessment as a process

I found this definition of formative assessment (FA), which I like, on the Association for Supervision and Curriculum Development's website (see references for details). The definition starts by reminding us that formative assessment is a process, not any one particular test; it is to be used by *both teachers and students, during instruction*; and its purpose is to provide *assessment-*

¹ Pre-session EAP courses prepare international students for a university course ('the session') given in English. The term can be juxtaposed with in-session courses that give language support to international students who have gained access to the university course ('the session') and run parallel to it. Access to the university course is generally dependent on the satisfactory completion of the pre-session course.

based feedback to teachers and students, so that teachers and students can make necessary *adjustments* in order to improve students' achievement of intended curricular aims.

“In this kind of feedback, the assessor would ‘explore’ the problem area with the assessee to discover how performance could be improved.”

More about intended curricular aims later; at this stage, it is more important to emphasise the inextricable relationship between FA and feedback. For this, I am going to lift a metaphor that Anthony Green takes from systems engineering: 'Feedback involves making use of information about the current state of a system to regulate and make adjustments to that system.' (Green, 2014: 90) Clearly, in the teaching/

learning process these regulations and adjustments are undertaken continuously during the course and cannot be set apart from it. Both the feedback and its consequences are the shared responsibility of teacher and learner. Feedback is not something only the teacher uses to regulate input or that the teacher passes on to the learners, leaving it to them to undertake the necessary steps to improve their ability and performance. A term Green uses to talk about this sort of feedback is 'mediation'. For Green, feedback has tended to focus on the end results of task completion, whereas mediation would focus on the development of underlying abilities (Green, 2014: 210). In this kind of feedback, the assessor would 'explore' the problem area with the assessee to discover how performance could be improved. This is the kind of feedback I am aspiring to.

Target language use

If FA and learning is a process, we need to have an idea of where this process is leading us and this is where the concepts of target language use (TLU) and intended curricular aims come in. Teaching/learning and FA need to be aligned with the TLU domain, which is the term Green uses to describe the settings beyond the current classroom in which learners are likely to have to use the language being learnt: the kind of language 'real life' will demand of them in predicted future circumstances. To give an example, many years ago I used to teach adults whose primary objective was to learn English for leisure travel.



So, international travel for recreational purposes could be understood as their TLU domain. In this case there was surely little sense in assessments that test written accuracy and grammatical control, yet that is precisely what they got. From today's standpoint, it is obvious that it would have made more sense to align assessment procedures with the TLU domain and place situational oral communicative criteria at the centre of assessment. But this was before the 1990s when, as Green (2014: 172) observes, classroom-based assessment had yet to become the subject of wide-scale serious attention.

Appropriate feedback

Obviously, an assessment process that in its criteria took more into account the communicative needs of cultured and mature adult learners who primarily intended to use English travelling as tourists (not only in English-speaking countries, for English was already quite well established as the language of international tourism) would have provided more appropriate feedback. Such an approach would have been more appropriate for the learners' TLU domain, rather than such matters as pointing out where the third-person 's' had been missed, their present perfect got mixed up with the simple past, words considered 'uncountable' in English were used in the plural or with an indefinite article, there was a need for a gerund rather than an infinitive, or their production of the /ð/ and /θ/ phonemes deviated too much from a native-speaker standard.

In those days, we thought we knew what English was: we tended to treat it as a single thing, valid in all situations. General English was king; the particular purpose princedoms were on the periphery of TEFL.

“One of my main concerns is that prospective students should get used to writing what is meaningful and valid to themselves ...”

Clear assessment criteria

The misalignment of formative assessment and TLU domain that was so common back then can hardly be said to have too many serious consequences because learning goals were vaguely defined and the stakes were low. And because there was less at stake, the need for clear assessment criteria was less urgent.

Today, it would be irresponsible not to align teaching and assessment with

projected TL domain uses, when so often, as in the case we are going to study, a high-stake exam scenario has so much depending on it. Success is so crucial to the student-assessee's life chances that it is of paramount importance for them to know exactly what is expected of them, to give them the best opportunity to demonstrate what they can do.

Case study

The material of the case study we are now going to look at is a static, cross-sectional, single-frame view, yet part of an assessment and learning process we will take care not to lose sight of. The context is a pre-sessional EAP course. Here the TLU domain is pretty clear-cut. It is academic writing for the attainment of a Masters degree in the field of Government and Society or International Development. But a more urgent goal was to gain access to the chosen MA course, and this depended largely on the successful completion of a 3,000-word 'academic research paper' which counted for some 80% of the entry requirement and was to be written by the end of the pre-sessional course, the crucial anteroom to the real TLU domain.

In order to align assessment feedback with the specific needs of the learners at a particular stage on the way to the TLU domain and provide assessment criteria that help the learners understand where they stand in relation to their goal, I have found rubrics and performance indicators to be effective, not least because they are common tools in current language evaluation and so familiar to students.

Nevertheless, I do not view the TLU domain uncritically and my criteria for good writing may differ from the requirements of the TLU domain, in which case I adapt the assessment criteria to make sure this divergence is transparent, leaving learners in a position to decide for themselves how to develop their own writing and take control over their own learning. One of my main concerns is that prospective students should get used to writing what is meaningful and valid to themselves, and that they resist the temptation to churn out texts that they think will satisfy their assessor/tutor/lecturer in the form of formulaic, academic-sounding texts of dubious content.

The writing task we are going to look at here was one in which, after attending a lecture on the ways in which a particular usage of language and other semiotic resources, such as images, sounds and bodily communication, could be related to their social and political contexts, with a focus on contemporary Britain², students were asked to write an essay

² The lecturer's website: <http://www.birmingham.ac.uk/staff/profiles/elal/bennett-joe.aspx>

with the title: *Do politicians in your country use semiotic resources in a similar way to UK politicians?*

I told them that these were the criteria by which their work would be assessed:

1. You show that you have understood the key concepts of ethos, and semiotic resources (linguistic and non-linguistic).
2. You include some original, interesting and relevant information about how politicians communicate with the public in your country.
3. The essay is of the 'right length': it answers the question fully but without irrelevant 'padding'.
4. Your use of language is sufficiently well controlled to get your meaning across.
5. The essay is well organised: easy to follow, with a 'logical' development and good linking between the introduction, the main arguments and the conclusion.

Rationale and observations

The criteria are supposed to make it clear to the assessee what I consider

important at this stage of the learning-assessment process, and a simple three-point rating scale tells them how well they have fulfilled the expectations of each criteria. If less than three points are awarded, a comment is added, explaining any shortcomings. This can be followed up in a one-to-one tutorial that is programmed into the course. The result is a points score, out of 15, of which more later.

Criterion 1 will give feedback on how well the assessee has understood and dealt with the main points of the lecture. Here is one example from a student's essay of the definition of ethos.

Ethos is a Greek word meaning 'character' that is used to describe the guiding beliefs or ideals that characterize a community, nation, or ideology. The Greeks also used this word to refer to the power of music to influence its hearer's emotions, behaviors, and even morals.

Early Greek stories of Orpheus exhibit this idea in a compelling way. The word's use in rhetoric is closely based on the Greek terminology used by Aristotle in his concept of the three artistic proofs.

I awarded one point for this response, taken, with acknowledgement, from Wikipedia, with its over-general definition of ethos and absence, incidentally, of any attempt to distinguish between linguistic and non-linguistic resources. The writer had taken some trouble to address the set criterion, but unfortunately the terms were defined and used in the lecture in quite a different way.

Criterion 2 makes clear the importance I am placing on the assessee relating what they have heard in the lecture to their own experience. I am asking for something more than just 'write a summary', although this might be appropriate for a task set at another stage in the process. Relating the topic to their own experience, I suppose, makes it more interesting for them to write, and certainly for me to read, and discourages them from simply regurgitating undigested facts and ideas that are still somewhat alien to them. By



way of example, here is a response that I considered fully met the criterion:

In China, the politicians seldom deliver impromptu speech, basically according to speech draft to read every single word. Moreover, the face seems extremely serious or solemn. One of the biggest signs that Chinese leadership speak slowly with gentle tone, each of speech word with exaggerated care, the golden mean invariably used in answering question. ... Jinping Xi is no exception. Nevertheless, the majority of west politicians speak naturally and casually. For example, Blair made a speech in the parliamentary, he leaned in and one of her arm on the table. This behavior is not desirable in the Chinese authorities. ... From another point of view, all the politicians increasingly trend to use nonlinguistic resources to show affinity. For example, Chinese president Jinping Xi with his wife utilize props to show appetency by 'camera'. Just like Putin shows his tough by muscle photograph.

The idea is to reward and encourage original and critical thinking on the topic. Penalisation of linguistic inaccuracies and infelicities would detract from this aim.

However, feedback is also given with an eye on the criteria of their final research paper on which successful course completion is so highly dependent, so criteria 4 and 5 assess language and organisation. Good organisation is a key to success both in the TLU domain and for pre-session course completion. However, at this stage language criteria are relatively played down. In their crucial final written paper, weighting for 'range and accuracy of language' counts for 50% of the total mark. Here the value is only 20%. At this early stage my focus is on other things I want to develop, such as focus and critical thinking, and finding a voice.

That the essay was set early in the course is also a reason for the criterion of length (criterion 3). This is to counteract the tendency either to write 'everything they know' in the hope that some of it will stick, or, alternatively, to valiantly plough on to get in the

required number of words even though they have run out of pertinent ideas.

As for the construction, blocks of vocabulary and the nuts and bolts of grammar may be dealt with separately elsewhere. The point is that the criteria are changed from task to task depending on the teaching focus and the needs of the learners, based on insights into the teaching/learning process arising from mediated feedback. So, a subsequent assessment focus on language range and accuracy may be deemed necessary, where these criteria are given a higher weighting, maybe even 100%. As the course progresses FA assessments will converge on the criteria of their crucial final paper.

“At this early stage my focus is on other things I want to develop, such as focus and critical thinking, and finding a voice.”

Scoring

The points score is frankly a compromise. It is the common currency of academic assessment. I include it because 'it is what is asked for'. In a vain attempt to escape the tyranny of numeric scores I remember my university awarded A-B-C scores that could get so complex as 'B+?', which much later I discovered corresponded to a score of precisely 67. In my rating scale, 3 points are awarded if I am satisfied with this aspect of the work presented, so a maximum score of 15 would mean the assessee is on course. However, I am aware it might be interpreted to mean the work is perfect and foster complacency. The major factor me, of course, is how the learner-assessee reacts to comments I

make, and in one-to-one tuition sessions we can take up any issues that arise. I am still uncertain as to whether I will have to reconsider the whole business of points giving. At the moment it is not very meaningful.

Conclusion

In view of the high stakes nowadays involved in gaining and above all demonstrating proficiency in English for a particular purpose, it seems that the role of assessment has become increasingly dominant in the teaching/learning process and with it the role of formative assessment has changed, requiring a more systematic approach than it might have had in the General English class of the not-too-distant past. We teachers can no longer trust our pedagogic, linguistic or interpersonal instincts and, in carrying out formative assessment, must align classroom interaction with real world (TLU domain) demands and parameters. This is what I am aiming towards in my described classroom practice and I would welcome feedback from any of my colleagues who share this concern.

References

ASCD (Association for Supervision and Curriculum Development). Accredited to: The Council of Chief State School Officers' (CCSSO) State Collaborative on Assessment and Student Standards (SCASS), Formative Assessment for Students and Teachers (FAST). <http://www.ascd.org/publications/books/108018/chapters/Formative-Assessment@-Why-What,-and-Whether.aspx>

Green A (2014) *Exploring Language Assessment and Testing: Language in Action* (Routledge Introductions to Applied Linguistics). Routledge.



Simon Andrewes regularly teaches EAP on pre-session courses to international students in Birmingham, UK. His main interests currently revolve around assessment and feedback, and communication in lingua franca (rather than native-speaker) contexts. He also occasionally teaches German and Spanish.