Capítulo 6

A four-dimensional rubric model to assess CLIL in primary and compulsory high school education

Alfredo Ibáñez, Oksana Polyakova

6.1. Initial considerations

This paper proposes a model of a rubric to assess Content and Language Integrated Learning (CLIL) that, besides considering the three-dimension descriptors of national curriculum contents, cognition and language (Massler, Stotz and Queisser, 2014), introduces a fourth one: categorising the rubric’s descriptors as input or output scaffolding assigning a Common European Framework of Reference for Languages (CEFR) level to each descriptor. The basis for this decision is that in CLIL instruction there are three stages (Coyle et al, 2010:96) that need to be assessed: meeting input, processing input and producing a response.
These stages require different cognitive skills and therefore need different language interaction. As a result of this, we can assign a higher level of language for input information following Krashen’s input+1 hypothesis (Krashen, 1985) and a lower one for output information (namely students’ productions). The CEFR we use is the "Flexibility in a Branch Approach” (1994–5 Swiss National Science Research Council Survey), edited by the Council of Europe (2001:35).

With this model, we intend that the rubric be a tool for assessment for learning, helping teachers to evaluate systematically the role of language in content scaffolding as well as providing them with a tool to monitor input and output of both language and content. This will facilitate the students’ reflection on the processes required in each sub-stage, raising their language awareness and fostering their understanding of how information is processed in a CLIL environment.

The model explains that in order to achieve language proficiency the CEFR levels in input and language activities do not need to be exactly the same (thus connecting it with the European Language Portfolio) on the one hand and scaffolding and Krashen’s input+1 hypothesis, on the other hand. We propose to assign different CEFR levels to the descriptors depending on whether they describe input scaffolding (allowing a higher CEFR descriptor) or output scaffolding (allowing a lower one).

To our knowledge, our work in this position paper marks the first approach explored specifically for the CLIL assessment rubric. Therefore, the study examines the role of input and output scaffolding in the context of primary and secondary compulsory education. The hypothesis being tested is concerned with whether a special four-dimensional rubric is needed to assess CLIL production and interaction in class.

The rubric model was first implemented in the English Training Courses at the Catholic University of Valencia - Universidad Católica de Valencia (UCV) in the winter of 2017.

6.2. Introduction

Assessment is an essential part of the teaching-learning process, and this is especially true in CLIL. According to Marsh et al. (2011:9), it is a process of "collecting and interpreting evidence" for measuring learning outcomes.

Since CLIL began to be implemented at the end of the 20th century, it has been a major concern the question of how to assess it. According to Kiely (2009), assessment practices that are successful in content classes may not apply to CLIL. One of the most distinguished characteristics of CLIL is its "added value" (Marsh, 2012:181), or in other words, the self-learning benefits that derive from learning in L2 since students must reflect more deeply on which methods are most suitable for each of them.
To better implement the above-mentioned self-learning benefits, Kiely (2008) and Broadfoot et al. (2002) consider that assessment in CLIL must be assessment for learning and not assessment of learning as well as the constant strive for learning progress and teaching development.

CLIL teachers need tools that allow them to assess not only content, cognition, culture and competences but also communication. One of the most distinguished characteristics of CLIL is its "added value" (Marsh, 2012:181), or in other words, the self-learning benefits that derive from learning in L2 since students must reflect more deeply on which methods are most suitable for each of them.

Several attempts have been made to tackle the effect of different assessment instruments. There exists research, such as the CLILA project undertaken by Massler, Stotz and Queisser (2014), which has shed light on how to assess all of these areas by making use of a three-dimensional project. This project came to light after analysing how Swiss and German teachers introduced contents in their countries in the L2.

In their research (CLILA project) Massler, Stotz and Queisser (2014) proposed a three-dimensional project which unfolded:

a. Subject skills and competences.

b. Subject-specific Themes.

c. Communicative Competence in the foreign language

However, despite the fact that CLILA is an ambitious process that deals with and manages to combine content and language, it does not reflect the different modes of interaction that language and content must perform according to the input/output stages. In other words and according to (Coyle et al, 2010:87) “in a CLIL classroom and in a CLIL medium-term planning we need to focus equally on how students meet the content (input) and what they do while learning (how they process the output)”. In our assessment practices at UCV we were able to observe how the linguistic descriptors that our students included in their rubrics were exactly the same regardless of whether the activity was introductory (input) or it was some kind of response (output) that the students had to give.

It was soon clear that a satisfactory CLIL could not be taught if there was no knowledge of the different linguistic modes that are needed for each of the stages of the activity. Reflection on this aspect was essential and since we accept that rubrics should be a tool for learning, according to Black & Wiliam (1998) assessment for learning or formative assessment should provide feedback to adapt teaching and learning activities to the students’ prior knowledge and competences.
In section 1.3 of this study, we will go into detail regarding how we were adapting the linguistic descriptors according to the activity described as input or output.

What became clear in this study is the need for an evaluation tool that contemplates different linguistic interactions according to the stage of application of the activity. This tool should be simple, flexible, standardized and supported by a reference framework common to the European Union, the CEFR.

By this, we do not mean that the linguistic reflection will dominate the delivery of the content. We believe, along with Coyle et al. (2010) that in CLIL courses language teaching pedagogy by itself does not guarantee success in teaching although language teachers’ awareness of how learners learn does.

What we propose in this study is a European model of a rubric that we call "four-dimensional" because it adds, to the three dimensions already collected by Massler, Stotz and Queisser (2014) a fourth one that is the consideration of whether the proposed descriptor corresponds to input or output scaffolding. In this way, we provide a "monitor" (Krashen, 1985) to assist teachers who plan a rubric and raise both teacher and students’ "language awareness" (ibid) when referring to the CEFR and European Language Portfolio (ELP).

6.3. Scaffolding the three sub-stages of CLIL in a lesson plan

CLIL aims at imparting content in a language that is not that of the students, but it does not pretend to emulate EMI (English as a medium of instruction). One of its main objectives is to organize and process input and to facilitate output in a way that, in addition to making it more easily attainable, provides students with insights into how they learn and how they can improve their learning.

In a world in which information and content have greatly exceeded the capacity that humans have to assimilate it, a methodology that is based more on the "how" than on the "what" becomes paramount. We believe that not only can CLIL benefit both teachers and students that learn contents in L2, but it can help them transfer this self-learning knowledge into all educational and personal areas.

To make this possible, it is essential to promote a systematic reflection of the processes that occur in the so-called "CLIL scaffolding" (Coyle et al, 2010:96), namely:

1. Meeting input
2. Processing input
3. Producing a response

When we talk about input and output, we refer to the content and the language or language for learning necessary to access those contents, although it should be noted that the focus of our analysis is on the content because according to Coyle et al (2010)
in CLIL language learning occurs even though the focus on language is peripheral, not intentional.

However, even if the focus is on content and attention on language is peripheral, following Krashen (1985) it is important to address the input +1 hypothesis that says the zone of proximal development (ZPD) set by Vygotsky’s (1986) is achieved when the input is the level of competence of the students +1 level (we take as reference the CEFR of the Swiss agency).

This leads us to consider the different zones of proximal development according to the skills. Receptive skills (reading and listening), which are often used in input activities, allow a higher level of CEFR than the productive skills (speaking and writing) that are used in output.

Systematic reflection on these aspects through the elaboration of descriptors including content, cognition, language skills, CEFR and input/output focus can contribute to introducing organised processes that when repeated frequently in association to content-oriented tasks can secure conceptual and language learning.

Citing Westhoff (2004:60), “our brain keeps track of the regularities in the occurrence of combination and frequency of these combinations. The frequency determines the ‘weight’ of the established connections”. These mental connections created through the above-mentioned neural network are activated in an easier way because of different circumstances. In a sense, a variety of input/output assessment tasks makes CLIL teachers reconsider their approach toward rubric design.

Introducing these concepts in a rubric is of key importance because it monitors the scaffolding process and emphasises those aspects that are essential for CLIL to be effective. In words of Blackwell, Trzesniwsky and Dweck (2007), “adequate scaffolding will take students from a “fixed mode” into a developmental mode”.

This scaffolding helps the teacher evaluate continuously how learners may initially comprehend new material and the role of the language in teaching. Reading Coyle et al (2010:96) again, as teachers’ monitoring input of content and language makes learning “accessible and challenging”.

6.4. A model of implementation of the 4D Rubric: from content-class rubrics to CLIL rubrics, analysis and reframing

The questions and issues discussed so far have emerged from the experience of working with CLIL trainee teachers at the Catholic University of Valencia and other institutions such as Educational Training, Innovation and Resource Centres - Centros de Formación, Innovación y Recursos Educativos (CEFIRE) in Valencia.
This experience was an ideal scenario to assess what our students were doing and to what extent this diverged from what should be done in a CLIL class. From the moment our students began to develop CLIL rubrics, the following information became clear:

- No criteria existed regarding how the language was combined with the content.
- There was no consideration about input-output scaffolding.
- The concepts of "language for learning" or "framework" were non-existent.

The following sample gives a glance of how our students began dealing with the descriptors in rubrics at the beginning of the course and how they eventually developed them recollecting the discussion points proposed in our classes.

**Figure 6.1. Student’s CLIL first rubric draft**

### 6.5. Analysis of the first students’ draft rubrics

The first rubrics made by our students, of which Figure 6.1 is an example, shows a typical situation of a CLIL beginner class. It is important to bear in mind that the assistants to this training course were teachers of content in L1 which makes our analysis more relevant as they had to "reframe" some misconceptions they had in order to "relearn" how to do it in CLIL.
In Figure 6.1, we can see a sample of some of the issues that needed revision:

a) The lack of order in the progression in the descriptors taken from the revised Bloom’s taxonomy (Anderson & Krathwohl, 2001) from Low Order Thinking Skills (LOTS) to High Order Thinking Skills (HOTS) to promote adequate scaffolding:

   In the proposed sample we can see that the first descriptor is “I can explain…”, “Explaining” is certainly part of output scaffolding and therefore not a suitable verb to start a sequence in a lesson plan, as it is HOTS. We would like to highlight the importance of the ordering of descriptors following the model of Bloom’s hierarchies as it paves the way for the rest of the scaffolding in the model of four dimensions that we propose.

b) The lack of indication of whether the descriptors are of Input-output scaffolding and the effects that this has on the skills sequencing (LRSW) and CEFR input +1:

   As we can see in the descriptor of Figure 6.1., there is no information about whether the descriptor belongs to input or output scaffolding. Nor does any information appear regarding the skills used and their sequencing or the Common European Framework of Reference (CEFR).

Through a Socratic questioning exercise, we suggested our students to focus their attention on the following issues when designing a CLIL rubric:

- Identify if the descriptor is defining input or output activities.
- Think about the most appropriate framework to present the content in question.
- Decide the skill in which it would be presented.
- Based on the above, assign a reference level of the Common European Framework of Languages (CEFR).

To facilitate these tasks, we proposed the following key:

L-Listening/R-Reading/S-Speaking/W-Writing

I-Input/O-Output

We have kept the students’ original writing, mistakes included, as this is an example of “realia” taken out from one of our classes.

These are some of the responses we obtained in the second draft:
Figure 6.2. Revised student’s CLIL rubric draft As

can be observed, the changes took place in (see Figure 6.2. above):

1. The sequencing of the LOTS to HOTS objectives (a).
2. The identification of the descriptor as output or input (b).
3. The writing of the language structures (language for learning) (c).
4. The skills used and their sequencing (d).
5. The appropriate CEFR level, depending on whether the activity is input or output (e).

This model does not intend to “prescribe” a particular use of the skills, but to increase awareness of how certain skills may be more or less convenient according to a variety of factors. Westhoff (2004:60) claims that “the frequency determines the “weight” of the connections”. Identifying the adequate CEFR level depending on the activities may have linguistic, pedagogical and institutional effects.

Linguistic effects become more visible because the four-dimensional rubric promotes teachers’ reflection on the students’ ZPD. The systematic decision making regarding the CEFR may develop in the teachers some mastery about the scaffolding of the contents in CLIL that currently does not exist or is not sufficiently developed.

Pedagogical effects emerge as teachers have to make decisions on the language skills that best fit the descriptors thus instilling in the students some procedures that they will embed into their learning.

From the institutional point of view, the systematic introduction of CEFR descriptors in rubrics at European level may help teachers and students following this approach converge on common linguistic and pedagogical backgrounds making it easy for policymakers, state holders and teachers alike to collect and analyse data and evidence.
6.6. A model of implementation of the 4D rubric

As we can see in Figure 6.3 the sequence starts with the selection of a suitable verb from the revised Bloom’s taxonomy (Anderson & Krathwohl, 2001). In order to provide a good scaffolding, the logical thing to do is start with LOTS and gradually increase the cognitive complexity to finish with HOTS. In the sample we have provided a triangle to mark this stage, an “L” or “H” would also be suitable to note whether the descriptor is LOTS or HOTS.

The second dimension, marked here with an oval symbol, indicates whether the descriptor belongs to input or output phase. This can also easily be marked with capital “I” (input) or “O” (output). Marking this second stage will place both teachers and students in the phase of the lesson they are working on and will raise their awareness on the skills used and the level of language they will be dealing with.

The third dimension (labelled with a diamond for the selection of skills, a sun for “input +1” and a double circle for zone of proximal development) indicates the decisions taken, hints at the reasons why they are taken and shows the results they hope to achieve.

The fourth and last dimension, marked here with a star, assigns to each descriptor a level of CEFR according to whether it is input or output scaffolding. This reflection will facilitate the adequate selection of materials by the teachers and will accustom students to deal with skills of different reference levels. It will also familiarise students with the levels of European convergence and raise their language awareness. If used by several institutions, it could be a common basis of interrelation and reference to create a common European evaluation tool.
This fourth dimension includes a sample of “language for learning” (labelled with a leaf here) which shows a sentence or structure the students can use, modify or alter and constitutes the “language framework” the students need to achieve the objectives.

<table>
<thead>
<tr>
<th>I can explain the differences among fundamental characteristics of different religions. (O)</th>
<th>I can explain the difference among different religions making use of comparative structures (“this religion” has the same features than ...). “this religion” have this similar/ different point than ...).</th>
<th>I know but I can’t use structures to expose and differentiate between fundamental characteristics of different religions however I can explain it in L1 but I’m not able to explain them in L2.</th>
<th>I don’t know and I can’t use structures to expose and differentiate between fundamental characteristics of different religions. I need to learn how to express similarities or differences to express expositive structures.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiate RL A2+</td>
<td>Explain SW A1+</td>
<td>Differentiate RL A2+</td>
<td>Explain SW A1+</td>
</tr>
</tbody>
</table>

Figure 6. 4. Explanation of the four-dimensional CLIL rubric

6.7. Conclusions and implications

Throughout this paper, we have highlighted the need to develop specific assessment tools for CLIL. Various authors and studies support this need. In our study we started from the level of 3 dimensions proposed by Massler et al. (2014) and add one more dimension which in turn included the two faces of CLIL and the generic umbrella described by Furlong (2009). These considerations can be collected in a simple way by means of a key code and imply the following advantages:

- It promotes a linguistic reflection on the order of the objectives according to revised Bloom's taxonomy.
- It relates skills with descriptors, depending on whether they are input or output scaffolding.
- It encourages teachers to describe a framework (language for learning) that will allow successful interaction of content and language.
- It places that framework within the CEFR levels applying the principles of ZPD and INPUT +1 of Krashen.
- It provides users with a common tool, easy to use and that can facilitate the convergence in the “assessment for learning” for several European and international schools and educational institutions.
The present study was designed to determine the effect of input and output on L2 scaffolding in a content and language class. Thus, the implementation of four-dimensional CLIL rubric represents a change of assessment paradigm shown by future CLIL instructors. Returning to the hypothesis on the 4D CLIL rubric necessity posed at the beginning of this study, it is now possible to state that the research reveals its usefulness both as a training and teaching tool.

For all the above and for the results obtained in the application of this learning tool with our students at the UCV and CEFIRE in Valencia, we invite educational institutions to study and provide suggestions for improvement based on the fact that the European evaluation assessment tool of the future should provide data and information to improve the areas of evaluation that are specific to CLIL and at the same time it fosters a culture of scaffolding that currently does not appear in any assessment tool.

6.8. References


