1. Introduction

Content and Language Integrated Learning (CLIL), where a subject is taught through the medium of a second language, is a growing trend in all levels of education in Europe and in other parts of the world. CLIL has always been part of organised schooling systems. Sometimes it has involved children from minority L1s (first languages) where education is not provided in that language. In other contexts it has been in the context of language maintenance projects, for example, children from English-speaking families in Ireland or Wales learning in Irish or Welsh medium schools. The growing trend in CLIL in Europe now however, involves learning through English communities in communities which speak other languages. There are many challenges in implementing CLIL. In this paper I focus on the issues around assessment in CLIL classrooms, schools and education communities.

Assessment in CLIL is a complex area for a number of reasons. First, there is the dual focus – language and subject – which inevitably means there are two assessment processes involved. Key issues here are the extent to which language and subject assessment are integrated, that is, they are assessed at the same time and through the same tasks and activities. Where they are integrated, the impact of the mode of integration on the assessment outcomes needs to be understood. For example, if a child in a primary school assessment task in geography performs poorly, is it because of her limited understanding of the geography concepts or details, because she has not understood the question or because she cannot express her understanding clearly?

Second, there is the purpose of assessment: a learning purpose which focuses on understanding and supporting learning, or an accountability purpose which demonstrates the success of the CLIL policy and implementation. Third, there is a complex set of practical issues, from tests, activities, standards, criteria and the teachers’ skills in bringing all these together in the classroom and in the wider school community. The key issue here is the basis on which a teacher, in relation to either the language or the subject, makes a judgement about achievements in learning (language and subject), or about a need for further work in a given area.

A further complexity is the innovative nature of CLIL and the range of ways in which it affects school life, and the work of teachers. This is particularly important in interpreting the outcomes of assessment processes: where the results are positive, we have a basis for continuing with the policy and practice. Where the results are unsatisfactory, a further set of questions need to be engaged: Are the results an outcome of unsatisfactory tests and assessment processes? Do they derive from problems in implementation in the classroom? Or do they reflect deeper problems with the CLIL policy in that context? In
addition to these innovation concerns, there is a question of expertise in assessment in CLIL. In many contexts, practices in CLIL from lesson design to assessment are novel. This means that the ways in which established assessment practices are valid may not apply to CLIL.

In this article I examine these issues – the dual focus of CLIL, and its implications for assessment, both in terms of assessment purpose and practical issues in making judgement about learning achievements and needs. I draw on perspectives from the field of assessment, teaching and learning, from the rationale and implementation of CLIL, and from personal experience, particularly as a team member and consultant evaluator of the Comenius-funded Pro-CLIL project (http://www.proclil.org/), which is currently looking at the implementation of CLIL at Primary and pre-primary levels in four European countries.

2. The dual focus of CLIL

CLIL is often implemented as a language pedagogy. The goal of the CLIL curriculum is effective competence in a second language (Mackenzie 2008). This has in the past been the situation in language maintenance projects, such as the teaching through Irish in English-speaking regions of Ireland in the 1950s and 1960s (my own initial experience of CLIL); and language imposition, where minority language users are educated in the dominant language (Coyle 2007). More recently, the basis of the communicative approach to language teaching owes much to the experience of people such as Henry Widdowson and Chris Brumfit teaching the school curriculum through English in contexts such as Bangladesh and Tanzania. For them the context of subject learning at school and college was the naturally communicative context in which language teaching and learning could be developed. Currently, as a response to the role of English in a globalised world, stakeholders such as parents, political leaders and employers are advocating the early integration of a second language – typically English – as core curriculum for the development of a 21st century workforce (Dafouz et al 2007). An additional motivation for CLIL is the development and maintenance of multilingualism in contexts where dominant and minority languages co-exist (Serra 2007). CLIL has emerged as a viable strategy for achieving such goals in Europe and beyond. A major element of the rationale here is efficiency in learning: two fields of learning – a school subject and the target second language can be progressed at the same time. The language development, viewed in a communicative framework as a means of understanding and sharing ideas, takes place through exploring concepts.

The theoretical rationale for CLIL is particularly clear and persuasive where the focus is on the target language. Subject learning activities provide a meaning context for the language use, and learning interactions push the developing language resources. In addition the reduced focus on language forms may assist with engagement and confidence. However, good language teaching is not necessarily good content teaching. Merrill Swain articulates clearly the potential tension of education in a developing second language:
Content teaching needs to guide students’ progressive use of the full functional range of language, and to support their understanding of how language form is related to meaning in subject area material. (Swain 1998: 68)

For the subject teacher of science, geography, art, etc., there are two issues: i) the extent to which the essential knowledge and concepts can be learnt as well by all pupils in the CLIL language as in comparable L1 classrooms; and ii) the extent to which the same range of learning opportunities, including activities which develop enthusiasm, motivation and confidence can be engaged in the CLIL classroom as in comparable L1 classrooms.

In contexts where teachers are implementing CLIL, there are many views on these issues. Research into and development of CLIL in many contexts however, focuses on process and achievements in language learning (Lasagabaster 2008). There are two reasons for this. First, the leaders and developers of CLIL initiatives tend to be second language teachers and researchers looking for novel ways of enhancing L2 learning in schools. Second, the CLIL initiative is likely to be innovative in terms of the language of instruction. The subject learning is not itself the focus on curricular change, or a context of dissatisfaction with educational stakeholders. The assumption in CLIL is that the subject curriculum does not change.

This dual focus is a major challenge both for CLIL organisation at school and curriculum level, and for the work of the teaching in the CLIL classroom. This challenge is particularly important in assessment policy and practice.

3. Assessment and CLIL

There are two major types of language assessment. First, there is language assessment as measurement, where the goal is to determine either the level of a student, or the extent to which specific language content has been learned. The former is typically used for programme admissions purposes (Is the language level right for entry to a specific programme?), or placement purposes (Which level class is suitable for a student?). Measures of attainment usually take place at the end of a course, and relate to the specific content and skills taught. Second, there is assessment for learning, a focus of research and development in recent years, which sees assessment practices as integrated into teaching, and oriented, not towards a statement of level, but towards enhanced learning. Assessment is thus viewed as an integral part of the process of teaching and the development of learning opportunities. In Ross’ characterisation of this type of assessment - Formative Assessment – the focus is on the role of the student and on interaction as characteristics of this form of assessment.

The key appeal formative assessment provides for language educators is the autonomy given to learners. […] Assessment episodes are not considered punctual summations of learning success or failure as much as ongoing formation of the cumulative confidence, awareness, and self-realisation learners may gain in their collaborative engagement with
The basic concept of assessment here has resonated across all subjects and contexts of learning. The term Assessment for Learning (AfL) has been used by Black and Wiliam in a broad-based initiative to enhance teaching and learning in all subject areas in British schools and beyond:

[Assessment for learning is] all those activities undertaken by teachers and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged.

Innovations which include strengthening the practice of formative assessment produce significant, and often, substantial, learning gains. These studies range over ages (from five year olds to university graduates) across several school subjects, and over several countries. (Black and Wiliam 1998)

Writing about the experience of children learning language (English) and subject content in British primary classrooms (known as English as an Additional Language (EAL)), Rea-Dickins shows how assessment practice encompasses all teacher work:

Teaching involves assessment. In making decisions about lesson content and sequencing, about materials, learning tasks and so forth, teachers have to determine the strengths and weaknesses of the alternatives available to them. They make selections based on their experience, on their understandings of learning, language development and of language proficiency itself, together with what they consider to be most appropriate and in the best interests of those they teach. Equally, as part of their professional practice, they are always involved in the observation of their learners, which leads to the development of insights about learner progress and judgements about specific learning outcomes and overall performance. (Rea-Dickins 2004:1)

The practices particular to assessment are clear here: teachers observe children as they do activities in the classroom, and they draw conclusions about i) their language and subject learning over time, and ii) the ways the performance of each child maps onto curricular frameworks and assessment bands.

Leung and Mohan, writing about similar classrooms in Britain and Canada describe a similar process, illustrating the extent to which assessment for learning IS teaching:

Assessment for learning should be part of effective planning of teaching and learning. A teacher’s planning should provide opportunities for both learner and teacher to obtain and use information about progress towards learning goals…. Assessment for learning should be recognised as central to classroom practice. Much of what teachers and learners do in classrooms can be described as assessment. That is tasks and questions
prompt learners to demonstrate their knowledge, understanding and skills. What learners say and do in then observed and interpreted and judgements are made about how learning can be improved. These assessment processes are an essential part of everyday classroom practice and involve both teachers and learners in reflection, dialogue and decision-making.

(Leung and Mohan 2004: 335-359)

The particular challenge for teachers in CLIL contexts is engaging with assessment in two fields, the subject and the language. And they have to do it in the context of innovative practice: for many teachers where CLIL is new, there is little local situated knowledge of how this can be done. A description I received recently from one such pioneering teacher implementing CLIL in a German primary school illustrates some ways in which teachers do such assessment through observation in the classroom:

I noticed that doing CLIL modules now for more than a year the pupils want to speak. They want to write already in year 2 (which they should not do in Germany). They want to read. I feel there is a difference to the years before.

Email from CLIL teacher

Here the teacher is describing her class as a group, referring broadly to behaviours which represent confidence and willingness to engage. Within this overview she has much more finely-tuned information on the attainment and progress in both English (the CLIL language) and Art. In her lessons, for example, she uses both English and bilingual worksheets, and directs individual children towards these, according to her assessment of what they are ready for. In her interactions with and feedback to children, she encourages and supports language use according to their individual capabilities.

In this context as elsewhere, because CLIL is new, many people want evidence of effectiveness, evidence that CLIL is a good way to for children to learn both subject and foreign language. The question of effectiveness takes us back to the first type of assessment: assessment as measurement of learning achievements. The CLIL teacher in Germany illustrates clearly the issue here:

I need to show whether the pupils made any progress for the Comenius Project. […] In the subjects I teach in German my pupils have to write "traditional" tests. At the end of the CLIL-units I often let the pupils draw or write or talk about the topic – this is the way I assess. But this is not assessing the progress. […] The problem how to find out where my pupils stand. What do they know? How to measure it?

Email from CLIL teacher

This teacher is skilled at doing assessment for her own purposes as a teacher. She accumulates valuable information on the way pupils make progress through the curriculum. She identifies achievements and difficulties of individual children, and uses differentiation to set appropriate tasks and recycle learning opportunities as required. She relates information gathered in these processes to normative requirements and achievements (traditional tests) for reporting to parents and school colleagues. What is
lacking is evidence to demonstrate the overall value of CLIL as a teaching strategy and policy. There are two gaps in our understanding of assessment in CLIL which still need to be addressed.

First there is the accountability question: how can those remote stakeholders in CLIL – education officials, school leaders, prospective parents, teachers thinking about implementing CLIL understand it as an appropriate strategy for this context? Second, how do we use assessment to manage an appropriate balance in CLIL practice between content and language, such that there is no fear that children achieve less where the learning is in L2? In the sections below I explore these questions further.

3.1 The accountability question

The focus is addressing this issue in CLIL has been on language gain. There are two reasons for this. First, as stated above, the initiative in many contemporary CLIL contexts is taken by experts in language pedagogy, and second, the rationale for the CLIL innovation is often enhanced language development. The tradition in our field for addressing the effectiveness of specific programmes has been language testing, typically with experimental and control groups. One such study in a CLIL context is Lasagabaster (2008): he compared proficiency gain results from two cohorts in secondary schools in Spain, one taught English through CLIL, and the other taught in traditional English as a Foreign Language classes. The results offer tentative support for CLIL, but as the students participating in the initiative were self-selecting, and there was no measure of gains or progress in subject learning, this answer to the accountability question is limited. Studies in Germany (Dalton-Puffer 2007) and Switzerland (Serra 2007) conclude that CLIL is good for language learning, though the results for the subject of CLIL are less clear – Serra for example included standardised mathematics test results, which indicated reduced (though not significantly so) progress in mathematics learning in CLIL classes. A school community or education policy groups would need more evidence before embarking on CLIL as a subject and language learning strategy.

There are many challenges to producing a clear picture of the learning value of initiatives such as CLIL through cohort-based summative assessment profiles. A core difficulty is documentation of the curriculum process: to what extent do aspects of the curriculum and classroom, other than the CLIL initiative, account for the difference in performance? Such factors include the use of L1, the L2 level and confidence of the teacher, the socio-educational status of the families of the participating children (can parents help with Science homework in English for example), and the range and quality of learning materials. Alongside such large-scale studies, we need detailed descriptions of CLIL in context, so that we can understand how teachers are implementing CLIL, and how assessment processes inform their practice.

One issue here is criteria where might consider the following questions:
To what extent are there fixed standards, stages or levels to which the teacher can reference her assessment judgements?
How are teachers interpreting these standards in classroom interaction and in responding
to pupils’ written work?
Where such standards are not fixed, how are teachers’ decisions informed and guided?
For example, if in a CLIL geography task where there are twelve planned vocabulary items on a topic when talking or writing about the relation between food, agriculture and industry in a given community, how does the teacher check comprehension and production of these? Does she use L1 as an aid? What if a child can only use nine of the words? Or if the pronunciation or spelling are inaccurate? Or if the child does not inflect verbs (say, 3rd person ‘s’) accurately, but otherwise uses the words with semantic precision and confidence?

Teachers do assessment like this all the time, and such assessment can serve the learning functions very well. A report by the teacher based on such assessments is likely to address the accountability concerns of local stakeholders such as parents and teachers taking the class the following year. The problem comes in relation to accountability: in order to constitute credible evidence, such judgements need to be located in a matrix of standards and progressive learning schemes. Where these are set out for a whole educational community, and teachers are skilled and confident in using them, then teacher assessment in classrooms can provide valuable information.

3.2 Subject and language balance

The discussion so far in this section has largely focussed on language. This is never enough in CLIL: a similar process of assessment is needed for the subject area. Learning in a subject such as geography needs to be characterised in a similar matrix, so that there is transparency for more remote stakeholders (managers, policy makers, researchers, project evaluators). For example in a subject like geography at primary level, there would be concepts (food, agriculture, industry), data processes (categorising what the children already know according to these concepts), and claim-evidence relationships. Such generic subject knowledge and skills would be specifically related to work on, for example, comparing aspects of life in the European country (typically the local community) with life in a community in an African or Latin American country. The geography learning is engagement with the similarities and differences between the two.

There are two options for the teacher (or more typically, the community of teachers) in carrying out this task: i) do it in the L1, the home language of the school community, or ii) do it in L2 the language of the CLIL initiative. The former is possible where there is interlingual or bilingual teaching, that is where the teacher ensure that in speaking or writing about the geography concepts, each child can do this in L1 and L2. In addition to the opportunities such a model of CLIL presents for assessment, there are two other important benefits. First, the bilingual model enables children to talk to their parents and carers about what they are learning in school, even if they are doing this in L2. Second, in many situations where continuity of subject in CLIL is difficult to guarantee, the bilingual approach enables children to cope with changes in the medium of instruction (that is, between L1 and L2) where this is necessary. In a school for example it may be possible to offer children in Years 3 and 4 a experience of learning geography through English, but not possible in Years 5 and 6.
The theoretical perspective on the roles of the two languages of CLIL have been changing slowly in recent years (Coyle 2007). Early prescriptions emphasised learning through exposure to the target language, which generated policies and practices which concentrated on exclusive use of the target language (Gajo 2007; Serra 2007). More recently an enhanced role for L1 has been part of implementation for a range of reasons from ensuring subject comprehension (de Graaf et al 2007) to facilitating flexible and creative work in classrooms (Coonan 2007).

An appropriate balance between subject and language learning is a key requirement for satisfactory CLIL implementation. This needs to be assured as part of the curriculum in schooling structures, and especially in the management of transitions within and across schools, and in classroom practice. This latter refers to the work of teachers: how they conceptualise good CLIL practice, plan their lessons, activities and worksheets, and in interaction, how their feedback supports sustainable learning in both language and subject.

4. Conclusion

In this article I have explored a range of issues related to assessment within CLIL. I have argued that the dual-focus of CLIL is a fundamental characteristic of the approach, and must guide planning and implementation. Assessment, understood both as periodic sampling and measurement, and ongoing planning for and provision of feedback, can provide a basis for understanding the characteristics of good practices in schools and classrooms. It can also establish evidence of the effectiveness of CLIL for more remote stakeholders in implementation contexts, and guidance for schools and educational communities considering whether to implement CLIL. The development of assessment practice should be based on a dialogue with teachers. This means the development of frameworks which guide teachers in their planning and pedagogic strategies, and supporting teachers in using these to shape schemes of work, lesson plans, worksheets, responses to written work, and especially in micro-interactions with students in CLIL classrooms.

References

Black, P. and D. Wiliam 1998 Assessment and classroom learning. Assessment in Education. 5 (1): 1-74
Coonan, C. M. 2007 Insider views of the CLIL class through teacher self-observation and Introspection. The International Journal of Bilingual Education and Bilingualism 10/5: 625-646
Coyle, D. Content and Language Integrated learning: Towards a connected research agenda for CLIL pedagogies. The International Journal of Bilingual Education and Bilingualism 10/5: 543-562
Dalton-Puffer, C. 2007 Outcomes and processes in CLIL: current research from Europe.
Bio-data

Richard Kiely has taught English and trained teachers in Zambia, France, Malaysia, Mexico, Hong Kong and Poland. He is a Senior lecturer in the Graduate School of Education, University of Bristol where he coordinates the masters in TESOL, directs the CREOLE Research centre, and teaches and supervises research on the EdD and MPhil/PhD programmes.

For more information, please see:
http://www.bris.ac.uk/education/people/academicStaff/edrnk
http://www.bris.ac.uk/education/themes/tesol/
http://www.bris.ac.uk/education/research/centres/creole
http://www.bris.ac.uk/education/programmes/masters/med/pathways/tesol
http://www.bris.ac.uk/education/programmes/doctoral