

Assessment in CLIL: Challenges and Opportunities



Assessment in Bilingual Education contexts where a Content and Language Integrated Learning (CLIL) approach is used is often seen as a problem by teachers and parents. This article aims to clarify and demystify some of the issues concerning assessment in CLIL, showing how assessment can be used not only to measure, but to support, the learning of both content and language.

 Tom Morton



University of London
t.morton@bbk.ac.uk
web: <http://www.bbk.ac.uk/linguistics/our-staff/associated-staff/tom-morton>
 @TM_Morton



Figure 1. Key questions in assessment



There is an old joke about a tourist in Ireland who asks a local person for directions to a well-known tourist spot. The local replies: “Well sir, if I were you, I wouldn’t start from here”. This is how I sometimes feel about worries and concerns expressed by teachers about assessment in Bilingual Education which uses a CLIL approach. I feel that we are not always starting from the right questions. Very often, it seems that the concerns are about the role of language in giving marks and grades, especially about whether language errors should be taken into account. In this article, I will argue that we need to start asking different questions about assess-

ment in CLIL, and this means that everyone concerned, teachers, school leaders, parents and students, needs to acquire some “assessment literacy”, that is, basic knowledge about the concepts, principles and good practices in assessing learning in contexts where content and a foreign language are taught and learned together. This will require a transformation in our thinking about assessment – from seeing it simply as a process of measuring what has already been learned, to seeing how it can be used to support learning.

The aim of this article, then, is to promote assessment literacy among all stakeholders in Bilingual Education and to begin this process of transforming our mindsets, to begin using assessment to support, rather than just measure, learning. The article is built around four key questions, which, when answered, help us to build a stronger foundation for assessing learners in CLIL and Bilingual Education programmes. These questions relate to what, why, how, and with what we assess (see Figure 1), and they should be considered in the order in which they are presented here. The model is inspired by the work of Kate Mahoney, whose book on assessing bilingual learners is in the recommended reading list at the end of the article.

What do we assess in CLIL?

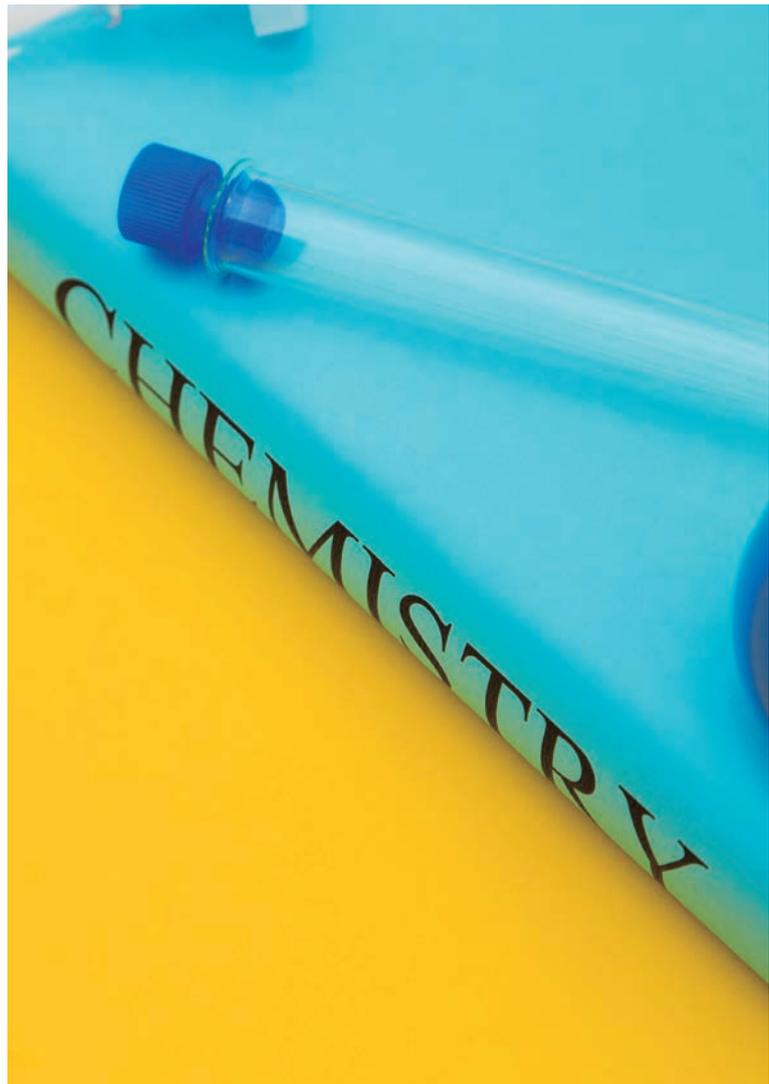
Before we can assess our learners, we need to be very clear about *what* we are assessing. This means that all instructional sequences (units and lessons) need to have very focused objectives. In CLIL, the learning objectives always start from the content, whether it is geography, history, science or mathematics. For example, a learning objective in primary science might be to identify parts of the digestive system and describe their functions. It is the achievement of this objective that the teacher is primarily interested in. If we look closely at the objective, we can see it is really a double objective and the students have to do two things which involve the use of language: identify and describe. Linguistically, identifying is

The key idea here is that we do not have to identify language that we assess as well as content, but we have to identify the language *in* the content

normally quite simple, for example the student may only need to label or match words in a diagram of the digestive system. However, describing is more complex linguistically. It may require the students to write (or say) a few sentences.

When students use language to, for example, identify or describe things in assessment tasks, we are not assessing “language” apart from the content. This language is part of the content. In fact, we can even say that it *is* the content. The key idea here is that we do not have to identify language that we assess as well as content, but we have to identify the language *in* the content. It is in this sense that asking if we should give or reduce marks for language in CLIL assessment is the wrong question. We do not need to add any more language to that which comes with the content and may often *be* the content.

One danger is that if we are not clear about our objectives in CLIL, language may become an “invisible” component of assessment. By this I mean that we may sometimes think we are assessing content when we are really assessing language. This happens when we are tempted to give better marks and grades to students who are more fluent when they speak or write with fewer spelling or grammar errors. In this way, we may be unfairly penalising those students who understand the content very well but have numerous “surface” errors in their writing or speaking. Conversely, we may inflate the grades of those students who have weaker control of the content but fewer language errors. If we want to assess language, then we should make sure it is the language



most closely related to the content – the specific vocabulary, sentence patterns and text types that are needed to express content knowledge – what we call *subject literacy*. We should also share these language objectives with students and support them in using this language, always in the context of content activities. One way to raise students' awareness of the content and language objectives we want them to achieve, is to show them samples of what quality work looks like and get them to share in the work of establishing success criteria.

Why do we assess in CLIL?

Once we know what we are assessing in CLIL, we can think about what use we are going to make of the results of that assessment – in other words *why* we are assessing. It is important to point out that assessment is not just a test we give stu-



CLASSROOM ACTIVITIES

Classroom formative assessment techniques

Objective: to introduce techniques for formative assessment in the Bilingual Education/CLIL classroom.

Context: any subject taught in Bilingual Education/CLIL (including foreign language)

Organisation: teachers will form teacher learning communities (see "Teachers' Agora"). Each teacher will choose one of the formative assessment techniques described in the book by William and Leahy (included in list of recommended texts). Three examples are given below.

Materials: lollipop sticks, coloured plastic cups, slips of paper or post-its.

Development of three example techniques:

- 1) "No hands up except to ask a question". Instead of allowing pupils to put their hands up to answer questions, ask questions at random by using lollipop sticks with students' names written on them. This is formative assessment because it allows teachers to get information about what all the pupils in the class know, rather than the ones who always put their hands up. It is good for language practice too as more students get the chance to speak in whole-class situations.
- 2) "Traffic lights". Give each student three plastic cups – one green, one yellow and one red when they are working individually, in pairs or in groups. They show the green cup when things are clear and they don't have any questions, the yellow cup when they have a question but they can keep working, and the red cup when they are stuck and need help. The questions can of course be about content and/or language.
- 3) "Exit tickets". About five minutes before the end of a lesson, display a question which really gets to the heart of what the lesson was about. If students can answer this question, they will really have understood the lesson. Give each student a slip of paper or a post-it to write their answer. Collect the answers and use them to give feedback in the next lesson. This can really show if members of the class have problems with the conceptual content, and/or the language they need to express it.

dents on one day. Rather, it is a process of eliciting information (evidence) of what students know about what we have taught them, analysing and interpreting that evidence, and using it to make decisions about the students' future, or our own future teaching. In this sense, assessment always has consequences. There are two main uses of the evidence we gather in the assessment process: summative and formative. I will discuss each one briefly in the context of Bilingual Education/CLIL.

When we use the information we have collected through assessment summatively, we award marks and grades to show the extent to which the students have achieved the learning objectives, normally at the end of a period of instruction. These grades are recorded and shared with other people who are important for the student – parents, other (future) teachers, other institutions (universities) and prospective employers. These

Formative assessment is

probably the most important

thing a teacher can do to

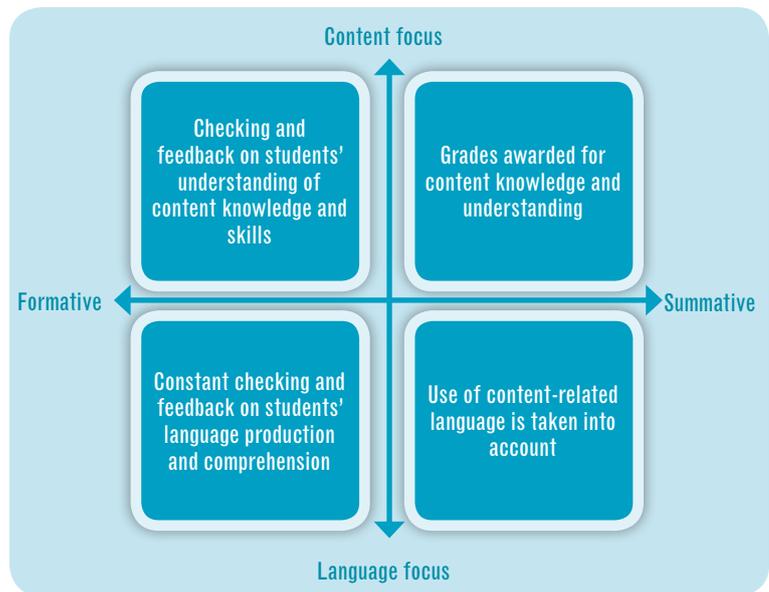
improve students' learning

can be "high-stakes" assessments in that they have important consequences for students' futures. In Bilingual Education/CLIL, assessment for summative purposes is usually focused on the content. We need to know that students who study science or history in bilingual programmes know as much as their counterparts who studied these subjects in non-bilingual settings. Of course, the students in the bilingual programmes are also likely to have learned more of the foreign language, but this can be assessed separately using commercially available language tests, if required.

When we use assessment information not to give marks or grades, but to help students to see how they can improve their work or how we might adjust our own instruction to make it more effective, this is a formative use of assessment. Formative assessment happens during the learning process and is planned and designed to support learning (and teaching), and the achievement of the objectives which will later be assessed summatively. Formative assessment is probably the most important thing a teacher can do to improve students' learning. Many research studies have shown that it is the strategy which has most impact on improving students' achievements as shown in test scores. Unfortunately, due to the mistaken belief that we should spend a lot of time preparing students for summative tests, formative assessment is not used nearly enough by teachers.

In Bilingual Education/CLIL contexts, implementing formative assessment techniques is an extremely important way

Figure 2. Why and what we assess in bilingual education



of avoiding language becoming an “invisible” component in our teaching and assessment. The three techniques shown in the “classroom activities” box are good examples of how we can check students’ understanding of content and discern whether any problems they have are related to content-related misconceptions or language issues. We can then take steps to remedy any problems before it is too late. We don’t want these problems to only emerge during the “high-stakes” summative assessment!

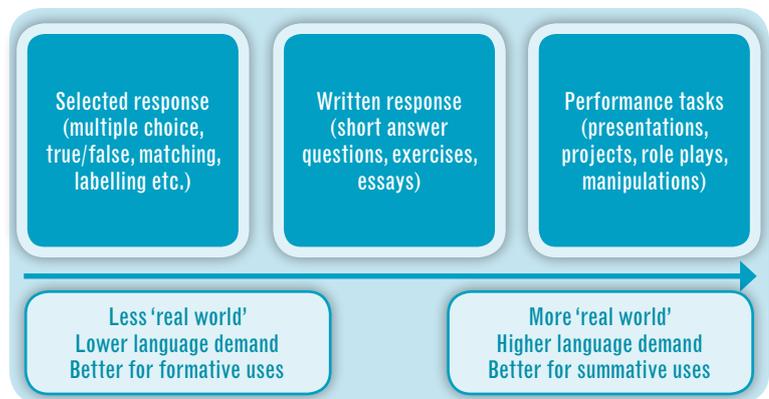
When we combine the content-language and formative-summative dimensions, we get a “menu” of four possibilities for assessment in CLIL (Figure 2). I suggest that a fully comprehensive approach to assessment in Bilingual Education/CLIL needs to pay adequate attention to all four components. For example, just focusing on the upper right-hand component will have several negative consequences: content-related misconceptions are not revealed until it is too late; language may be an “invisible” component of summative content assessment; students do not get formative feedback on either their understanding of conceptual content or the language they need to use to express it.

How do we assess in CLIL?

Having decided what we are going to assess, and how we want to use the evidence we gather, we now need to decide *how* we are going to collect that evidence. We need to choose an appropriate assessment *method*. While the choice of assessment method is important for all educational contexts, it is especially important in Bilingual Education/CLIL because of the language dimension. The choice of assessment method may either enable students who are studying in a second or foreign language to demonstrate what they know about the content or obstruct them from doing so.

Assessment methods can be seen as being more or less ‘real-life’ or authentic (Figure 3). On the right-hand side of the scale, performance tasks ask students to

Figure 3. Assessment methods in bilingual education



perform whole activities, such as giving an oral presentation, creating an artistic object, or making a video. In the middle, written response activities are a very common method in exams. Students have to write short answers to a set of questions. On the left-hand side, selected response items are also a very popular testing method. They are considered more objective and practical (they are quick and easy to mark).

As we move along the scale from left to right, the methods get more ‘realistic’ but also more demanding in terms of language. With performance-type tasks, it may be very difficult to separate students’



TEACHERS' AQORA

Creating teacher learning communities to adopt new assessment practices

Objectives: To enable teachers to support each other in introducing new assessment practices in a Bilingual Education/CLIL programme over a year.

Development: Teachers will form groups of between four and six. The teachers do not need to teach the same subjects, as new assessment techniques can be applied across all subjects. It will be positive if each group includes at least one English language teacher.

The aim is to introduce new assessment techniques over one academic year. At the beginning of the year, teachers should receive input to increase their assessment literacy. It can be on the areas covered in the four questions in this article, or more specifically focused on the strategies and techniques of formative assessment (the book by Wiliam and Leahy in the list of recommended reading is an excellent source of ideas).

It is important that individual teachers choose the areas they want to work on (e.g. establishing and sharing content and language objectives, using specific techniques for questioning students or providing feedback, developing assessment activities with appropriate language demand etc.). Teachers should pair up and observe each other's lessons to give and receive feedback, and the group should meet monthly to share experiences and identify new priorities.

mastery of the content from their language or communication skills, for example in oral presentations. In terms of the use of the assessment, it is better to use more 'authentic' tasks for summative purposes. This is because in complex tasks there are many elements to consider at the same time (both in terms of content and in terms of language and communication), and this makes it more difficult to give detailed, specific and focused feedback to improve performance. If we want to provide more focused feedback on content or language, it is better to break the task up into smaller steps or chunks, get students to practice them and give feedback on specific content and/or language errors.

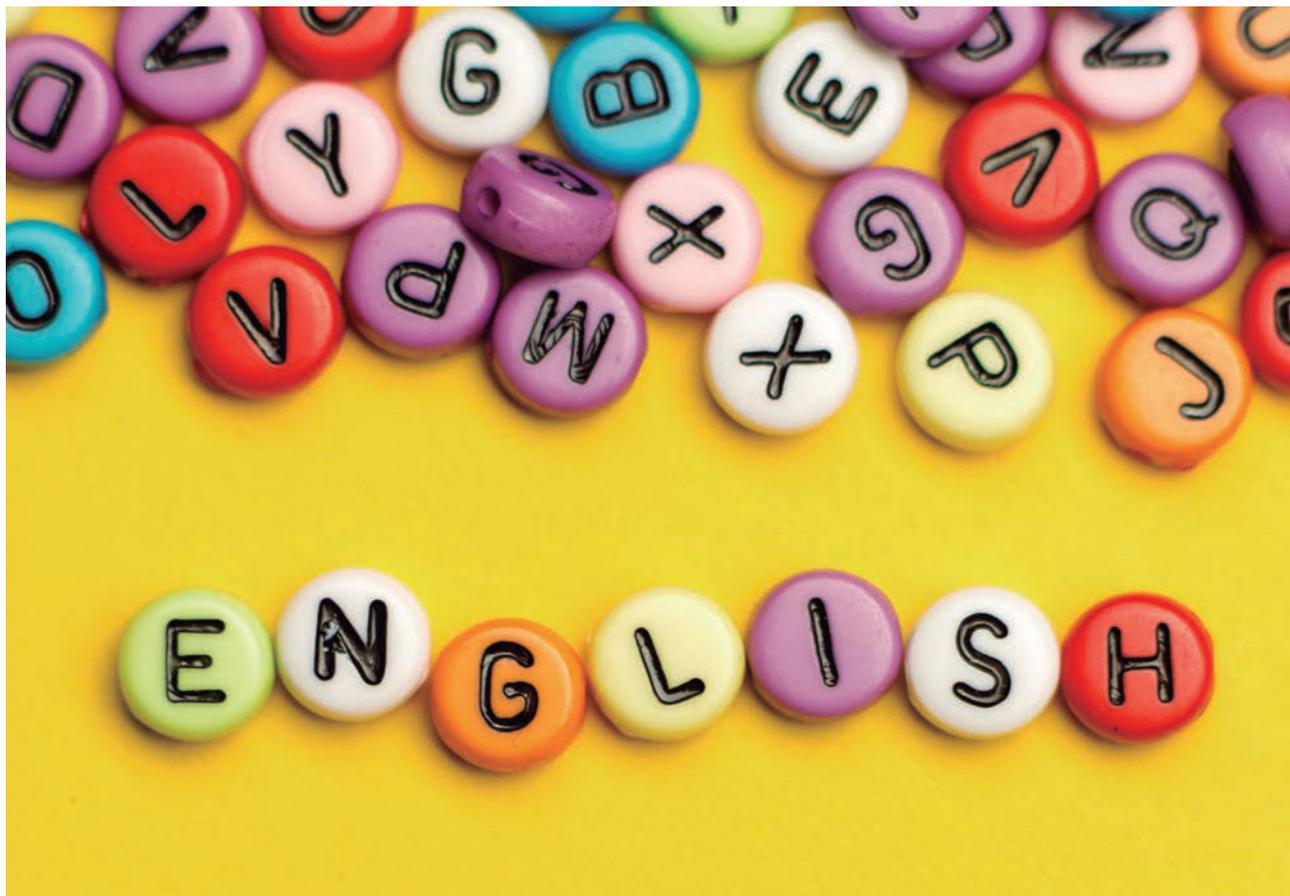
In choosing assessment methods and tasks, we should use the "Goldilocks" principle. This means that the content and language demands should not be too high or too low. The language demand should be "just right" for the students to express their knowledge of the content. For example, if the clearest and best way to assess some content is to use a selected response test (true/false, multiple choice), then there is no point in asking students to produce written responses. This only

increases the risk of language becoming an "invisible" component, which may interfere with the students' ability to show what they know about the content. If, on the other hand, the content knowledge requires an extended written task and/or spoken performance, then the students should receive explicit teaching of the language they need, and this can be incorporated in the learning objectives and success criteria. As we saw in the first question (what do we assess?), this is not assessing language separately, but assessing content through the language that is necessary to express it – the subject literacy. It is assessing content through language, not assessing language through content.

With what do we assess in CLIL?

Once we have decided what, why and how we are going to assess some learning objective, we need to select or create the instruments which will help us to carry out the assessment effectively, fairly and reliably. If we decide to use a multiple-choice quiz, then the instrument is relatively simple: it is the set of questions and the responses that the students select (e.g. by circling, underlining, ticking, or clicking if the test is online). In order to mark this type of test, we will not need any further instruments, except perhaps for a template or answer-sheet. If the test is online, it can be marked automatically.

However, if we decide to use a more "open" type of assessment task (the ones in the middle and the right-hand side of the scale in Figure 3), we will need to select or design instruments to help us. This is because, with open assessment tasks, there is a risk that the assessment process can be subjective, for example with different teachers giving very different grades for the same piece of written or oral work. This is particularly problematic in Bilingual Education/CLIL because language may be an "invisible" component with some teachers confusing language performance (good or bad) with (more or less) content knowledge. This is clearly unfair for students and must be avoided.



The instruments most commonly used to make the assessment of open tasks more reliable are checklists and rubrics. These contain the “success criteria” – what students have to do as evidence that they have met learning objectives. Checklists have the advantage that they are quite simple. They can consist of a list of elements (some can be content-related language elements) with a column for ticking “yes” or “no” depending on whether or not they have been achieved. It is a good idea to add an additional column where we can record what evidence we have that the item was achieved. Rubrics are more complex instruments as they divide the learning objectives into different sub-categories (again, some can be language-related) and also specify different levels of performance (usually four levels). Each level needs to have a descriptor which clearly describes what performance at that level looks like.

In Bilingual Education/CLIL, as in education generally, rubrics are useful for summative purposes. They allow us to put students in categories or “boxes” according to pre-determined levels of achieve-

It is important to bear in mind that changes to assessment practices may be initially resisted by teachers, students and parents and take time to be firmly established

ment. In Bilingual Education/CLIL they have the additional advantage that they can allow us to identify content-related language criteria which we can take into account. However, they are not so good for formative purposes, as the descriptors are often vague, or use language that is not familiar or helpful for students. They can tell a student if he or she is level 3 or 4 but cannot tell the student how to get from 3 to 4. If we want to share learning objectives and success criteria with students, rather than rubrics, it is better to show real examples of other students' work, as we advised in considering the first question (what do we assess?).



WALKING TOGETHER

“Parent Talk”

Objective: to alleviate parents' concerns and to share with them the reasoning behind and the benefits of new approaches to assessment.

Development: “Parent Talk” is an idea developed by the world-leading expert on educational assessment W. James Popham. It asks teachers to imagine scenarios where they have to speak to parents about assessment issues and to prepare what they will say in explaining the reasoning behind their practices and the benefits of the changes they make.

For each assessment strategy or technique which teachers plan to adopt, they should prepare a protocol for a conversation with parents which will explain clearly the reasoning behind the technique and provide evidence of its benefits. To prepare these scenarios, teachers can use a resource like Popham's popular book on classroom assessment (included in the list of recommended reading). Once these protocols have been prepared (this can be done in teacher learning communities as described in the “Teachers' Agora”), they can be used in parent-teacher meetings, and in ongoing communication with parents who may have concerns about the ways in which their child is being assessed.

It is important to bear in mind that changes to assessment practices may be initially resisted by teachers, students and parents and take time to be firmly established. It is for this reason that it is important to move slowly and gradually when introducing new assessment techniques and strategies and to ensure the reasons behind the changes are clearly explained, especially to parents. For example, if schools decide, rightly, to place more emphasis on formative assessment they need to explain its benefits to parents, who may be initially unhappy at



Spanish version



TO KNOW MORE

- CHRISTODOULOU, D. (2017). *Making good progress: The future of Assessment for Learning*. Oxford: Oxford University Press.
- GOTTLIEB, M. (2016). *Assessing English language learners: Bridges to educational equity: Connecting academic language proficiency to student achievement* (2nd ed.). Thousand Oaks, CA: Corwin.
- MAHONEY, K. (2017). *The assessment of emergent bilinguals*. Bristol: Multilingual Matters.
- POPHAM, W. J. (2017). *Classroom assessment: What teachers need to know* (8th edition). Boston: Pearson Education.
- WILLIAM, D., & LEAHY, S. (2015). *Embedding formative assessment: Practical techniques for K-12 classrooms*. West Palm Beach, FL: Learning Sciences International.

We are assessing content

through language, not

assessing language through

content

their child's work being graded less often. For these reasons, the types of activities described in “Caminando juntos” and “Ágora de profesores” are essential to ensure that a change in mindset about assessment in Bilingual Education/CLIL is translated into actions that will benefit all students.

Conclusion

In this article I have argued that the anxiety about assessment felt by many teachers in Bilingual Education/CLIL programmes may be based on misconceptions and a lack of knowledge about the focus, uses, methods and tools of assessment, what is called ‘assessment literacy’. It is hoped that the article has helped to clarify the relationship between content and language in CLIL assessment, the implications of formative and summative uses of assessment for content and language integration, and the importance of calibrating language demand in CLIL assessment activities •



We have been
talking about

**Bilingual education; CLIL (AICLE);
formative assessment; summative
assessment; assessment literacy.**

Este artículo fue solicitado por PADRES Y MAESTROS en septiembre de 2018, revisado y aceptado en diciembre de 2018.