

The evaluation of in-house Computer Assisted Language Learning materials: *InGenio Online FCE Course & Tester*

Antonio Martínez-Sáez¹, Ana Sevilla-Pavón¹, Ana Gimeno-Sanz¹

¹CAMILLE Research Group, Department of Applied Linguistics
Universidad Politécnic de Valencia, Spain

Abstract. Nowadays, finding, choosing and developing or adapting materials is a very important component of education as well as a key element for the success of scientific and teaching activities. According to Tomlinson (2003: 1) “every language teacher is a materials developer”, and the materials development process should be “based on universal principles” (Tomlinson, 2010: 72) so as to ensure that local needs and wants do not dictate decisions and that important learning principles are not forgotten. In the case of materials for Computer-Assisted Language Learning (CALL), the same as with other kinds of language learning materials, a very thorough and pedagogically-sound set of guidelines should be followed prior to the design of the materials. Following that, a careful evaluation should take place so as to ensure that prior to the use of these materials the guidelines have been followed and that, as a result of this, the materials are valid, efficient and effective. This is especially important in the case of in-house CALL materials development. This paper describes the basis for the evaluation and validation of the *InGenio FCE Online Course & Tester*, which will take place once the design and development phases of these materials have concluded. The *InGenio FCE Online Course & Tester* are two language learning in-house CALL materials for students of English, developed by the CAMILLE Research Group at the Universidad Politécnic de Valencia (UPV) in Spain. A large part of the focus of this paper will be on the aspects concerning the evaluation of the two previously mentioned materials according to well-established frameworks and guidelines from the most relevant literature.

Keywords: CALL, in-house materials, software development, evaluation.

1. Introduction

Nowadays, finding, choosing and developing or adapting materials is a very important component of education as well as a key element for the success of scientific and teaching activities. When it comes to language teaching and learning, according to Tomlinson (2003: 1), “every language teacher is a materials developer”, and materials development processes should be “based on universal principles” (Tomlinson, 2010: 72) so as to ensure that local needs and wants do not dictate decisions and that important learning principles are not forgotten. The process of Computer-Assisted Language Learning (CALL) materials design usually takes place in several steps in which it is important to make the right decisions on different task-related aspects such as topics and actions, participants and mode (Chapelle, 2003). Once the materials have been created, different quality-related aspects should be borne in mind:

1. Accessibility, user friendliness and validity.
2. Multimedia features of the materials.
3. Clear learning objects and relevant contents.
4. Content presentation, organisation and order.
5. Pedagogical use and orientation of contents.
6. Use of the materials in different training and teaching processes.

Coll & Engel (in Barber àet al., 2008: 65)

2. The evaluation of in-house CALL materials

Chapelle & Douglas (2006) highlight the important role of student assessment, as students usually worry a lot about what is expected from them and try to fulfil the teacher's expectations as far as the tasks and activities they have to complete are concerned, especially if their partial or final grades depend on their performance. Because of this, it is very important to provide students with information on what would be assessed and how, as well as on the type of tasks they will be asked to do in such a way that they know how a correct or incorrect way of solving problems and completing activities will affect their grades. A good way of providing this information is through an interface specifically devoted to the students' grades and individual reports such as the ones shown below, taken from the *InGenio FCE Course and Tester*.

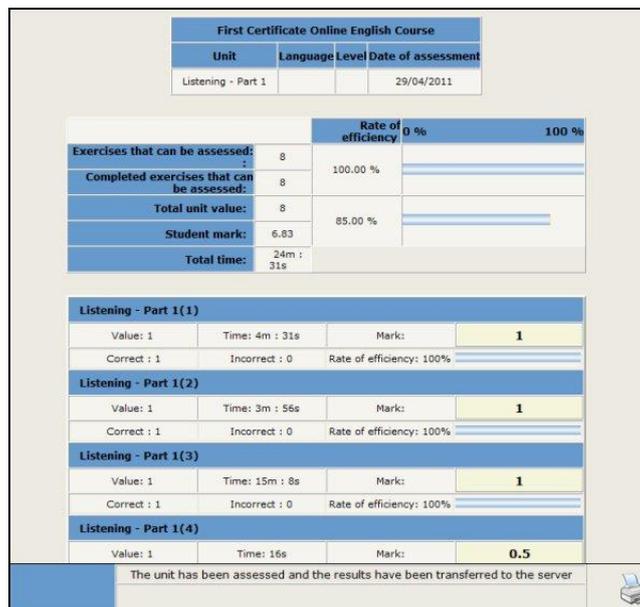


Figure 1: *InGenio* student progress report.

According to Dowsing (1999), designers should ensure that test management, which is usually very time-consuming, is kept to a minimum. He also advises that the interface should be general in such a way that, if necessary, it could be used by a large number of institutions in a greater variety of contexts. When using CALL materials, another aspect to be taken into account is whether the interface will be visible or not, which is a decision the designers have to make. This decision will depend on whether the fact that the interface is visible or invisible threatens to cause variations in the students' performance, which in turn would depend on the kind of tasks to be completed or in what is being assessed (Chapelle & Douglas, 2006). Although the processes of assessment and evaluation usually take place once the design process has been completed, both processes are often complementary and so interconnected that they may even overlap (Levy & Stockwell, 2006). The main reasons why designers and academic staff rely on evaluation are that they need to "assess students' attitudes and perceptions in a learning environment where technology is used, make sure that technology is working the way they had predicted and test whether methodologies and strategies are viable and effective, with a view to improving teaching practices" (Levy & Stockwell, 2006: 40).

There is a varied range of ways to collect data when conducting the evaluation process, the most basic ones being the checklists and surveys, whereas some more complex methodologies comprise multifaceted and longitudinal studies. In the case of the *InGenio FCE Course & Tester*, the designers have chosen to use multiple kinds of data in such a way that quantitative and qualitative approaches are combined. According to several authors, this combination can be a good means of increasing the validity of a CALL study (Harless et al., 1999; Trinder, 2003). As for already-existing evaluation frameworks to be followed, Hubbard (1996) could be a very good starting point, as this author provides some of the key principles that delimit formative and summative/accumulative evaluation. These principles are very useful when evaluating teaching projects and the materials developed as a result of them, such as in the case of the *InGenio Online FCE Course & Tester*. The formative evaluation has taken place during the process of development of the materials in an attempt to obtain as much information and improvements as possible from the process itself; whereas the

summative evaluation will be put into practice once the whole project is finished and the materials which have been developed -the *InGenio Online FCE Course & Tester*- are fully implemented and used by students at the UPV. The questions arising from the formative modality of evaluation have been set out by content designers and writers in such a way that the feedback emitted and the results obtained have a positive impact on learning. The information gathered through the different evaluative approaches is very useful for content writers and designers, who make use of it during the process of adaptation and improvement of the materials. The general aims of the evaluation studies concerning CALL materials such as the *InGenio* materials are listed below (Levy & Stockwell, 2006):

1. To prove that the materials which have been developed are useful and that they respond positively to their specific purposes.
2. To bring them about by taking into consideration specific decisions.
3. To design the contents included by bearing in mind a well-defined target group of users.
4. To obtain practical results.
5. To make sure that not only the evaluation process but also the evaluation product provide useful data and information.
6. To check whether something works or not as it had been planned.

Another important procedure within the evaluative framework used is known as third-party evaluations, which present two important challenges to the evaluator (Levy y Stockwell, 2006): the need to make right decisions as far as appropriate evaluation criteria are concerned and the need to know the programmes or materials to be evaluated as well as their different contents and resources, their options of use in different contexts and with different students, etc. Although the approaches can differ from author to author, according to Burston (2003), there are some requirements that have to be fulfilled by any type of programme. These are: pedagogical validity, curricular adaptability, efficiency, effectiveness and pedagogical innovation. Once content writers or authors have checked that those requirements are fulfilled, they may look at the following categories (Hubbard, 1996):

1. Technical characteristics: reliability, ease of use, etc.
2. Activities (procedure): design of activities and the different typologies.
3. Importance of suiting needs and methodologies followed by teachers (approach): linguistic and learning presuppositions.
4. Importance of suiting students' needs and demands (design).

According to Hubbard (1996), it is advisable to make use of a methodological framework for CALL evaluation based on previous methodologies and on other existing frameworks for the evaluation of the language teaching process. In this way, one of those older frameworks that are still valid and that could therefore be used is the one provided by Richards and Rodgers (1986). These authors proposed a flexible framework which presented an open structure which had the additional advantage of not being limited by a particular language or by a specific teaching or learning process. Chapelle (2001) provided another important framework based on theory and focused on tasks, which provided a new view based on an interactionist perspective (Levy & Stockwell, 2006). Chapelle's (2001) framework comprises the following principles:

1. The evaluation of CALL is determined by a specific situation.
2. Two main perspectives when dealing with CALL evaluation are a critical analysis of software and planned activities as well as an empirical analysis carried out by students.
3. Criteria on the quality of the activities developed should be based on theory and research on SLA (Second Language Acquisition).
4. Criteria should be applied in compliance with the aims of the activities.
5. Learning potential should be the main criterion when evaluating CALL.

According to Levy & Stockwell (2006), the first of the abovementioned principles is related to the fact that the results obtained from a study should not be seen as indicators of proven effectiveness, as the study - as well as the parameters which are involved in the whole process- is much more complex. Concerning the

second principle, Chapelle (2001) mentions two interrelated levels, starting from both levels of critical analysis: CALL software and CALL activities planned by a teacher. In this case, a programme is evaluated in a non-contextualised way, the abovementioned checklists being one of the possible ways to carry out this analysis and evaluation. The second level would be based on an empirical analysis phase by taking into consideration the students' performance while bearing in mind the context provided by the teacher who plans and organises the implementation of a particular language learning programme. A third level based on an empirical and contextualised analysis involves gathering data related to the use that students make of a specific programme while trying to obtain evidence of negotiation of meaning during the interaction process. Bearing this in mind, Chapelle (2001) established a list of criteria to be followed when designing CALL tasks:

Table 1: Criteria for CALL Appropriateness (Adapted from Chapelle, 2001)

Language Learning potential	The degree of opportunity present for beneficial focus on form
Learner fit	The amount of opportunities for engagement with language under appropriate conditions given learner characteristics
Meaning focus	The extent to which the learner's attention is directed toward the meaning of the language.
Authenticity	The degree of correspondence between the CALL activity and target-language activities of interest to learners out of the classroom
Positive impact	The positive effects of the CALL activity on those who participate in it
Practicality	The adequacy of resources to support the use of the CALL activity

The last option, based on a third-party evaluation, could be very positive if conducted by external evaluators with a wide experience in the design and use of CALL materials and programmes such as the *FCE Online Course & Tester*. The data gathered would provide content designers and writers with useful information about the strengths and limitations of the materials created and developed. However, there is also a risk that these external evaluators do not take into consideration the whole spectrum of possibilities, as they would not have taken part in the whole design and development process.

3. Conclusion

Some of the modalities of evaluation and assessment presented throughout this paper have been already put into practice during the design and development processes of the materials in the *InGenio Online FCE Course and Tester*, while some others have been established in the form of a framework that will be used when conducting the final evaluation stages of these materials. Since the phases devoted to the design and edition of contents have already been completed, and so have the activities and tasks, the materials are ready to be used by students at our institution. At this point, the evaluation and validation phases will take place bearing in mind all the parameters listed throughout this paper.

4. Acknowledgements

Acknowledgements are due to the Valencian Regional Government (Generalitat Valenciana) for funding Ana Sevilla-Pavón and Antonio Martínez-Sáez's research grants (FPI).

5. References

- [1] E. Barber à T. Mauri & J. Onrubia (Coord.) *C ómo valorar la calidad de la ense ñanza basada en las TIC – Pautas e instrumentos de an álisis*. Barcelona: Gra ó, 2008.
- [2] J. Burston, J. Software selection: A primer on sources and evaluation. In *CALICO Journal*, 2003, 21 (1): 29-40.
- [3] C.A. Chapelle. *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. Amsterdam: John Benjamins Publishing, 2003.
- [4] C. A. Chapelle & D. Douglas. *Assessing language through computer technology*. Cambridge: Cambridge University Press, 2006.
- [5] R. D. Dowsing. The computer-assisted assessment of practical IT skills. In S. Brown, P. Race y J. Bull (Eds.), *Computer-Assisted Assessment in Higher Education*, Kogan Page Ltd, London, 1999: 131-138.
- [6] W. Harless, M. Zier & R. Duncan. Virtual Dialogues with Native Speakers: The Evaluation of an Interactive Multimedia Method. In *CALICO Journal*, 1999, 16 (3): 313-337.
- [7] P. Hubbard. Elements of CALL Methodology: Development, Evaluation, and Implementation. In Pennington, M. (ed.) *The Power of CALL*. Houston: Athelstan, 1996.
- [8] M., Levy & G. Stockwell. *CALL Dimensions: Options and Issues in Computer-Assisted Language Learning*. New Jersey: Erlbaum Associates, 2006.
- [9] J. C. Richards & T. S. Rodgers. *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press, 1986.
- [10] B. Tomlinson. Introduction: Are Materials Developing? In: B. Tomlinson (Ed.) *Developing Materials for Language Teaching*. London & New York: Continuum, 2003.
- [11] B. Tomlinson. Principles and procedures for self-access materials. In *Studies in Self-Access Learning Journal*, 2010, 1(2): 72-86.
- [12] R. Trinder. Conceptualisation and development of multimedia courseware in a tertiary educational context: juxtaposing approach, content and technology considerations. In *ReCALL*, 2003, 15: 79-93.