Toward an Understanding of Content and Language Integrated Learning Assessment (CLILA) in Primary School Classes: A Case Study

Προς την κατανόηση της αξιολόγησης, στο πλαίσιο της Ολοκληρωμένης Εκμάθησης Περιεχομένου και Γλώσσας σε τάξεις του Δημοτικού σχολείου: μια Μελέτη Περίπτωσης

Makrina ZAFIRI and Keti ZOUGANELI

This article focuses on the assessment of students’ learning in a Content and Language Integrated Learning (CLIL) context and presents the findings of a case-study conducted in a Primary School in Greece, where subjects from the general curriculum are taught in English, by EFL teachers. Based on the results of a qualitative exploratory study, the article provides evidence of the teachers’ practices for the assessment of students’ knowledge, abilities and understanding. In the context of the article, the basic theoretical assumptions which underpin CLIL are presented and assessment is discussed as an integral part of language teaching and learning and as an act of safeguarding that the aims of teaching are monitored and achieved. Assessment in CLIL for young learners is presented and discussed as a process which should account for the goals and objectives of two different areas (content and language) and at the same time retain the principles of validity, reliability and appropriacy for the young learners’ context. The article concludes with suggestions towards the development of an assessment framework which encompasses CLIL assessment and methods that exploit existing resources in Greece and in Europe.
sužητάει την αξιολόγηση ως αναπόσπαστο μέρος της διδασκαλίας της γλώσσας και της εκμάθησής της, καθώς και ως δράση μέσω της οποίας διασφαλίζεται η παρακολούθηση της επίτευξης των στόχων της διδασκαλίας. Η αξιολόγηση των μικρών μαθητών, στο πλαίσιο της CLIL, παρουσιάζεται και συζητείται ως μια διαδικασία κατά την οποία πρέπει να λαμβάνονται υπόψη οι σκοποί και οι στόχοι δύο διαφορετικών περιοχών (του περιεχομένου και της γλώσσας), ενώ παράλληλα διασφαλίζεται η αρχή της εγκυρότητας, της αξιοπιστίας και της καταλληλότητας για το συγκεκριμένο πλαίσιο διδασκαλίας σε μικρούς μαθητές. Τέλος, γίνονται προτάσεις για την αξιολόγηση στο πλαίσιο της CLIL με μεθόδους που αξιοποιούν τις υπάρχουσες πηγές στην Ελλάδα και στο εξωτερικό.

**Key words**: Assessment, CLIL, young learners, primary school, L2 learning.

---

**1. Introduction**

Content and Language Integrated Learning (henceforth CLIL) has been introduced as a means to achieve the 1+2 policy aim put forward in the 1995 White Paper on Education and Training by the European Commission, i.e. that all EU citizens should master two community languages in addition to their mother tongue” (Nikula, et al., 2013, pp. 70-71). This need for a ‘multilingual European society’ or the “[...] willingness to communicate (WTC) in the L2 [...]” and “[...] the social nature of L2 acquisition [...]”, as Dörnyei (2001, p.51) names it, has led many European counties to reevaluate and reform their foreign language curricula in an attempt to “[...] nurture a feel good and can do attitude towards language learning in general” (Marsh, 2000, p. 10), thus improving students’ language proficiency. There are scientists and teachers who strongly believe that the advantages of the application of CLIL are many more compared to the disadvantages, more specifically[...]the fact that CLIL is still increasing in popularity as an educational measure suggests that its aims must be important to many people around the globe” (Dalton-Puffer & Smit , 2013, p. 547). Lasagabaster and López Beloqui (2015, p. 55) estimate that “[...] taking into account the prevailing need to learn foreign languages in addition to the mother tongue(s), CLIL may become an effective way to engage students in language learning”.

The development of CLIL over the past decade is dynamically manifested in primary and secondary education. The final report of the European Survey on Language Competences (ESLC) (2012, p.174)\(^1\) shows that CLIL is offered, most often, in secondary education schools in Belgium, Estonia and Malta and least often in schools in Croatia, France and Greece. Analysis of the data for Greece, which appears in the National Report (Dendrinos et al, 2013, p.97)\(^2\), shows that the percentage of the sampled secondary education schools which offer

---

\(^1\) The European Survey on Language Competences (ESLC) was designed to collect information about the foreign language proficiency of students in the last grade of lower secondary education in 16 countries which provided a sample of 54,000 students. The survey comprised language tests and questionnaires which provided contextual information.

\(^2\) Greece participated in the ESLC with 112 lower secondary education schools and a sample 1,594 students of English and 1,378 students of French. The National Report includes a detailed analysis of the language test results and the data from the contextual questionnaires as well as a discussion of the findings and recommendations for policy measures at an in-country level.
CLIL is very low. In Primary education, CLIL provision is favourably discussed as a means towards an early start to reinforcing language learning through content teaching which links topics across the curriculum and includes a variety of subjects (Massler et al., 2014, p.138).

At a classroom level, practitioners often wonder what CLIL really is and what it demands from them. Do teachers restrain themselves from applying it because they see themselves as foreign language experts not content experts, so they feel unsafe applying something which they themselves have not mastered or do they apply CLIL because it is in vogue? Most importantly, what is it that CLIL offers to students and how can the CLIL students’ progress be measured? Assessment of the outcomes of CLIL instruction to students’ progress and achievement appears as a “thorny” issue, due to the dual focus of the method on content and language which requires consideration of the goals of two different subject areas including knowledge, skills, competences and attitudes for both language and content.

The present article deals with the issue of assessment of students’ progress in CLIL and presents the assessment practices of four EFL teachers who implemented CLIL at the Third (3rd) Model Experimental Primary School of Thessaloniki in Greece. It refers to an exploratory qualitative study which aims, on the one hand, to provide insights into when, why and how EFL teachers integrate assessment in CLIL and, on the other, to propose suggestions towards the development of an assessment framework which encompasses assessment, and methods, which exploit existing resources.

2. Theoretical considerations about CLIL

In educational settings, CLIL is rapidly establishing itself as a new educational approach which promotes learning innovation in teaching methods. As Coyle, Holmes and King (2009, p.6) point out, “It [CLIL] encompasses a variety of teaching methods and curriculum models and can be adapted to the age, ability, needs and interests of the learners”, in this way making CLIL a very ‘student friendly tool’ for language teaching and learning.

The CLIL method broaches the subject of foreign language teaching and learning using content subjects, but it is not only that. In CLIL, language learning and content learning are tightly interwoven and integrated and neither seems to dominate the other even though greater emphasis may be placed more on one and less on the other at a particular point in time when there is a specific need. According to Dalton-Puffer and Smit (2013, p.546), “CLIL can be seen as a foreign language enrichment measure packaged into content teaching”.

The theoretical framework for the implementation of CLIL, in primary education, has been shaped by theories which pertain foreign and second language teaching and have influenced relevant pedagogies. An insightful discussion of the impact of Krashen’s Comprehensible Input Hypothesis, Swain’s Output Hypothesis and Long’s Interaction Theory is presented by Mattheoudakis et al (2013, p.218), who also discuss the role of cognitive and constructivist learning theories in the development of a “robust” theoretical base for CLIL. Additional support is offered by Kiely (2011, p.27) with reference to task-based learning, advocated by Communicative Language Teaching. He presents it as an approach, which provides “[…] a degree of conceptual fit between communicative language teaching and the pedagogy of other subjects” but incorporates the risk of prioritising language development over content
knowledge. In order to facilitate parity between language and content Kiely (ibid) suggests that the 4Cs (Content, Cognition, Communication and Culture), framework coined by Coyle (2007) has to be considered. Through the 4Cs, learners construct their knowledge and skills as well as their identity as learners, in a context culturally shaped by two languages. This enhances their understanding of both own and other cultures (see also, Korosidou & Griva, 2013), and promotes their communicative abilities, social skills and motivation to learn the foreign language.

From a socio-cultural perspective foreign language learning for primary school students involves the process of socialization. A child learning English, for example, does not only learn grammar and structure. A wide range of knowledge and skills are also developed such as “learning how to make meaning for communication”, “learning the discourse of the EFL classroom” or “learning the discourse of content areas of the curriculum”. Focusing on the children’s need “to learn the specific discourses of subject content areas such as science and social studies”, McKay (2006) stresses:

“Young learners are already engaging at an early age with beginning versions of discourse of specific content areas [...] as they progress through the elementary years, the content areas become more specialized, and the language used to talk about the content becomes more linguistically complex and academically demanding.” (McKay, 2006, p.33)

In relation to the above Johnstone (2000) introduces the notion of “embeddedness in the flow of events” and suggests that in the primary classroom the learners’ knowledge and experience, gained through subjects across the curriculum, can be linked to foreign language and appear in activities which encourage learners to “draw” knowledge form their L1 and expand it through to L2.

“This natural flow of events in which the foreign language pops in and out of relevant classroom activity reflects a view of the elementary school curriculum in which the universe of children’s knowledge is not divided into discrete areas called “subjects” but is organised more holistically into broader areas that allow children to integrate a variety of different experiences.” (Johnstone, 2000, p.129).

According to a case study of two bilingual students learning English as L3 (Papalexatou, 2013), Johnstone’s notion can expand to having the subjects use all languages interchangeably. Following Brown’s principles of learning, in the above study, there have been several instances of activities relevant to the learners’ interests that built on previous learning and contributed to ‘meaningful learning’, indicating, in this way, that such learners drew knowledge from various experiences and attributed specific roles to different individuals in different situations.

When engaged in CLIL, students, sometimes, have to respond to content meaning making needs which are beyond their present state of knowledge in the foreign language. For example, a child may be able to communicate effectively with the teacher or peers in the language classroom, but may find it difficult to use in L1 specific terminology related to the subject taught in L2. Research from bilingual settings (Cummins, 1987, 2000 in Kiely (ibid)
and in Mattheoudakis et al., 2013; Papalexatou & Zorbas, 2015a; Papalexatou & Zorbas, 2015b; Papalexatou & Zorbas, 2015c; Zorbas, Papalexatou & Griva, 2016) stresses the distinction between basic interpersonal communication skills (BICS) and cognitive academic language proficiency (CALP) and highlights the contribution of CLIL to the development of students’ critical thinking and meaningful use of the foreign language. With reference to the CALP component of language proficiency, Papalexatou (2013, p. 21) argues that “minority children, in particular, must have the common underlying proficiency well developed before entering the classroom, in order to cope with curriculum processes”. This is in line with Kiely (ibid) who makes an interesting point about the role of the students’ L1 particularly when CLIL is implemented in primary education:

“A key strategy in meeting the CALP challenge is continuing development of the pupil’s first language, specifically in terms of the subject language used, in order to understand and explain subject concepts and processes […] so that pupils are able to discuss their subject learning in L1 with parents and carers at the end of each day and in L2 in classroom or formal assessment contexts.” (Kiely, 2011, p. 30)

Commenting on the changing role of the first language in CLIL he goes on to suggest that, recently, L1 implementation is enhanced “for a range of reasons from ensuring subject comprehension to facilitating flexible and creative work in the classrooms”. In the same vein, Papalexatou and Zorbas (2015b) and Zorbas et. al (2016) suggest that teachers should help children retain their L1, by communicating messages about the value of learning foreign or additional languages; thus, highlighting the intellectual and linguistic value of bilingualism. Students are the ones to provide teachers with a ‘knowledge bank’ which the latter can make use of by linking various topics to students’ personal experiences in order to enrich their classrooms both culturally and linguistically. This is also in line with Stathopoulou (2015) who suggests that when teachers facilitate the use of the mother tongue or other languages, brought into the classroom by the students, the latter are encouraged to understand that there can be different levels of proficiency in different languages, used in different situations and for a variety of purposes.

3. Assessment in young learners’ language learning

Teaching is inextricably linked to assessment, which is a means for the documentation of knowledge, skills, attitudes and beliefs and is usually carried out in a measurable way. Defined by Genesee (2001, p.145) as “[…] that part of evaluation that includes the collection and analysis of information about student learning” assessment focuses on understanding student performance in class, identifying students’ specific needs, monitoring the teaching process and providing information about individual students’ progress.

Assessment happens in class continually and is usually discussed in terms of the purpose and the use of information that is provided through its processes. In the classroom it may be formative and summative. Formative assessment is carried out as part of the teaching process, and is central to effective teaching. It is also linked to the notion of “assessment for learning” (Black & William, 1998, in McKay 2006, p.140) and includes a broad range of tools for information gathering, such as self-assessment, peer assessment, performance
assessment and portfolio assessment. The information gathered from formative assessment may complement the input for summative assessment. Summative assessment aims to assess what has been learned at the end of a unit or a period of study. It may be constructed by the teacher, as a set of tasks implemented in the classroom or may come in the form of a formal test. Its outcomes are used to report to others (e.g. parents) about the individual learner’s achievement, for scoring purposes and/or for promotion to the next grade. Summative assessment is not a feedback of the teaching process. On the contrary, it is high-stakes and can have an adverse “wash back effect” to the young learners’ motivation to participate in the language classroom.

Despite the general tendency to define classroom based formative assessment as low-stakes, in comparison to high-stakes testing, Rea-Dickins (2000, p.237) warns us that “there may be cases when high-stakes decisions are made on the basis of a student’s performance in class, which will negatively influence the attention the student gets from the teacher or provision of assistance”. Therefore, when developing formative assessment activities teachers have to pay close attention to the aims of assessment as well as to how these aims will be achieved and how the assessment results will be interpreted and communicated. Formative assessment is frequently an additional element of a valid and reliable assessment plan.

An assessment is valid when it measures what it claims to measure. Reliability refers to the extent to which an assessment is consistent. As formative assessment concerns improving learning and is embedded in the classroom there is reconsideration of the way validity and reliability are examined in classroom based assessment according to McKay (2006). More specifically, she (ibid: 116-117) suggests, that validity and reliability should be “contextualised in the realities of formative assessment in the classroom” and goes on to encourage teachers to “keep a close eye on the characteristics of usefulness as they go about their formative assessment”. The idea of teachers’ self-inquiry when deciding on a formative assessment schedule, as implied by McKay, is discussed under the notion of “Fairness” or “Equity” by Cameron (2001, p. 226). More specifically, she mentions that (ibid: 226): “Equity principles require that children are given plenty of chances to show what they can do and that their language learning is assessed through multiple methods”.

The issues raised so far, reflect theoretical underpinnings to assessment in foreign language learning and address the need to form the base for assessment in the context of CLIL.

4. Assessment in CLIL

The dual focus of CLIL on content and language implies, for language teachers, that they have to teach academic content which they themselves may not have mastered. Consequently, teachers are expected to assess students’ development of language skills and comprehension of the content of the subject matter. Assessing content bears the characteristics of assessing non-language subjects and differs from the modes adopted to assess language proficiency. Usually tests in nature, the latter measure linguistic and communicative competence as well as accuracy, thus focusing on basic language skills necessary to respond to everyday social communication needs (BICS) as coined by Cummins (ibid) in the BICS/CALP distinction. CLIL, however, involves academic language (CALP) found
in subjects and requires use of language in a specific, formal context which does not resemble the way language is used for communication in social informal contexts. Students, who may be fluent speakers and who may have developed interpersonal communication skills, may not be equally proficient in their academic skills, which demand cognitive processes and take longer to develop.

It is pointed out that in CLIL the foreign language is the medium for mastering content (Coyle et al., 2010) therefore both in teaching and in assessment, content must be the focus. Integrating assessment of language and content is a crucial issue, which teachers who implement CLIL have to manage. The issue of integrating content knowledge with language competence in assessment is broached upon by Short (1993, pp. 629-630) who, referring to secondary bilingual contexts, suggests that some types of assessment instruments, such as reading comprehension and writing, involve both content knowledge and language skills. At the same time, she highlights the problem that arises for the teacher about how to assess each element separately.

The situation is not different in EFL contexts and it becomes more complicated with young learners involved in CLIL in the first years of primary education. These children are still in the process of developing their first language (L1) and they may face difficulties in understanding special discourse related to the content of different subject areas such as history, environmental studies or physical education, and furthermore in communicating using the appropriate discourse patterns in the foreign language (L2). Teachers should account for this when devising assessment tasks as well as when providing feedback about progress in content knowledge. The suggestion made by Kiely (ibid) that a bilingual (L1 and L2) approach to CLIL assessment can be adopted in order to ensure a balance between subject and language seems to find fertile ground. According to the suggested model, each child can choose whether to use L1 or L2 when speaking or writing about content concepts.

The idea of allowing alternation of languages, for purposes of monitoring comprehension during assessment, is critically discussed by Coyle (ibid, p. 118) as a problematic one “for both practical and pedagogical reasons”. On a practical level, it can fail because the input for content has been provided through the CLIL language, so it may be unknown in the L1. On a pedagogical level it fails to adhere to the basic aim of CLIL which is “to build capacity to cope fully in an additional language, which includes finding strategies to communicate and developing thinking as far as possible in that language.”

It could be argued here, that a balanced combination of L1 and L2 use in CLIL creates, for the students, a framework of “translanguaging”. The development of such a framework through FL programmes that support linguistic diversity and promote inter-/pluricultural competence, as well as plurilingual competences, in other words “[…] competences in a number of languages from desire or necessity, in order to meet the need to communicate with others […] (Coste, Moore & Zarate, 2009, p.17), including “translanguaging skills” which are strongly suggested by Stathopoulou (ibid, p.214). More specifically she explains that in today’s multilingual contexts, being able to cope with multiple intercultural experiences and to mediate effectively seem to be a prerequisite for an individual’s successful participation in such contexts.
It is up to the teachers to decide which approach to CLIL will be adopted but what needs to be stressed is that each approach requires an appropriate strategy to assessment. When planning assessment in the context of CLIL, the materials, the teaching and learning aims, the teaching method, the instruments of assessment and above all the students’ cognitive and language level have to be considered carefully.

5. The study

5.1. The aims of the study

The present study (conducted in the school year 2015-2016) focused on the case of the 3rd Experimental Primary School of Evosmos in Thessaloniki and investigated the student assessment methods and practices followed by EFL teachers who implemented CLIL through teaching a curriculum subject, other than the foreign language, to students who learn English as part of their curricular studies. The aim was to provide baseline data on CLIL classroom assessment, within the context of young learners. As no previous empirical study of this nature had been conducted at the time, this was an exploratory study (Check & Schutt, 2012, p. 11) which attempted to lay the groundwork for future studies in the area of CLIL assessment in primary education in Greece. The case-study was based on the following research questions:

- When and how do CLIL teachers plan learner assessment?
- Do CLIL teachers focus learner assessment on content or on language?
- What is assessed, in what ways and through what tools?
- What is the role of L1 in CLIL assessment?
- How are the outcomes of assessment ‘put together’, expressed and communicated?

5.2. Methodology of the study

In the context of the study both quantitative and qualitative data was collected in order to safeguard validity of the study (Cohen & Manion, 1997). In particular, online questionnaires, in Greek, were administered to CLIL teachers, through which data concerning the context, as well as the teachers’ experience in CLIL instruction were collected. The research methodology involved mainly face-to-face semi-structured interviews in order to collect direct and accurate information and to identify variables in the teachers’ assessment purposes and practices. Additionally, non-participant observation of two classes took place, in order to facilitate clarification of the teaching process and the assessment practices used during the CLIL sessions.

Overall, four teachers participated in the case - study, three female and one male, all qualified EFL teachers, who implemented CLIL lessons for two hours per week, to classes ranging from grade three to grade six (Table 1).
5.3. Presentation and analysis of the data

5.3.1. The context for CLIL implementation

The 3rd Experimental Primary School of Evosmos, which is supervised by the School of English, Aristotle University of Thessaloniki, is the only public school in Greece that provides intensive English language instruction from the first grade (see Table 1) and one where CLIL was introduced in 2010, on a pilot basis, and has since expanded (Mattheoudakis et al, ibid, p.223).

<table>
<thead>
<tr>
<th>EFL teachers</th>
<th>CLIL subject area</th>
<th>Class level</th>
<th>CLIL taught hours per week</th>
<th>EFL taught hours per week</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>History</td>
<td>3rd grade (sections 1 &amp; 2)</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4th grade</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>T2</td>
<td>Environmental education</td>
<td>3rd grade</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>T3</td>
<td>Religious Education</td>
<td>5th grade</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6th grade</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>T4</td>
<td>Environmental education</td>
<td>4th grade</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Geography</td>
<td>5th grade</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 1: CLIL subjects, class level and teaching time and learners’ exposure to EFL.

Analysis of the data drawn from the contextual questionnaire show that the CLIL teachers were given some training before the introduction of the project, but most of their knowledge was the outcome of self study and cooperation with the content teachers as well as other EFL teachers who also teach CLIL classes. According to their responses, CLIL instruction in their classes focuses mainly on providing knowledge of the subject matter and on promoting L2 skills and communication. They also involved, in their aims, development of intercultural awareness and self-knowledge. CLIL teachers also reported co-operation with the content teacher and other language teachers for the development of the CLIL syllabus in order to exchange teaching ideas and suggestions concerning the evaluation of the CLIL project.

During the interviews, CLIL teachers explained that they designed the CLIL syllabus and developed their own material, taking into consideration the educational goal and the learning aims of the respective subject area, as they are described in the primary school curriculum. They stressed that this was a laborious task and highlighted the need for resources. Their syllabus design did not include assessment as a distinct area of concern nor the development of a set of guiding assessment principles or criteria.

For the needs of syllabus development, CLIL teachers selected, from the course books, written in Greek, areas which could be adapted into English, so that the language meets
their learners’ level in L2 or areas in which the content implied use of specific to the subject vocabulary to learn:

“[…at the beginning I had a look at the book and the syllabus suggested by the ministry, and I saw that some of the things could be left out. My syllabus was designed on the basis of what I can leave out. But I see that during the year you have to be flexible and reorganize according to the needs of the students”.

“I chose from the book chapters that included more scientific terminology and I looked at the description and the aims from the ministry, what the book included […]”

Occasionally, the priority for selecting content concerned its compatibility with the teacher’s knowledge, as well as the resources available and the estimated gains for the learners:

“[…] there are instances in these chapters with information that even I don’t know. I decided to do things about Greece, because it would be more motivating and I have lots of pictures to show them, the vocabulary is also very good, because they will find it sometime in the future in their English classes”.

The practice of ‘teacher made teaching materials for CLIL’ appears common in many EFL contexts. Steiert and Massler (2011, p.100), presenting an example from relevant practice in the context of the PROCLIL project, refers to it as a challenging task because it raises demands for the systematic correlation and integration of content and language learning with the selection of texts and information as well as their methodological design from the beginning.

All teachers in this study reported to use both printed material and ICT applications in their CLIL classes and stressed that multimodality in resources (also see Demace & Zafiri, 2010), facilitated their teaching and increased learner motivation and their participation. Carefully selected materials in CLIL instruction help the integration of content and language and according to Guerrini (undated: 82) they can be scaffolding tools for learning. ICT applications, in particular, connect the CLIL classroom to the students’ everyday realities and practices and facilitate the development of digital literacy skills. In the context of the present study, ICT was also reported as a means towards the enhancement of the teachers’ opportunities for assessment of the teaching process and of the learning processes and outcomes.

5.3.2. Learner assessment

5.3.2.1. Types and purpose of assessment

CLIL teachers in our study claimed to adopt classroom formative assessment in order to collect information about the students’ learning and the teaching procedure: “The children’s answers show me where we stand. How much of the content has been assimilated”.

The main areas assessed formally are content knowledge as well as competences in L2, use of L2 for communication purposes including the purposes of reception and production of the
written form of language and development of mediation skills. *Summative assessment* was also reported as a means towards assessing the students’ progress in content knowledge at the end of a semester or at the end of the year. All teachers in this case-study claimed to keep a balance between assessing content knowledge and L2 development. According to the questionnaire findings students are assessed during every class formatively and after the end of a teaching unit, through a test. The teacher usually assesses students’ understanding of the content, during the class, and they are also given opportunities for self and peer assessment.

5.3.2.2. Assessment tasks and techniques

During the process of devising assessment tasks to check comprehension of content, a serious point of concern seems to be the CLIL students’ level of competence in L2. As the CLIL teachers in this study were EFL practitioners, they were aware of the progression of the L2 level of their CLIL students throughout the school year, so they adjusted the level of difficulty of their formative assessment activities and tasks, accordingly. The teacher of History at grade three explained:

“In the third grade, when we start History they haven’t completed their phonics books, which teach them reading and writing, so I can’t do much as far as reading content is concerned. So the first test they are taking after the first unit in Mythology, which is about the creation of the world, is a test in which they have to put pictures in the right order, so that I know that they know what came first and what followed. As the lessons proceed and they develop their phonics I give them simple matching tasks […]”.

At the beginning of grade three, students had developed only aural/oral skills in L2 and were familiar with identifying facts and characters presented in pictures. So, they worked on an ordering activity which enabled them to show their knowledge of the content without requiring L2 production. Sometime later in the year, they were given a matching activity (Figure 1) that required reading at the word level, which they had developed in L2. It also combined image and language which facilitated scaffolding of both input and output.
In Environmental Studies, in grade three, at the beginning stages of CLIL instruction, comprehension of content and L2 development were checked through arts and crafts. Students were given a map of Greece and, guided by the teacher, had to colour the geographical features.

“[...] they had to colour the mountains brown, the lakes and the rivers blue, the islands red. The instructions were read to them [...] and I checked their comprehension of the words “islands” [...] “mountains”, “plains”.

Additionally, they were given a black and white picture of a mountain with trees, bushes and a river and were asked to colour it and then present it to the class. Later on, as students’ literacy skills in L2 advanced, they worked on reading comprehension activities which integrated knowledge of the content and language (Figure 2 and Figure 3).
The teacher’s process of grading the difficulty and varying the type of the activities according to the students’ cognitive development reflects the idea of “embeddeness” and the effort to “warm-up” in terms of the topic, the ideas and the language that students will need (Johnstone, ibid) in order to respond to the requirements of their tasks. It also provided the context for fair assessment.

CLIL teachers of grade three also referred to summative feedback, collected through activities, which were given as mid-year tests and involved production of written language in order to test content knowledge. The tests were marked and the results were communicated to the students and were available for their parents. According to the teacher of Environmental Education the mid-year test aimed at revising content knowledge and expected from students “[...] to write some sentences from the text which they had to study at home” in order to respond to open ended questions such as: “What do we need to do in order to help the planet?” to which the expected answer was: “we need to save energy”. The teacher explained that some students had difficulty in forming complete sentences and some others provided answers such as: “save energy”, “turn the computer off” or “plant more plants”. All these answers were accepted as correct because they provided evidence that students had understood the question and had transferred information from the subject matter. The students’ answers contained spelling errors, which, however, did not affect their final mark. As the teacher pointed out all answers were accepted “as long as I could understand that they [the students]comprehended the question and their answer gave me what I wanted [...] the information I mean”.

Although this activity functioned as a progress test, which provided a mark, the teacher highlighted its formative value in helping learners to focus on content information and language. Namely, she described a post - test process during which the class went through the answers to the test questions and reviewed language. Coyle (ibid: 2010, p.120) introduces the term “language clinic” and describes this process of reflecting upon language and content as “a necessary step to support better communication of content”. She suggests that it is a useful version of corrective feedback “which undermines content confidence”.

An activity in History (Figure 3) given to grade three learners at the end of the year, in order to collect summative feedback, is indicative of the CLIL teacher’s monitoring of the progress of L2 competences as well as of the effort made in order to scaffold content knowledge so that children will feel self secure.

Figure 3: Activity in History – grade three
The activity was based on the production of written discourse and focused on the composition of a personal booklet for each student which involved pictures given by the teacher on one of the heroes studied within the subject matter of Greek Mythology. Each student had to provide content information in L2 using knowledge gained through CLIL History and through subjects from the general curriculum. The format of the assessment task facilitated the generation of ideas and stimulated information about the content thus, allowing learners to respond in the most direct way, according to their L2 capacity.

As is clearly coined by Coyle (ibid 2010, p.123) “such format activates and organizes thinking to support maximum demonstration of knowledge, thus forming part of the process of working within a student’s zone of proximal development”. In this test, the expected output was quite demanding for the learners as concerns their L2 competency. The CLIL teacher explained that although accuracy in L2 is expected from the students at the end of the school year, keeping a balance between testing knowledge of the content and examining language skills is the key priority: “I’m lenient as far as language is concerned. As long as you can make out the meaning [...].”

Teacher-learner interaction was a technique adopted so as to acquire formative feedback on the students’ understanding of the content. As the CLIL teacher of Environmental Education in the third grade comments: “Because up until Christmas, they couldn’t read or write [...], in every class, I asked them questions, different questions from what we had covered, up to a point, and I checked their comprehension.

Referring to teacher-learner interaction Coyle (2007, p.556) argues that it is a means towards engaging learners cognitively and it generates new language use. Classroom observation, in the present case-study, provided evidence of questioning as a formative assessment strategy and also of strategies such as focusing on content, in order to identify content words, and elaboration when the teacher noticed that some students had not understood a word, a question or what was required by a task. Such strategies are also included in the findings of Tsagari and Michaeloudes (2013) who researched on the formative assessment patterns adopted by CLIL primary school teachers in Cyprus and concluded that ‘questioning’ was the main strategy teachers used to assess content and language. It was used to motivate learners and encourage them to use the target language.

“Understandability”, that is the degree to which students have understood content, is mentioned as the main assessment criterion, by the CLIL teachers, in this study. A second one refers to the ability to understand and use the specific language or terminology that is included in the CLIL content areas. As a third grade teacher comments: “[...] if they managed to understand the terminology”. At the level of input facilitating understanding of content specific language was practiced through teacher simplifications and interaction, as commented earlier in this section.

The issue of language output appropriate to the subject matter was raised by another teacher: “[...] the aim is not if they write correct English, but after all it is science language that they are trying to use, so they should be rewarded for that, they should be encouraged [...].”
Therefore, assessment of language and literacy in content areas involves assessment of the learners’ ability to use the language specific to each subject and, what is more, to use it appropriately for the purpose of the text and the context of communication. As children progress further into the content area the requirement for language appropriacy increases. With upper level grades (i.e. five and six) the complexity and/or sensitivity of the content as well as the aims set by the CLIL syllabus also affect the CLIL teacher’s choices for student assessment. For example, in the context of Religious Instruction the CLIL teacher incorporated “Life skills” among the educational aims. This seems to have influenced both the type of assessment and the way through which assessment feedback was collected: “Let’s not forget that they are learning life skills and that I’m an English teacher. I’m interested in the language, this special language which is academic language, in a sense. […] So, I have to explain. […] What I do is give them examples from life. […] they watch videos […] then we talk about other religions”.

The teacher described a visit to a worship place and explained that the students spoke with people and learned about differences, to conclude: “There is an amazing amount of information that they picked up. How do you assess that? Well, you come back to class and have a discussion. […] we draw idea maps sometimes and I see if they have understood the link between ideas and practices”.

The interaction and elaboration of subject matter content, ideas and meanings, as implied by the aforementioned procedures, act as a scaffolding assessment process which provides formative feedback and stimulates participation, interest and the generation of knowledge.

5.3.2.3. CLIL assessment and project work

Teachers in our study referred to project work which they use in order to assess both content and language. They stressed the contribution of project work in assessing output in a differentiated manner, which allows each learner to show what has been learned according to his or her cognitive development and language abilities.

Referring to project work3, at grade five, on the theme “Love your neighbor” the CLIL teacher of Religious Instructions commented: “The more competent students made a comic strip with a lot of language […] a student who’s excellent at drawing made a beautiful picture with some language. I could see that she had grasped the basic idea and she had some basic language. This is what I expected from her”.

Similarly, grade five students demonstrated their content knowledge in Geography through projects4 which they elaborated upon individually or in groups. These projects provide abundant formative feedback about the students’ learning of the content and their skills in using language specific forms in order to describe objects or facts.

---

3 those interested can see the students’ projects uploaded on the school website, at: http://padlet.com/nkdimos/Godislove.
4 those interested can see the students’ projects uploaded on the school website, at: http://padlet.com/ziakaioa/meet-Greece.
Assessing student knowledge and language development through project work informs the teacher about cognitive strategies and behaviours which are involved in learning and also helps to make instruction more responsive to the learners’ needs.

5.3.2.4. CLIL assessment and L1 vs. L2

In a CLIL context, it is sometimes necessary to mingle L1 and L2 in assessment tasks and activities in order to help students, whose L2 is not sufficient to express content knowledge, to provide evidence of their progress. To this end, an activity in History constructed for grade four learners (Figure 4), presented its instructions in Greek (L1), so that the input language would not be a barrier.

Similarly, the priority to encourage learners, whose L2 skills are not sufficient, to produce oral output relevant to the content in L1 is stressed by the teacher of Religious Instructions: “If you need [...] you can say it in Greek, but I want to hear what you think about it”.

The teacher explained that in mixed ability classes, slower students became intimidated by their classmates who were acquainted much earlier with the CLIL methodology and were eager to communicate in L2. In line with the teacher in our study, Massler (2011, pp.121-122), drawing from research and experience gained from the PROCLIL, EU funded project clearly states:

“[...] in case students lack adequate L2 skills to do so, (show what they have learned in L2) they should either be supported to respond through non-verbal means or allowed to use their L1. Sometimes a mixture of L1 and L2 may be allowed in order for the students to express their content knowledge and so as not to put weaker students at a disadvantage. Overall, it could be argued that pre-primary and primary school children in CLIL programmes should be allowed to choose the language in which they respond to an assessment task [...].

Figure 4: Activity in History-grade four (source: Koutalakidou, 2014)
information provided to the teacher by the student’s choice of language can provide valuable information as to their foreign language competence”.

5.3.2.5. ICT and CLIL assessment

CLIL teachers, in our case - study, claimed to make use of the opportunities that ICT offers for raising students’ motivation to work on and explore the subject matter as well as for facilitating teachers to monitor and assess their teaching: “Learner assessment [...] comes every step of the way actually. I use a lot of ICT tools. I use a lot of games and I use them both to give practice, opportunity for practice for the students, but also as feedback for me to see what students understand, what I need to revise, what I need to go over again [...]”.

The teacher also pointed out that interactive games in educational platforms motivated learners to use their knowledge of the CLIL content area in order to proceed to another game or do a crossword: “... they have to answer questions (based on content) before they get to play”. The challenge of providing an appropriate answer led to the learners’ reflection upon content knowledge and raised their awareness of their progress. Thus, they were informally involved in a process of self-assessment. Interactive games in educational platforms were also been commented upon as useful sources of feedback about when and how many times students played a game and their scores. The same games, played in class can show “[...] how they worked at home and how well they know their material”. Moreover, as electronic games can be played at home, parents were able to see what their children had learned and could follow their progress.

6. Discussion of the study’s findings and suggestions for further practice

Generally, the teachers in this study make a clear effort to integrate content and language in their assessment practices. They use formative assessment as a continuous process which is inherent in their teaching process, is linked to learning and to educational goals and appears as the outcome of interaction between the teacher and the learners or the learners and the learning content. Summative assessment occurs in the form of teacher made tests, the marks of which are communicated to the students and to their parents together with general comments about the students’ overall performance. Thus, tests have both a summative and a formative function. McKay (2006, p.68) points out the distinction between formative and summative assessment is “blurred” for the teachers.

It can be argued that the assessment practices, recorded in the context of this exploratory study, do not appear as part of an assessment schedule, which would link to the aims of the syllabus and would incorporate clearly defined criteria for student performance and progress and a variety of types of assessment and more particularly alternative assessment. Such a schedule would reinforce validity and reliability of the formative and summative assessment processes. An example of assessment criteria, which is separate for content and language, is provided by Calabrese and Rampone (2009) on the Theme “Growing” (Table 2). These criteria are presented in the form of Can Do Statements.
Theme: Growing

<table>
<thead>
<tr>
<th>Content criteria</th>
<th>Language criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>After completing the unit, students will be able to:</td>
<td>After completing the unit, students will be able to:</td>
</tr>
<tr>
<td>• Distinguish living things from non-living ones</td>
<td>• say what living things can do</td>
</tr>
<tr>
<td>• Identify the characteristics of living things</td>
<td>• say what non-living things cannot do</td>
</tr>
<tr>
<td>• Sort and classify according to chosen criteria</td>
<td>• recognise simple words and match them with pictures</td>
</tr>
<tr>
<td>• Identify and describe living things in a work of art</td>
<td>• describe and complete a picture according to instructions</td>
</tr>
<tr>
<td></td>
<td>• use content specific language</td>
</tr>
</tbody>
</table>

Table 2: An example of assessment criteria

Such criteria can be organised in a one-sheet table of descriptors that will be used by the teacher for both formative and summative assessment. The table should involve a column for the teacher’s rating, which can be presented through expressions such as ‘very well’, ‘well’, ‘unsatisfactory’.

It was quite clear, by the teachers’ stance in this study that they strongly believed in alternative assessment and practiced it informally. Alternative assessment techniques offer advantages, since they can help meet the needs of various learning styles, involve criteria which provide detailed feedback of what students can do and allow student involvement in self and/or peer assessment.

A self assessment instrument on the Theme “Growing”, presented above, could entail descriptors similar to the ones for the teacher, which have been adapted linguistically to meet the needs and understanding of young learners (Table 3).

<table>
<thead>
<tr>
<th>WHAT I CAN DO</th>
<th>😊</th>
<th>😊😊</th>
<th>😊😊😊</th>
</tr>
</thead>
<tbody>
<tr>
<td>I can classify animals according to the characteristics which they have in common.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can identify similarities and differences between animals.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can describe some animals in English.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can say in English what animals eat.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can say the names of some animal and their young ones in English.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can put pictures and phrases in order to show the growth process of a frog or a butterfly.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can write a comic story about the growth process of a frog.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
<tr>
<td>I can tell stories about animals.</td>
<td>😊</td>
<td>😊😊</td>
<td>😊😊😊</td>
</tr>
</tbody>
</table>

Table 3: Example of content and language descriptors for self assessment

Portfolio assessment can also be used as a tool of alternative assessment for summative assessment purposes. A student’s portfolio usually involves samples from his or her work over the year as well as tests and self-assessment forms.
An answer towards a framework for assessment, which integrates content and language, is attempted by Barbero (2012, p.42). It is based on Mohan’s (1986) knowledge framework which considers knowledge in relation to language at the levels of: 1) classification/concepts 2) principles/processes 3) evaluation/creation and their language manifestations: 1) description 2) sequence 3) choices. Additionally, this “conceptual” framework involves thinking skills in the form of lower-order processing (e.g. defining, identifying) and higher-order processing (e.g. explaining, hypothesizing).

The development of frameworks for content and language assessment needs further research in order to assist teachers’ understanding of the discourse features of content tasks and to enhance the validity and reliability of assessment.

7. Concluding Remarks

This study focused upon the assessment of students’ learning in a Content and Language Integrated Learning (CLIL) and presents the findings of a case-study conducted at the Third (3rd) Model Experimental Primary School of Thessaloniki in Greece. The teachers of the aforementioned school designed their CLIL syllabus and developed their own material, taking into consideration the educational goals and the learning aims of the respective subject area, as this is described in the primary school curriculum. They, also, made a clear effort to integrate content and language in their assessment practices. They used formative assessment as a continuous process which was inherent in their teaching process, was linked to learning and to their educational goals and appeared as the outcome of interaction between the teacher and the learners or the learners and the learning content. Summative assessment occurred in the form of teacher made tests, the marks of which were communicated to the students and to their parents together with general comments on the students’ overall performance. Thus, tests had both a summative and a formative function.

However, the CLIL syllabus design did not include assessment as a distinct area of concern, nor did it foresee the development of a set of guiding assessment principles or criteria. It can be argued that the assessment practices, recorded in the context of this exploratory study, do not appear as part of an assessment schedule which would link to the aims of the syllabus and would incorporate clearly defined criteria for student performance and progress and a variety of types of assessment and the particularly alternative assessment. Such a schedule would reinforce validity and reliability of the formative and summative assessment processes.

The data collected from this study clearly shows that, assessment in the context of CLIL is a challenge for the teachers who are obliged to develop their own materials rather than have access to materials designed for CLIL instruction. Moreover, the development of frameworks for content and language assessment needs further research in order to assist teachers’ understanding of the discourse features of content tasks and to enhance the validity and reliability of assessment.

Nevertheless, this is a small-scale study, and despite the positive feedback of the teachers who participated, it is necessary that more research be conducted in the field, with more
teachers applying the CLIL and its assessment in their teaching process, before we can come to any safe conclusions.

References


(Eds), Cultural Diversity, Equity and Inclusion: Intercultural Education in 21st century and beyond. International Association for Intercultural Education.

Dr Makrina Zafiri (m_nzafiri@yahoo.gr) holds a PhD in First and Second Language Teaching from the University of Patra, an MA in Theoretical and Applied Linguistics from the Aristotle University of Thessaloniki, a second MA in Studies in Education from the Open University of Greece and a third MA in the field of Sociology, Philosophy and History of Law, from the Aristotle University of Thessaloniki. She has two degrees, one in Law and a second degree in English Language and Literature, both from the Aristotle University of Thessaloniki. She has worked with students of all language levels and age groups and is currently working as a permanent ESP/EAP language teacher at the University of Thessaly in Greece. She also teaches at the Open University of Greece and more specifically for the program ‘Master in Education in Teaching English to Speakers of Other Languages’ and trains teachers of other disciplines in ASPAITE.

Keti Zouganeli was an experienced EFL teacher, Teacher Trainer and educational researcher. She held a BA in English Language and Literature from the University of Athens and an MA in Teaching English to Young Learners from the University of Warwick in England. Her main interests were in the training of teachers of English and her publications referred to issues of didactics and assessment. Keti passed away in September 2016.