



Universidad  
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES  
DE POSGRADO Y DOCTORADO

**ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL**  
(FOR EVALUATION OF THE ACT DOCTORAL THESIS)

Año académico (academic year): 2017/18

DOCTORANDO (candidate PHD): **OTTO CANTÓN, ANA ISABEL**

D.N.I./PASAPORTE (Id.Passport): \*\*\*\*374E

PROGRAMA DE DOCTORADO (Academic Committee of the Programme): **D404 LENGUAS MODERNAS**

DPTO. COORDINADOR DEL PROGRAMA (Department): **FILOLOGÍA MODERNA**

TITULACIÓN DE DOCTOR EN (Phd title): **DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ**

En el día de hoy 05/03/18, reunido el tribunal de evaluación, constituido por los miembros que suscriben el presente Acta, el aspirante defendió su Tesis Doctoral **con Mención Internacional** (In today assessment met the court, consisting of the members who signed this Act, the candidate defended his doctoral thesis with mention as International Doctorate), elaborada bajo la dirección de (prepared under the direction of) ANA HALBACH.

Sobre el siguiente tema (Title of the doctoral thesis): **ASSESSMENT IN CLIL. THE BALANCE BETWEEN THE CONTENT AND THE LANGUAGE: MADRID BILINGUAL SECONDARY SCHOOLS AS A CASE STUDY**

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL<sup>3</sup> de (no apto, aprobado, notable y sobresaliente) (After the defense and defense of the thesis, the court agreed to grant the GLOBAL RATING (fail, pass, good and excellent): **SOBRESALIENTE**

Alcalá de Henares, a 5 de marzo de 2018

Fdo. (Signed): .....

Fdo. (Signed): .....

Fdo. (Signed): .....

FIRMA DEL ALUMNO (candidate's signature),

Ana Otto

Fdo. (Signed): .....

Con fecha 19 de marzo de 2018, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado, a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- Conceder la Mención de "Cum Laude"
- No conceder la Mención de "Cum Laude"

La Secretaria de la Comisión Delegada

<sup>3</sup> La calificación podrá ser "no apto" "aprobado" "notable" y "sobresaliente". El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad. (The grade may be "fail" "pass" "good" or "excellent". The panel may confer the distinction of "cum laude" if the overall grade is "Excellent" and has been awarded unanimously as such after secret voting.)



Universidad  
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES  
DE POSGRADO Y DOCTORADO

En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 19 de marzo, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por OTTO CANTÓN, ANA ISABEL, el día 5 de marzo de 2018, titulada ASSESSMENT IN CLIL. THE BALANCE BETWEEN THE CONTENT AND THE LANGUAGE: MADRID BILINGUAL SECONDARY SCHOOLS AS A CASE STUDY, para determinar, si a la misma, se le concede la mención "cum laude", arrojando como resultado el voto favorable de todos los miembros del tribunal.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado **resuelve otorgar** a dicha tesis la

**MENCIÓN "CUM LAUDE"**

Alcalá de Henares, 5 de abril de 2018  
EL PRESIDENTE DE LA COMISIÓN DE ESTUDIOS  
OFICIALES DE POSGRADO Y DOCTORADO



*F. Javier de la Mata*

F. Javier de la Mata de la Mata

**Copia por e-mail a:**

Doctorando: OTTO CANTÓN, ANA ISABEL  
Secretario del Tribunal: THOMAS MORTON  
Director/a de Tesis: ANA HALBACH



Universidad  
de Alcalá

ESCUELA DE DOCTORADO  
Servicio de Estudios Oficiales de  
Posgrado

DILIGENCIA DE DEPÓSITO DE TESIS.

Comprobado que el expediente académico de D./D<sup>a</sup> \_\_\_\_\_  
reúne los requisitos exigidos para la presentación de la Tesis, de acuerdo a la normativa vigente, y habiendo  
presentado la misma en formato:  soporte electrónico  impreso en papel, para el depósito de la  
misma, en el Servicio de Estudios Oficiales de Posgrado, con el nº de páginas: \_\_\_\_\_ se procede, con  
fecha de hoy a registrar el depósito de la tesis.

Alcalá de Henares a \_\_\_\_\_ de \_\_\_\_\_ de 20 \_\_\_\_\_



Fdo. El Funcionario

ANA HALBACH, Profesora Titular de Universidad del Departamento de Filología Moderna de la Universidad de Alcalá, tiene a bien informar:

Que el trabajo realizado por Dña. Ana Otto Cantón titulado *Assessment in CLIL: The Balance between the Content and the Language. Madrid Bilingual Secondary Schools as a Case Study* reúne los requisitos científicos de originalidad, metodología y documentación necesarios para poder ser presentado y defendido como Tesis Doctoral. La formulación de los objetivos, la revisión bibliográfica y la descripción del contexto son adecuados, y dan paso a un acertado diseño de la investigación así como a un posterior análisis de datos y evaluación de los mismos que permiten extraer conclusiones significativas sobre el papel otorgado a la lengua en la evaluación en las asignaturas de contenido impartidas en lengua extranjera en las secciones bilingües de los Institutos de Enseñanza Secundaria de la Comunidad de Madrid. Los resultados obtenidos, en opinión de la abajo firmante y directora de la tesis, son interesantes, y contribuyen al desarrollo del conocimiento de este aspecto específico de la enseñanza bilingüe.

Para que así conste, y surta los efectos oportunos, firmo el presente informe en Alcalá de Henares a 15 de noviembre de 2017.





CRISTINA TEJEDOR MARTÍNEZ, Coordinadora de la Comisión Académica del Programa de Doctorado en Lenguas Modernas: investigación en lingüística, literatura, cultura y traducción

**INFORMA** que la Tesis Doctoral titulada *Assessment in CLIL: The Balance between the Content and the Language. Madrid Bilingual Secondary Schools as a Case Study*, presentada por D<sup>a</sup> Ana Isabel Otto Cantón, bajo la dirección de la Dra. Ana Halbach, reúne los requisitos científicos de originalidad y rigor metodológicos para ser defendida ante un tribunal. Esta Comisión ha tenido también en cuenta la evaluación positiva anual del doctorando, habiendo obtenido las correspondientes competencias establecidas en el Programa.

Para que así conste y surta los efectos oportunos, se firma el presente informe en Alcalá de Henares a 29 de noviembre de 2017





Universidad  
de Alcalá

**Programa de doctorado en lenguas modernas**

**ASSESSMENT IN CLIL: THE BALANCE BETWEEN THE  
CONTENT AND THE LANGUAGE. MADRID BILINGUAL  
SECONDARY SCHOOLS AS A CASE STUDY**

**Tesis doctoral presentada por  
ANA ISABEL OTTO CANTÓN**

**2017**



**Programa de doctorado en lenguas modernas**

**ASSESSMENT IN CLIL: THE BALANCE BETWEEN THE  
CONTENT AND THE LANGUAGE. MADRID BILINGUAL  
SECONDARY SCHOOLS AS A CASE STUDY**

**Tesis doctoral presentada por  
ANA ISABEL OTTO CANTÓN**

**Directora: Dra. Ana Halbach**

**Alcalá de Henares, 2017**





## RESUMEN

Debido a la variedad de modelos CLIL/AICLE en Europa y a la ausencia de pautas específicas sobre como evaluar en los programas bilingües, los escasos estudios sobre evaluación en CLIL muestran que existe una gran disparidad en cuanto a la manera en la que ésta se lleva a cabo en la práctica. Además, la lengua vehicular desempeña un papel primordial en la integración curricular puesto que las expectativas lingüísticas suelen ir integradas en los criterios de evaluación aun cuando el nivel de competencia en lengua extranjera no se pueda equiparar al nivel de competencia en L1.

Este estudio presenta una doble dimensión. Por un lado, pretende identificar las dificultades que el profesorado de asignaturas no lingüísticas (ANL) encuentra en el ámbito de la evaluación. Por otro lado, el estudio se ha concebido con el objeto de analizar las prácticas de evaluación por parte de este profesorado de las Secciones Bilingües de la Comunidad de Madrid; es decir, los instrumentos de evaluación que se usan y, en concreto, si el nivel de inglés de los alumnos tiene algún efecto en la calificación de estas asignaturas.

Siguiendo esta doble dimensión, el estudio se divide en dos grandes bloques en relación, en primer lugar, con el marco teórico de CLIL y la evaluación y, en segundo lugar, con el contexto de la investigación. Para el análisis empírico de este último, se ha optado por una metodología mixta consistente en el uso de cuestionarios, entrevistas, *focus groups* y el análisis de un corpus de instrumentos de evaluación. Los hallazgos de la investigación arrojan datos relevantes en relación con la formación del profesorado de centros bilingües y la ausencia de evaluación formativa en el contexto del estudio. Así, se observa que el examen escrito prevalece como el instrumento de evaluación más frecuente y que, en efecto, el nivel lingüístico de los alumnos se tiene en cuenta a la hora de calificar la asignatura. A partir de estos resultados se formulan una serie de recomendaciones para el profesorado de las Secciones Bilingües de la Comunidad de Madrid.



First of all, I would like to express my most sincere gratitude to Dr. Ana Halbach, for her endless support, trusting the research even when it was an initial stage, and dedicating a tremendous amount of time. I am also very thankful to Professor Christine Penman, who invited me to complete my international stay in Edinburgh, and to my friend Xabier San Isidro, who made me feel home out of home.

No informants, no research. This work would have been impossible without all the anonymous professionals at Bilingual High Schools in Madrid. I am especially grateful to Helena Massó, Elena del Pozo, Manuel Blázquez, Marcia Loma and Cristina Tilve, because their compromise with bilingual education is the proof that excellence is here to stay.

My gratitude also goes to my colleagues at Nebrija University: Dr. López-Medina and Dr. Estrada, and to the staff at Nebrija Library for their encouragement and the ‘human touch’ in the last phase.

A special thanks to my friends Soledad Hidalgo, and Mercedes Heras, who endured my endless monologues about the thesis.

This work is dedicated to my husband, Jesús, who never hesitated to follow me and my dreams, even if that meant abroad, and because everything I did was because of his patience. Also, to my son, for his love and support throughout the process.



## LIST OF GRAPHS

## LIST OF TABLES

## LIST OF ABBREVIATIONS

## INTRODUCTION

## SECTION 1: BACKGROUND

CHAPTER 1: Content and Language Integrated Learning	1
1.1. What is CLIL?	2
1.2. CLIL in Europe	6
1.3. CLIL in Spain	17
1.4. The role of language in CLIL	22
CHAPTER 2: The CAM Bilingual Project	49
2.1. Origins and structure of the CAM Bilingual Project	49
2.2. Teachers and their training	61
2.3. Challenges in the implementation of the CAM Bilingual Project	66
2.4. External evaluation	68
CHAPTER 3: Assessment	73
3.1. What is assessment?	73
3.2. Types of assessment	75
3.3. The impact of assessment in education: Washback effect	83
3.4. Assessment regulations in Spain	85

3.5. Assessment in CLIL	97
SECTION 2: THE STUDY	
CHAPTER 4: Research Methods	129
4.1. Research questions	129
4.2. Research design	130
4.3. Participants	133
4.4. Research phases	134
4.5. Limitations of the study	150
CHAPTER 5: Analysis of results	153
5.1. Teachers' questionnaire	153
5.2. Focus groups	176
5.3. Students' questionnaire	189
5.4. Students' interviews	198
5.5. Analysis of assessment tools	203
CHAPTER 6: Discussion	219
6.1. Language-related issues in CLIL	219
6.2. Integration of content and language in CLIL assessment	222
6.3. The balance between content and language skills.	223
6.4. The aspects of language taken into account in CLIL assessment	225
6.5. Most common assessment tools and main reasons to use them	229

6.6. Recommended CLIL assessment tools	234
6.7. Assessment quality in the context of the study	237
CHAPTER 7: Conclusions	239
7.1. Main Conclusions	239
7.2. Recommendations for improving assessment in CLIL	241
7.3. Suggestions for future research	250
REFERENCES	253
APPENDICES	291
Appendix 1: Rubric Model in relation with CEFR levels	293
Appendix 2: Rubric used by different teachers	294
Appendix 3: Rubric for 4 C's	296
Appendix 4: Cover letter to introduce a portfolio	298
Appendix 5: Descriptors used in portfolio work for Geography	299
Appendix 6: Questions for portfolio reflection	300
Appendix 7: Teachers' questionnaire	301
Appendix 8: Focus group questions	305
Appendix 9: Students' questionnaire	306
Appendix 10: Students' interview questions	309
Appendix 11: Action plan for language mistakes	310
Appendix 12: Constructed response test-type	312

Appendix 13: Visual support in CLIL exams	313
Appendix 14: Scaffolding in CLIL History exam	315
Appendix 15: Diary Entry	316
Appendix 16: Newspaper article	317

## LIST OF GRAPHS

- Graph 1. Student support schema for BICS and CALP
- Graph 2. Halliday and Matthiessen's metafunctions of language
- Graph 3. The types of language in CLIL
- Graph 4. Bilingual Education in Madrid in the academic year 2016-2017
- Graph 5. Coordinating assessment and evaluation
- Graph 6. Bailey's washback model
- Graph 7. Criterion-based assessment in Spain
- Graph 8. Assessment in the LOMCE
- Graph 9. CLIL content-language continuum approach
- Graph 10. WALT and WILF posters
- Graph 11. Steps in the research
- Graph 12. CLIL teacher training
- Graph 13. Assessment format according to teachers' questionnaires
- Graph 14. Use of guidelines according to teachers' questionnaires



Graph 15. Criteria for oral presentations

Graph 16. Exam format according to teachers' questionnaires

Graph 17. Use of rubrics by teachers

Graph 18. Guidelines for language assessment in CLIL

Graph 19. Weight of English in CLIL assessment according to teachers

Graph 20. Treatment of language mistakes according to teachers

Graph 21. Aspects of language taken into account according to teachers

Graph 22. Compensation strategies

Graph 23. A comparison between teachers' and students' answers in questionnaires

Graph 24. Students' questionnaires: Treatment of language mistakes

Graph 25. Students' questionnaires: Weight of English in the final grade

Graph 26. Students' questionnaires: Aspects considered in language assessment

Graph 27. Students' questionnaires: Treatment of Language Mistakes

Graph 28. Students' questionnaires: Common Guidelines in CLIL Assessment

Graph 29. Analysis of assessment tools: Test type in relation to subjects

Graph 30. Analysis of assessment tools: Assessment criteria in bilingual schools

Graph 31. Analysis of assessment tools: Assessment in non-bilingual and bilingual groups

## LIST OF TABLES

Table 1. The four Cs in CLIL

Table 2. Content-obligatory and content-compatible language

Table 3. Cognitive discourse functions in CLIL

Table 4. Raising awareness on CLIL language

Table 5. CLIL Subjects in Bilingual Sections in ESO

Table 6. Types of assessment

Table 7: Traditional versus Formative Assessment

Table 8. Language competence descriptors in CLIL

Table 9. Ernst's error typology

Table 10. Polias' model for text analysis

Table 11. Analysis of assessment tools

Table 12. Types of mistakes in CLIL subjects

## LIST OF ABBREVIATIONS

AfL: Assessment for Learning

AICLE/AICOLE: Aprendizaje Integrado de Contenidos y Lenguas Extranjeras

BICS: Basic Interpersonal Communication Skills

CALP: Cognitive Academic Language Proficiency

CAM: Comunidad Autónoma de Madrid

CBI: Content-based Instruction

CDF: Cognitive Discourse Functions

CDI: Prueba de Conocimientos y Destrezas Indispensables

CEFR: Common European Framework of Reference for Languages

CF: Corrective Feedback

CIMNE: Cursos de Inglés para Maestros no Especialistas

CIPNE: Cursos de Inglés para Profesores de Secundaria no Especialistas

CLIL: Content and Language Integrated Learning

CLT: Communicative Language Teaching

CUP: Common Underlying Proficiency

EA: English Across the Curriculum

EC: European Commission

EFL: English as a Foreign Language

ESO: Compulsory Secondary Education

EvAU: Evaluación para el acceso a la universidad

FL: Foreign Language

FoF: Focus on Form

HOTS: Higher-Order Thinking Skills

ICT: Information and Communication Technologies

IGCSE: International General Certificate of Secondary Education

K-W-L-H: Know, Want to Know, Have Learnt, How I will Learn

L1: First Language

L2: Second Language

LA: Language Assistant

LAC: Language Across the Curriculum

LGE: Ley General de Educación

LOE: Ley orgánica de educación

LOGSE: Ley orgánica de ordenación general del Sistema educativo

LOMCE: Ley orgánica para la mejora de la calidad educativa

LOTS: Lower-Order Thinking Skills

MECD: Ministerio de Educación, Cultura y Deporte

MET: Marco de evaluación Trilingüe

MM: Mixed Method

OECD: Organization for Economic Cooperation and Development

PAU: Prueba de acceso a la universidad

PIRLS: Progress in International Reading Literacy Study

PISA: Program for International Student Assessment

SFL: Systemic Functional Linguistics

SI: Students' Interviews

SLA: Second Language Acquisition

SQ: Students' Questionnaires

TALIS: Teacher and Learning International Survey

TIMSS: Trends in International Mathematics and Science Study

TFG: Teachers' Focus Groups

TLA: Teacher Language Awareness

TQ: Teachers' Questionnaires

ZPD: Zone of Proximal Development



## INTRODUCTION

Bilingual education, also referred to as CLIL, has become a revolution over the last decades. Since the coining of the term in 1994, CLIL has evolved from a way to increase foreign language exposure into a pragmatic approach renovating classroom practice (Ting, 2011). Today, CLIL is present in nearly all European countries (Eurydice, 2017), and it is likely to shape language teaching in the near future as it promotes multilingualism through a new conception of language teaching and learning offering significant language exposure through real-life curricular contexts.

In Spain, CLIL programs have also become widespread since the implementation of the MEC-British Council Bilingual and Bicultural Project, which paved the way for other models such as the Plan de Fomento del Plurilingüismo in Andalusia and the CAM Bilingual Project in Madrid, to name just a few. The CAM Bilingual Project is a state-funded program which started in 2004 in primary schools, and was made extensive to the secondary level in 2010. As in other bilingual programs across Europe, the CLIL approach was adopted to teach non-linguistic subjects, except for Mathematics and Spanish Language, using English as a vehicular language.

However, a conceptual framework for content and language integration needs to go hand in hand with the adoption of new educational approaches and methodologies. Despite its rapid growth, and the significant involvement of educational authorities, teachers and families, bilingual programs in Europe are still object to improvement concerning aspects such as teacher training, methodologies, the use of appropriate materials, and the way assessment is conducted. When it comes to assessing students' learning, which is one of the most controversial issues in CLIL, the most common debate arises in the attempt to identify the nature of CLIL assessment (Coyle et al., 2010; Kiely 2009; Järvinen 2006), and the way teachers deal with the integration of content and language. Other aspects are

related to the methods and tools which are best suited to assessment in CLIL, the best way to measure previous knowledge and/or progression, skills and processes, cognition and culture (Coyle et al., 2010), and the role of language in assessment (Llinares, Morton & Whittaker, 2012) among others. However, due to the variety of CLIL models, the relative novelty of this integrated educational approach, and the lack of established assessment criteria, the small number of studies completed on CLIL assessment (Serra, 2007; Serragiotto, 2007; Hönig, 2010; Wewer, 2014 and Reierstam, 2015) show evidence of significant disparity among the assessment practices conducted in CLIL programs. With the analysis of assessment practices in the Bilingual Sections of the CAM Bilingual Project, this study aims to address this gap in the CLIL literature, and thus, to analyze the impact that assessment has on teaching and learning.

### **Objectives of the study**

This study has several objectives:

- To know to what extent secondary teachers are aware of CLIL assessment procedures;
- To know to what extent content teachers in Bilingual Secondary Schools in Madrid are aware of the need to take the dual focus of CLIL into account when assessing;
- To find out about teachers' perception on the balance of content and language in assessment;



- To know about teachers' preferences regarding assessment tools and tasks, and whether, apart from standardized tests, they use other alternative tools (portfolio, checklists, essay writings, oral reports and interviews among others);
- To know about students' perception of what teachers expect from them in assessment concerning content and language knowledge, and to what extent they consider they are being assessed on English proficiency;
- To assess the balance between content and language in tests and other tools.

Following the objectives mentioned above, the study has a twofold dimension. As assessment in CLIL is a controversial topic and an ambiguous issue that commonly frustrates teachers, the project aims, on the one hand, to identify the challenges teachers are faced with in relation to assessment in CLIL contexts. On the other hand, the project has also been conceived to investigate teachers' current practice in CLIL assessment in Bilingual Secondary Schools in Madrid, i.e., the assessment tools used and, more specifically, whether and to what extent linguistic achievement in the foreign language influences the grade students obtain. The identification of teachers' difficulties and practices in relation with assessment does not only prove beneficial within the CLIL dimension, but it might also serve well in larger frameworks of education such as Spanish secondary education, where content teachers often lack training in CLIL (Halbach, 2010; Salaberri Ramiro, 2010; Fernández & Halbach, 2011; Olivares Leyva & Pena Díaz, 2013; Pavón & Ellison, 2013), and Formative Assessment has been recently implemented. Following this twofold dimension, the study has been organized into two different main blocks concerning the background and the context:

First, there is a need to take a deep look at the literature available to identify the main problems in CLIL assessment, the balance between language and content and the convenience of assessing language in content areas.

Second, to offer suggestions which could serve CLIL in general, and the context of the study in particular, it is also necessary to know first-hand about current assessment practice in the Bilingual Sections. Thus, the empirical part of the study is aimed at investigating how assessment in CLIL is carried out in Bilingual Secondary Schools in Madrid, placing particular emphasis on the balance between content and language and the tools being used by teachers for assessing students' performance. In this sense, and taking into account Spanish regulations regarding assessment, both Formative Assessment and Summative assessment will be considered. It is hoped that looking at existing practices, we will be able to suggest possible solutions for assessing content subjects in different contexts. Furthermore, since one of the challenges for assessment in CLIL is to obtain descriptors for CLIL competences and different levels, the last part of the research is specifically aimed at comparing the results with those obtained in other projects, developing specific competence descriptors for CLIL contexts, and suggesting practices which could prove effective for teachers.

### **Organization of the study**

Chapter one of the current study provides a detailed look at Content and Language Integrated Learning (CLIL). First, theoretical rationale for CLIL is presented by focusing on the 4Cs Framework, Communicative Language Teaching (CLT), and Cummins' notion of Common Underlying Proficiency. Then, bilingual education in Europe is

analyzed, placing special emphasis on the different CLIL models and programs in Spain. Finally, the role of language in CLIL is discussed to further contextualize the study.

Chapter two narrows the scope of CLIL in Spain to the focus of the study, the CAM Bilingual Project. It draws attention to the origins and structure of the program, teachers and their training, most common challenges as well as its external evaluation.

Chapter three explores the concept of assessment as opposed to evaluation and analyzes different types of assessment, the role of assessment in education, and assessment regulations in Spain. Afterwards, CLIL assessment is introduced by focusing on several aspects, namely possible ways of assessing, Formative Assessment, quality standards, best CLIL assessment tools and exam questions and task design.

Chapter four provides a description of the research carried out. It looks at the research questions, the methodology, the design of the study, and the procedures used to analyze the data as well as the limitations of the study.

In chapter five, the results of the study are explained in depth, providing insight into current assessment practice in Bilingual Sections in Madrid. Along with the data from questionnaires, interviews, focus groups and the analysis of assessment tools, direct quotes from both teachers and students have been included to illustrate their views and expectations. This analysis of results leads onto chapter six, where the questions guiding this research are answered.

Finally, chapter seven presents the conclusions obtained from the study, some recommendations for both content and language teachers, and suggestions for further research.



## SECTION 1: BACKGROUND



## **CHAPTER 1: Content and Language Integrated Learning**

In the formation of the *Europe of knowledge* and of European identity, languages have been of paramount importance as one of the *sine qua non* conditions for mobility and understanding among countries. In the 1990s, as the European Union was gaining strength, many European governments were questioning the effectiveness of language teaching and learning, and recognizing the value of languages for intercultural communication. However, it was not until 2003 that the Council of Europe designed several education policies aimed at promoting language learning and respect for languages by means of reinforcing Plurilingualism, linguistic diversity, mutual understanding, democratic citizenship and social cohesion. According to these policies, plurilingualism is the potential and actual ability to use several languages to varying levels of proficiency and for different purposes to interact effectively with people from other languages and cultures (Council of Europe, 2007). As a concept opposed to monolingualism, plurilingualism refers to the different repertoires of varieties which individuals may use apart from their mother tongue, be it a co-official, minority, second or foreign language. Besides, in contrast with multilingualism, which deals with the presence of more than one language in a geographical area, plurilingualism places a clear focus on individuals as social users proficient in several languages, and experiencing several cultures “for the purposes of communication and to take part in cultural interaction” (Council of Europe, 2001: 168), and not solely on the languages. As stated by the Council of Europe (2003), all European citizens should have the skills to communicate with their neighbors that is, they should be able to communicate in at least two languages apart from their mother tongue. From this humanistic viewpoint, and to ensure the linguistic objectives aimed to foster European cohesion were achieved, educational authorities supported bilingual education under the name of CLIL as a natural

way to improve language instruction, and for its value and contribution to the development of the child (Baker, 2002).

### 1.1. What is CLIL?

CLIL is the acronym for ‘Content and Language Integrated Learning’. This acronym is used as a general or “umbrella term” (Mehisto, Frigols & Marsh, 2008) which designates different types of bilingual or immersion education. CLIL refers to the kind of education where a co-official, regional, minority or foreign language is used to teach non-language subjects, i.e. subjects other than languages and their literatures or cultures. Using this innovative educational approach, students learn school subjects while at the same time they have the opportunity to practice and improve language skills. This combination of subjects and language learning in the curriculum offers students better preparation for life and European mobility (Eurydice, 2012), and a higher level of relevance and authenticity in language learning, (Marsh & Frigols, 2012) which was absent in traditional methodologies and approaches. Besides, opportunities to interact with different language communities and cultures, and additional exposure to the foreign, regional or minority language are also significant in this methodological approach.

CLIL can be broken down into content i.e. the academic discipline in question, communication as the language through which students acquire content knowledge, cognition and culture, as expressed in the chart below:



Table 1. The four Cs in CLIL

Content	Integrating content from across the curriculum through high-quality language interaction
Cognition	Engaging learners through creativity, higher order thinking and knowledge processing
Communication	Using language to learn and mediate ideas, thoughts and values
Culture	Interpreting and understanding the significance of content and language and their contribution to identity and citizenship

(Adapted from Coyle, 2005: 5)

However, the most distinguishing feature of CLIL is probably the development of cognitive skills originally laid down in Bloom’s taxonomy of educational objectives (1984), and later adapted by Anderson & Krathwohl (2001). Cognitive skills enable the learner to engage in meaningful activities requiring high-order thinking skills through the focus on enquiry, information processing and problem-solving (Coyle, Holmes & King, 2009: 7). Recently, a fifth C, standing for competence, has been suggested by Ball, Kelly & Clegg (2015). This competence refers to ‘procedural knowledge’ or the “skills used to work on the concepts” (2015: 52) i.e. the ability to deal with challenging cognitive tasks as students progress on specific thinking skills.

### 1.1.1. Theoretical Rationale of CLIL

Apart from Bloom's taxonomy of cognitive skills we referred to in the previous section, CLIL benefits from other sources which can help explain its theoretical rationale. To start with, the Communicative Approach or Communicative Language Teaching (CLT) raised awareness of language learners' needs, and the functional and communicative potential of language (Wilkins, 1972). Later, Krashen's (1981) distinction between learning and acquisition highlighted that second language acquisition results from a subconscious process, which is similar to that of first language acquisition requiring meaningful interaction between speakers in the target language. This theory of second language acquisition constituted a novelty back at that time since it left behind traditional language teaching and learning based on memorization of grammatical rules and drills later used for language production. Instead of focusing on form (FoF) or language structures, natural conversation among speakers concentrates on the meaning, and it provides relevant input for the learner, a typical feature in CLIL lessons. Currently, both Krashen's acquisition principle and CLT have been reinforced and updated by the consideration of language use as action oriented, and the learner as a social agent and user of the language(s) advocated by the *Common European Framework for Languages (CEFR)* (2001), and the Council of Europe:

Language use, embracing language learning, comprises actions performed by persons who, as individuals and as social agents develop a range of competences, both general and in particular communicative language competences. They draw upon the competences at their disposal in various contexts under various conditions and under various constraints to engage in language activities involving language processes to produce and/or receive texts in relation to themes in specific domains activating those strategies, which seem most appropriate for carrying out their tasks to be accomplished. The monitoring of these actions by the participants leads to the reinforcement or modification of their competences (Council of Europe, 2003: 9).

In fact, social activities such as education are considered to shape language use and the way that language itself constructs knowledge. CLIL is thus, based on the following integrating theoretical approaches: First, Systemic Functional Linguistics (SFL) or the social-semiotic theory of language as a meaning-making activity. SFL considers that all choices speakers make from both the lexical and grammatical systems of a language are shaped by the social activities in which they are involved, and consequently, they should be considered as appropriate or inappropriate in relation with language use rather than language structures (Halliday, 1994; Halliday & Matthiesen, 2004). Following this systemic functional approach to language (1994) in CLIL is thus, translated as the attempt to build on language learning starting from the specific domains of the academic subjects as a way to develop both content knowledge and language competence. Second, the sociocultural theory of learning developed by Vygotsky (1978, 1986) states that language use and interaction with other speakers constitute the essential mediating tool in cognitive development. Third, recent work in second language acquisition (SLA) abandoned individual cognitive processes in favor of broader perspectives considering second language learning as inextricably linked with social activities between individuals (Block, 2003; Firth & Wagner, 1997; Hellerman, 2008; Kasper, 2009; Pekarek Doehler, 2010).

Another principle on which CLIL is based on is Cummins' Common Underlying Proficiency (CUP) (Cummins, 1980, 1981 as cited in Baker, 2011). CUP refers to the cognitive proficiency underlying academic performance in the student's L1 and L2. When learning a language, the child acquires a set of skills and implicit metalinguistic knowledge that can be later transferred to the learning of another language. Thus, this underlying proficiency does not hinder first language proficiency, but in fact, it helps in second/foreign language acquisition and learning as it allows development for both languages.

Finally, the concept of scaffolding, and the Zone of Proximal Development (ZPD) as the distance between what a student can achieve in his/her leaning process and the next step s/he can reach with the teachers' or peers' assistance have also been relevant in CLIL provision. The concept of scaffolding was initially developed by Brunner (1976), and later revisited by Wood et al. (1976) as a sociocultural theory portraying the interactive social nature of learning. According to the latter, scaffolding consists of breaking up the lesson or task into manageable chunks so that students are directed to move onto learning with the help of the teacher, this help being removed when it is no longer necessary. This type of guidance is required in education in general but even more urgent in bilingual education due to the cognitive challenge that CLIL activities can pose for learners. First, for deeper understanding and learning, CLIL tasks require from students more complex tasks than just retrieving and summarizing information. Second, the students' limited language proficiency if compared to native speakers' also requires from students to deal with more tasks to achieve similar competence. Consequently, to achieve those higher degrees of cognition such as analyzing, comparing or reflecting critically upon what they have learned, students will need some help to move from the easiest stages or prior knowledge to the highest degrees of competence in content learning. For scaffolding to be effective, it should be just beyond the level the student can manage on his/her own (Olson & Platt, 2000) and assistance "should always be just enough and just in time" (Walqui, 2006: 165).

## 1.2. CLIL in Europe

Schools offering CLIL provision or the teaching of particular subjects through the medium of a foreign, regional or minority language have a long tradition in Europe. In fact, although the Canadian, American and Australian immersion models have influenced

in the upsurge and development of European CLIL, the attempt to offer integrated instruction of content and language in order to facilitate language acquisition, and promote better language proficiency is not new in Europe as it can be traced back to 1990 (Marsh, 2000; Coyle, 2005).

European CLIL can be thus, considered as the result of the combination of the influence of several factors. First, the Canadian immersion model which paved the way for the introduction of significant language exposure through content subjects. Second, due to the lack of success of previous methodologies of language learning, an attempt to change language provision emerged to provide genuine communication opportunities for students to make learning meaningful. In fact, CLIL was regarded a new force in language teaching to overcome some of the most common weaknesses in traditional methodologies, namely the lack of real and relevant input outside the language class, and motivating the process of learning since the content lesson could successfully integrate the focus on form and the focus on meaning (Muñoz, 2007). Third, European legislation favored the upsurge of multilingualism and plurilingualism as key issues in European cultural identity (Council of Europe, 2007). Furthermore, the information age has led to changes in the way we conceive education, and as a result, innovative working models moving towards integration across disciplines and real-life skills have consequently affected the position and significance of teaching foreign languages (Marsh & Frigols, 2012). Accordingly, language teaching has also learned from the rise of different approaches, and English as a lingua franca is no longer considered as central in the mainstream curriculum but a core skill that facilitates content knowledge (Graddol, 2006).

Before the 1970s, this type of educational provision consisting of the teaching of one or more additional languages was limited to linguistically distinctive regions, which were

close to national borders or used two languages, especially in capital cities. That was, for instance, the case of the emergent European Schools. The first European School was established in Luxembourg (1953), and they rapidly spread across Europe. They were not conceived as immersion or subversion models nor did they have a strictly dual immersion approach (Lorenzo, Trujillo & Vez, 2011). On the contrary, European Schools were funded with the purpose of providing high quality multicultural and multilingual education for the children of staff in European Institutions and thus, promoting the development of plurilingual students so that they could be ready to live in multilingual and culturally diverse societies. Although no explicit attention was first devoted to the integration of content and language, students were encouraged to join in different activities with other international students during weekly 'European Hours' promoting extended foreign language exposure in a multicultural context.

Another precedent of CLIL is Language Across the Curriculum (the LAC movement) which started in London in 1966 as an attempt to promote discussions in the English lessons. In 1975, the Bullock Report raised awareness on the central role that language plays in all learning, and it later spread through England, Australia and Canada (Pokrivckáková et al., 2015: 3).

During the 1970s and 1980s, CLIL development continued to be greatly influenced by the Canadian immersion model. In 1978 the European Commission (EC) suggested the teaching of school subjects through other vehicular languages, and in 1983 the European Parliament challenged the EC to forward a new program improving the quality of foreign language teaching (Marsh, 2002: 52). But it was not until the 1990s that the acronym CLIL (Content and Language Integrated Learning) started to be widely used. First, it was believed that a better knowledge of modern European languages is vital to facilitate communication and integration among European citizens, and it will thus promote

mobility and mutual understanding. Consequently, and due to the need to convert the rich heritage of languages and cultures in Europe from a barrier to communication into a source of mutual understanding (Council of Europe, 1982), several documents were later decisive in the creation of the future CLIL. The first of these documents is the 1995 Resolution of the Council. It refers to the promotion of innovative methodology for language learning and, in particular, to the teaching of foreign language lessons for disciplines other than languages, to provide bilingual teaching. Likewise, it suggests the need to improve the quality of language teacher training. At the same year, the White Paper on education and training (*Teaching and learning; towards the learning society*) stated that everyone, irrespective of his/her academic training, should gain proficiency in two languages apart from their mother tongue so that they were able to communicate in those languages.

The *Common European Framework of References for Languages* (2001) was also relevant as it provided a practical tool for establishing standards and evaluating language knowledge within the European Union at an international level. These European actions along with the European education and training programs had a catalytic effect and promoted mobility actions under Socrates and Comenius programs from 2000-2006 aimed at staff wishing to teach in a foreign language. Besides, the European Year of Languages in 2001 helped to draw attention to language learning and to promote linguistic diversity. The Barcelona European Council in 2002 also addressed the need to ensure teaching of at least two foreign languages from an early age, and specific reference was made to CLIL as a pedagogical tool effectively contributing to the European Union's language learning goals. Finally, the 2003 Action Plan for the Promotion of Language Learning and Linguistic Diversity 2004-2006 by the European Commission highlighted the importance of CLIL provision in mainstream education as necessary for innovation

in language teaching- from theoretical knowledge to practical skills and learning how to learn. This Action Plan highlighted the importance of competences i.e., it is not only what we know but also what we can do with that knowledge (Lorenzo, Trujillo & Vez, 2011). It also recommended CLIL to improve the quality of language teaching, and to introduce different perspectives enhancing European citizenship.

During the May 2005 Education Council, one of the main conclusions addressed the importance of CLIL type provision at different levels of school education, and recommended to encourage teachers to receive specific CLIL training. CLIL benefits for better language skills were later confirmed and reinforced by the Eurydice study (2006) about CLIL in European countries and researchers from the ECML (European Centre for Modern Languages) in Graz such as Maljers, Marsh & Wolff (2007), Marsh and Wolff (2007), Mehisto, Marsh & Frigols (2008) who add data from CLIL implementation and development in more than 20 European countries. Today CLIL is present in all European countries in primary and secondary levels, Denmark, Ireland, Greece and Turkey being the last ones to welcome it into their classrooms (Key data 2012). Additionally, since it has been recently implemented in tertiary education, related research on this subject matter has increased (Järvinen, 2007; Fortanet Gómez, 2013; Smit & Dafouz, 2012; Doiz, Lasagabaster & Sierra, 2013; Cenoz & Etxague, 2013).

#### 1.2.1. Types and models in CLIL

The fact that CLIL is often referred to as an ‘umbrella term’ (Mehisto, Frigols & Marsh, 2008) describes a plethora of situations and contexts which have paved the way to classifications of different CLIL programs. Regardless of the different CLIL varieties and



models, and depending on the languages used as a vehicle of instruction to teach content or non-language subjects, Coyle, Hood & Marsh (2010) distinguish two types of CLIL:

Type A CLIL: Non-language subjects are taught through the medium of a foreign language. The number of the non-language subjects offered through the foreign language may vary depending on the school, region and/or country. There are schools in which all non-language subjects are offered through the foreign language. In others, two languages are used to teach non-language subjects in the curriculum: some non-language subjects are offered through the medium of the foreign language while some others are taught through the language of the governing body or administrative body of the school.

Type B CLIL: Non-language subjects can be taught using a regional or minority language, a non-territorial language, a state language where in a country we can find more than one, and a second language, which may be any language other than the mother tongue. Nevertheless, it is interesting to notice that in some schools, in addition to the two languages which are commonly present in school life, a third language can also be used as is the case of multilingual communities in Catalonia, Galicia or The Basque Country in Spain.

Variations in the number of hours offered through the vehicle of the target language have led to coining the terms of hard CLIL, modular CLIL and soft CLIL (Ball, 2009; Bentley, 2010). Hard CLIL is also referred to as partial immersion and is subject-led since about half of the curriculum is taught through the medium of the target language. Modular CLIL is also subject-led, but a smaller proportion of a curricular area is delivered through the target language. In this model, it is the schools and teachers' responsibility to decide on the parts of the curriculum, which are taught using the CLIL approach. Soft CLIL is, on the contrary, language-led and, some curricular topics are selected to be included in the language classroom. Dale & Tanner (2012: 4-5) add another category to CLIL models

regarding the types of teachers involved in the program - language teachers and subject teachers- and the extent to which they have been trained to teach CLIL subjects. Thus, they differentiate between subject lessons taught by CLIL subject teachers and language lessons taught by CLIL language teachers. Finally, we can distinguish between early or late immersion in CLIL depending on the age of the students when they access the program (Eurydice, 2006).

### 1.2.2. Common features with immersion models

Considered as an evolving methodological approach (Bentley, 2010) which has resulted from the lack of success in previous language teaching methodologies throughout history, CLIL is not a new phenomenon. In fact, other names were coined in the 1980s and 1990s to refer to the same reality or similar approaches to language teaching (Byrnes, 1998; Brinton, Snow & Wesche 2003; Grabbe & Stoller 1997; Richards & Rodgers, 2014). Among them, we find Content-based Instruction (CBI), Content-based Language Teaching (CBLT), English Across the Curriculum (EAC) and Languages Across the Curriculum (LAC), to name just a few. Other terms commonly used in the European context to refer to CLIL are: Bilingual Studies, Bilingual Education, Subject Teaching in the Target Language and Immersion Education or teaching. CLIL is often compared with Immersion programs mainly because they share a common objective, i.e. to ensure that students acquire curricular subject matter knowledge at the same time they develop competence in a second/foreign language. For that purpose, the L2 - considered as the language *through, for* and *of* learning (Coyle, Hood & Mehisto, 2010)- is used as a vehicle or medium of instruction. This language is taught through content subject using the communicative approach to create effective communication in a learning environment, which is motivating for students (Lasagabaster & Sierra, 2010: 370). The L1 receives

support as a language itself, and an essential component of the curriculum, sometimes being used as a subject and sometimes as a medium of instruction in other subjects. As in immersion programs, the starting point in European CLIL is additive bilingualism as opposed to subtractive bilingualism. In additive bilingualism, the students' native language is developed along with its culture at the same time that a second or foreign language is added, but not at the expense of the mother tongue (Cummins, 1991). Thus, at the end of the program students' L1 proficiency ideally could be paralleled with those who have studied in non-bilingual programs (as was demonstrated in the case of the Trilingual Education Framework<sup>1</sup>/MET program in the Basque Country) so that L2 proficiency is not obtained at the expense of the L1, a common fear and complaint on the part of parents. As regards the curriculum in CLIL programs, it is usually the same as in non-bilingual schools. It thus guarantees that bilingual students study the same content and develop the same skills that in non-bilingual students so that the language of instruction is the only variable that changes. However, in some models such as the Bilingual and Bicultural Project by the British Council-MEC agreement in Spain, we can find a joint curriculum, which enriches instruction.

### 1.2.3. Key Features of European CLIL

Although the terms CLIL and Immersion Education are sometimes used interchangeably, European CLIL has its own specific features, which must be taken into account. In the following section, the most noticeable ones are described as related to the language of instruction, the teachers involved in the process, students' starting age, learning materials, main language goals and the profile of students' entering each type of bilingual education

---

<sup>1</sup> Author's translation.

(Lasagabaster and Sierra, 2010). Other aspects analyzed in this section deal with official teaching time devoted to CLIL, teachers' profiles, the choice of CLIL subjects, and the main barriers the CLIL methodological approach encounters.

To start with, the language used as a medium of instruction is not spoken in the community, but a foreign language is used in many CLIL contexts. Therefore, despite of the current availability of Information and Communication technologies (ICT), second language (L2) exposure is limited to the classroom context and not present in society. In fact, in countries like Spain, the foreign language has little presence in society since practically all films are dubbed, there is a strong reliance on traditional methodologies, and thus, students are not used to practicing the language in real contexts.

As for the teachers involved in the process, unlike teachers in immersion programs, CLIL teachers are not usually native speakers of the L2. In general, teachers are required to have an advanced knowledge of the target language, as they are responsible for both content and language-related aspects of teaching. As stated in the 2017 Eurydice report, ideally they should prove a very good knowledge of both the content subject and the language in which it is taught, and be familiar with the requirements of CLIL (Eurydice, 2017: 91). Thus, in relation with teachers' qualifications, the extent to which teachers are bilingual in the students' L1 and L2 or language of instruction greatly differs depending on the country and region. The same accounts for recruitment criteria; in general, the basic qualification required relates to non-language subjects but certified language skills ranging from B2 (for instance in the Andalusian region) to C1 levels are also a prerequisite. However, it is important to highlight that specific CLIL training or the kind of teaching and/or methodological skills peculiar to CLIL are not always a prerequisite to access the program in countries such as Lithuania and some Autonomous Communities in Spain (Eurydice, 2017: 92).

Although some countries and regions are planning to implement bilingual education at an early age (Consejería de educación Comunidad de Madrid, 2016) some others still resemble the immersion programs in Canada as bilingual education do not usually start until the primary or secondary level.

Maybe one of the most distinguishing features of CLIL as compared to immersion programs are language goals. This aspect is, in fact, of paramount importance in order to clarify and establish reasonable language objectives, which cannot be paralleled to native-like proficiency (Lasagabaster & Sierra, 2010: 372) and thus, avoid misconceptions about bilingual education amongst parents and educators. In this sense, the learning materials used in CLIL contexts are usually adapted for non-native speakers especially at the initial stages of the program.

As regards to the students' profile, not all the students join the program with similar levels of L2 proficiency especially in primary education. The levelling of students regarding their language proficiency would facilitate the adoption of a curriculum and pedagogy suited to students' needs (Baker & Jones, 1998), as in the case of Latvia, which establishes conditions governing access to CLIL instruction. However, conditions to access bilingual education greatly differ among countries. In this respect, critical voices have raised concerns about whether the good results of bilingual programs depend on the selection of students, and the appropriateness of bilingual education programs in certain educational stages such as the secondary level (Bruton, 2011; García López & Bruton, 2013).

As for the official teaching time devoted to CLIL subjects or the minimum amount of time recommended for teaching in the target language, it varies depending on the country and in some cases, differences are also made visible from one region to another and also on a school basis. In the Czech Republic, for instance, CLIL instruction depends on the

institution and the subjects concerned while in Spain, variations relate to the particular Autonomous Community and the school itself.

In relation with the subjects taught through a vehicular language, and partly because, in some contexts, it is possible for schools to choose the subjects taught via CLIL from the entire curriculum, there is no clear preference for any particular subject in primary CLIL. However, in secondary education, there is a tendency to select science subjects or those related to Social Sciences as happens in 12 countries (Eurydice, 2006). As for the presence of the L2 culture, although it is treated as fundamental in CLIL instruction (Coyle, Mehisto & Marsh, 2010), in contrast to immersion models, this does not necessarily mean that culture is integrated into the mainstream curriculum of all the CLIL models.

Regarding the main barriers CLIL encounters, and which prevent it from becoming more widespread, there are the shortage of specialized teachers, especially at the initial stage of the program (Eurydice, 2012), finding appropriate teaching materials adapted to the specific CLIL contexts in question (Ruiz de Zarobe & Lasagabaster, 2010; Halbach, 2012), legislative restrictions, and the high cost of program implementation.

Finally, with regard to the main criticisms, the following are commonly highlighted: First, the lack of collaboration between content teachers and language teachers (Cross & Gearon, 2013, Pavón & Ellison, 2013). Second, the extent to which CLIL provision is implemented in some countries and regions, which may lead to CLIL programs being limited to the elite (Bruton, 2011) as was pointed out before. Third, the conflict between content and language-related aspects, and the most appropriate way to deal with language aspects, which commonly frustrates teachers (Llinares, Morton & Whittaker, 2012).

### 1.3. CLIL in Spain

This section aims to give a general view of the variety of CLIL scenarios in Spain. But before we deal with that, it is interesting to explore the recent history of foreign language provision in Spain to understand CLIL implementation in depth. In Spain, students traditionally started learning a foreign language as a compulsory subject at the age of six. In fact, it was not until the arrival of the law currently in force, the LOMCE (2013), that students started learning a foreign language at an early age - three years of age - in infant education<sup>2</sup>. This foreign language -usually English- is a compulsory subject from primary to Compulsory Secondary Education, i.e. from six to 12 years of age, and it continues as an obligatory subject in the curriculum for students in upper secondary education until they reach 18. In compulsory education, students may also choose a second foreign language as a core curriculum option, but unlike other European countries, learning a second foreign language is an entitlement rather than an obligation (Eurydice, 2017).

About the certificate awarded on completion of compulsory education, unlike those in other countries, the one issued in Spain does not refer explicitly to foreign language learning. Likewise, those students taking part in bilingual education programs do not receive a special certificate indicating they studied certain subjects through the vehicle of the foreign language, but the same general certificate all students receive at the end of upper secondary education or “Bachillerato”, which corresponds to a B1 level according to the Common European Framework for Languages (CFR).

---

<sup>2</sup> It is necessary to take into account the fact that although compulsory education starts at six, most children in Spain attend school from 3 years of age. Thus, before the LOMCE, pupils already had contact with the foreign language in the second cycle of Infant Education, which favored the learning of the foreign language at an increasingly early age.

### 1.3.1. Learning foreign languages in Spain. Why don't our students speak English?

Despite the fact that Spanish students start learning a foreign language at a relatively early age, it seems that foreign language proficiency remains as the “unfinished business” or an aspect of education that still has to be addressed in our society. The main reasons that could explain the failure of language teaching in Spain are the following. To start with, and as was pointed out in the introduction and the section devoted to the differences between CLIL implementation in Spain and immersion programs, although since the arrival of the LOGSE in 1990 more attention was given to foreign languages, their provision has not traditionally been implemented at an early age. Besides, except for bilingual programs, the context in which foreign languages are learned relates more to foreign language study rather than acquisition i.e. learning takes place exclusively in the classroom with only a few hours devoted to the foreign language in the curriculum. In this context, there is no real linguistic immersion, and the foreign language is not used in real-life communicative situations in which the students have a genuine need for using the foreign language so that learning can be meaningful (Ausubel, 1968; Novak, 1998). Also, for decades the teachers' levels of foreign language proficiency have been quite low mainly due to the lack of constant teacher training by educational authorities and the lack of promotion of oral skills. In fact, foreign language instruction in Spain has been traditionally based on grammar, vocabulary and some writing skills (Romero Lacal, 2011: 5, Rubio & Tamayo, 2012), usually using Spanish as the vehicle of communication. This lack of emphasis on communication provoked an approach focused on teaching facts about the foreign language- learning about the language to use it rather than using the foreign language to learn it, which does not facilitate language acquisition. As Marsh (2009: 2-3) points out, “successful language learning can be achieved when people have



the opportunity to receive instruction, and at the same time experience real-life situations in which they can acquire the language”.

A clear example of the focus on writing and reading skills over oral skills in Spain is the fact that to-date, the University Entry Examination does not include a listening comprehension activity or oral tasks/questions. The English exam only focuses on grammar and vocabulary activities along with reading and writing tasks about topics previously dealt with in the official curriculum. This lack of oral skills is, on the one hand, opposed to European recommendations about the four skills having equal standing at the end of compulsory education and on the other hand, it might indicate a weakness in the exam design. As the Spanish curriculum is clearly product-oriented due to the prevalence of exams in the educational culture, syllabi tend to be exam driven. Although the absence of oral tasks has raised complaints amongst educators, there are still no evidences that the situation will change in the near future. Finally, Romero Lacal (2011) also points to the Spanish widespread inferiority complex when pronouncing foreign words appropriately, and the lack of original version broadcasting in the Spanish media as part of the problem. Due to the lack of presence of the foreign language in society, students do not have numerous and varied opportunities to use the language in different contexts and with different aims (Swain, 2006). Today, the current law of education gives priority to oral language -listening and speaking- and bilingual schools are encouraged to present novelties regarding quality-oriented actions. It is hoped that these measures are actually implemented and a real emphasis on oral skills is seen in the near future.

### 1.3.2. CLIL in Spain. A variety of scenarios

In this context of unsuccessful traditional language pedagogies throughout history, CLIL emerged as a promising pedagogy in the Spanish territory. Spanish CLIL- also called AICLE (Navés Muñoz, 1999) or AICOLE in the Madrid region- has been commonly described as a microcosm inside the macrocosm of CLIL (Frigols in Megías Rosa, 2012) in that it reflects different European CLIL realities, and where each Autonomous Community has developed its own model. Spanish CLIL resembles other European models in several ways. First of all, in the use of a foreign language as the vehicle of instruction to teach content. Second, regarding the teachers' profiles and training, (Frigols, in Megías Rosa, 2012) and the lack of specialized materials especially in the early stages of the program implementation (Halbach, 2012; Ruiz de Zarobe & Lasagabaster, 2010). Finally, it is similar to other programs in that more sound research is still needed especially concerning assessment and the role of language in CLIL. Maybe one of the most noticeable features of Spanish CLIL is that some regions offer CLIL provision in which three languages are used to teach the curriculum, including the state language and one regional language (Eurydice, 2017: 56). However, the foreign language input is still considered very limited as compared to other European models. A clear example is Madrid Autonomous Community; even though this region is the one with the largest number of hours taught through the medium of English, it does not reach 50% of the total number of hours in school. Finally, as related to curriculum planning, although CLIL Spanish programs start in primary education, CLIL does not always find a continuation in higher stages of education- it is sometimes continued in lower secondary education, but it is not the norm in upper secondary education.

Currently, nearly 40% of students in mainstream education participate in CLIL programs in Spain (Ministerio de Educación, 2016), and this figure is hoped to be increased in the

near future. However, although CLIL has been successfully established as an innovative form of education throughout Europe, no standard criteria have been agreed upon nor do legal regulations exist even within the same country or region. CLIL in Spain is represented by a diversity of scenarios, and as such, it varies greatly concerning the language or languages being used, the subjects through which students learn the language, and the time devoted to language exposure among other factors. The following are the main three scenarios for CLIL in Spain (Frigols, 2008). First, we find those programs promoting bilingualism in a monolingual community such as the CAM Bilingual Project, which constitutes the focus of this research. Second, those fostering multilingualism in an already bilingual community- Catalanian and Basque programs. Finally, the ones improving competence in English through a bilingual and bicultural project as in the case of the Bilingual and Bicultural Project by the MEC-British Council Project described below, and which served as a precedent for the CAM Bilingual Project.

### 1.3.3. The MEC-British Council Bilingual and Bicultural Project

The Bilingual and Bicultural Project also known as the MEC-British Council Project or the Bilingual Education Project was designed as a cooperation agreement between the Ministry of Education, Culture and Sport (MECD) and The British Council. Signed in 1996 and renewed in 2013, it extends to ten Autonomous Communities in Spain. The main goal of this project is to offer bilingual education through an integrated Spanish-British curriculum comprising both linguistic and cultural aspects, a condition which makes it unique if compared with the rest of bilingual projects in Spain. Apart from this joint curriculum, there are other noticeable features. Unlike other programs starting in primary education, this program is implemented at a very early age (the second cycle of infant education) and extends up to the end of Secondary Compulsory Education (ESO).

One of the main benefits that this early exposure to foreign language promotes is the opportunity to focus on literacy in English from a very early stage. This emphasis on literacy is considered to help children with the acquisition and recognition of sounds, and thus to promote reading skills (Halbach, 2012; Johnston & Watson, 2005; Coyle et al., 2014).

Although the amount of the curriculum delivered through English as a vehicular language was not specified in the agreement, it was made effective to cover a significant exposure roughly equivalent to 40% of each week at school, allowing pupils to learn some content subjects through English such as science, history and geography. Other distinguishing features apart from the awareness of the diversity of the Spanish and the British cultures integrated into the curriculum relate to the use of modern technologies in learning the foreign language, teacher training and access to materials, and the international certificate students obtain (Frigols, 2008).

As regards teachers' training, special emphasis is laid on both language skills and methodologies. From the beginning of the project, schools benefitted from the help of British Council specialists and teacher trainers acting as linguistic and educational consultants to ensure the correct implementation of the integrated curriculum. Besides, teachers are also encouraged to design and use their own materials when it is needed, and to implement ICT to suit students' needs.

Finally, the joint curriculum provides students who continue in the bilingual project during secondary education the opportunity to obtain the double certification, i.e. the Spanish secondary certificate and the one by the British educational system with the International General Certificate of Secondary Education (IGCSE) exam. This allows the students participating in the project to access upper-secondary studies in any of the two educational systems- the Spanish and the British.

The British Council-MEC Project was implemented in infant education in 1996. It was made extensive to some high schools implementing Linguistic Sections where a minimum of three subjects are taught through the medium of the vehicular language, and the amount of exposure to English as a Foreign Language was incremented to five weekly hours. In Madrid, the project covers ten infant and primary schools, which continue offering bilingual education under the auspices of the joint agreement and curriculum. As for secondary education, the English Linguistic Sections were conceived as a temporary measure to offer bilingual education to those students coming from the first cohort of the British Council-MEC primary schools. Later on, in the academic year 2013/14, they were assimilated to the new Bilingual High Schools in the framework of the CAM Bilingual Project when the latter became mainstream.

#### 1.4. The Role of Language in CLIL

##### 1.4.1. The Types of Language in CLIL Contexts

Language is present everywhere as one of the most visible expressions of culture, the tool we have to communicate our thoughts and wishes, and to access content knowledge. In fact, even if the language used for academic instruction is our mother tongue, school learning involves lots of cognitive effort since the main tools we use for that purpose are based on language i.e. reading, writing, listening and speaking (Clegg, 2007: 144). However, due to the students' possible lack of fluency in the vehicular language, language in CLIL should not be merely considered as the vehicle of instruction for content knowledge or as an added value to subject matter, but also as a 'core skill' on its own (Graddol, 2006) i.e a basic component in CLIL along with content knowledge and skills. Therefore, apart from being a key issue in any classroom, the language in CLIL

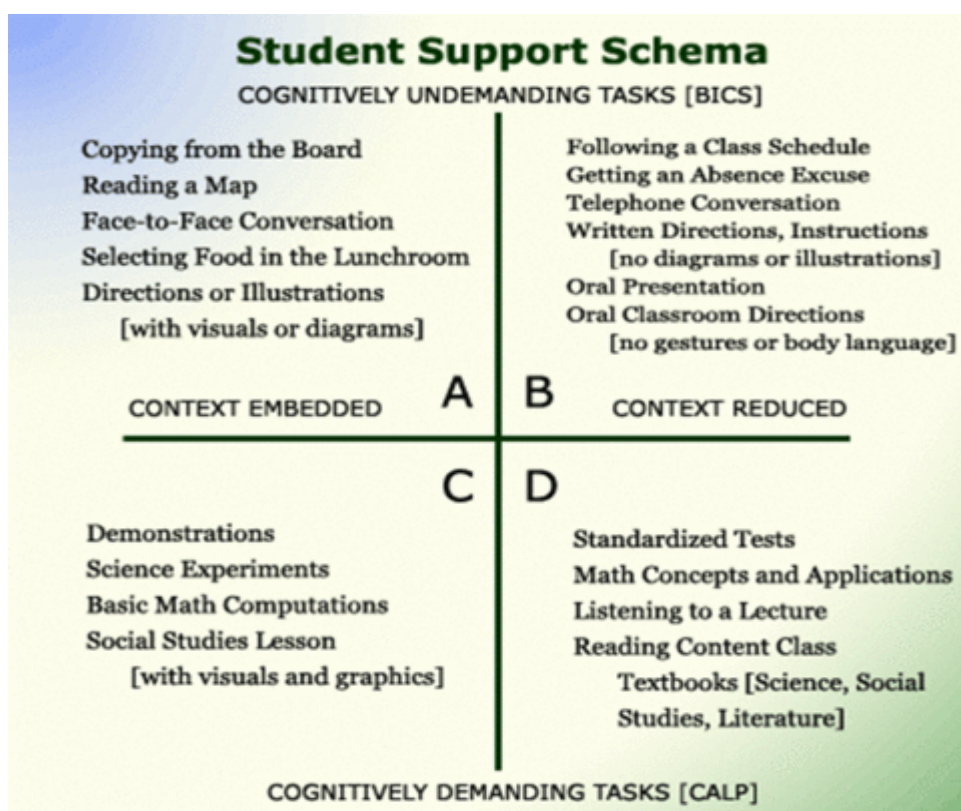
complicates the learning and teaching process. In fact, the language demands extra effort from learners, and challenges teachers to provide students support in understanding new concepts and academic jargon in another language, because “academic language is no one’s native language” (Coyle, 2015).

In broad terms, the language in CLIL refers to grammar, vocabulary, discourse markers, thinking skills/language functions and language skills (Ball, Kelly & Clegg, 2015: 66). In this section, the role of the vehicular language and the different manifestations it can adopt in CLIL will be analyzed in three steps. First, the types of language in CLIL (Coyle et al., 2010) and the concepts of content-compatible and content-obligatory language will be described taking Cummins’ (1996, 2000) classifications of Basic Interpersonal Skills (BICS) and Cognitive Academic Language Proficiency (CALP) as the basis for the analysis. Second, Dalton-Puffer’s Construct of Cognitive Discourse Function (CDF) (2013) will also be discussed as paramount to identify the discourse functions typical of the specific language in the CLIL subjects. Last but not least, in order to understand the relation between language and content knowledge and skills, I will also comment on the three levels at which language can be integrated in CLIL (Llinares, Morton & Whittaker, 2012).

The distinction between BICS and CALP was introduced by Cummins (1979, 1981a) in order to highlight the difficulties that students of a second language might find as compared to native language speakers in mainstream education. BICS refers to context-embedded language or “conversational fluency” in everyday uses of the language whilst CALP refers to context-reduced language or how different registers vary depending on the academic field i.e. “students’ ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school” (Cummins, 2008: 71). This academic language, which tends to be represented in the written form deserves

particular attention in bilingual programs since it takes longer to be acquired (Cummins, 1996, Mc Kay et al, 1997), it complicates the tasks, and forces students to engage in more cognitively challenging activities. The following graph shows different types of tasks depending on their cognitive demand along with the extent to which language is context-embedded or content-reduced.

Graph 1: Student support schema for BICS and CALP



From [http://jillrobbins.com/au/540/cummins\\_quad.pdf](http://jillrobbins.com/au/540/cummins_quad.pdf)

As can be observed from the graph, some tasks such as copying from the board and following directions or understanding illustrations require everyday language usage which is context-embedded and thus, easier for students. On the contrary, some other tasks such as conducting an experiment and listening to lectures deal with the type of

language which is context-reduced, and which inevitably requires more cognitive effort and academic language proficiency on the part of the learners. While the language and literacy-related demands of the curriculum - the registers of school - are unfamiliar to a greater or lesser extent to all children when they start school, native children learn these new concepts and registers through the medium of their mother tongue. Thus, native speakers can build on the foundations of their first language unlike second language learners in an English-medium school, who cannot rely on that prior knowledge. Children who are learning through the medium of their first language, and who come to school already having acquired the core grammar of this language, and the ability to use it in a range of familiar social situations, have a head start in learning to use the academic registers of school (Gibbons, 2002: 5).

Academic language consists of different levels in which meaning is organized, namely the lexical and morpho-syntactic level (vocabulary and grammatical patterns) and social-pragmatic level i.e. the functions or uses of academic language and its text structure. Along with these two levels, academic communicative competence implies the ability in the communicative uses of language in all productive and receptive skills, and the ability to process this academic or subject-specific language as well as the language for interaction in classroom situations (Lorenzo, Trujillo & Vez, 2011: 139). First, communication in the CLIL classroom is built through an additional language other than the dominant language in the community. Second, this academic language is demanding and complex since it refers to very specific domains and it might also include higher cognitive skills such as classifying, evaluating and synthesizing.

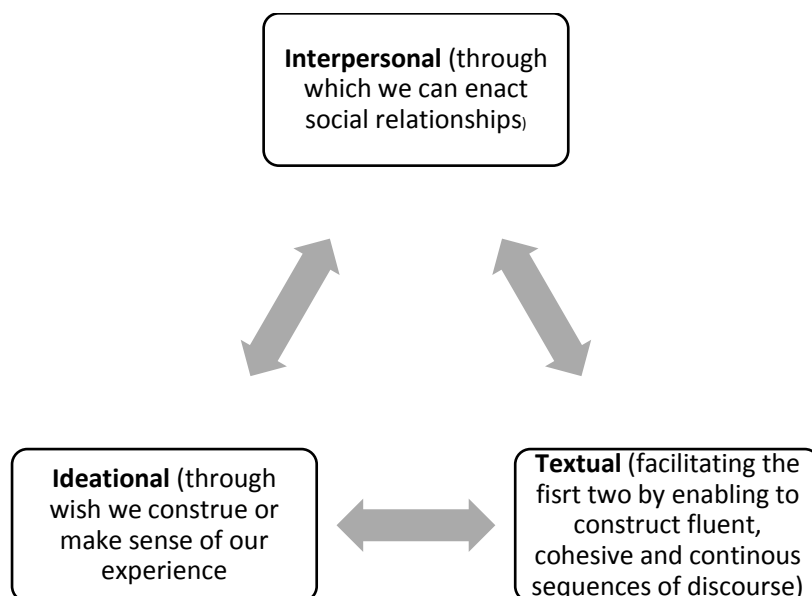
As for the most common challenges students deal with in relation to language issues, these can be organized into three levels: the discourse, sentence and word levels. The discourse level refers to the challenges affecting learners' ability to interpret or produce



longer pieces of spoken and written information. Second, the sentence level refers to the difficulties students encounter for instance with grammatical structures in sentences such as problems with tenses, word order, question forms, comparisons, pronouns, linking clauses and sentences amongst others. Third, the word level deals with the difficulties in relation with the meaning or form of words as in words having different meanings in different subjects, using figurative language, and the meaning of Greek and Latin-based words to name just a few (Dale & Tanner, 2012: 44). These linguistic challenges require some action on the part of the teacher to modify or adapt academic texts in order to suit to students' needs, and even to consider language alternation or "translanguaging" i.e. to allow students to use their L1 in order to work through complex ideas (Swain, 2006). The resulting language is seen as a type of interlanguage, which should be judged taking into account communicative intentions regarding language functions rather than students' language accuracy or grammatical correction in the effort to express complex content knowledge with a relative lack of linguistic resources.

According to the Systemic Functional Approach (SFL) we mentioned as one of the theoretical approaches on which CLIL is based, the choices speakers and writers make from the lexical and grammatical systems of a language draw on the following three types of meaning or meta-functions of language:

Graph 2: Halliday and Matthiessen's metafunctions of language

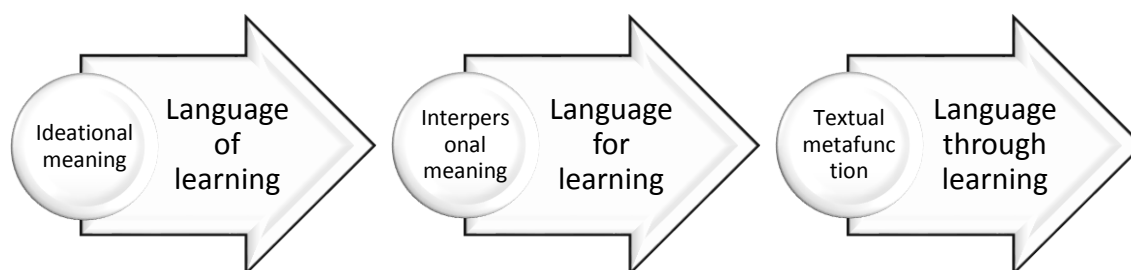


(From Halliday and Matthiessen, 2004: 29-30)

These meta-functions of language are seen as more general and intrinsic to language use than the functions of individual examples of language use (Halliday and Matthiessen, 2004: 31), and they provide an excellent framework to understand the functions of language in CLIL or how language is used in CLIL contexts (Llinares, Morton & Whittaker, 2012: 10). Considering Coyle et al.'s (2010) model of language in CLIL, the language in CLIL plays the following roles: First, as the vehicle through which content knowledge is transmitted (*language of learning*) which can be paralleled with the ideational meaning as it deals with the way content knowledge is represented through talk and other communication modes in the classroom. Second, as the instrument for peer interaction (*language for learning*) or interpersonal meta-function managing social relationships in the educational context, and how different stances of content are

expressed. Finally, language is also regarded as the instrument to access cognitive skills (language *through* learning) or the textual meta-function considering how specific content is constructed and put together in texts, and how teachers can help students construct this content expression as they move from spoken to written forms of discourse.

Graph 3. The types of language in CLIL



The language *of* learning or the language needed for learners to access basic concepts and skills related to subject theme or topic (Coyle et al, 2010) is, for instance, the language of science, which is specific depending on different disciplines, and which usually presents challenges for students, e.g. using the past tense or recognizing past markers in History. The language *for* learning is described as the kind of language needed to operate in a foreign language environment or BICS, i.e. language for pair work, group work, debating and discussing topics, expressing your own view, etc. Finally, the language *through* learning is based on the principle that effective learning requires active involvement in language and thinking processes, “a level of talk, of interaction and dialogic activity which is different to that of the traditional language or content classroom”, e.g. the ability to grasp emerging language *in situ* (Coyle et al. 2010: 37).

Another distinction in terms of the types of language in CLIL, which can help teachers identify the language students need for CLIL subjects, and make them visible for students (Llinares et al., 2012) is that of content-obligatory language (COL) and content-compatible language (CCL), the latter being also referred to as ‘peripheral language’ (Ball, Kelly and Glegg, 2015).

Table 2. Content-obligatory and content-compatible learning

<b>Content-obligatory language</b>	<b>Content-compatible language</b>
Necessary to learn the key content concepts	Expands language beyond academic forms and functions
Primary, usually generated first	Provides extra or additional language
Content or discipline specific, more academic in nature	Include more communicative forms and functions
What-oriented, the “What” being the content	How-oriented, more than what
Required to learn for successful assessment	Complement and supplement the content-obligatory language

(Adapted from Fortune & Teddick, 2016)

Both content-obligatory language and content-compatible language assist the content teachers in making them aware of high and medium frequency words or the most commonly used vocabulary both in general English and academic English, and of the most frequent collocations used in curricular contexts (Bentley 2010: 15).

To highlight the presence of language in CLIL contexts, Dalton-Puffer suggests a taxonomy of language functions dedicated to the verbalization of the cognitive processes that are central in CLIL subject curricula (Dalton Puffer, 2013; Mohan, 1986). Although the Construct of Cognitive Discourse Functions (CDF) is still under study and discussion, and further empirical research is needed to support the author's intuition, it can be helpful in the following ways. First, it can establish "a zone of convergence between content and language pedagogies", and it can be used to "help teachers to demonstrate the students how rational and deliberate thought works" (2013: 16-17), i.e. how to deal with Higher Order Thinking Skills (HOTS) effectively. As can be seen in the table below, the cognitive language functions, considered illocutionary speech acts and verbalizations of cognitive processes to communicate content knowledge, are the following: classify, define, describe, evaluate, explain, explore and report. For teachers to be able to connect these cognitive language functions with practical activities or tasks in the CLIL class, the actions that students are asked to perform have been translated into the typical verbs used for each of the cognitive discourse functions (e.g. classify, compare, contrast, match, structure, categorize, subsume).

Table 3. Cognitive discourse functions in CLIL

F1: Classify	<i>Classify, compare, contrast, match, structure, categorize, subsume</i>
F2: Define	<i>Define, identify, characterize</i>
F3: Describe	<i>Describe, label, identify, name, specify</i>
F4: Evaluate	<i>Evaluate, judge, argue, justify, critique, comment, reflect</i>
F5: Explain	<i>Explain, reason, express cause/effect, draw conclusions</i>
F6: Explore	<i>Explore, hypothesize, speculate, predict, guess, estimate</i>
F7: Report	<i>Report, inform, narrate, present, summarize, relate</i>

(Adapted from Dalton-Puffer, 2015: 69)

Finally, apart from the types of language and the discourse functions specific to CLIL contexts, language can be integrated with content and skills at three different levels (Llinares et al., 2012) as regards classroom interaction, curriculum integration and the way the language is intertwined with content in learners' language development and assessment. First, the language used in classroom exchanges refers to the opportunities that can be created for CLIL learning through different types of interaction namely teacher-student and student-student interaction. These communicative exchanges differ from other real communicative situations- language *for* learning- in that they present a clear focus on the instructional and regulative register- language *through* learning. That is, communicative systems are used with the pedagogical goal of guiding students towards the appropriate use of language in CLIL contexts. Secondly, content and

language integration in the curriculum refers to subject literacies as dealing with genre and registers. Subject literacies deal with the types of language that are relevant to CLIL subjects - language *of learning*- and the way we can help students in learning and using those types of languages i.e. the genre or types of texts whose function and structure students need to learn. Third, content and language integration is also visible in learners' language development and assessment in the following categories: To start with, as the language for academic content or the vocabulary and grammatical choices used in order to construe meaning, which is adequate for the genre, register and type of activity- language *of learning*. Also, as the interpersonal language or language for socializing, operating in the classroom, dealing with specific genres and personal experience i.e. language *for learning*. Finally, as the language students need to move from speaking to writing representations of discourse such as those strategies activated in dealing with nominalization or the language *through learning* (Llinares & Whittaker, 2006).

#### 1.4.2. Every CLIL teacher is a language teacher

In this context of integration of content knowledge, language and skills, the teachers' role in accommodating language-related issues in the subject curricula is of paramount importance. The conception of every teacher as a language teacher originally dates from the Bullock report (1975), which attempted to review the status of the teaching of the English language in the UK. The report reflected on the fact that most schools did not explicitly focus on English literacy but rather expected students to learn academic vocabulary and reading and writing skills on their own, without specific time being devoted to those aspects. The report also aimed to improve the level in English by making language visible and explicit to students, and ensure that all of them left school with job skills such as being able to read and write proficiently. Likewise, and as a natural follow-

up of the recognition of the paramount role of language skills in mainstream education, it focused on need for teachers to take responsibility for language skills development – more specifically reading skills - in their subjects, and for making language content and academic vocabulary (CALP) more accessible in core subjects. In this sense, the Bullock Report voiced concerns about Teacher Language Awareness (TLA), which later have been translated to other contexts such as CLIL.

Appreciation of language-related issues on the part of the teachers or the so-called Teacher Language Awareness (TLA) is described as a deep insight into how the language works, how to use it, and how to show a high level of proficiency (Edge, 1998: 10). It refers to the knowledge that teachers have of the underlying systems of the language that enables them to teach effectively (Thornbury, 1997; Andrews, 2007), namely the knowledge about the language demands of the curriculum and the students (Gibbons, 2006; Glegg, 2007). This knowledge implies empathy for the students' difficulties in acquiring and learning a vehicular language (Andrews, 2003), and understanding the support students need in order to accomplish subject tasks through a second language (Clegg, 2007). The language demands of the curriculum are dependent on the specific subject and register, and engage teachers in several processes. The first of them will be the analysis of textbooks and other teaching materials in order to identify aspects such as the spoken language demands and the types of texts students will require to read and/or listen to. This close analysis of materials is also necessary to consider the written text types and their schematic structure, the most significant grammatical aspects e.g. the use of tense that the topic demands, and the appropriate content-obligatory and content-compatible lexis for the subject matter (Gibbons, 2002: 22), before these aspects can be made visible for students.



As for the focus on students' difficulties, since the CLIL approach does not necessarily require students to have reached a specific level in the vehicular language, teachers must also be attentive to students' language needs. Thus, teachers can identify what might be challenging for them beforehand, and plan accordingly in collaboration with the language teacher. This emphasis on language pedagogies typical of foreign language instruction is, in fact, beneficial in CLIL contexts for a variety of reasons. On the one hand, because the specific focus on form through an overt focus on language (Pérez Vidal, 2007) might help students at the upper levels in the acquisition of the foreign language as opposed to second/foreign language acquisition in young learners who still seem to lack the abilities to contrast the vehicular language and their mother tongue. On the other hand, the dangers of not being aware of language in CLIL could minimize the correction of students' errors or even result in the overlooking of errors with the subsequent lack of accuracy over fluency in the foreign language that some immersion programs in Canada were criticized for (Lyster, 2007). Finally, in identifying language objectives in bilingual education, teachers can relate them to competences and "assessable learning outcomes" and thus, contribute to cross-curricular teaching as will be discussed below.

#### 1.4.3. Curricular Integration and Teachers' Roles in CLIL Contexts

Curricular integration or the 'interdisciplinary approach' (Jones, 2009) is a complex educational challenge which requires attention to content-specific disciplines and language objectives (Dalton-Puffer, 2007; Lyster, 2007; Pica, 2008) along with new organizational principles regarding the way we conceive teachers' roles and subjects. From this perspective, it is necessary to define the role of content teachers and language teachers so that they can work in successful cooperation with each other. On the other

hand, we also need to consider subjects from a more holistic viewpoint to envision competences and contents not as pertaining to individual subjects but to knowledge areas, and to engage in cross-curricular work which enriches instruction (Savage, 2011; Westwood, 2006). The following are desired changes to be implemented for successful integration of content, language and skills according to Lorenzo, Trujillo & Vez (2011). First, in relation with term and concept unification, the incorporation of linguistic concepts into other curricular subjects requires sharing these linguistic terms, which need to be added into each subject. Contrarily to what is commonly thought, and as will be explained in detail when we refer to content teachers' role, this does not entail that the content specialist teaches grammatical concepts. Instead, integration should include higher levels in the linguistic system such as those related to social and discourse functions i.e. working with argumentative texts in History or being able to design a scientific poster in Biology to promote higher literacy levels in students. Secondly, in order to help student advance in their language skills, it is also highly recommended to add some flexibility in schedules in order to organize groups of students taking into account linguistic levels, and thus make the most of students' participation and interaction in class. Third, as for term and concept integration, the integrated syllabus should be designed in such a way that linguistic goals and objectives are incorporated into the content subjects without content-related goals being compromised. In this sense, integration is considered not as an end in itself but as a medium to achieve general and specific goals so that non-linguistic and linguistic competences are aligned. Finally, regarding assessment, it is necessary to point out that changes in assessment practice derive from the changes observed so far. As will be described in the chapter devoted to assessment in CLIL, successful assessment results from the teachers' awareness of the elements they take into account, only content-related issues or also language, and to what

extent incomplete mastery of the L2 can interfere with the expression of content knowledge. Likewise, when assessing language components in CLIL subjects, teachers need to know if their assessment criteria conform to the principles of second language acquisition and use so that the language is graded in a fair way, i.e. if language mastery is considered incomplete and subject to change as opposed to native-like proficiency (Lorenzo, Trujillo & Vez, 2011: 164).

#### 1.4.3.1. Cooperation among Teachers

Teacher cooperation is considered as a way to share and compare teaching practices. It can lead to the development of more effective responses regarding students' misbehavior and disruptions in class, and it also stands out as a powerful engine for change and quality development in educational institutions (Teaching and Learning International Survey, 2014). More specifically, in CLIL provision, cooperation is vital as it helps create communities of learners in the ecosystem of the bilingual school (Lorenzo Trujillo & Vez, 2011), and integrate content and language-related issues in an efficient manner (Bertaux et al., 2010; Marsh et al. 2010; Pavón & Ellison, 2013). In this context of collaboration, the joint efforts of content teachers and language teachers do not only become essential in that they contribute to the efficient integration of the vehicular language with content issues as was pointed out before but also to help to construct knowledge holistically. However, it is interesting to point out the fact that content teachers- especially those who have not received specific training in language pedagogies- commonly complain about the lack of time for dealing with both content and language focus in the class (Bigelow, 2010: 37). These complaints might be due on the one hand, to teachers' limited perception of their role as content teachers since they tend to think they should not be made responsible for language-related issues. On the other hand, it might also respond to common misconceptions about CLIL pedagogies and the role of vehicular languages

assuming that CLIL subjects are taught *in* English, i.e. that the vehicular language is used only as the medium of instruction, not as an essential component of it. Furthermore, content teachers might feel insecure about their language level, their roles and the best way to deal with language aspects (Otto, 2017 b), probably due to the lack of previous training on CLIL and language pedagogies, which is common in some CLIL contexts (Eurydice, 2017). Consequently, an analysis of teachers' needs is still necessary (Halbach, 2010; Alejo & Piquer, 2010) and prior training in both language skills and a pedagogical approach to CLIL become paramount (Halbach, 2010; Salaberri Ramiro, 2010).

Teachers participating in CLIL instruction come from a variety of disciplines. On the one hand, language teachers can be certified to teach other non-linguistic subjects or have a double degree, which is quite common in some European countries such as Belgium and Sweden (European Commission, 2012: 88). On the other hand, content teachers can obtain some language certification allowing them to teach subjects through a vehicular language as happens in the Spanish territory. However, as we commented in section 2.2 (CLIL in Europe), in the latter case, teachers' initial training can involve various levels of language ability. Besides, in those countries in which the foreign language is also a subject in itself, the most common situation is to have a language specialist teaching English as a foreign language while content specialists are responsible for teaching non-language subjects. Regardless of teachers' profiles, and despite similarities in language and content teachers' responsibilities- like promoting and maintaining discipline and helping each other in identifying goals and objectives amongst others- content teachers and language teachers might be confronted regarding their responsibility for language-related issues or when their functions have not been defined in advance.

As Coyle, Holmes & King (2009) point out, the swiftest solution in order to overcome language restrictions on the part of content teachers is usually to work in collaboration with a language specialist. CLIL teachers are unlikely to work on their own but rather through teamwork, sharing responsibilities for teaching and learning across subjects, and developing a broader perspective on curriculum design (Coyle, Holmes & King, 2009: 17). However, this is not always a frequent practice, mostly due to the dual profile of teachers, in secondary schools, teachers do not always benefit from joint planning (OECD, 2014).

Ideally, in a CLIL program, there should be language specialists together with teaching assistants or language assistants working in collaboration with subject specialists as they can exchange subject knowledge and broaden their skills and understanding across different areas. As was pointed out before, CLIL teams are commonly successful when there is a common vision of shared goals across subject disciplines. Furthermore, content teachers also benefit from the contact with language teachers who are more open to the communicative approach to foreign language teaching and thus, do not consider themselves as mere transmitters of knowledge about language but as embracing more flexible roles as communicators, organizers and facilitators of knowledge (Abelló Contesse et al, 2013: 12).

Still, some CLIL practitioners seem to be insecure about how to organize their CLIL sessions and more specifically as to whether they should give over responsibility for her/his subject(s) to the language teachers. In fact, collaboration between language teachers and subject teachers does not imply that ESL teachers are subordinated to the content areas and that imbalance is found regarding curriculum authority and importance (Davison, 2006: 456). Instead, this collaboration aims to balance the development and

support of language learning and the development and support of content learning (Coyle, Holmes & King, 2009).

But how is this collaboration made effective in daily practice? The underlying conceptual framework is thus, that of collaboration where subject area specialists and language specialists work together to design the course. Collaborative groups between members of the same didactic department are also a good option in that they function as an ecosystem for the creation of class material and other teaching resources. Besides, they can also help fellow teachers analyze learners' needs, develop efficient task-sharing and support colleagues in using innovative methodologies (Bertaux et al. 2010: 8).

#### 1.4.3.2. The Role of Language Teachers

As was pointed out in relation to teachers' profiles, both content and language teachers come from a variety of profiles. For practical purposes regarding the nature and purpose of the study, language teachers will be referred here as those specialists in language subjects regardless of the fact that they might also be specialized in other non-language subjects. Consequently, language teachers are responsible for the subject they teach, i.e. English as Foreign Language and not the content subjects. However, as regards the correct integration of content, language and skills in bilingual education, the following are necessary actions to accomplish by language teachers when they work in bilingual contexts. The English teacher is the one who must assume the responsibility for language, following the content teacher's observation about linguistic needs and deficiencies that need to be corrected, and identifying language demands in content areas so that students are able to understand and learn academic content (Davison, 2006: 462; Pavón & Ellison, 2013: 68). To do so, language teachers can help to establish clear-language focus, and analyze language demands for content areas. Second, they can help content teachers to plan instruction, and contribute to reporting on students' foreign language development

by analyzing and considering aspects such as students' prior language level in relation with the CEFR, and students' language challenges and difficulties. Third, language teachers can help by collecting useful materials and strategies for class support, foster cross-curricular language awareness in students, and help content teachers to use FL strategies (Davison, *ibid*: 462). Finally, in relation to assessment, language teachers will assess language in their subject whereas content teachers must focus on content by offering students a range of varied assessment tools so that they can show their acquisition of knowledge without their marks being biased by a higher or lower language proficiency.

#### 1.4.3.3. The Role of Content Teachers

Ideally, in CLIL contexts, content teachers would be bilingual and/or multilingual since these language skills allow them to make content knowledge comprehensible for students as well as offer them the necessary linguistic support. However, this is not always the case. To start with, as was pointed out in the previous section, teachers' linguistic level can vary depending on the country and region. For instance, in some Autonomous Communities like Andalusia in Spain, content teachers only need to certify a B2 level to access teaching in bilingual education contexts. Besides, content teachers' requirements do not necessarily entail prior training in the CLIL approach or language pedagogies (Eurydice, 2017), especially in some contexts where secondary teachers are subject teachers who certified the required language level to teach CLIL subjects. Consequently, as content specialists, content teachers focus mostly on content objectives rather than linguistic ones, and they might not feel responsible for the latter, a fact that has a deep impact in some educational stages:

One of the challenges of CLIL education at the secondary level, in contrast to primary education, concerns teacher profiles. While teachers in primary education have a dual profile (content and language) most teachers in secondary are content experts with certified knowledge of the target language. Therefore, some of the greatest efforts from the administration are focused on both ensuring teacher competence in the foreign language as well as raising their awareness of the specific language demands and characteristics of the different subject disciplines. (Llinares & Dafouz, 2010: 100)

Regardless of their training, for CLIL to be effective some actions need to be taken on the part of content teachers. In general, following the recommendations from the *CLIL Teachers' Competence Grid* (2010), the content teacher should be able to take the following actions: To start with, adapt the course syllabus in order to include content, language and learning skills outcomes. Second, integrate the language and content area curriculum so that content is supported by language-related goals and vice versa. Third, guide students in the processing of both BICS and academic language (CALP), and select the language needed to provide rich input, and to ensure students' learning in both content and language. Finally, for learning to be meaningful, content teachers need to help students develop meta-cognitive awareness, and therefore, deepen their understanding of content subjects (Bertaux et al., 2010: 4).

On the other hand, by concentrating on the CLIL language, the content teacher becomes a language user and language promoter (Coonan, 2013), incorporating a procedural ('knowing that') and declarative ('knowing how') dimension of language in subject matter knowledge (Andrews, 2007: 31). In becoming fully aware of language processes in CLIL, content teachers should engage in several actions apart from taking responsibility for students' content and language knowledge and skills development. First, they need to consider their own level of language awareness in the subject (s) they teach (Davison, 2006, Marsh et al., 2001) as well as revise their language proficiency



regularly. These language abilities relate to the competence to master sufficient target language knowledge and pragmatic skills of the vehicular language. Second, the CLIL teacher needs to master some linguistic principles or what s/he terms as “theory”, i.e. the comprehension of the differences and similarities between the concepts of language learning and language acquisition so that s/he can deal with language-related issues effectively (Marsh et al. 2001: 78-80). In this sense, it is also relevant to establish clear learning goals so that language is visible for students. These learning goals can be presented in the form of vocabulary, the four language skills, grammatical structures, functions of language and learning strategies so that learners are able to acquire them in a successful way (Baecher, Funsworth & Ediger, 2014: 118). Third, the content teacher should be prepared to deal with learning strategies adapted to both content and language issues in the classroom. This knowledge of foreign language methodology and pedagogy is usually what teachers lack, and what leads them to focus on content-related objectives and forget about the foreign language (Arkoudis, 2006). Under this category of foreign language pedagogy, we consider, on the one hand, the ability to notice linguistic difficulties, recognize students’ interlanguage, and be able to use communicative and interactive methods facilitating the understanding of meaning or subject knowledge. Among these methods, we find repetition and echoing for correction and modelling good language usage, and use dual-focused activities, which can cater for both language and subject aspects (Marsh, 2007). Furthermore, regarding the learning environment, as Marsh et al. point out (2001) the content teacher should also have the ability to work with learners of diverse linguistic/cultural backgrounds. Last but not least, the content teacher must also be responsible for the creation and development of materials suiting the students’ needs and the CLIL purpose, and for assessment issues (Bertaux et al., 2010). In terms of assessment, content teachers need to be able to develop and implement

Formative Assessment tools appropriate for the CLIL scenario, including the following actions. On the one hand, the ability to make connections between planned outcomes, learning skills and processes, actual outcomes, planning and negotiating strategies for future learning. On the other hand, to use self and peer-assessment tools as recommended by Formative Assessment or Assessment for Learning (AfL). Another important feature is to be able to maintain a triple focus on language, content and learning skills, and use CLIL-specific characteristics of assessment. (Bertaux et al., 2010: 8). As is recommended by the same authors, these specific features include the following: First, to use the language for various purposes. Second, to work with authentic materials, and regular communication with speakers of the CLIL language in order to promote ongoing language growth, and some level of comfort in experimenting with language and content. Finally, to distinguish content and language errors while carrying out assessment in the target language.

Another big challenge regarding assessment is to offer students the appropriate assessment tools so they can show what they do know in terms of concepts and skills rather than catch them out on the things they are not able to do (Assessment Reform Group, 2002). Finally, practitioners should be able to prepare students for formal examinations including high-stakes examinations.

#### 1.4.4. Making language salient in the content class

Once I have analyzed content and language teachers' roles, it is time to focus on the strategies at the teachers' disposal to make language visible in content subjects. The language in CLIL has to be highlighted and not simply taken for granted as something students will simply "catch up by osmosis" (Llinares, Morton & Whittaker, 2012: 14).

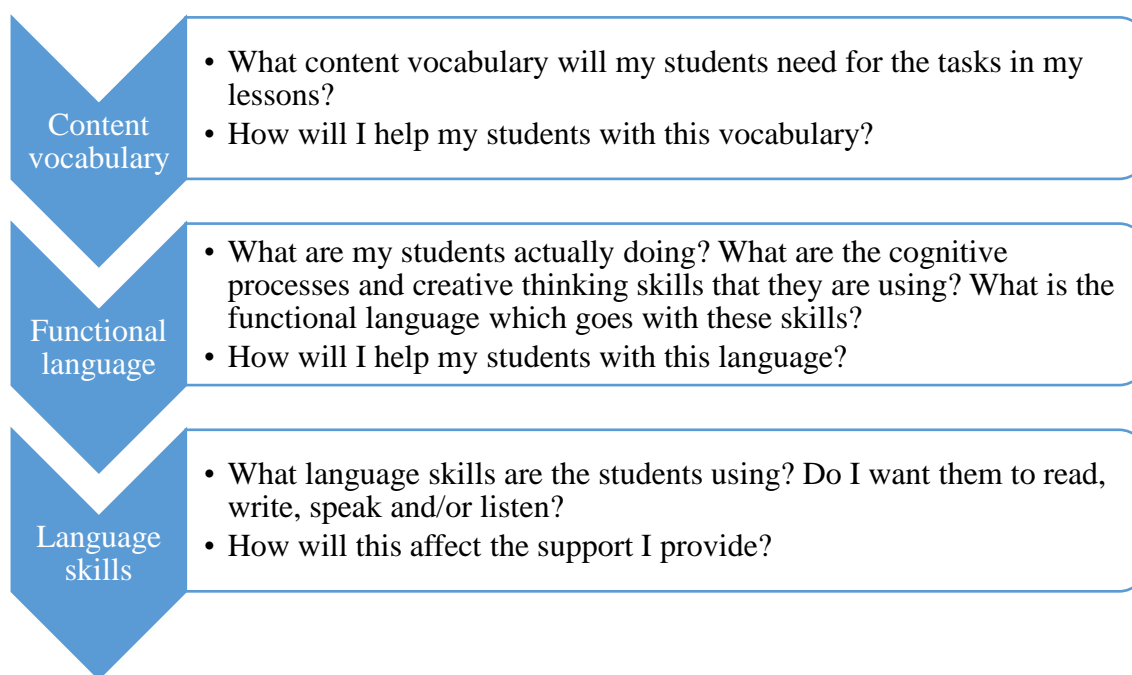
This focus on language is especially relevant since students' foreign language might be insufficient in order to express content knowledge and skills in CLIL subjects as compared to proficiency in their mother tongue or students' language proficiency in immersion contexts. Thus, in introducing the language in CLIL, teachers need to bear in mind that student learning progression in bilingual education requires more time than the mere acquisition of BICS in L2 learning (Hulstijn, 2015; Cummins, 1981). Besides, apart from the content vocabulary, academic language has special features, which deserve closer attention: It is more precise than BICS, it avoids slang, and has its own style and tone.

But how can this language visibility be done in practice? To start with, language objectives in a lesson or didactic unit must be made visible and explicit by teachers as referring to both the language demands of the curriculum and those of the students (Gibbons, 2006) as was pointed out in the previous section. To do so, the teachers themselves need, on the one hand, to consider language as an essential step in the planning of the lesson, and raise awareness on the language that students will need by taking into account that language functions vary from one register to another. On the other hand, they should reflect on the level of illiteracy that students show in the foreign language and plan accordingly. The biggest challenge, however, is to make teachers aware of the importance of language and literacy in their subject (Morton, 2016) so that the curriculum is not as demanding for them to devote some time and effort to deal with language issues (Airey, 2013). It might also be the case that teachers find it difficult to identify language objectives (Llinares & Whittaker, 2006: 28) for a variety of reasons. First, maybe because content teachers often confuse language objectives with language activities and thus, need to work further on the first, and to consider how specific they wish them to be (Baecher, Funsworth & Ediger, 2014: 131). Introducing language objectives in the form of

functions, grammatical structures, micro-skills, specific vocabulary and the associated learning strategies in the CLIL class results in a new form of language interaction or discourse. It presents distinguishing features compared to other forms of discourse in Second Language Acquisition, and, consequently, it requires several strategies by content teachers, such as the conversion of an ideational text into a didactic one to name just a few (Flowerdew & Peacock, 2001). Second, these language objectives might be blurred because there are few easily available frameworks for the integration of content and language (Morton, 2016). Finally, it might happen, as was pointed out before, that content teachers do not see themselves as language teachers and therefore, they do not consider they should be made responsible for language-related aspects and deal with students' use of the language (Airey, 2012). In fact, some of them also point out that they feel they might not be prepared to deal with language in content lessons as they can make occasional mistakes themselves.

As content teachers are not often trained on how to raise awareness of language in academic subjects and identify language levels and support language in CLIL, Chadwick (2012) suggests to consider the following questions when dealing with language challenges in relation to content vocabulary, functional language and language skills:

Table 4. Raising awareness on CLIL language



(From Chadwick, 2012: 4)

Besides, apart from raising awareness on the CLIL language, a profound analysis of the genres in CLIL can help content teachers adapt the tasks accordingly, and design a linguistic inventory for their subjects. This will enable teachers to go beyond academic vocabulary for each topic, and provide additional information on the grammatical and discursive features from the activities and genres that students need to master in the foreign language to produce good oral and written texts (Llinares & Whittaker, 2006: 28-29).

Like texts, the types of tasks vary depending on the subject. In science, for example, learners need to know to hypothesize, observe experiments, and describe different procedures. In History and Geography, learners read source materials, recounts, reports and case studies, and produce written and oral texts highlighting causes and effects, to name just a few. In Art and Music, learners read and write descriptions and explanations.

Thus, regardless of the different genres and academic disciplines, by working with different, texts and the tasks associated with them, content teachers can prepare students for meaningful learning in CLIL.

## CHAPTER 2: The CAM Bilingual Project

### 2.1. Origins and Structure

In the academic year 2004-2005 Madrid's Regional Government created a network of 24 Bilingual Schools, in a context in which bilingual education was still considered elitist and confined solely to the renowned private and subsidized institutions or the schools in the British Council project. The Regional Government of Madrid, being conscious of the importance of mastering at least two foreign languages, was determined to create a fully competent in English community by using English and Spanish as vehicular languages to acquire content knowledge and skills: "The Comunidad de Madrid is convinced that bilingualism is imperative for students to be effective and integrated participants in the European Union, and that the Ministry of Education is demanding renewed efforts to achieve this goal<sup>3</sup>" (Consejería de Educación, Juventud y Deporte, 2009: paragraph 1). The initiative rapidly expanded, and soon numerous families became interested in the project. In the academic year 2009-2010 there were already a total of 206 bilingual schools in the Madrid area, a number which has dramatically risen in the whole community. The project was also extended to subsidized schools in the academic year 2008-2009, and it has coexisted in the region with the initial 10 Bilingual Primary Schools from the MEC-British Council Program. Additionally, other institutions such as the BEDA Program (Bilingual English Development & Assessment) by FERE-CECA Madrid and Colegios Bilingües Cooperativos, to name just a few, contributed to the expansion of the bilingual program.

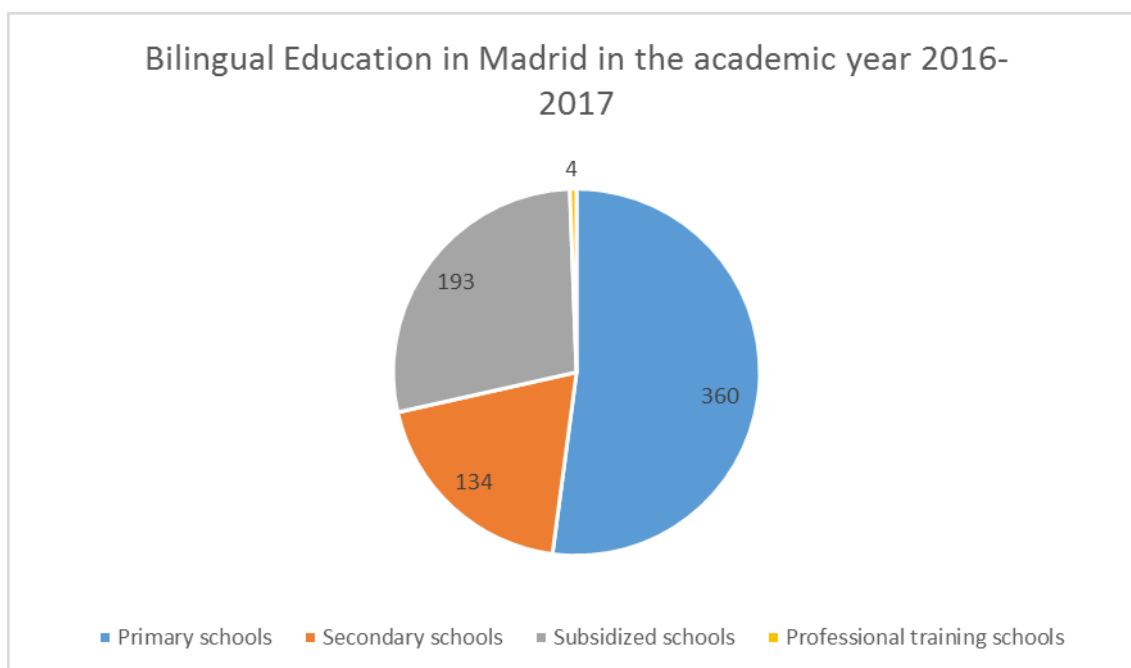
Bilingualism is not considered a passing fad but a beneficial approach for students. In fact, in the effort to ensure the continuity of bilingual programs, a quality necessary for

---

<sup>3</sup> Author's translation.

the success of the program (Llinares & Dafouz, 2010; Lasagabaster & Ruiz de Zarobe, 2010) and students' access to better job opportunities, Madrid's Regional Government has made it extensive to other educational stages. In the academic year 2016-2017, the Madrid region has implemented bilingual education in professional training, and to-date bilingual schools in the Madrid area reach 45% of the state sector. As can be seen in the following graph (Comunidad de Madrid, 2016), the 498 state schools comprise 360 primary schools, 134 secondary and upper secondary schools and 4 professional training schools along with 193 subsidized schools.

Graph 4: Bilingual education in Madrid in the academic year 2016-2017



One of the most distinguishing features of the CAM Bilingual Project was its implementation in the whole school and not only sections in the primary stage, as is a common procedure in other bilingual programs in Europe. Regarding legislation and organization, Madrid's Regional Government has been adapting its regulations as a result



of the experience gained throughout years of implementation. Today, a minimum of 30% and a maximum of 50% of the curriculum is taught in English. This means that in Bilingual Primary Schools a minimum of five hours is devoted to EFL lessons and three hours to subjects taught through the medium of English. In Bilingual Secondary Schools, the time allotted for English lessons and English as a vehicular language depends on whether the students are following the Bilingual Section or Bilingual Program itinerary, as will be explained in depth in the section devoted to Secondary Bilingual schools. All the subjects that are taught using English as a vehicular language should be entirely delivered in English but as the administration suggests, to ensure basic academic vocabulary is acquired in English and Spanish, it is the responsibility of each school to determine the strategies and resources used for that purpose (Royal Decree 89/2014). As in the case of the MEC-British Council Project, a specific syllabus was designed for primary education (Royal Decree 89/2014, 24<sup>th</sup> July for primary education), and an advanced level syllabus in the subject of *English as a Foreign Language* was introduced in secondary education after the first stage of the program implementation (Order 29/2013, 11<sup>th</sup> January for secondary education). However, no integrated curriculum for secondary education has been suggested so far.

Back in 2004, the first bilingual schools were organized in terms of co-tutoring. Each group counted with a Spanish tutor, responsible for those subjects taught in Spanish and an English tutor, usually an English teacher, responsible for the subject of English as a foreign language and Social and Natural Sciences, these subjects being the ones selected to be taught and delivered through English. The underlying principle for using languages in the class was that of “one face one language” (Grosjean, 2010) and the layout of classrooms depended on the language; that is, one class being used for English and

subjects taught through English and another class for subjects taught through Spanish so that students could easily associate languages with different people and physical spaces.

The program was implemented in the first grade of primary education, and it then continued to the rest of the grades. However, conscious of the benefits that an early exposure to foreign languages can have in young learners (Cummins, 2001; Halbach, 2010; Ferjan Ramírez & Kuhl, 2017), efforts are being increasingly made in some schools in order to raise the exposure to the English language in infant education, and it is hoped that the project will be extended to the early stages of schooling in the academic year 2018-2019 (Consejería de educación and Institute for Learning and Brain Sciences, 2016). This increased exposure in early stages is usually made possible with the help of English language teachers in the form of “language showers” or regular short periods of language exposure during a limited time ranging from 15 to 30 minutes several times a week. Apart from using both Spanish and English as the languages of instruction, the main goal of the program is to create a bilingual community where the students perceive English in a natural environment. For that purpose, the school displays all information and notices in English, and the teachers who teach English or CLIL subjects should always use English as the medium of communication with students.

#### 2.1.1. Bilingual Primary Schools

The Bilingual Project for primary schools is developed in state schools of infant and primary education which are selected by “Consejería de Educación, Juventud y Deporte”. For a school to become eligible to be bilingual, the principal has to submit an educational project supported by the majority of the teaching staff and the school board members to guarantee that the whole teaching community is in favor, and commits to being involved

in the program implementation. Other prerequisites are related to the feasibility of the program implementation regarding the school's experiences in other programs, teachers' specialization and proficiency in English and the school resources.

According to the current legislation (Order 5958/2010, December, 7th for Bilingual State Schools in Madrid), all Bilingual Primary Schools teach the following subjects using English as a vehicular language: *Social Science*, *Natural Science* and at least one more subject like *Arts and Crafts* or *Physical Education*. Thus, along with the teaching of *English as a Foreign Language*, any subject except for *Mathematics* and *Spanish Language* can be taught through the medium of English. The number of subjects that are taught using English as a vehicular language depends on the number of teachers officially certified to teach in a bilingual school by means of linguistic capability certification (Eurydice, 2017) or "habilitación lingüística".

### 2.1.2. Bilingual High Schools

The first state Bilingual High Schools were created in the academic year 2010/2011 along with the previously established Bilingual High Schools in the British Council-MEC Project, which were incorporated into the CAM Bilingual Project. In order to ensure a smooth transition from primary to secondary bilingual education, and although no specific recommendations were followed by Secondary High Schools, some of the participant schools imitated good practice from Linguistic Sections by the British Council-MEC Project. Thus, they adopted some of the following measures: Invite and welcome the secondary teachers into the primary school in order to spend some time in the classroom and observe teaching practice, inform secondary teachers about the pupils' level, and comment on their most common challenges in terms of language proficiency,

and finally organize welcome sessions into the secondary school so that prospective students and families could meet some of the teachers, and get to know the facilities (*Hand in Hand Magazine*, 2005: 14).

As in primary schools, for a secondary school to be eligible to access the CAM Bilingual Project, the school community commits with the “Consejería de educación” to the involvement in the implementation of the project and the teachers’ proficiency in English. Contrarily to primary schools, however, teaching in Bilingual Secondary Schools in the CAM Bilingual Project is not implemented in the whole school but schools organize the teaching into two separate categories or itineraries: The Bilingual Program and the Bilingual Section.

During the first years of implementation, for a student to be eligible to pursue Bilingual Secondary Education, s/he was required to have completed Bilingual Primary Education and pass the external language examinations at the end of primary education. Those students who had not pursued primary education in a bilingual school needed to certify a minimum level of B1 according to the CEFR to access the first and second grades of Compulsory Secondary Education (ESO) in the Bilingual Program, and a minimum of B2 (CEFR) to access the third and fourth grades of Compulsory Secondary Education in the Bilingual Section. However, the reform completed by the LOMCE in the academic year 2016/2017 established that the admission criteria for CLIL programs would have to be the same across the country, and that language skill requirements would no longer be allowed. Thus, students wishing to access the Bilingual Program no longer need to certify their level. In case students do not possess any language certification, Madrid’s educational authorities can administer a test measuring communicative competence in all four skills- speaking, listening, reading and writing- the student needs to pass, and which is usually conducted in the High School.

The range of subjects offered in the CLIL vehicular language in the Bilingual Program track depends on the number of certified teachers and the resources in the school, so these subjects might vary from one year to another, and from school to school. Thus, in Bilingual Secondary Schools, the same subject can be taught in English and Spanish, depending on the itinerary students pursue, being common for example to study one subject in Spanish in the Bilingual Program and the same subject in English in the Bilingual Section. In the lack of an integrated curriculum for these subjects taught in English, schools tend to adapt to the guidelines for subjects taught in Spanish, a fact that might lead to the unification of content subjects regarding methodologies and assessment criteria.

Finally, it is necessary to point out that students are allowed to change the itinerary they pursue from the Bilingual Program to the Bilingual Section and vice versa when difficulties arise, the student can benefit from it, and after consultation with the high school guidance department. These measures promoting students' levelling in terms of language proficiency are in fact one of the main strengths of the Project. As research on bilingual education has shown, for a bilingual program to be successful, previous linguistic knowledge and a relatively similar level in communicative competence of participants in the classroom among other factors are of crucial importance. (Baker, 2011; Baker, Lewis & Jones, 2013)

#### 2.1.2.1. The Bilingual Program

In the Bilingual Program, *English as a Foreign Language* is taught five days a week with a one-hour session each day. As for other subjects taught through the medium of English, in the first cycle (1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> grades) at least one of the following subjects should be taught using English as vehicular language either from Specific Core Subjects (“Materias específicas obligatorias”): *Biology and Geology*, *History and Geography*, and *Physics*

and Chemistry, from the subjects of *Technology, Software Programming and Robotics*, or from Specific Optional Subjects and Elective Subjects except for *Second Foreign Language, Remedial Math* and *Math Extension: Problem Solving*. In the 4<sup>th</sup> grade of the Bilingual Program, at least one subject from the Specific Core Subjects, Specific Optional Subjects and Elective Subjects mentioned above, except for *Second Foreign Language* should be taught through the medium of English.

#### 2.1.2.2. The Bilingual Section

Considered as the real bilingual project for secondary schools regarding the time devoted to the use of English as a vehicular language, the Bilingual Sections in the CAM Bilingual Project are the focus of our research. As in the Bilingual Program, *English as a Foreign Language* or the so-called *Advanced English Curriculum* is taught five days a week with a one-hour session each day. The inclusion of this subject in the mainstream curriculum is probably one of the main strengths as well as one of the most distinguishing features of the Bilingual Sections. This subject substitutes *English as a Foreign Language* in the first, second, third and fourth grades of Compulsory Secondary Education, and it is aimed at providing students with advanced language skills by covering both English language and literature. As for other subjects taught through the medium of English, the teaching of the *Advanced English Curriculum* with the rest of the subjects taught in English takes at least one-third of the weekly schedule. The following Core Subjects are taught using English as a medium of instruction in Bilingual Sections:

Table 5. CLIL Subjects in Bilingual Sections in ESO

1st grade ESO	2nd Grade ESO	3rd Grade ESO	4th Grade ESO
<i>Biology and Geology</i>	<i>Geography and History</i>	<i>Biology and Geology</i>	<i>Geography and History</i>
<i>Geography and History</i>	<i>Physics and Chemistry</i>	<i>Geography and History</i>	<i>Another subject</i>

In the second grade of Compulsory Secondary Education, and as we commented about the distribution of subjects in a Bilingual High School, the subject of *Physics and Chemistry* can be taught through the medium of English provided that there are available qualified teachers in the school. Besides, in the fourth grade of Compulsory Secondary Education, along with the subject of *Geography and History*, schools must ensure the teaching of another subject in English from the following: Core Optional Subjects, Core Specific Subjects or Core Elective Subjects except for *Second Foreign Language*. Thus, students following the Bilingual Section itinerary can pursue all the subjects in English except for the following: *Math, Spanish Language and Literature, Latin, Second Foreign Language, Remedial Spanish Language, Remedial Math and Math Extension: Problem Solving*.

The tutoring sessions will also be taught entirely using English as a vehicular language in all the grades of Compulsory Secondary Education in the Bilingual Section. Under no circumstances can the same subject be taught using two vehicular languages. i.e. English and Spanish in the same itinerary. However, unlike Primary Bilingual Schools, no specific recommendations are offered regarding the acquisition of academic vocabulary or CALP in Secondary bilingual education. Finally, there is an additional teaching session available in the first and fourth grades of Compulsory Secondary Education, which can

be added optionally to any of the subjects taught through the medium of English except for *English as Foreign Language*. Again, this additional session is taught provided that the school has available resources.

#### 2.1. 2.3. “Bachillerato”

Students having completed Compulsory Secondary Education in a Bilingual High School can continue bilingual education in Baccalaureate (“Bachillerato”). Bilingual High Schools, prior consent from the educational authorities in the Comunidad de Madrid, offer the following options (which are not mutually exclusive):

- a) The teaching of *English as a Foreign Language* five days a week with a one-hour session each day.
- b) The teaching of the *Advanced English Curriculum* providing there are available certified teachers and a minimum of 15 students who can certify a B2 level (CEFR). Exceptionally, students coming from the Bilingual Program can be part of this group if the teaching team agrees their level is good enough to make the most of the teaching of this subject.
- c) The teaching of any subject in the curriculum providing there are available certified teachers except for those subjects established in the first provision of Royal Decree 52/2015, 21<sup>st</sup> May. These subjects will follow the official curriculum in the Madrid Region.

Although bilingual education has naturally extended to “Bachillerato”, it is quite significant that the entry exam to access Tertiary Education is still generic for all students in the region of Madrid and, as such, it is still conducted in Spanish. Maybe for this reason, and being conscious of the importance of this exam for their academic and professional lives, some students and families prefer not to continue studying through the medium of



English in “Bachillerato”. Instead, they choose the non-bilingual path where Spanish is used as the vehicular language (personal communication with students and teachers, academic year 2015-2016).

### 2.1.3. Resources and programs

The lack of specialized CLIL resources has always stood out as an important issue for both teachers and academics (Meyer, 2010; Banegas, 2012; Pavón & Rubio, 2010) in bilingual education in general, and more specifically, since the implementation of the program in the Madrid area (Fernández & Halbach, 2011). Concerns are usually raised on the one hand, about the difficulty of finding appropriate materials which are not mere translations of the traditional Spanish textbooks, and on the other hand, on content teachers' possible lack of the professional competences required for materials adaptation, supplementation and design (Coonan, 2007: 628). To facilitate teachers' access to CLIL materials contextualized and adapted to the Spanish context, when the program started, Madrid's educational authorities launched an easy collection of materials for Social and Natural Science for the fifth and sixth grades of primary education, which was later complemented with some audiovisual tools on the regional television. Today additional online resources covering other subjects in primary education are already available, and publishing houses are making an effort to find and adapt materials to suit the needs of bilingual schools. However, the challenge still lies in searching for appropriate resources for non-native speakers and some subjects such as Spanish History and Geography in secondary education. Other issues concerning resources are the need to standardize strategies to leverage CLIL materials, and the creation of platforms and (virtual) discussions to access and share them to name just a few. In this sense, and to facilitate new schools access to bilingual education, Madrid's Regional Government relies on the

Centers for Innovation (CETIF), the Regional Center for Innovation and Training CRIF Acacias and Educational Consultants in the different geographical areas. The main role of these centers and professionals is to work as networks for active communication among participant schools favoring mutual help and consultation between the main reference schools and those ones being welcomed into the program every year. An example of good practice by CRIF Acacias is the organization of a workshop to share educational experiences in bilingual schools, and the recognition of teachers' attendance and participation in some bilingual events. It is hoped that active communication among schools continues in the future so as to result in sharing ideas and materials during school visits and organized periodical meetings.

Other measures such as twinned schools and international programs are aimed to help create an international atmosphere, and promote the acquisition of foreign language skills in a wider context. In primary education, each school is assigned a twinned school in the United Kingdom. The twinned schools work together so as to exchange projects, experience and activities, correspondence and organize visits and students' and teachers' exchanges. Although they were conceived as essential for students to develop an international personality and be part of a bigger institution sharing a common language facilitating foreign language acquisition, the efforts were not always effective in practice (Senise, 2012). Other programs offered in order to broaden students' knowledge of the foreign language and foster respect and appreciation of other cultures and societies (Subdirección general de programas de innovación, 2014: 2) are the following: Global Classrooms, Global Cities, Theatre in English, International Public Speaking Competition, and IGGY (Connecting and Challenging the World's Brightest Young Minds). Additionally, under the auspicious of the *Comenius Regio* Program, the CAM Bilingual Project participates in the program for "Supporting a Cross Curricular

Approach to Second Language and Content Learning” to support research in Second Language Teaching and promote primary teacher’s collaboration between Madrid and Edinburgh. Lastly, in order to inform teachers and families about the CAM Bilingual Project, the educational authorities in the Comunidad de Madrid (CAM) regularly publish a basic guide including the main goals of the program, and a brief description of the AICLE or AICOLE approach<sup>4</sup> (Consejería de educación, juventud y deporte, 2015: Madrid, A Bilingual Community).

## 2.2. Teachers and their training

In promoting multilingualism, a bilingual school is an ecosystem which features a broad range of human, material and spatial resources. Regarding personal resources, the school requires special engagement from the different professionals in the institution, namely the principal, the bilingual coordinator, the teachers and the language assistants. They all have a relevant role in that they need to work collaboratively with each other so that language, content and skills can be effectively integrated into the curriculum.

The principal - a key figure in the school according to the educational law in force (LOMCE 2013) - is responsible for the bilingual project in the school although s/he can delegate the coordination on others such as the school secretary or the bilingual coordinator. Among her/his responsibilities are to direct and coordinate all the school activities, lead pedagogic instructions, promote educational innovation, and drive forward plans and actions aimed at fulfilling the objectives in the School-based Education Project (PEC).

---

<sup>4</sup> A name coined by educational authorities in the Madrid region.

The bilingual coordinator or bilingual program coordinator is, along with the principal and the rest of the school management team, one of the most important agents for the success of a bilingual program. Every school has a bilingual program coordinator, who is appointed by the principal. The bilingual program coordinator advises the principal and the rest of the management team in questions related to the elaboration of the Curricular Plan, the Yearly General Plan (PGA), and the Yearly Final Report, and s/he supervises the successful implementation of the academic program of CLIL subjects. In other communities, such as Andalusia, they also take responsibility for the elaboration of the PLC or “Language Project” which includes all the vehicular languages in the school along with the actions undertaken to promote them, and in some cases even a reading promotion plan. Other responsibilities include the following: to ensure that a weekly meeting is held, and that it welcomes all members in the bilingual program; get in touch with the language assistants assigned to the school, welcome them and guide them through their main functions and responsibilities; ensure there is coordination among teachers and the language assistants; hold a coordination meeting every week with all the members of the bilingual program; encourage critical reflection about teaching practices; promote new pedagogical and methodological initiatives and finally, serve as a link between the school and institutions, and with other bilingual schools.

Teachers’ profiles in Spain are dependent on the educational stage in question. In primary schools, the typical profile of a teacher is a generalist teacher who holds a degree in primary education or a language specialist who majors in English as a Foreign Language. Contrarily, and unlike the situation in other countries such as Italy where content and language teachers can share the responsibility of teaching CLIL subjects, secondary teachers in Spain are specialists in the subject(s) they teach, be it a content subject or a

language subject. Thus, CLIL subjects are the responsibility of the content experts whilst English is taught by a language specialist.

Regardless of the different profiles in bilingual schools, all teachers need to possess an advanced level of English; this effort is recognized by educational authorities who grant both primary and secondary teachers in Bilingual schools a financial bonus in accordance with the number of hours taught in the bilingual program. In Bilingual Primary Schools, teachers are expected to obtain the Linguistic Capability Certification or “habilitación lingüística” to teach subject areas in English. Those English Foreign Language specialists who have not received the required Linguistic Capability Certification are only allowed to teach English as a Foreign Language. This Language Certification along with the certificate to teach the “Advanced English Curriculum” in secondary education deserve special attention. The first allows content experts and foreign language teachers to teach subjects through a foreign language, a prerequisite in Madrid’s CAM Bilingual Project according to the recent Order 1317/2015, May 7<sup>th</sup>, and the latter is for those teachers of English as a Foreign Language who wish to teach the “Advanced English Curriculum”. These certificates measure teachers’ linguistic skills as part of the competences they need to teach content through a vehicular language. Throughout time, the requirements to access them have been adapted to suit the school’s needs. Therefore, during the initial stage of the program, and due to the urgent need to certify teachers to be part of the bilingual teaching staff, teachers were required to have at least a B2 level. To complement this intermediate language certification, some additional measures were later implemented such as immersion courses in British universities or yearly language courses for the whole teaching staff in bilingual schools (CIMNE courses for primary schools and CIPNE courses for secondary schools). Today, to access teaching in a bilingual school and although conditions might vary from one year to the next one, teachers need different

certification depending on their profile: content teachers need to pass an advanced language exam - a C1 according to the CEFR- along with an interview, and English teachers need to document the required level in English language, Culture and Literature by either presenting the corresponding certificates of advanced language skills, that is at C1 level, or a didactic unit which is defended publicly in front of a group of experts from Madrid's Regional Government. Apart from this advanced level of English, no methodological training is apparently required to access teaching in bilingual education in Madrid. Nevertheless, it is interesting to highlight that the oral interview for content teachers usually includes a number of questions on the CLIL approach and related methodologies, so it seems that some CLIL knowledge is asked from candidates apart from language skills (author's personal communication).

Apart from content and language teachers, language assistants also play a relevant role in the development of the project. They are primarily conceived as providing both language support for all the teaching staff and cultural insights to teachers and students while they benefit from being immersed in the culture of the host country. Although the role of the language assistant slightly differs depending on whether they work in a Primary or Secondary Bilingual School, their main functions in the CAM Bilingual Project are to help teachers to plan lessons and assist them providing language support in the class, reinforce students' oral skills, encourage and promote students' motivation and interest in the language assistant's language and culture. Along with the rest of professionals in a bilingual school, language assistants should work in collaboration with other teachers. However, since their function is never that of a teacher, they are not responsible for content teaching or grading exams. There is a recent interest in the role and efficient use of language assistants in bilingual education programs these issues being a priority action by researchers and educational authorities in Madrid. Consequently, examples of good

practice have started to be shared in professional communities and seminars such as the *II International Conference on Bilingual Education in a Globalized World*, 2016, the *Revista Digital EducaMadrid* (Buckingham, 2016) or the *I Seminar on Bilingual Education: Collaborative Strategies between Teachers and Language Assistants*, 2017.

Apart from the different professionals in bilingual schools, teacher training stands out as a primary concern among teachers in the CAM Bilingual Project. As was pointed out by Fernández & Halbach (2011), after the initial stage of the program implementation, teachers revealed the need to be supported through specific training in the following areas: teaching reading and writing skills and teaching Science in English, designing materials, and locating and having access to resources among others. Another common concern during the first years of implementation was the lack of specific linguistic training, which educational authorities tried to compensate by offering language courses to the whole bilingual community, and more specifically to content teachers. Back in 2004, the first teachers to take part in bilingual schools received training in language and some methodological advice. Training from 2009 onwards focused on foreign language skills as they avoided any methodological or subject-specific training (Olivares & Pena, 2013: 89). In fact, specific training in CLIL was not implemented in the CAM Bilingual Project by the regional government in collaboration with European institutions until 2012, and then it was aimed at a very limited number of participants. Today, as there are already an increasing number of certified teachers in language skills, training needs have evolved. Teachers currently demand more training courses aimed at their regular teaching practice, and the need to expand their methodological capabilities (Cabezuelo & Fernández, 2014: 61) since the challenge, they recognize, is teaching through the vehicle of English, and

not merely “teaching the subject in English”<sup>5</sup>. In-service training for teachers in the CAM Bilingual Project has been extended to cover both linguistic and non-linguistic demands and even specific training for school boards and education leadership skills. The 2016 Training Program in Foreign Language Teaching consisted of the following courses, which are held during the academic course and in the month of July: Leadership training for school boards, CLIL training, Language Capability Certification or “Habilitación”, Language upgrading courses, Classroom-management training, CLIL resources, IT, Working with Language Assistants, Training sessions, seminars and conferences (Consejería de educación, juventud y deporte, 2016). Still, teachers, and more specifically content secondary teachers, consider the training offered by the educational institutions as insufficient, especially in terms of class management and methodological aspects (Cabezuelo and Fernández, 2014: 62) as has also been observed in the teachers’ questionnaires in this research, and during informal interviews and conversations (2015-2016).

### 2.3. Challenges in the implementation of the CAM Bilingual Project

The implementation of bilingual programs goes beyond the inclusion of subjects taught through the medium of English, additional hours devoted to foreign language teaching, and welcoming language assistants into the bilingual school. In fact, teaching in a bilingual program implies certain changes regarding the methodologies being used, teachers’ needs when teaching both content and language subjects, and specific training for teachers. Finally, it is also necessary to focus on the use of assessment techniques favoring learning and “avoiding the dichotomy between national standard examinations

---

<sup>5</sup> In fact, this was a common misconception among teachers during the first stage of the project probably due to the lack of specific training on the CLIL approach.



and bilingual education curricula” (Johnstone 2012: 21). Although the CAM Bilingual Project was well received by the Administration Departments of the schools in which it was implemented (Laorden Gutierrez & Peñafiel Pedrosa, 2010: 326) and despite its rapid growth in the region, there are still needs which need to be met in the future. First, as far as methodologies are concerned, it is necessary to use compatible models in syllabus organization (Lorenzo, Trujillo & Vez, 2011: 163), so it is considered holistically, and pay particular attention to literacy and phonics at the primary level (Halbach, 2008: 458). In this sense, as was stressed in relation to teachers’ roles, content teachers at all educational stages need to consider the CLIL vehicular language as an integral part of the learning of non-linguistic subjects, revise their perceptions of the role of language and literacy in teaching and learning content subjects (Morton, 2016), and include second/foreign language pedagogies they might not be familiar with into their daily teaching practice. On the other hand, for the integration of contents and skills in practice, teachers will also need to adapt techniques for classroom instruction such as simplifying the textbook and other materials so that the language is less demanding for weaker students, allowing students to have some extra time before they answer, simplifying and adapting teachers’ speech, using strategies to activate background knowledge- brainstorming, K-W-L-H charts, etc. - and to clarify comprehension, and increasing student-teacher and student-student interactions so as to activate communication strategies in the foreign language (Reiss, 2005: 84). Other desired actions would be encouraging students to activate schemata through the use of bottom-up and top-down strategies in order to infer the meaning of new words and thus being able to work with challenging texts and vocabulary, and promoting discourse functions which will help students with cognitively demanding tasks (Halbach, 2012: 28-29). However, CLIL subject pedagogies are not the only ones which need to accept the methodological

challenge. In fact, along with extended exposure to foreign language, teachers of English should also consider some methodological changes in terms of the inclusion of functional language in the curriculum (Lorenzo, Trujillo & Vez, 2011: 135). This would help to avoid a closed list of language contents, foster the selection of authentic materials and tools, and prioritize the use of oral and written activities which encourage real-life communicative tasks. Furthermore, regarding the consideration of the curriculum in holistic terms, more efforts should be made to favor coordination with L1 teaching and to improve and adapt human, material and spatial resources with a special emphasis on the need to free the teachers taking part in the bilingual program from some of their teaching hours. This is thought to contribute to teachers' work so that they can efficiently prepare and adapt materials in the absence of them, and have more time for coordination among teachers (Laorden Gutiérrez & Peñafiel Pedrosa, 2010: 333) as was pointed out in the chapter devoted to collaboration among teachers in CLIL contexts. Finally, assessment practices need to accommodate the dual focus of CLIL teaching on language and content. Assessment issues will be discussed in depth in subsequent chapters as assessment constitutes the main focus of this research.

#### 2.4. External evaluation

External evaluation refers to the process of collecting, analyzing and interpreting information in order to make informed decisions about students' development, the success of educational programs (Baehr, 2010: 441), and the subsequent interpretation of that evidence concerning stated or desired objectives in the curriculum. In this context, Madrid's Regional Government conducts yearly external evaluations. Unlike other bilingual programs such as the evaluation in The Netherlands, or the evaluation conducted under the auspice of the British Council or the MET Evaluation in The Basque Country,

Madrid has focused so far on evaluating students' language proficiency, not content knowledge or skills. The tests measuring language skills in English are conducted in both primary and secondary education. During the first stages of the program implementation (academic year 2008-2009), the evaluation was carried out in the fourth grade of primary education with the help of external evaluators, namely the Trinity College along with a group of experts from the Official Schools of Languages in the CAM. Today, Cambridge Language Assessment is in charge of measuring students' language skills in the second, fourth and sixth grades of primary education, and in the second and fourth grade of ESO. In the second and fourth grade of primary education, the evaluation focuses on oral skills while in the sixth grade of primary education, students are tested on the four skills-reading, writing, listening and speaking- as a prerequisite for those who wish to pursue education in a Bilingual High School. As for secondary education, students also have to be tested in the four language skills. Results from the academic years 2013-2014 and 2014-2015 reveal that the success rate in the exams in the Bilingual Sections lies around 75% as compared to those students in the Bilingual Program in which the rate of success is limited to 60%. Furthermore, according to recent research comparing students' level in bilingual and non-bilingual schools, the latter seem to be at a significant disadvantage to the first (Shepherd & Ainsworth, 2017). Although no priority actions have been taken as a result of external language assessment so far, it is clear that the Comunidad de Madrid is determined to adopt a position on multilingual policies through, for example, the implementation of School Language Projects in the near future (Jornadas "Elaboración del Proyecto plurilingüe de centro", 2016).

Language evaluation in the CAM Bilingual Project is also combined with external content-knowledge evaluation in the Madrid region in both primary and secondary education. External evaluation measuring content knowledge takes place at a national

level in the third and sixth grades of primary education (LOMCE Exams) and at an international level (International PIRLS) in the fourth grade of primary education. Thus, apart from students' linguistic proficiency, other indicators from external evaluation in mainstream education - PISA 2009 and PISA for schools 2013- are used to analyze students' results, and to compare them with results from students in non-bilingual programs (Ruiz, 2016; Sotoca & Muñoz, 2015; Tamariz & Blasi, 2016; Montalbán, forthcoming). This was the case, for instance, of the CDI tests - a test conducted until the academic year 2016-2017 in the fourth and sixth year of primary education measuring basic contents and skills in reading and writing in *Spanish, Mathematics, and Science and Technology*. These external tests are common to both bilingual and non-bilingual schools in the Madrid area. In general, as can be deduced from the first stage of analysis of the CAM Bilingual Project by Madrid Autonomous Community (2016), it seems that the lower results in Bilingual Primary Schools (Brindusa, Cabrales & Carro, 2016; Ruiz, 2016) in the last year of primary education in Science as compared to non-bilingual schools are compensated in the secondary stage. In fact, in the long term, bilingual education seems to have a positive influence as it does not reduce the level of content knowledge and skills in any subject, it improves significantly the level in the foreign language (Tamariz, Blasi & Planck, forthcoming) and improves other educational aspects such as the promotion of reading habits (Montalbán, forthcoming) without lowering students' results in the university entry examination (PAU)<sup>6</sup> (Comunidad de Madrid, 2016). To give well-rounded soundness to the project, and to avoid common misconceptions and criticism about bilingual programs compromising students' content knowledge and skills in favor of language proficiency, it is hoped that more research on linguistic and non-linguistic aspects is conducted in the future to analyze results obtained,

---

<sup>6</sup> The university entrance exam was renamed EvAU in the academic year 2016-2017.

for instance, the *Trends in International Mathematics and Science Study* (TIMSS 2015), PISA 2015 and the *Progress in International Reading Literacy Study* (PIRLS 2016).



## CHAPTER 3: Assessment

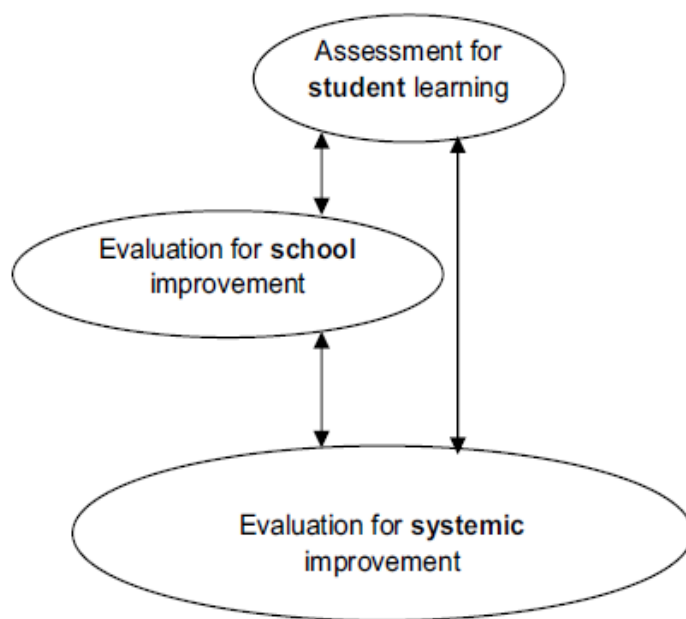
The main aim of this chapter is twofold. First, to define assessment, its scope and main types, and to reflect on the impact it might have on educational practice. Besides, to provide a theoretical framework for the research, current assessment regulations in Spain will be described in detail. Second, the chapter is aimed at dealing with the main issues in CLIL assessment: Its nature and dual focus, how to assess language and skills, the importance of Formative Assessment, quality standards and CLIL assessment tools.

### 3.1. What is assessment?

As Baehr (2010) points out, educators use two distinct processes to give students feedback about their performance, and help them build lifelong learning skills: assessment and evaluation. However, in the assessment literature the terms assessment and evaluation are commonly used interchangeably. This misconception could be due, on the one hand, to the fact that the term evaluation, is used to describe both processes in some languages such as Spanish, and on the other hand, because depending on the source of reference, they can be treated as synonyms or different concepts (Scanlan, 2012). Throughout this study, assessment is used to refer to internal evaluation or classroom assessment of learners. As Rea-Dickins puts it, it is the students' evidence or the gathering of data from planned learning activities, usually made by teachers in order to make informed decisions about knowledge and skills or the process of monitoring or keeping track of students' progress in order to gather evidence of their learning (Rea-Dickins, 2000: 376). Evaluation, on the contrary, refers to the measurement of instruction by determining the level of quality of a performance or outcome and enabling decision-making based on the degree of quality demonstrated. In this sense, evaluation goes beyond this process of gathering information on student learning as it implies the analysis, reflection upon and

summary of assessment data along with the making of judgments and decisions usually based on multiple sources of information. Evaluation is not conducted often by teachers but by external institutions with the purpose of judging the quality of a program or parts of the program, and it coexists with assessment in most educational systems. Therefore, the information gathered at both the internal and external level is thought to help identify strengths and weaknesses, and to design strategies for future improvements (OECD, 2008: 6). The relation between assessment and evaluation is portrayed in the graph below:

Graph 5: Coordinating assessment and evaluation



(From OECD, 2008: 5)

Another common misconception in defining assessment refers to the tendency to describe it as formative in nature (Harris & Brown, 2009). Assessment is generally depicted as a process to guide instruction and drive students' learning, which is process-oriented, on-



going and flexible (Apple & Krumsieg, 1998) as opposed to evaluation, which is considered judgmental or summative (Kizlik, 2017). However, the term assessment should not be restricted to formative measurement but be open to a wide range of interpretations as will be explained in the section devoted to types of assessment. In this research, assessment is a broad term which compiles a range of procedures such as measurement, testing, and grading but is not exclusive to them (Lynch, 2001: 358). Thus, assessment implies “anything a teacher does to gather information about a student’s knowledge or skill regarding a specific topic” (Marzano, 2010: 22). A test on the other side, is one of the methods to gather evidence along with observation, interviews, and checklists to name just a few, while grading is described as “a formal, summative, final product-oriented judgement of overall quality” (Scanlan, 2012). Finally, it is important to notice that in current educational systems assessment is usually conducted to know whether or not learning has taken place, and plan strategies to guide the learning process.

### 3.2.Types of assessment

There are many types of assessment. The following table shows the parameters taken into account to classify assessment types regarding their function, scope, the referent used to make comparisons, the time in which it is set, the tools and context, the focus or goal of the assessment and the agents taking part in the assessment process.

Table 6. Types of Assessment

Function	<ul style="list-style-type: none"> <li>•Diagnostic (starting point)</li> <li>•Summative (record learning achievements)</li> <li>•Formative (adjust and regulate teaching practice)</li> </ul>
Scope	<ul style="list-style-type: none"> <li>•Global (the whole subject)</li> <li>•Partial (parts of the subject)</li> </ul>
Referent	<ul style="list-style-type: none"> <li>•Norm-based (a normative group)</li> <li>•Criterion-based (assessment criteria)</li> </ul>
Time	<ul style="list-style-type: none"> <li>•Initial (beginning of academic year or course)</li> <li>•Continuous (along the academic year or course)</li> <li>•Final (end of academic year or course)</li> </ul>
Tools and context	<ul style="list-style-type: none"> <li>•Formal (at specific times using tests and other tools)</li> <li>•Informal (class observation, interviews, etc.)</li> </ul>
Goal	<ul style="list-style-type: none"> <li>•Product (final outcome)</li> <li>•Process (what is completed through time)</li> </ul>
Agent (s)	<ul style="list-style-type: none"> <li>•Assessment conducted by others (the teacher(s))</li> <li>•Self-assessment (the students)</li> <li>•Peer-assessment (the students)</li> </ul>

(Adapted from Casanova, 1998: 13)

Considering the function of assessment, we distinguish between diagnostic, summative and formative assessment. Diagnostic assessment is conducted at the beginning of the academic year or course, and its primary aim is to determine students' levels before instruction. Consequently, it provides a baseline against which we can assess progress over time, and help teachers to adapt their teaching practice accordingly. Summative Assessment is usually considered to be opposed to Formative Assessment. However, both types of assessment can be used in combination with each other as when teachers use Summative Assessment to inform students about their strengths and weaknesses, and to change instruction to adapt it to learners' pace and needs (Harlen, 2012). Due to the

relevance it has for innovative teaching practice in general and CLIL in particular, Formative Assessment is analyzed in depth below.

Attempts to change current assessment practice have become widespread in recent years. Studies and volumes such as the recent *La Constante Macabre* (2009) and the *Diploma Disease* (2007) criticizing the Exam Culture, Howard Gardner's Multiple Intelligences Movement, Daniel Goleman's Emotional Intelligence and the influential papers by the Assessment Reform Group (2002, 2012) have stressed the fact that assessment practices might be outdated, and exam results and standardized tests might be artificial as they fail to reflect students' skills and abilities. In fact, the paradigm shift in the twenty-first-century advocates for new visions of education resulting from the advent of Post-Positivism, a reformed vision of the curriculum, and cognitive and constructivist learning theories. All these changes have naturally been reflected in assessment issues as in the focus on contextualization, integration, the specific, understanding, diversity and the process rather than the result among other aspects (Jacobs and Farrell, 2001: 2), and the consideration of classroom assessment as central to teaching practice (Shepard, 2000: 8). In this context of rapid change, Assessment for Learning (AfL) or Formative Assessment, appeared in response to the need to use evaluation to detect learning gaps and compensate for them (Black & William, 1998), align learning and assessment, integrate the assessment practice with the curriculum goals, and build up a constructive relationship between teaching and learning processes (McNamara, 2000). As has been broadly defined by the Assessment Reform Group, AfL is "the process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go to, and how best to get there" (Assessment Reform Group, 2002: 2). Thus, what distinguishes AfL from other forms of assessment, and specifically from summative assessment, is the focus on the nature of the evidence collected, and how

it is used by both teachers and students, on the one hand to inform teaching and learning practice rather than necessarily associate it with evaluation purposes, and also to identify learning needs in time, and act accordingly.

AfL is aimed at promoting learning rather than evaluation for several reasons. To start with, assessment is embedded in the process of teaching and learning and as such, it is considered as an essential part of the process. Second, it is mainly aimed at helping students know and recognize the standards they are aiming for and for that purpose, it involves sharing goals with students, engaging and involving them in the process of self-assessment, and providing feedback leading to students recognizing their next steps and how to take them. Finally, Formative Assessment is underpinned by confidence that every student can improve, and it involves both teachers and students reviewing and reflecting on assessment data (Assessment Reform Group, 2002: 13).

The features mentioned above link Formative Assessment with innovative teaching practice for a variety of reasons: First, because it places the focus on the students' performance while traditional assessment focuses on the teachers' view of students' performance. Second, because learning contents are regarded in a holistic way and as such, they are integrated or contextualized. Third, because assessment is no longer considered as an end in itself but as part of the teaching and learning process in which students are welcome to participate. The table below shows the differences between traditional assessment and Formative Assessment by looking at the tools used for the data gathering, the type of judgment, the focus of assessment, its aim(s) and the agents involved in the process of assessment:

Table 7: Traditional versus Formative Assessment

Traditional Assessment	Formative Assessment
Memorization	Contextualization
Tools: Multiple choice, matching, true or false, essay questions, etc.	Tools: Project-based work, portfolios, debates, discussions, class quizzes, etc.
Judgment is based on objective data gathered and interpretation of its grading	Judgment is based on observation, subjectivity and interpretation of facts
Focused on the grading of the student as compared to other students	Focused on the grading of the student in relation to her/his own learning process and stage of development
The students' knowledge and skills are presented by means of a score	The students' knowledge and skills are presented by means of timeline assessment in terms of progression
Assessment is an end in itself	Assessment is used in order to inform future instruction
Emphasis on the product	Emphasis on the process
The teacher and/or other external forces are the agents involved in the assessment process	The students can participate as active agents in the assessment process

(Adapted from Mateo, 2000: 7)

Nevertheless, the distinction between Traditional and Formative Assessment is not always clear-cut, and some researchers (Harlen, 2012, Maxwell, 2004) prefer the concept of “progressive assessment” or “good Formative Assessment”, the latter being defined as a combination of “good judgement by teachers about students’ progress and levels of attainment and Summative Assessment which provides feedback which can be used to

help (and inform) learning” (Harlen, 2012: 115). In any case, the formative-summative tension (OECD, 2008) can be solved by using summative tests in a formative way and thus, prevent the “teaching to the test” (Rea-Dickins, 2001) frequent in some educational environments.

Regarding its scope, assessment can focus on the whole subject - global assessment - or on the contrary, on parts of it - partial assessment - depending on the aspects highlighted or considered most relevant, or the time at which the assessment is conducted. Global assessment has traditionally been associated with summative evaluation at the end of the academic year or the entry exams to access higher education, while partial assessment is usually done at the end of each term. The latter is usually found in combination with continuous assessment, described below.

As for the point in time in which evidence is gathered, initial assessment refers to the gathering of data at the beginning of the academic year or course and thus, it is commonly associated with diagnostic assessment. Continuous assessment aims to gather evidence periodically, generally at the end of the term or after each unit or number of units, and it is one of the most frequent forms of gathering evidence from students in today’s educational systems. Final assessment, on the contrary, is conducted at the end of the academic year or course regardless it covers the whole subject- global assessment- or parts of it.

Another distinction relates to the context and tools used for the assessment and in this sense, we distinguish between formal and informal assessment. Formal assessment usually requires systematic collection of data to measure students’ progress through tests and other assessment tools. Contrarily, informal assessment evidence tends to be gathered in the context of the classroom, usually through class-observation techniques and other data typical from class activities and assignments. Informal assessment is frequent among

teachers at all levels and in fact, many decisions on students' development are taken on the basis of routine class activities, which allow teachers to determine the next stage in the process, and to what extent students are ready to move forward (Rea-Dickins, 2001: 434).

Norm-based assessment (or norm-referenced assessment) consists of determining the relative position of a student (her/his knowledge and skills) in relation to the level of a group or class, which is considered the norm (Heredia Manrique, 2009). The assessment is based on some norms or reference scales measuring students' performance through standardized tests. The results of these tests are expressed in grades or averages which allow the comparison among students (Popham, 1983). Criterion-based assessment, on the other hand, was suggested as an alternative to norm-based assessment by authors such as Popham (1983, 1995). This type of assessment is based on the close link between basic competences, learning goals and assessment criteria. The process is as follows: First, basic or related competences are selected from the subject curriculum- for example, reading comprehension in Biology. Second, specific learning goals are selected in relation to these competences and set to be accomplished by students at the end of an academic period. Finally, those objectives serve as the basis to establish clear assessment criteria and achievement indicators. Criterion-based assessment is thus based on clear assessment criteria expressed in terms of competences or can-do descriptors which are observable in the student and/or the group. These assessment criteria are thus external, explicit, specific and transparent (Fullana & Pallisera, 1995; Gómez Arbeo, 1990), and can be adapted to each student's mode of learning disregarding the comparison with the rest of the members of the group.

Regarding the aim of assessment, teachers can concentrate on the product or the process. Product-based assessment aims to judge the final product the student creates without

taking into account the process or the intermediate steps the student has gone through in the acquisition of knowledge and skills. This type of assessment tends to be associated with traditional and summative forms of assessment. On the contrary, process-based assessment considers the process of acquiring skills and content knowledge as part of the evidence. Therefore, in process-based assessment errors are regarded as natural consequences of the learning process and any output as an approximation which is subject to change and improvement. Consequently, process-based assessment is very much in line with Formative Assessment as it focuses on the individual, and gives feedback for both the teacher and the student, this feedback being aimed at improving the teaching/learning process.

If we consider the agents involved in assessment, we have the following main categories: Teacher-assessment, self-assessment and peer-assessment. Teacher-assessment or assessment by other professionals is the most traditional form of assessment in which a person other than the student - usually the teacher(s) - collect(s) evidence through exams or other tools. Self-assessment is assessment conducted by the student of her/his work. It is considered as good practice by advocators of Formative Assessment since it improves motivation, allows the student to reflect on and take an active role in the learning process, and identify her/his strengths and weaknesses (Lorenzo, Trujillo & Vez, 2011: 277). Peer-assessment is the process of mutual assessment between peers. As in the case of self-assessment, it allows the student(s) to reflect on and take responsibility for their learning process by examining critically and appreciating other students' work in a constructive way. Although self and peer-assessment assessment have been recently welcome in many educational contexts, some teachers still do not consider them efficient tools for assessment probably due to the students' difficulties to adopt a critical view and value their own work objectively. Thus, for both self-assessment and peer-assessment to be



effective, some prior work on the part of the teacher is necessary in order to help students in the assessment process, and to establish clear criteria for good work.

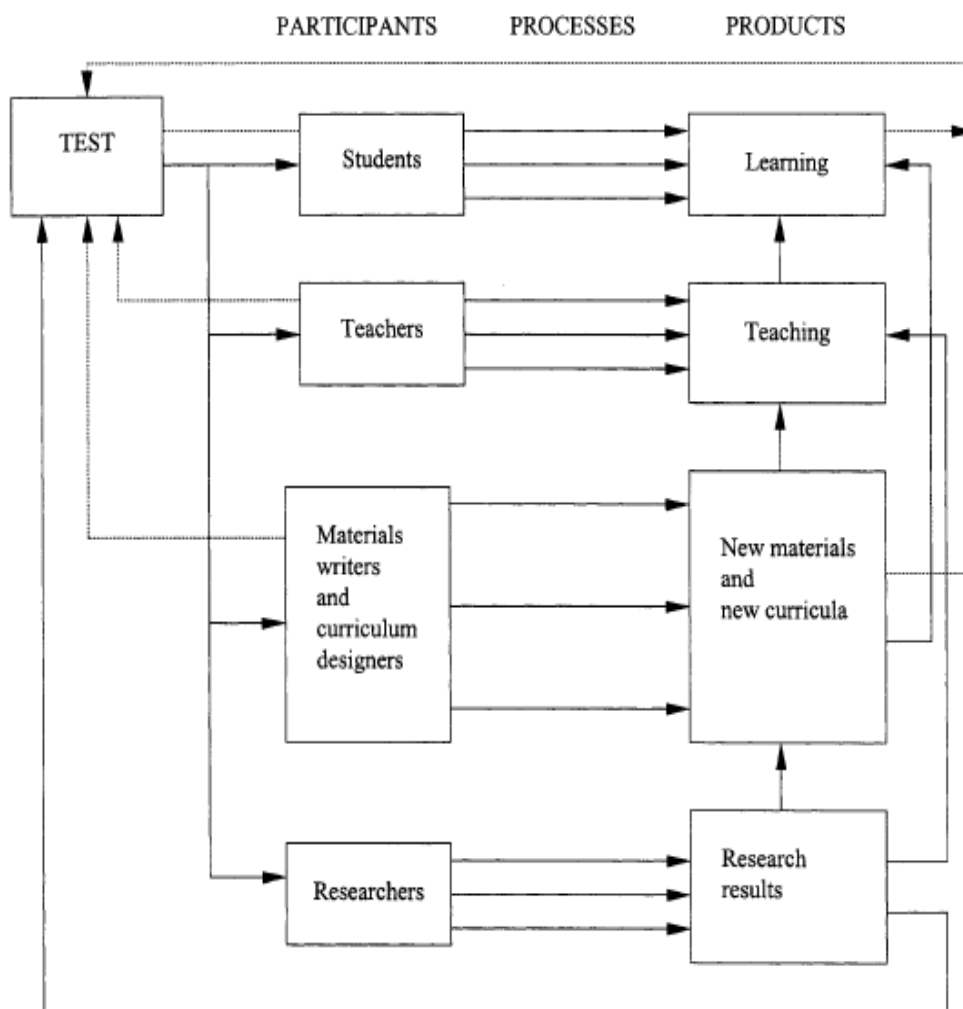
### 3.3. The impact of assessment in education: Washback effect

This section aims to explore the impact that assessment may have on students' performance both in positive and negative ways by regarding the extent to which it can favor or impede the achievement of educational goals (Bailey, 1996). Assessment is never neutral, but it can affect and shape what and how we learn as well as it also has a direct impact on the curriculum (Gardner, Harlen, Hayward & Stobart, 2008: 13). This impact is known as washback or backwash effect (Hughes, 1989). According to Alderson and Wall (1993: 120-121) tests having important consequences usually influence teaching and learning regarding the following: What and how teachers teach, what and how students learn, the rate and sequence of teaching and learning, the degree and depth of teaching and learning, and attitudes referring to content, methodologies, etc. of teaching and learning. For example, traditional instruction focusing on content knowledge transmission and thus, undervaluing student interactions and promoting a testing culture which favors scores over feedback is thought to have a negative impact on teaching and learning (Harlen, 2012: 67). Conversely, assessment which does not have important consequences will have no washback at all.

There is extensive research about the impact and effects of standardized testing on students' outcomes and motivation among other factors (Alderson & Wall, 1993; Alderson & Wall, 1996; Cheng, 1997; Bailey, 1996; Chapelle & Douglas, 1993; Cheng, Watanabe & Curtis, 2004; Shohamy, 1993). However, more research is still needed into the topic to ascertain the ways assessment issues in general, and not solely testing in

particular, might be influencing the teaching and learning process, and how to prevent the sometimes, inevitable process of narrowing, simplifying or adapting the curriculum or “teaching to the test” (Rea-Dickins, 2001). Although in the assessment literature there has been a tradition to refer explicitly to the power or influence of tests on the teaching/learning process, I will adopt a broader scope and thus refer to assessment in general since in assessment washback, the participants, the process(es) and the product(s) are closely linked and relate to each other in a variety of ways (Hughes; 1993; Bailey, 1996). As can be seen in the graph, the nature of assessment can affect the participants’ perception and attitudes towards their teaching and learning tasks. In turn, these perceptions and attitudes can also affect participants’ actions in the process of carrying out their work, which includes the practice of those items which are likely to be assessed, thus affecting their outcomes and the product of that work (Hughes, 1993: 2).

Graph 6. Bailey's washback model.



(From Bailey, 1996: 264)

### 3.4. Assessment regulations in Spain

The following section aims to analyze the assessment regulations in Spain by looking at the main types of assessment. Since the data in this research were gathered between 2014 and 2016, two educational laws - the LOE (2006) and the LOMCE (2015) - need to be taken into account. Thus, Spanish regulations regarding assessment refer basically to the LOE, but some insights on the LOMCE are also provided, especially those relating to the

adaptation of competences to fit the assessment criteria, and the way they are suited to the new “assessable learning standards” introduced in the LOMCE. These standards are, therefore, included so that they can be compared with traditional can-do descriptors used in CLIL assessment. Attention will be also paid to the concept of competence and its impact on current educational practice - from the focus on teaching to the focus on learning - and more specifically, how they are related to CLIL assessment.

Over the past 40 years, we have witnessed several educational changes regulated under specific laws which have had a big impact on the ways and uses of assessment. The advent of the “General Law of Education” (LGE) in 1970 was a significant step forward in evaluation as it was the first mention of assessment in a Spanish educational law; the LGE introduced the need to implement continuous and guiding assessment practice, and passed the responsibility of assessment to the teachers. However, as this law focused on objectives at the end of the course or educational stage, it tended to be summative or focused on the final product, and thus, it emphasized final results rather than process assessment. With the passing of the “Law for the General Organization of the Educational System” (LOGSE, 1990), education changed sharply in all aspects, including assessment issues. To start with, the formative character of assessment was stressed and the agents in the evaluation were no longer only teachers but students were also given the possibility to assess themselves and their peers; second, contents were widened taking into consideration concepts, procedures and attitudes and not just theoretical knowledge; third, norm-based assessment was substituted by criterion-based assessment, and finally, the tools and strategies to collect students’ evidence were widened as the exam was no longer considered the only valid tool for assessment.

### 3.4.1. Assessment in the LOE

The LOE (2006) considers external and internal evaluation as central in teaching. In fact, one of the main innovations in this law is the inclusion of the evaluation of the whole educational system including curriculum planning, organization, teaching and learning processes as well as students' results. Another innovation regarding the law's basic principles is the recognition of the students' capacity to regulate their own learning, the trust in their abilities and knowledge, and their potential to develop creativity, personal initiative and enterprising spirit<sup>7</sup> (LOE, 2006: article 2), all of these being factors recognizing the students' commitment to lifelong learning. Assessment in secondary education is described as continuous (LOE, 2006), global and systematic (Madrid Royal Decree 89/2014), and it varies depending on the subject in question, so teachers are free to choose the assessment tools which are more appropriate for their discipline(s). The continuous assessment process is carried out regularly during the academic year and the whole learning process, and it is hoped to enable teachers to detect difficulties and successes in students' academic lives when they occur. The formative character of assessment is also considered in the law although no specific recommendations are given in this regard. As in the previous law, the LOGSE, assessment must conform to the division of the subject contents into concepts, procedures and attitudes, and to assess the learning process and not just final results, teachers need to take decisions regarding the situations, strategies and instruments which are more appropriate for the assessment of each subject. The balance between the three types of contents (concepts, procedures and attitudes) varies widely depending on the subject. Thus, in Social Science, a big percentage of the final degree tends to rely on conceptual contents while in other disciplines such as Arts and Crafts and Physical Education there is a larger proportion of

---

<sup>7</sup> Author's translation.

the final grade which corresponds to practical activities. Finally, a minimum of three periodical assessments corresponding to the three class periods are conducted in the academic year, and parents or tutors are informed about the students' progress accordingly. As has been pointed out on several occasions throughout this research, assessment regulations depend on different Autonomous Communities, but no additional specifications were made in the LOE concerning bilingual education in Spain in general or the Madrid area in particular. Therefore, CLIL subjects follow the assessment criteria in force for subjects taught through an official or co-official language. As for assessment tools, the LOE states that they need to fulfill the following conditions: Be varied, in order to evaluate different abilities and curricular contents; use different codes (verbal, numeric, audiovisual, graphic, etc.); be applicable to more or less structured situations of the learning activity in question; allow the assessment of the functionality of learning; assess the transference of learning to different contexts so as to measure competences and abilities; and offer concrete information for the subject and activity in question. Likewise, particular attention has to be given to informal observation such as that of students' class work and the analysis of class notebooks - a compilation of students' work and practical activities organized by units which teachers supervise on a regular basis, usually at the end of the term. These class notebooks are in everyday use in the Spanish educational system especially since the implementation of the LOGSE's procedural or practical contents, which were to be translated to competences later.

### 3.4.2. Assessing competences

As was pointed out in the previous section, performance-based and innovative/alternative assessment<sup>8</sup> are thought to allow educators to focus on the individual or group ability to demonstrate content knowledge through practical skills along with the students' attitude when applying those capacities and skills to authentic and meaningful tasks. This change from the emphasis on teaching to the emphasis on learning has been made possible due to a new conception of learners as active participants who are responsible for their own learning process (Vygotsky, 1978, 1986), a change that was already made visible in the LOE, thus conforming to European guidelines on lifelong learning. Although curricular contents are still relevant, the assessment of competences has gained ground in educational systems (Meyer, 1992; Hargreaves, 1994; Eurydice, 2012). As a result, learners need to become conscious, competent and critical citizens who do not only possess adequate knowledge about different topics and aspects of today's society but who are also able to use their cognitive-linguistic, communicative and social skills effectively and successfully (Eurydice, 2002).

Key competences date back to the Delors Report (UNESCO, 1996) which established the four main pillars of life-long learning in the 21st century- "learning to know", "learning to do", "learning to be" and "learning to live together"- and which were later stressed in the DeSeCo Project (OECD, 1999). According to the Organization for Economic Cooperation and Development (OECD), from the implementation of PISA, success in a students' life depends on the acquisition of a wide range of competences. A competence

---

<sup>8</sup> Although some authors consider alternative assessment as similar to authentic assessment and both are usually associated with "performance-based assessment" and assessment through competences (Martínez López & Cantero García, 2014), it is necessary to clarify that the term "authentic" refers to the practices based on real-life tasks and activities whilst alternative only refers to those practices other than traditional exams (Meyer, 1992; Álvarez Valdivia, 2005).

is defined as the ability to integrate knowledge, skills and attitudes in a practical way to solve problems and react appropriately in a variety of contexts and situations, i.e. the ability to integrate theoretical and practical knowledge in real-life situations outside the academic contexts (European Parliament and Council of Europe, 2006). Competences in education are defined more precisely as “the group of skills, knowledge and attitudes according to the context that every student of each educational stage should reach for their fulfillment and development, as well as for their active role in society, their social integration and employment”<sup>9</sup> (Ortega Martín, 2012: 71).

In Spain, competences were already mentioned in the “Organic Law for Quality Education” (LOCE, 2002), a law which was never implemented, but they were not introduced until the arrival of the “Organic Law for Education”, LOE (2006) with the name of “Basic Competences”. Competences are described in the curriculum of this law as the body of knowledge, skills and attitudes that students should reach at the end of Compulsory Education to achieve personal development and fulfillment, exercise their citizenship, join adult life in full and be able to go in lifelong learning. These competences were made extensive to all academic subjects along with some other innovations such as the minimum learning objectives, and they were later introduced in the current academic law for the Improvement of Quality in Education (LOMCE, 2015). The seven basic cross-curricular competences are to be developed by all pupils, and need to be understood in the context of education as a continuous lifelong learning process, within the capabilities of the majority of students, common to many areas of everyday life, and essential for personal, social and professional development (LOE, 2006: Annex 1). Although some regions adapted the list of competences to introduce specific ones in their curricula, Basic Competences according to Madrid’s Regional Government are the same as in the LOE:

---

<sup>9</sup> Author’s translation.



Competence in Linguistic Communication, Mathematical, Scientific and Technological Competence, Competence in Processing Information and the use of ICT, Competence in Social Skills and Citizenship, Cultural and Artistic Competence, Learning How to Learn, and Autonomy and Personal Initiative.

One of the most important competences students need to acquire and develop is the competence in linguistic communication or communicative language competence, which is inextricably linked with all the subjects in the curriculum and which, in the case of CLIL, deserves a closer look. In Spain, the original European key competences of “Communication in the mother tongue” and “Communication in a Foreign Language” were blended into “Competence in Linguistic Communication”, which involves communication in the mother tongue (the official and co-official language, where it exists) and in at least one foreign language. Competence in Linguistic Communication is described as the use of language as a means of oral and written communication and also as a learning tool which enables the user to self-regulate her/his thinking, emotions and behavior. This competence is considered to contribute to the development of a positive self-image and helps forge a constructive relationship with other individuals and the environment (LOE, 2006: Annex 1). Furthermore, communication is also paramount in establishing social links and approaching and understanding other cultures in order to coexist peacefully, and as the vehicle of communication in non-linguistic subjects, as will be analyzed in depth in the section devoted to CLIL assessment.

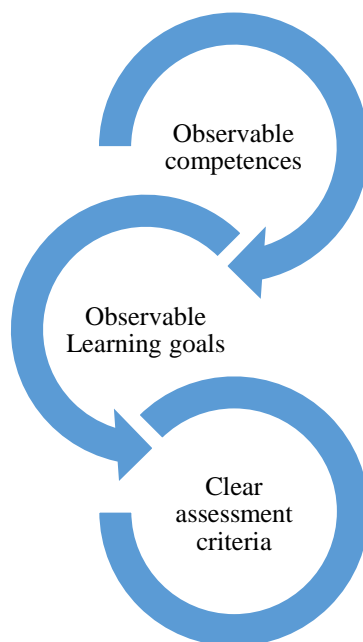
Finally, it is also worth mentioning that the introduction of competences in any educational system presupposes methodological changes: On the one hand, in relation with the way learning processes are conceived- from knowing concepts to being competent in using them- which require a more practical approach to learning. On the other hand, in relation with the assessment tools in line with this practical approach such

as performance-based tasks, which allow to assess content and language skills in an integrated way.

### 3.4.3. Assessment in the LOMCE

The implementation of the LOMCE in 2013 brought a number of changes to the Spanish curriculum: To start with, assessment criteria are now established by the central government and complemented by the Autonomous Communities. Also, for the first time in the history of Spanish educational laws, the concept of Formative Assessment is explained in detail. Assessment is now considered to have a “formative character” which, in combination with continuous assessment, the previous measures (LOGSE, 1990, LOE, 2006) for reinforcement and remedial programs, and the Advisory Education Council at the end of each course in Compulsory Secondary Education (ESO), contributes to educational improvement. Besides, a more integrated and integrative approach to the curriculum in general and subjects, in particular, is implemented in mainstream education. On the one hand, the concept of competence stays but from a holistic and integrated viewpoint since in the LOE, competences are considered as included in subjects while in the LOMCE, they are considered as pertaining to different subjects in a cross-curricular mode. On the other hand, objectives or learning goals are also measured from an integrated conception according to educational stages and not individual subjects. These learning goals are in close relation with the “assessable learning standards”, which will be described in the next section, and help to implement the criterion-based assessment lacking in precedent educational laws, as depicted in the graph below:

Graph 7. Criterion-based assessment in Spain



Criterion-based assessment is based on clear assessment criteria expressed regarding competences or can-do descriptors, which are observable in the student and the group. These assessment criteria are thus, external, explicit, concrete and transparent (Fullana & Palisera, 1995; Gómez Arbeo, 1990), and they focus on each student disregarding the comparison with the rest of the members of the group.

Furthermore, the integrated approach is also reflected in a new conception of the role of teachers i.e. they must be able to integrate the basic competences with the rest of the elements in the curriculum, and they must integrate their teaching competences- curriculum planning, classroom management and assessment- in their teaching practice. Finally, following recommendations by the OECD about the aspects which are subject to improvement in the Spanish educational system, standardized exams are conducted at the primary and secondary level - a measure which has been recently amended. This external

evaluation is based on core competences, and it does not affect students' academic register in the case of primary education. Contrarily, in secondary education, it includes competence-based tests and greatly influences students' academic life.

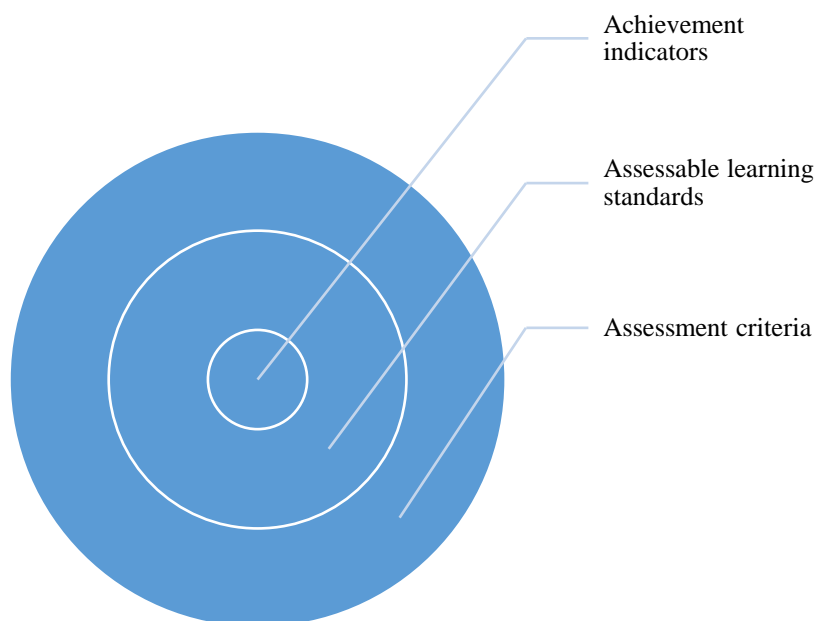
#### 3.4.4 Assessable Learning Standards

Traditionally, educational systems have been considered as non-transparent, mainly because educational criteria are not established, and schools tend to operate in an isolated and closed manner from the rest of society (Schleicher & Zoido Lobatón, 2013: 46; Martínez Rizo, 2013: 67). In response to this lack of transparency, assessable learning standards arise in the current law as indicators of objectivity, certainty and comparability of results. According to the LOMCE (2013), a learning standard is a specification of assessment criteria which allows defining learning outcomes, and which helps to specify what the student should know, understand, and be able to do in each subject. Conceived to be subject to evaluation, operationalized, measurable, public, clear, precise and rigorous (LOMCE, 2013), they are meant to contribute and facilitate the process of assessment along with the design of comparable and standardized tests and thus, make assessment practices as transparent as possible.

#### 3.4.5. Achievement indicators

Apart from the assessment criteria, which are dependent on the subject in question, and the assessable learning standards, which enable the assessment of the achievement of objectives in each subject, the LOMCE (2013) includes the concept of “achievement Indicators”. As shown in the graph below, they are considered as intrinsic to assessable learning standards and assessment criteria:

Graph 8. Assessment in the LOMCE



Since competences are considered to be rather generic, to specify and express them in detail so that they can be used as reference for educational action, indicators show students' degree of achievement as referring to competences in a clear way. These performance indicators also serve as a guidance to identify students' learning outcomes, and as such, they will prove useful in the creation of rubrics as will be commented in the chapter devoted to CLIL assessment tools.

As for the last changes implemented in the field of assessment, following the LOMCE recommendations to base assessment on the competences acquired by students and taking into account the formative character of assessment we pointed out before, assessment activities and tools will conform to the following guidelines. To start with, priority is given to problem-solving which imitates those activities in real-life contexts, and thus mobilizing students' knowledge, skills and attitude. Second, strategies allowing the

students' participation in the assessment of their achievements such as self-assessment and peer-assessment are also encouraged and promoted. Furthermore, collaborative activities are welcome as they are conceived to have a significant impact on the teaching and learning process.

The measures mentioned above aim to help students organize their thinking favoring critical reflection, elaboration of hypotheses and research work in a project in which they must take responsibility for their learning applying their knowledge and skills to real-life actions. Thus, through an action-oriented approach, in which all areas and subjects are integrated, the students mobilize a vast repertoire of knowledge, skills and personal attitudes, i.e. the elements combining the competences mentioned above (Order ECD/65/2015: 18). Finally, informal assessment procedures such as the systematic observation of students' work, class work, the portfolio and checklists along with oral and written homework must be designed as to allow integration of all the competences in a coherent assessment framework. This, in fact, marks a departure from traditional assessment practices based mainly on written tests and the assessment process conducted on the part of the teacher - not the students - and goes in line with the recommendations for CLIL assessment as it will be explained in detail in the following chapter. Again, it is important to point out that, as was the case of the former law, assessment regulations are common to basic education in Spain, and no additional assessment specifications concerning Bilingual Secondary Education were made in the LOMCE, so CLIL subjects follow the general assessment criteria in force for the rest of the subjects.

### 3.5. Assessment in CLIL

Assessment is one of the most controversial topics in CLIL probably because of the duality between content and language, and the lack of official guidelines and research on this subject matter (Kiely, in Ioannau-Georgiou & Pavlou, 2011: 114). Furthermore, as CLIL is an umbrella term (Mehisto et al., 2008) portraying different realities, it is essential to consider the educational contexts in which this methodological approach is set. These various settings make each CLIL program unique concerning general aspects such as the educational level, the amount of exposure to the foreign language, the students' age and level in the foreign language, and the different subjects being taught through it. Likewise, the different contexts also shape the way assessment is conducted regarding the following aspects. First, the type of assessment done in the country and region might hamper the adoption of Formative Assessment practices. Second, in those countries favoring summative assessment, we need to take into account the extent to which CLIL is accommodated in mainstream education, and whether it is implemented following standard assessment criteria, a fact which constraints Formative Assessment. Third, the teachers taking part in the bilingual model also play a significant role since their profiles and training in CLIL related pedagogies have a deep impact on the way assessment is conducted.

Finally, another typical concern refers to the way skills are treated in assessment. The common debate around CLIL and bilingual education deals with the integration of content and language and, therefore, skills are usually disregarded or restricted to the subject of English as a foreign language (Pavón, 2016: 103). This might be due, on the one hand, to the content teachers' tendency to focus on theoretical knowledge rather than skills and task-based work, and to use lower order thinking skills in exam questions. On the other hand, this absence of emphasis on skills can be due to content teachers' lack of training

in CLIL pedagogies, which might lead them to think that CLIL is merely “teaching in another language”. Besides, cognitive and practical skills are clearly noticed in some subjects, but not as easily recognizable in others, a fact which complicates the assessment practice. In subjects like History and Natural Science as opposed to Physical Education or Technology for instance, students might still be assessed in terms of traditional content-knowledge rather than skills, which has an inevitable impact on assessment.

### 3.5.1. Possible ways of assessing in CLIL

The primary goal of CLIL assessment is content (Coyle et al., 2010) and thus, assessment in CLIL is more related to assessment in non-linguistic subjects rather than in foreign languages. However, as was pointed out before, the dual focus of CLIL might complicate the assessment as teachers commonly doubt whether to place the focus on both content and language issues. In fact, due to the relevance of language in CLIL as the vehicle to express content knowledge and skills, language-related issues are one of the most contested aspects of the CLIL literature (Llinares, Morton & Whittaker, 2012; Massler, 2011; Kiely, 2009, 2011; Serragiotto, 2007). When it comes to deciding whether and how to assess language in CLIL, the following are common questions which arise among practitioners. First, in CLIL assessment, do we assess content, language or both? Do we sometimes assess one and not the other? If so, which and when? Why and how? (Coyle, 2010). Second, should we assess the language in CLIL (Llinares, Morton and Whittaker, 2012); if so, which aspects of the language should be assessed, and who is responsible for that: the language teacher/the content teacher or both? Third, research has also focused on how to compensate for limited language proficiency, i.e. what happens with those students who are weak in language skills but good at content. In this regard, questions are posed to whether students should be allowed to use their mother tongue as a



communication strategy (Coyle, 2010; Kiely, 2011), the effect this might have on their grades (if any), and whether an overt focus on form favors language skills (Pérez Vidal, 2007; Pica, 2002).

Contrarily to foreign language teaching, where language objectives are at the forefront, attention given to language in CLIL can vary among practitioners depending on their profile, the teachers' expectations, and its relative priority within CLIL objectives (Coyle et al., 2010). Consequently, concerning the treatment of language-related issues, we find two approaches to assessment: discrete assessment and integrated assessment.

#### 3.5.1.1. Discrete assessment

Discrete assessment (Barbero & Clegg, 2005; Järvinen, 2009; Serragiotto, 2007), which is the most popular approach to CLIL assessment (García, 2009 in Wewer, 2014; Mohan, 1986), considers language and content separately. According to advocates of discrete assessment, language should be given special attention so that it is not downgraded in the subject. Thus, since language inevitably interferes with content as the vehicle of expression, it is important to distinguish the language-related aspects from the disciplinary ones to prevent 'muddied assessment' (Weir, 1990). Muddied assessment results from the overlapping of tasks, for instance, as when the performance of one task depends on language skills such as understanding a reading or listening extract. Therefore, "assessment must be structured in such a way that there remain no doubts as to whether missing elements or mistakes are linguistic-oriented, content-related or both" (Serragiotto 2007: 271). Another interesting debate which frequently arises among those in favor of assessing the language aspects as separate from content ones is whether the language should be taken into account in the grade. In this sense, Frigols, in Megías Rosa (2012) asserts that foreign language proficiency should be kept apart from the content proficiency and skills so that it does not contaminate the grade or is marked down in the

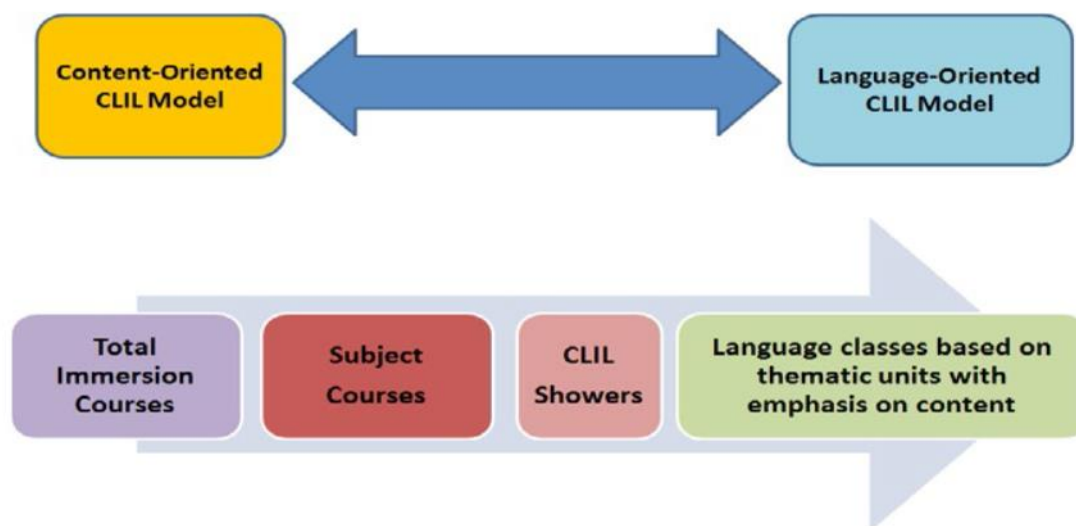
task/exam. She advocates then for assessing both content and language, and to inform students about the language they need to focus on to improve: “We should not assess or mark down content in the subject of English as well as we should not assess or mark down English in Math or Science” (2012: 13) (author’s translation), she concludes.

### 3.5.1.2. Integrated assessment

On the other hand, we also find the integrated assessment recommended in *The CLIL Compendium* (2001) where both content and language are assessed simultaneously. In this type of assessment, language is used as an instrument through which learners can show “the breadth of their knowledge and skills in relation to both content and language” (Marsh, Marshland & Stenberg, 2012: 12). In this sense, Coyle et al. (2010) consider that language objectives may serve several functions as related to content objectives. First, they might relate to the effective communication of content or include notions- specific vocabulary (CALP) from the unit - or functions such as the ability to communicate and use language to conduct practical discussion on the subject. Second, language objectives might also focus on form but in relation with the type of academic discourse in question— like the ability to use tenses correctly depending on the subject and discipline. Following this instrumental approach to language issues, language is used to improve content communication, i.e. to ensure the message in the foreign language is clear enough, and it fulfills its expected function in the academic discourse of the subject. Besides, language-related skills are necessary to “make the language more visible and give students the chance to progress in academic language (Mc Kay, 2006: 34, in Massler, 2011). Thus, although students need to master the language allowing them to express skills and knowledge in content subjects, language-related issues are measured in relation to content objectives.

In any case, regardless of teachers' approach to language, teachers should be clear about why they are assessing language as well as content, how they would like to do so (Coyle, Hood & Marsh, 2010: 11), and the changes they need to implement Formative Assessment when it is not present in mainstream education. Besides, as was stated before, considerations about assessment in CLIL need to take into account several factors such as the model of CLIL, which shapes the amount of language present in the curriculum and program, and the students' level in the foreign language. In immersion programs or high exposure or hard CLIL for instance, there is a significant prevalence of both content and language or content only, which facilitates the focus on content-related issues. Contrarily, low exposure or soft CLIL models are more language focused (Bentley, 2010 in Wewer, 2014) as is shown in the graph below, and thus, teachers tend to give more prominence to linguistic aspects:

Graph 9. CLIL content-language continuum approach



(From Ioannau-Georgiou & Pavlou, 2011: 16)

In general, and despite suggestions by researchers (Coyle et al., 2010), national recommendations with regards to CLIL tend to highlight the language proficiency that students are meant to acquire over content knowledge (Eurydice 2006: 56). Nevertheless, for those assessing language-related aspects, the biggest problem as Cushing Weigle & Jensen (1997), Hönig (2010) and Wewer (2014) point out, lies in the lack of a CLIL curriculum specifying the role and weight of language in CLIL assessment. This curriculum could help to determine “the extent of English language exposure in subjects other than language, the subjects which follow the CLIL curriculum, the contents instructed through the foreign language, and the desired level of English in all four skills plus cultural skills” (Wewer, 2014: 234). To compensate for the problem of the lack of a CLIL curriculum, official regulations and established criteria, Cushing Weigle & Jensen (1997) suggest anchoring the proportion of target language in CLIL (say 25%). This way, practitioners could have a rule of thumb or an approximate idea of the weight that should be given to the target language, i.e. 25%, and proceed accordingly. Other authors such as Gottlieb (2006) recognize the need to parallel language proficiency and academic achievement in such a way that content objectives can help us define the academic language required for achieving content standards. In this sense, teacher collaboration about the aspects that should be considered, and the weight they are given (if any) can facilitate the content teachers’ work and make language visible in the content class. Likewise, as Bentley (2010: 124) explains, in considering linguistic aspects, we contribute to narrow the focus of assessment depending on the subjects, and help in the design of assessment instruments which pinpoint essential language features for the topics and subjects in question. For instance, subjects like Art require the least level of language production while in Information and Communication Technologies (ICT), where language is needed for the correct expression of content knowledge, both content and

language-related issues are subject to assessment if the teacher decides to assess language at all. In Music, to take another example, the teacher can decide whether to include only content-related issues, and ask the students to perform a musical instrument so that they can show their musical abilities or skills. Another possibility in this subject could be to answer questions for example, using comparatives and superlatives appropriately to describe the qualities of sound, and write about pitch, duration and intensity, which require proficient language use (Bentley, 2010: 127).

After having described the different options to deal with language in CLIL, and in the absence of standard assessment criteria, the vision supported here advocates for the assessment of language issues depending on the CLIL model, context and subjects in particular. In hard CLIL, and those subjects requiring least language production, the focus should be on content, and language should be assessed as integrated with content knowledge, (Coyle et al., 2010). Contrarily, in soft CLIL, and subjects demanding more language production, the language should be treated as a separate component. Regardless of the choice, an appropriate treatment of language following a functional approach, and highlighting the role of language in the construction of academic discourse is still essential. This way, we can avoid language becoming an invisible part of instruction (Llinares, Morton & Whittaker, 2012), and use it to inform students about how to overcome language mistakes in a near future (Frigols, in Megías Rosa, 2012).

#### 3.5.1.3. “Functional” assessment

In this section, a model to assess language registers following a functional approach is provided. For this purpose, language will be considered from the view of SFL, and other important issues such as the students' level in a foreign language depending on the different skills, and the choice of code will be also considered. It is hoped that the

suggested model applied to different contexts can serve as an aid to teachers in assessing language in CLIL subjects.

Language is an essential part of CLIL instruction and as such, it should be devoted specific attention as the primary evidence that teachers use to judge students' achievement in certain subjects. Nevertheless, for the integrated assessment of content and language, a new vision of language literacy, emerging from the systemic-functional model of language (Halliday, 1985), is required. As was pointed out in chapter one, SFL considers language not just as a way to communicate and function in society, but also as a resource for creating meaning in a range of contexts (Coyle, 2009). In defining the language in CLIL, SFL helps us to consider how each subject makes use of different genres along with academic vocabulary to serve academic discourse and thus, to express content knowledge. SFL also helps to ascertain how the language in CLIL can be assessed by taking into consideration specific domains and genres. However, because of the CLIL nature, the way language proficiency is considered deserves closer attention. In CLIL contexts, students do not need to master the vehicular language before instruction, and thus, this new language literacy should be viewed as limited if compared to native-like proficiency in monolingual and immersion contexts (Lasagabaster, 2010), i.e., a type of interlanguage on the road to successful lifelong learning (Council of Europe, 2001; Lasagabaster & García, 2014; García & Lin, 2014), which is incomplete and subject to change.

In fact, due to the students' limited language proficiency, learning a subject through the vehicle of a foreign language is not the same as learning it in a first language. If the student is not able to express herself/himself in this foreign language, the grade s/he receives might be lower than the one by the student who is more proficient. Thus, as language expectations are often embedded in the assessment criteria, when language is not assessed

appropriately, it can threaten the validity of assessment, and fail to provide an accurate picture of students' content knowledge and skills (Boscardin et al. 2008: 4). To prevent this, the language needed for the competent performance of content learning needs to be clearly visible. First, it should be linked to the achievement of content-based learning objectives (Coyle et al. 2010; Llinares, Morton & Whittaker, 2012). Second, language goals should be expressed regarding external language standards from the CEFR (Llinares, Morton & Whittaker, 2012: 284-5), and adapted to students' proficiency levels to determine the desired level of English (Wewer, 2014: 234). Third, these language-related goals should be shown to students. In this sense, teachers need to be aware on the one hand, of the students' language proficiency, and be familiar with the different levels in the CEFR. On the other hand, teachers are also encouraged to know about the specific language competence descriptors intrinsic in CLIL. The following are the main aspects of language competence content teachers need to take into account when assessing language in CLIL:

Table 8. Language competence descriptors in CLIL

The ability to recall subject-specific vocabulary
The ability to operate using functions, i.e. appropriate language structures and forms- to discuss, disagree, ask effective questions and for clarification, etc.
The ability to listen and read for meaning
The ability to present or discuss effectively
Demonstrate thinking/reasoning in the CLIL language
Show awareness of grammatical features of the language

(From Coyle, Hood & Marsh, 2010: 119)

Among the descriptors displayed above, productive skills such as the ability to recall academic vocabulary, operate using functions, presenting, discussing, and reasoning in the vehicular language demand a high level of language proficiency on the part of the learners. To overcome and compensate for limited language skills which can compromise (some) students' scores, Massler (in Ioannou-Georgiou & Pavlou, 2011) suggests that teachers try to use the most direct method of assessment which uses the least language such as completing grids, and drawing diagrams or pictures to boost students' comprehension. However, although reducing the level and amount of language present in assessment tasks can be beneficial for pre-primary and primary students, for higher educational levels, cognitively challenging content requires more advanced language use and skills supporting content expression. So, if CLIL is aimed at developing both content and language skills, diminishing the presence of language in assessment tasks does not seem to succeed in the long-term especially in those subjects and contexts in which productive skills prevail over receptive ones<sup>10</sup>. For students to be language-competent in CLIL, they need to be able to express themselves in both the written and the spoken form along with any specific aspects of foreign language grammar and vocabulary helping them to communicate that content knowledge (Hargett, 1998). Regardless of the weight given to language aspects in CLIL, if teachers decide to assess it, they should define the construct or specify what aspects of language should be assessed. According to the CLIL Compendium (2001), for students to be able to function in CLIL contexts, they first need to improve their overall target language competence; second, develop communicative skills and third, deepen an awareness of both their mother tongue and the target language.

---

<sup>10</sup> In the Spanish scenario, for instance, probably due to the lack of oral tasks in the PAU and EvAU exams, students need to demonstrate their ability in relation to mastery of subject-specific vocabulary and grammatical features of a language (lexical-grammatical knowledge).



The problem arises when students fail to improve the target language competence, the language they produce is not adequate or correct for the context in question, and the teacher doubts as for the type of mistakes s/he would correct (if any). In this regard, Mohan & Huang (2002) suggests that since language is not learnt separately from content knowledge in CLIL, mistakes should not be considered regarding grammatical correctness/incorrectness but in functional terms. As they point out: "the question is not whether a language form is grammatically correct but whether a form is used appropriately to convey a meaning in functional contexts" (Mohan & Huang, 2002: 240). Although an overt focus on form is believed to have a positive impact on the development of students' interlanguage and linguistic production in immersion programs and CLIL contexts (Pérez Vidal, 2007; Pica, 2002), language mistakes should be judged differently as compared to EFL mistakes i.e. as taking into account their communicative intention in terms of language functions rather than language accuracy or grammatical correction. Thus, contrarily to traditional practice in a foreign language lesson, the question of assessment in CLIL does not deal with the students' ability to use a linguistic form correctly but to use the appropriate form to express meaning in the particular academic context. For instance, in History, we need to focus on whether the student was successful in using factorial explanation, causal language and simple language forms to express degree of certainty (*The war was probably caused by...*) rather than focusing on accuracy and spelling in verb tenses (Llinares, Morton & Whittaker, 2012: 294).

In relation with limited language proficiency, it is interesting to consider the type of errors/mistakes<sup>11</sup> students make since their treatment would be different depending on

---

<sup>11</sup> According to Corder (1967, in Ellis, 2008: 961, 971), an error is a deviation in the student's interlanguage, which results from lack of knowledge of the correct rule. Contrarily, a mistake is viewed as a lapse reflecting processing problems, which occurs when the learner fails to use the system correctly. As in the second language literature, the terms error and mistake are often used interchangeably and, as I will not reflect on

their nature. Errors are important in that they help differentiate among CLIL assessment practices. In fact, individual differences usually lie in the approach teachers take to error correction, which inevitably has a profound impact on how students perceive assessment. It seems that in general, a big number of CLIL teachers tend to assess language with an apparent prevalence of lexical errors over pronunciation ones, which are usually ignored (Dalton-Puffer, 2008) or regarding the use of target academic vocabulary (Fuentes Arjona, 2013). However, a closer look at different practices in CLIL usually reveals that decisions about whether to assess language-related issues or not and if so, the best criteria to assess language in CLIL, greatly depend on individual teachers and not departments or institutions.

Currently, errors are considered as part of the process of acquiring a language and as such, teachers have to undertake specific pedagogical procedures to reduce their number, and promote reflexive attitudes with their students to help them develop their language skills. The approach to errors is, consequently, different to that of mistakes as they should be corrected in such a way that they do not interfere with communication while encouraging students, and providing clear feedback and correct models (Council of Europe, 2001: 27).

Regarding error typology, Ernst (1995, in Hönig, 2010) divides them into the following categories in the context of error correction typical in form-focused instruction (FFI)<sup>12</sup> (Pawlak, 2014). This typology can help teachers to identify the kind of errors which should be corrected in the CLIL context as depending on the extent to which

---

their nature, hereinafter, in the second part of the study, I will use the term “mistake” to refer to both students’ errors and mistakes.

<sup>12</sup> FFI is broadly understood as any attempt on the part of the teacher to encourage learners to pay attention, reflect and gain control over targeted language features, whether they are grammatical, phonological, lexical or pragmalinguistic in nature, in a planned or spontaneous way (Pawlak, 2014: 2).

understanding is impeded or impaired, i.e., considering language as the vehicle for expressing content knowledge:

Table 9. Ernst's error typology

Phonological, morphological, syntactic, semantic or pragmatic errors
Pragmatic or register errors
Errors of form
Errors in content specific terminology

(From Ernst, as cited in Hönig, 2010: 28-9)

The first type refers to phonological, morphological, syntactic, semantic or pragmatic errors that impede or impair understanding which, in the context of CLIL assessment, should be corrected and assessed. The second type of errors is pragmatic errors or errors of register which are considered inappropriate to both culture and situation, and which should be corrected. Third, errors of form, i.e., deviations from grammar rules which do not impede understanding and that could be treated differently than in language lessons as will be explained below. Finally, errors in content-specific terminology- particularly those previously dealt with in class- which impede understanding and prevent students from progressing in content subject knowledge due to the absence of specific academic vocabulary or CALP, which should be corrected and assessed (Hönig, 2010: 29).

Finally, about the choice of code in CLIL assessment, and again due to the lack of clear guidelines or specifications about CLIL assessment in general, and the use of L1 in particular (Lin, 2015), options vary among CLIL practitioners. Regardless of the fact that

instruction should be mediated in English, the teacher should be open to using the L1 moderately, and allow students to do the same occasionally (Gablasova, 2014; González & Barbero, 2013; Massler, 2011 and Höning, 2010). This moderate use of the students' L1 is especially recommended in monolingual contexts, and when they need to engage in "exploratory talk" to co-construct knowledge and understanding of the topic, check comprehension, and promote interlingual work by exploring the two languages (Kiely, 2011: 62), and thus, support learning. By giving students the choice of using their native language or the language of instruction, they benefit from the explicit clear and plurilingual approach in deepening awareness of both the target language and the mother tongue, and develop plurilingual interests and attitudes (Marsh, Marshland, & Stenberg, 2001). The use of the L1 is particularly relevant in some CLIL contexts such as Primary Bilingual Schools in the Spanish CAM Bilingual Project, in which official guidelines recommend the reinforcement of academic vocabulary in both Spanish and English.

In an attempt to assess learning in subject matter, the model proposed by Polias (2006), based on the SFL (Halliday and Hasan, 1985; Halliday & Matthiesen, 2004; Bachman & Palmer, 2010), can be useful for teachers to assess language effectively in specific CLIL genres (Cope & Kalantzis, 1993; Hasan & Williams, 1996; Martin & Veel, 1998; Unsworth, 2000; Whittaker et al., 2006). This model is functionally organized as to operate in all three manifestations of register – field, tenor and mode -, this register being what distinguishes different types of genre. The genre refers to the text type and structure, i.e. the purpose, stages, organization and phases in the text. The field deals with the type of lexis, i.e. how varied it is and its degree of technicality and abstraction. The tenor describes whether the text is consistent with the roles taken on by the language user, i.e. the degree of expertise and objectivity the text shows. Finally, the mode refers to whether the information in the text is organized in a coherent and cohesive way along with spelling

and punctuation patterns (Polias, 2006: 59). According to Polias, the more able students are to operate successfully in the register continua, the better and more appropriate their production becomes.

Table 10. Polias' model for text analysis

<b>Genre</b>
- Are the stages and phases of the text organized logically and according to the genre and the task?
- Have all the stages and phases been included?
- Do each of the stages and phases achieve their purpose?
<b>Field</b>
- Does the text include all the field knowledge expected?
- How varied is the students' vocabulary?
- How well has the student expanded the nominal groups?
- Is the level of technicality and/or abstraction in the text appropriate?
<b>Tenor</b>
- Does the student show the appropriate level of expertise?
- Is the appropriate level of uncertainty used?
- Is the appropriate level of objectivity used?
<b>Mode</b>
- How well does the student choose theme (orientation)? In the text?
- How well are conjunctions used?
- Is the text cohesive?
- How accurate are the grammatical elements such as tense, articles and word order?
- Is the spelling accurate?
- Is the punctuation accurate?

(From Polias, 2006: 59)

Perhaps one of the strengths of the model is that it can be used not only for product-based assessment- the type of information which focuses on the final product, namely essays, project work and oral presentations- but also for the process-based assessment recommended in CLIL contexts. That is, for instance, the case of portfolio work, in which students can reflect on their work at distinct periods of time, and thus comment on their improvements. In fact, the focus on long-term work can be helpful for students in the first years of secondary education who often lack academic language or Higher Language Cognition (HLC) (Hulstijn, 2015) to produce high-quality academic explanations in subjects like Science and History (Aguirre-Muñoz, Park, Amabisca & Boscardin, 2009), and whose production should be thus judged following a process-based approach.

### 3.5.2. Formative Assessment and CLIL

In section 3.2, the concept of Formative Assessment or AfL was introduced as part of the educational innovations in the 21<sup>st</sup> century, and as a response to the exam-driven culture focusing on summative assessment practice, which demotivates students (Scriven, Bloom, Dweck, Sadler & Cohen as cited in Coyle, 2010: 112-113; Harlen, 2005; Harlen & Crick, 2003 as cited in Llinares, Morton & Whittaker, 2012: 282) and prevents them learning from mistakes. In this section, Formative Assessment and the relation it has with CLIL will be explored. However, although Formative Assessment is recommended in CLIL, it is necessary to point out that it can also be used along with Summative Assessment, as is still present in some educational contexts. In fact, the combination of both Formative and Summative Assessment can benefit the latter especially when Formative Assessment is based on rigorous planning and uses robust instruments and tools suited to CLIL subjects, leading to more soundly based assessment process (Llinares, Morton & Whittaker, 2012: 282).

Formative assessment has been described as planned, reactive and reciprocal (Popham, 2008, as cited in Llinares et al. 2012). First, planned formative assessment deals with the need for educators to identify the learning blocks necessary to achieve learning objectives, and to sequence them in terms of learning progression (Heritage, 2008). In the case of CLIL, this could be translated as the need to parallel concepts, knowledge and skills with the language needed to express them (Coyle, 2010; Llinares, Morton & Whittaker, 2012: 287). Secondly, Formative Assessment is reactive in that teachers need to constantly revise their teaching practice according to the information they gather from students on a regular basis, to identify learning gaps, and plan the following teaching sequences accordingly. This also helps teachers to adjust scaffolding techniques according to students' development in both content and language-related objectives and the practical skills they are required to demonstrate. Finally, the reciprocity of assessment refers to the students' active role in receiving feedback from the teachers, other peers or themselves, identify weak areas in content, language and skills, and thus, reflect and adjust their learning skills when they fail to be effective (Heritage, 2008).

In Formative Assessment, both students and teachers focus on both the process and the product, and students become the protagonists of their own learning process. This active role on the part of the students involves the capacity of analyzing their own work and those of their peers', and work independently to close the gap between present achievement and desirable goals (Black & William, 1998: 25). As Heritage points out:

The purpose of formative assessment is to provide feedback to teachers and students during the course of learning about the gap between students' current and desired performance so that action can be taken to close the gap. To do this effectively, teachers need to have in mind a continuum of how learning develops in any particular knowledge domain so that they are able to locate students' current learning status and decide on pedagogical action to move students' learning forward. Learning progressions that clearly articulate a progression of learning in a domain can provide the big picture of what is to be learned, support instructional planning, and act as a touchstone for formative assessment (Heritage, 2008: 2)

Apart from the teachers' ability to close this gap, for assessment to be effective, students need to be able to engage in reflection about their own learning process i.e. their strengths and weaknesses, and be prepared to receive feedback. This means to be attentive, motivated and familiar with the form in which the feedback is offered, and be able to record, organize and personalize it, and interpret and integrate the information received as useful feedback (CEFR, 2001: 186). Second, Formative Assessment practices considers assessment from the beginning of instruction, i.e. in "backward design" (Wiggins & McTighe, 2005), and allows teachers to organize the teaching sequence in three different steps ranging from the identification of desired results or learning outcomes to decisions on what could be termed as acceptable assessment evidence until the eventual planning of lessons and activities. Third, because even if we conduct final assessment at the end of the process or the teaching sessions, Formative Assessment can help teachers to motivate students to develop a positive attitude towards content along with a simultaneous improvement in the vehicular language performance (Ball, Kelly & Clegg, 2015: 226). Thus, because of its task-based nature, and the wider variety of classroom interaction that it promotes (Ball, Kelly & Clegg, 2015: 213), Formative Assessment is the most appropriate method of assessment in CLIL.in combination with other student-centered methodologies, and it can stand out as an empowering force in the debate of innovative educational trends (Llinares, Morton & Whittaker, 2012).



### 3.5.2.1. Activating self and peer-assessment

Throughout this chapter, several allusions are made to the need of sharing success criteria with students for effective learning in CLIL (Coyle et al., 2010), and to add transparency to the assessment process so that all the participants are informed, and criteria are agreed on by teachers (Ball, Kelly & Clegg, 2015: 252). By engaging students in self-assessment, teachers promote the kind of reflection which helps them analyze their strengths and weaknesses, and take responsibility for their own learning process. As for peer-assessment, when students are encouraged to discuss their learning and development with others (Massler, 2011), they engage in meta-cognitive thinking, activate skills which make learning successful (Massler, 2011), those processes leading to better self-assessment (Coyle et al. 2010).

Such learner training or awareness of the stage(s) in which s/he is at particular moments can be activated and reinforced, for instance, with WALT and WILF Posters. WALT (“We Are Learning To”) and WILF (“What I’m Looking For”). These posters are pedagogical tools teachers can use in the classroom to help students to be actively involved in their learning, create a positive learning environment, foster peer-assessment and self-assessment, and reflect on what has been learned during the lesson and unit (Heather, 2012: 160). The first one is used to make students aware of the main objectives of the lesson or session, and the second is aimed at showing and sharing success criteria.

Graph 10. WALT and WILF posters



(From [www.twinkl.co.uk](http://www.twinkl.co.uk))

Other possibilities facilitating self and peer-assessment by allowing students to become aware of good work in their own tasks and their peers' include other wrap-up techniques such as outcome sentences: *I wonder, I discovered, I still want to know, I still don't understand, etc.* (Echevarría, Vogt & Short, 2013: 217) used as exit slips for students after the lesson, TIB ("This Is Because") signs or WAGOLL ("What a Good One Looks Like") posters. The first one explains certain features of texts or discusses why students are learning certain topics, and the latter helps to show and share exemplary work. These techniques prepare students for later work in analyzing their work and their peers' with the help of rubrics.

### 3.5.2.2. Rubrics in Formative Assessment

Apart from the techniques mentioned above, rubrics can also help to make assessment as transparent as possible, and help students reflect on their own work. According to Mertler

(2001), rubrics are descriptive scoring schemes developed by teachers or institutions to guide the analysis of students' outcomes regarding processes, the final product and skills. Due to the lack of established assessment criteria in CLIL, rubrics have become a valuable method for evaluating content subjects as they stand out as authentic assessment tools for teachers and students. In the Spanish educational context, they are also gaining grounds partly based on the teachers' need to improve the quality of assessment by offering objective methods which can be shared by both teachers and students. Besides, the implementation of assessable learning standards and achievement indicators in the recent law on education (LOMCE, 2013) can help identify the descriptors used in the rubrics concerning outcomes or achievements.

There are two types of rubrics: analytic and holistic. Analytic rubrics are made up of different parts which the teacher can score separately, and then calculate a total score. Proponents of assessing content separate from language objectives consider these are the type of rubrics which probably suit the purpose of CLIL best since they provide two separate scores (content plus language) which can afterwards be summed up to obtain a total score. According to Foran (2012), analytic rubrics provide more detailed information, and are useful in planning and improving instruction. However, analytic rubrics also have some downsides. On the one hand, articulating its different components can be time-consuming. On the other, teachers can find it difficult to write language-related descriptors in case they have not been specified in advance and/or have not been paralleled with content objectives, which unfortunately might be common in some CLIL contexts. Another downside can be in relation to the need to write a rubric which is clear enough to define performance levels effectively. For instance, using the classifications "very clear" and "very organized" to refer to content expression may hamper the scoring

since the text may be clear but not organized or vice versa, and matching different levels of proficiency (Mertler, 2001).

On the contrary, holistic rubrics require a total score of the overall process on the part of the teacher. As criteria are summarized for each score level, they take less time to create but do not provide detailed information about student performance in specific areas of contents or skills, and consequently, students may exhibit traits at two or more levels at the same time (Foran, 2012). These holistic rubrics are best suited for those who advocate for measuring language as integrated with content, and who do not wish to assign language a different grade.

As Marzano (2010) states, with rubrics criteria are established which allow teachers and students to determine the performance quality of students' tasks. Rubrics are helpful for CLIL for a variety of reasons: First, rubrics help define assessment criteria- i.e. what the teacher considers good performance. On the one hand, this favors peer and self-assessment, as they can analyze their work and their peers', and promotes quality in students' work. On the other hand, they help to share criteria, which is beneficial for planning learning sequences, and to reduce anxiety since students and families can know what is expected from them in advance. Second, rubrics allow considering both linguistic and content aspects so that the student can have independent feedback by the same teacher- or by both the content and language teacher, if they decide to assess collaboratively. Finally, rubrics can be agreed upon with students in class so that they are asked about what they consider a good product- and analyze mistakes as in what Coyle et al. (2012) calls "The Language Clinic" or just at the beginning of instruction. For that purpose, rubrics should ideally be translated into student-friendly language (Marzano, 2010) so that they can notice their strengths and weaknesses in the task in question, reflect on their learning and progression and use the rubric in a satisfactory way for self and peer-

assessment. By doing so, the students can focus on analyzing the objectives to be achieved in each unit or piece of work with the help of the teacher and other students.

As a response to the absence of clear guidelines and instructions in CLIL assessment, Díaz Cobo (2009) offers a rubric which combines content and second language assessment criteria as referring to different levels in the CEFR (Appendix 1). Maybe its strength lies in the fact that, as the author points out, criteria are not explicitly divided into content and language areas. Contrarily, they deal with less conventional assessment aspects such as understanding the use of L2 grammar structures as a pragmatic tool to optimize the message reception, thus, considering language functions rather than language accuracy or correction.

Rubrics can also foster teacher collaboration, as in Appendix 2, which could be used by different teachers. This rubric includes both content and language criteria, and the option to write comments about the linguistic aspects.

Analytic rubrics can also be divided into different categories following the 4 C's framework, as in Appendix 3. The aim here is to provide useful feedback on areas of strength and weaknesses so that the criterion can be weighted to reflect the importance of each dimension.

Good rubrics should ideally allow the students and teachers to include comments on critical aspects. The problem is, however, that students are not used to correcting their work or their peers'. Thus, some previous work on the part of the teacher is necessary to make the students participant of the processes needed in self and peer-assessment (Escobar Urmeneta, 2006: 91), help them understand the meta language typical in the rubric descriptors and, as was pointed out before, translate them into language which students can easily understand.

### 3.5.3. Quality standards in CLIL assessment: Validity, reliability and fairness

Quality assessment requires alignment between learning and assessment practices, i.e. standards, curricula, assessment and instruction (Anderson, 2002; Resnick et al, 2000). However, this relation is not always that straightforward, and cannot be taken for granted. In fact, leaving aside the common debate about whether to assess language-related objectives separate from content knowledge and skills, assessment in CLIL should conform to the same quality criteria in any kind of assessment regarding validity, reliability and fairness (Järvinen, 2009). Validity and reliability of assessment are context-dependent, i.e. they are aimed to measure whether all assessment tools are appropriate for the context. Validity is based on the assumption that a test should measure what is supposed to measure. For example, in CLIL, it would entail to assess only the content if the focus was not on language throughout the lesson, or taking into account students' language skills at varying degrees at the initial stages of learning. Reliability is the degree to which an assessment tool produces stable and consistent results (Phelan & Wren, 2005) as for instance when assessing Art portfolios taking a rubric to measure the achievement of certain standards and thus, avoiding subjectivity. Finally, fairness deals with giving all students the opportunity to demonstrate their knowledge regardless of their learning style, and proficiency in the foreign language.

Other factors ensuring quality in CLIL assessment concern the use of multifaceted assessment methods (pertaining to content, all four language skills and cultural skills), the adoption of criterion-referenced inferences (sustained by pre-defined criteria), and the showing of evidences of foreign language proficiency paralleled with the CEFR. Finally, quality in CLIL is also related to the need for conducting assessment in a regular or continuous way so that there is sufficient information and feedback, and using the most appropriate assessment tools for CLIL (Wewer, 2014).

#### 3.5.4. CLIL Assessment Tools

In this section, CLIL assessment tools will be described. Although, as was mentioned earlier, Formative Assessment methods should be given preference in CLIL, they do co-exist with more traditional tools such as exams. This is reflected in the following paragraphs, which will first focus on AfL through portfolios and other alternative tools, and then describe suitable exam formats. The suggested assessment tools are thought to balance the cognitive and language demands, and be responsive to students' difficulties that is, to differentiate whether students have grasped essential knowledge, they are hindered by language barriers, or both (to varying extents) (Lo &Lin, 2014: 98). Likewise, for successful assessment in CLIL leading to informed instruction, assessment tools should be aligned with skills/competences, and measure students' progress with performance-based tasks. So, "the more directly a test assesses these skills, the more likely it is to encourage them" (Stobart, 2008: 115).

##### 3.5.4.1. AfL assessment tools

AfL involves classroom-based assessment, which is part of every teacher's practice, and other alternative tools, such as the portfolio, used as a support to assessment. In fact, much of the information gathered by teachers comes from informal class observation, which allows them to make decisions on students' knowledge and progress, and might influence in determining what is taught next, and how it is going to be taught (Rea-Dickins, 2001: 434). Besides, classroom-based assessment can help to "balance the overemphasis on scores over the useful advice by teachers in the line of Formative Assessment, as well as the tendency to emphasize quantity over quality, which is sometimes common in Summative Assessment" (Black & William, 2001: 5).

There are many varieties of alternative assessment which include performance-based tasks meant to provide students regular feedback on their progress: portfolios, journals, projects (Coyle et al., 2010; Short, 1993), handheld devices, such as clickers and other performance tools, vocabulary journals for particular subject areas, non-print reviews which favor other types of learners apart from visual ones, and games (Echevarría, Vogt & Short, 2013: 222). In this section, Short's assessment matrix (1993) will be taken as an example of the assessment methods which best serve CLIL instruction. As Short points out, due to the infinite type of tasks, content teachers might need to adapt assessment tools to suit the different contexts and the subject (s). Besides, to avoid muddled assessment (Weir, 1991), and thus, to prevent language and content objectives overlap, it is necessary to select the assessment tools carefully, and focus on the objective, i.e. to be sure what we want to assess- language, content or both-, and select the appropriate method accordingly.

The matrix Short (1993) proposes considers assessment as viewed holistically and measuring the following objectives: Problem-solving, content area skills, concept comprehension, language use, communication skills, individual behavior, group behavior, and attitude. These areas can be assessed through the following tools: Skill checklists and reading and writing inventories, anecdotal records and teacher observations, student self-evaluations, portfolios, performance-based tasks, essay writing, oral reports, and interviews (Short, 1993: 636). Likewise, as Formative Assessment recommends to assess groups along with individuals, the matrix also distinguishes between individual and group work assessment.

Finally, it is important to note that, apart from the methods suggested by Short, twenty-first-century assessment practices should include innovative tools and digital media that are suited to students' needs and interests. First, teachers should consider the use of



technology-based testing (for example through computer simulations). Second, task-based testing, for instance, in the form of web quests would also be welcome. Third, online portfolios and projects that can be stored digitally are also good alternatives as opposed to more traditional forms (Wewer, 2014).

#### 3.5.4.1.1. Portfolio work

The portfolio is a representation of a students' most valued work through evidence, collected over a considerable period, of a student's knowledge, skills and of progress made in achieving learning outcomes. Portfolios are highly regarded as an example of authentic and alternative assessment (Maggi, 2012) to measure students' progress in CLIL-related competences, and present content in a natural way to parallel activities in real-life contexts (Järvinen 2009, Coyle, Hood & Marsh 2010, Hönig 2010). In using portfolios in CLIL, we need to consider aspects like the selection criteria for student work, and its possible contents and organization. First, to favor reflection about one's work and development, students' work must be carefully selected by students with the help of teachers (Mehisto, Marsh & Frigols, 2008). This selection is, in fact, what distinguishes the portfolio from traditional compilations of students' work through class notebooks. In this sense, Escobar Urmeneta (2006) suggests the adaptation of traditional class notebooks or units into portfolio work by introducing students to content organization and selection in tutorial sessions, and sharing assessment criteria with them. This would encourage teachers who are often reticent to change and implement new teaching methods.

There are many types of evidence to be included such as an introductory letter, essays, presentations, etc. The letter in Appendix 4 is an example of a cover letter explaining the purpose of the portfolio and the activities/pieces of work that will be included in it. It can

be written by the teacher or the student and its main goal is to explain the main purposes of the portfolio. Other materials might be the following:

- A table of contents
- Updated articles and news written by the students, adapted to their linguistic level and preferences
- Videos, essays, and mind maps to assist students in organizing their thinking processes, look for key vocabulary and summary purposes
- KWLH Charts, posters, and presentations requiring a deep analysis, synthesis and evaluation of information by the student
- Plans for the upcoming work, assessment criteria concerning both content and language objectives and descriptors, and rubrics used in the subject as in Appendix 5.

Once the selection has been made, the portfolio requires some reflection on the part of the students. This reflection is the key element regarding Formative Assessment since it allows students to move forward and take steps in their learning process. In this regard, it is important to introduce and train students in the use of the portfolio so that they can make the most of it. Thus, teachers guide learner's reflection and help them to identify the main goals they tried to achieve, the best pieces of work, the degree of difficulty of tasks, the parts they enjoyed doing, the aspects they could improve and how, etc. Appendix 6 exemplifies the typical questions that can be included in a portfolio, by emphasizing students' reflection at the beginning and the conclusions.

As we can see from the suggested examples, portfolio assessment can prove useful in making informed decisions on our students' skills and abilities if it is accompanied by self and peer assessment along with teacher support. However, education in the twenty-first

century also demands tools which are updated, adequate to students' needs and suited to their preferences. The use of electronic portfolios is already becoming a common practice among CLIL educators worldwide, forcing schools to abandon paper-based activities in favor of storing files on the internet, which makes students' work easier to share, organize and access. Nevertheless, working with e-portfolios is not as straightforward as it might seem at first sight. For successful implementation of digital tools, we must consider, firstly, whether the online tool suits our needs, and to what extent the students' work can be made public. Second, whether it is possible for students to view and comment each other's work, and if the teachers can offer individual feedback. Third, it is interesting to take into account if the platform allows for multiple file formats (documents, sound files, videos, etc.) to organize data by date, course or content, and whether its cost is worthwhile among others (Hertz, 2009). Finally, in considering portfolio work, it is also important to decide on its total score and the value it has in relation to other assessment tools. In this sense, Mehisto, Marsh & Frigols (2008) recommend assign it a score up to 20% out of the final grade.

#### 3.5.4.2. Designing exam tasks and questions

Apart from classroom-based assessment and alternative assessment tools, CLIL subjects might involve the type of Summative Assessment which is typical in some educational contexts. Consequently, the way exams are designed also deserves closer attention regarding the type of questions and tasks, the use of receptive or productive skills, the level of cognition required to answer questions, and the scaffolding techniques used to facilitate content expression. To start with, the type of questions and tasks included in the exam can be oriented to focus on content, language or both, depending on the aspects the teacher needs to highlight or the students' needs. It is also interesting to consider whether tasks demand the use of receptive or productive skills, the latter being the most difficult

to assess due to the students' difficulty in expressing their knowledge and skills through an additional language. Also, the exam questions might involve LOTS - and ask for retrieval of content knowledge- or HOTS- in which students are asked to deal with more complex tasks, thus requiring a higher level of cognition, and language proficiency. Scaffolding strategies can be aimed at content knowledge and language skills. These strategies are especially beneficial for productive tasks requiring higher cognitive effort, and for students with limited language proficiency (Lo & Lin, 2016: 98). Besides, they also help educators who wish to be aware of students' progression and development as they facilitate the assessment of skills progression in process-based assessment as opposed to product-based assessment. The following are strategies to facilitate students' output during testing and written assessment: paraphrase or simplify language, add an example or model about what the students have to do, use synonyms and simple rather than long complex sentences, add visuals or diagrams, use bullet points to present information, design questions which require closed answers, use matching, labelling and fill-in-the-gaps, allow the use of glossaries and give extra time to complete the task (Bentley, 2010: 140) .

## SECTION 2: THE STUDY



## CHAPTER 4. Research Methodology

In this chapter, the research design of the study will be described in detail, including the research questions, the research methods, the participants, and the different research phases. Finally, the main limitations of the study will also be dealt with.

### 4.1. Research questions

This study aims to investigate content teachers' practice about CLIL assessment in Bilingual Secondary Schools in Madrid, and more specifically, whether and to what extent linguistic achievement in the foreign language influences the grade students obtain. For that purpose, and taking into account both assessment regulations in the CAM Bilingual Project and recommendations for CLIL assessment, this study has been conceived to answer the following research questions:

- Are teachers aware of the need to take content and linguistic aspects into consideration in assessment?
- Is there an attempt to separate language and content while assessing? How?
- Do teachers give a higher weight to the content since the main area of their teaching is content-related? Do they share their learning intention with students?
- What kind of assessment tools do teachers prefer and why? What type of tests-written or oral?
- Which are the best tools to assess CLIL? How could assessment methods be changed to make them more suitable for CLIL instruction?

In the course of the research, after having verified that language is assessed by content teachers in the CAM Bilingual Program, the following research question was added:

What aspects of language are taken into account?

#### 4.2. Research design

This is a mixed method (MM) research combining quantitative and qualitative data. The main advantage of the MM research is that it provides a fuller picture and a complete understanding of the data collected (Johnson, Onwuegbuzie & Turner, 2007) by two means. On the one hand, it presents an in-depth look at the general context, the processes and interactions, and on the other hand, the precise measurement of perceptions and attitudes (Lodico, Spaulding & Voegtler, 2006: 284). In this study, the combination of methods can help to present an objective reality, i.e. the real weight that linguistic objectives have in relation to content objectives in CLIL assessment, and relate it to other more subjective realities such as the teachers' and students' perceptions about the balance between content and language, and the way assessment of the language is really carried out in practice. Following an explanatory design, quantitative data were collected and analyzed first, and then qualitative data were collected to follow up or refine results from quantitative data. Quantitative data are collected through teachers' and students' questionnaires, while qualitative data are mostly provided through the students' interviews, the focus groups, and the analysis of assessment tools<sup>13</sup>. To assure the reliability of the study, the tools used in the methodological triangulation were built around focused, concise and relevant questions, which make the study apt to be replicated

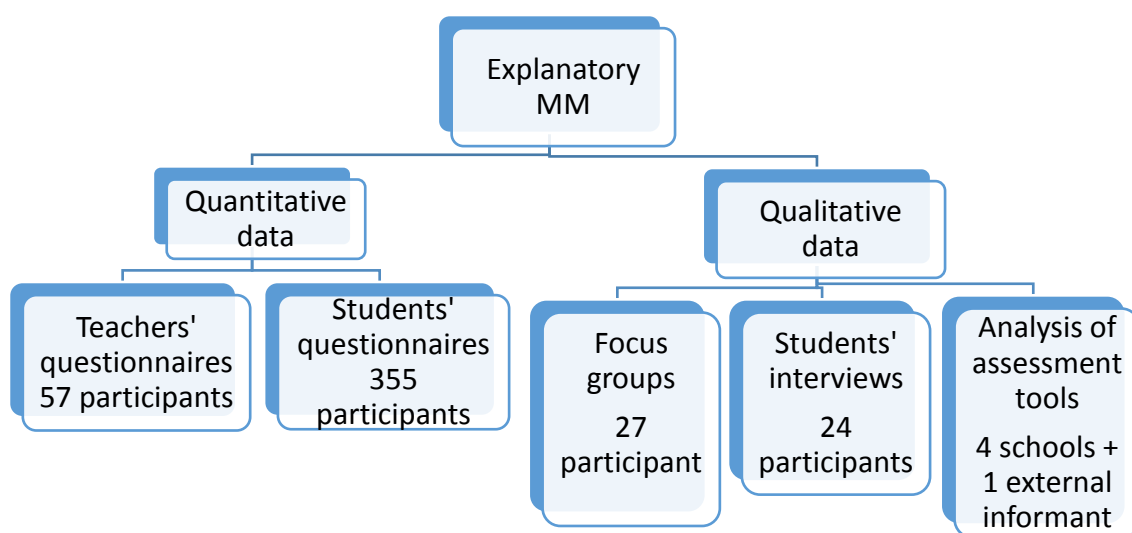
---

<sup>13</sup> As will be pointed out later, the teachers' and students' questionnaires combine closed and open questions, which can offer both quantitative and qualitative information.



with similar results. Although the study is hoped to be representative of the actual conditions in which assessment in CLIL is conducted in secondary settings in the Madrid region, it does not attempt to generalize and transfer results to other CLIL contexts. Given the complexity of the mixed-method approach, visual support is offered below to depict the different steps in the research in a clear way:

Graph 11. Steps in the research



To start with, questionnaires in language research offer the possibility to collect data in a more amenable way than other discursive elicitation techniques, and to combine qualitative and quantitative information through the use of open and closed questions. The teachers' questionnaire was aimed to answer research questions in relation to the main instruments which are used in CLIL assessment, the weight of language and content, the sharing of their learning intentions with students, and the kind of tests which are commonly used. As for the students' questionnaire, the main goal was to confirm

teachers' opinions in relation with the aspects mentioned above, and also to offer deeper insights into the balance between content and language, and the type of mistakes which are usually assessed and/or penalized.

Second, the students' interviews were meant to clarify some aspects from the students' questionnaire, and thus to observe on the one hand, to what extent language-related issues are present in CLIL assessment, and on the other hand, whether teachers share linguistic intentions with students and explore the students' perceptions about linguistic aspects in CLIL assessment. Interviews can be described in terms of the degree of formality they present ranging from unstructured to semi-structured or structured (Nunan, 2002). In this study, structured interviews were conducted with students as the questions were previously decided on the part of the researcher, who worked through a list of questions in a determined order. However, due to the young age of the respondents, and the fact that the interviews were held at school, every effort was made in order to welcome the students in an informal way, adapt the vocabulary to the students' real use of language, avoid technical terms, and encourage them to express their views using their own words. Besides, students were given the possibility to answer the questions in pairs, ask for clarification if needed, and/or they discuss the questions in pairs before providing a final answer.

Third, teacher focus groups (TFG) were conducted in order to clarify aspects about the main assessment tools teachers use, the weight of language, and the difficulties of using the same assessment tools which are common in non-bilingual subjects. Focus group interviews are excellent to complement other quantitative and qualitative research methods as they bring depth into the research, allow the researcher to verify findings from surveys and questionnaires (Vaughn, Schumm & Sinagub, 1996), and because they can help to shed light on aspects which were left unclarified in previous research or stages of

the research. In this study, the focus groups used the phenomenological approach, i.e. to understand the topic of assessment through the perspective of the everyday knowledge and practice of the participants, with the main purpose of making the most of the synergy created in the groups, which is thought to contribute to the free expression of thoughts. The questions used in the focus groups were organized in the light of the data gathered in the previous steps of the research.

Finally, a range of different tests and other assessment tools such as students' homework and project work were analyzed to observe the following: First, the way skills and processes are assessed in practice. Second, the linguistic corrections made by teachers, and whether they affect the final grade. Third, the presence of guidelines regarding the use of language in Bilingual High Schools in Madrid. This compilation of data can also help to validate the data from the previous data analysis and on the other hand, to offer deeper insights into the balance between content and language in CLIL subjects by taking a look at language corrections and the treatment of errors. Likewise, in analyzing the corpus, special emphasis was laid on observing whether specific scaffolding strategies are used to minimize students' limited language proficiency.

#### 4.3. Participants

The participants in this research are content teachers working in high schools in the CAM Bilingual Project and secondary students in the Bilingual Section of the project. Teachers are specialists in the subject(s) they teach- Music, Technology, Robotics, Biology, History and Geography, Physical Education and Arts and Crafts-, and mostly Spanish native speakers who have certified a minimum of a C1 level of English proficiency which

allows them to teach their subjects through English<sup>14</sup>. As for their training and experience, they come from different backgrounds, and have different levels of experience, being some of them novice interim teachers recently arrived in a bilingual school, and some others veteran teachers coming from the first bilingual high schools in the MEC-British Council Project or from other schools who became bilingual in the recent years. Student participants came from 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup>-grade classes and ranged in age from 11 to 16 years old. The students were mostly native Spanish speakers along with some other nationalities such as Colombian, Ecuadorian and Romanian, and most of whom enrolled in the bilingual program in primary education. Regardless of their grades in content subjects, students in the Bilingual Section have accredited a good level of English in the four skills- reading, listening, speaking and writing- by the external exam conducted at the end of primary education. Those students coming from non-Bilingual schools need to certify a B1 level to access the first grade and a B2 level to access the third grade in the Bilingual Section in the CAM Bilingual Program. Exceptionally, a level test can be conducted by “Dirección de mejora de la calidad de la enseñanza” for those students who wish to access the program without having certified the required language level

#### 4.4. Research phases

##### 4.4.1. The teachers' questionnaires

Teachers' questionnaires (TQ) (see Appendix 7) were designed with the purpose of establishing the first contact with teachers, and gathering initial information about the

---

<sup>14</sup> Although English is not the only vehicular language used in the CAM Bilingual Project, the participants were chosen among schools teaching CLIL subjects in English, as they were the pioneers in the program and thus, they were supposed to be more experienced in bilingual education.

balance between content and language in CLIL assessment. Both open and closed questions were used to elicit responses and analyze data from a qualitative and quantitative perspective. In this sense, it is interesting to highlight the fact that although open questions are more difficult to analyze in statistical terms, they also allow greater freedom of expression and enable the gathering of richer responses than merely quantitative questions can yield. For this reason, in multiple choice question, teachers were also encouraged to add their own comments if necessary. The problem of overabundance of data typical of qualitative research was avoided due to the relatively small number of participants: 57. The questionnaires were written in Spanish, usually the mother tongue of the participants, with a view to enabling them to share their views with as much confidence as possible.

The questionnaire started with some preliminary questions in which teachers were asked about their school (name, city and when the bilingual program was implemented) the subject(s) they teach in the bilingual program, years of expertise in bilingual education and whether they have been trained in CLIL. This was followed by a set of questions related to the balance between content and language, which were divided into two different sections, firstly focusing on their CLIL assessment practice and secondly, about the most frequent problems teachers find in assessment.

The first section (questions 1.1 to 1.13) dealt with general questions about CLIL assessment: the frequency and the tools teachers use to assess students, whether they use any guidelines from educational authorities or any other institution, the criteria for assessing oral presentations, the preferred format(s) for written exams, whether there are clear guidelines to assess CLIL subjects and who, in the school, is in charge of assessing students' content knowledge and language skills in CLIL subjects (content teacher, language teacher or both). Other questions in this section covered relevant issues in the

literature about CLIL assessment, such as the weight teachers give to language (if any), which aspects of language they take into account, and whether they correct language mistakes and/or take them into consideration for the final mark. As Spanish regulations about assessment at that time (LOE) did not specifically distinguish between Formative and Summative Assessment, but only determined that assessment had to be continuous, no questions were included on this topic.

The second set of questions (questions 2.1 and 2.2) takes up one of the most frequent concerns in the literature on assessment in CLIL contexts: the role and weight of language. This section was aimed at finding out how teachers deal with the use of students' mother tongue, the strategies and /or support for weaker students and those with lower English skills, and finally, which are the main obstacles teachers find in assessing CLIL subjects.

The questionnaire was initially piloted by three teachers from a Bilingual High School in Rivas-Vaciamadrid (a district in the city of Madrid) in September 2014. Unfortunately, the return rate was rather low since, after three weeks, I only received two of them. However, two interesting comments arose from these teachers. One of them suggested to take into consideration the fact that some teachers might give the language a different weight compared to the ones suggested and thus, I would probably have to reformulate the initial question in such a way that it covered a bigger range:

*I'd suggest a change in question 1.9 about the percentage of the weight of English. It's kind of closed (for instance, what option is better for a teacher giving 10% of the final score if that option is not present). It would be better for respondents to do it following different stretches. Like:*

- *None*
- *Less than 25%*

- *Between 25% and 50%*
- *Over 50%*

The other teacher commented on the suitability of the questionnaire for those teachers in ‘Sección Bilingüe’. As she pointed out, some of the questions were not applicable since students in this kind of programs are supposed to have acquired a good level of English before entering, and therefore, English is the only curricular language, and students are not allowed to use their mother tongue in class or oral/written assessment.

Taking the last comment from the piloting stage into account, it was decided not to delete the question about the use of the L1 in the class for instruction, management and feedback (Lorenzo, Casal & Moore, 2010) even if English is the only curricular language according to the teachers’ current practice. The reason is that in spite of recommendations for the use of the foreign language at all times, variations might exist among teachers in real practice. In fact, letting students use their mother tongue occasionally is common practice in some educational contexts, and could help to make assessment fairer and more accessible for students who struggle with English without changing the construct being measured. (Eurydice, 2012; Massler, 2011; González & Barbero, 2013; Gablasova, 2014)

The process of gathering data was difficult at first. The online questionnaires were sent to all Bilingual Schools in Madrid in late October 2015, and after two months, only seventeen questionnaires had been completed. The fact that I sent them to be completed online did not facilitate or accelerate the process but rather hindered it. This was probably due to the great amount of information high school teachers usually receive from the school board, the administration itself and other institutions and colleagues, which meant that they probably paid little attention to my request for collaboration. Therefore, I undertook the whole process again, this time by sending it through Facebook, and

introducing myself to some Head of Studies, Directors and Bilingual Coordinators and taking with me printed-out questionnaires. On some occasions, initial interviews were conducted during the time allotted in the schedule for Bilingual Coordination before I gained access to the High School and the teachers in the Bilingual Section. However, some centers were still reluctant being interviewed about assessment procedures; after having agreed on a first contact, three out of seven high schools failed to provide access to the teaching staff or even complete the printed questionnaires. Conversely, those few ones who seemed genuinely interested in the research, also asked for some training in CLIL Assessment as they expressed their concerns about the lack of clear guidelines from the local administration and educational authorities about assessment in CLIL in general, and the role of content teachers in bilingual education in particular. Finally, after a two-month period, in which I could access and distribute some questionnaires, 57 were completed and some interesting informal interviews and two focus groups were conducted.

#### 4.4.2. The work with focus groups

Two focus group interviews (Appendix 8) were conducted consisting of 12 and 15 teachers each. The focus groups were carried out in order to refine and explore in depth some of the information gathered in the previous step of the research- the teachers' questionnaire. Initially, the focus groups were designed to be conducted in the second academic year of the research. However, due to the tremendous difficulties I found in getting teachers to answer the questionnaires, once I had contacted some schools who agreed to participate, I decided to make most of that opportunity and tried to conduct the focus group interviews. The bilingual coordinators, being conscious of the importance of CLIL assessment for school life, invited all the members of the bilingual team- content



teachers, language teachers and language assistants- in the first focus group, and content and language teachers in the second focus group, with the main goal to facilitate teacher cooperation, draw further conclusions, and comment on future suggestions of improvement. However, as the teachers' questionnaires had previously made it clear that content teachers were the only ones assessing content subjects, the questions were strictly designed for them. The focus group interviews were conducted in Spanish so that teachers would benefit from a relaxing atmosphere and could feel free to express their own views. The discussions were focused but some scope for individual perspectives was also considered beneficial, according to what Krueger (1994) calls "the interview guide" which provides subject areas and the possibility of freely exploring, and asking questions, depending on the participants' answers. Responses were analyzed focusing on the key questions driving the study, but attention was also paid to additional comments by teachers as they help to understand their views and keep their conversation going smoothly. After the two groups were conducted, an abridged transcript was created with the most relevant and useful portions of the conversations. These transcripts were then analyzed using the constant comparative analysis (Krueger & Casey, 2009), to identify the most important trends or ideas by participants about the topic of assessment. Likewise, the questions were organized to move smoothly from one topic to the next, and special emphasis was laid on using non-technical vocabulary to promote teacher interaction at all times.

Group interaction was based on a list of topics pertaining to the main obstacles teachers find in CLIL assessment, the instruments they commonly use, the way informal class observation is used as a tool to measure students' progress and whether language competence has a direct influence on the grade they assign to a student, piece of homework/test or any tool they may use for assessment. Attention was also given to the

way teachers deal with the absence of CLIL assessment guidelines- a common complaint according to data obtained from teachers' questionnaires and informal conversations i.e. whether they communicate with colleagues in their department and/or at school to know how to deal with assessment issues, and whether they have coordinated in that matter or have reached any agreements so far on topics such as the role of the foreign language in CLIL assessment or the aspects that could be penalized (if any) in assessing the language. Other questions covered topics such as the treatment teachers give to language mistakes and the use of rubrics among other assessment tools.

#### 4.4.3. The students' questionnaires and interviews

The students' questionnaire (SQ) (Appendix 9) was conducted after the teachers' questionnaires and the focus group interviews with the purpose of double-checking the information provided by teachers, and shed light on those aspects in which the teachers' opinions seemed to be divided. As with the teachers' questionnaires, both open and closed questions were included with the aim of gathering qualitative and quantitative data. In this sense, open questions were optimal as they allowed students better freedom of expression about those topics which needed to be clarified. For this reason, students were encouraged to add their own opinion or comments to some multiple-choice questions. This time the problem of overabundance of data typical of qualitative research was avoided by the limited amount of questions offering this double option (two out of seven) and the fact that students' answers were predictably very much alike and proceeded in the same direction. As with the teachers' questionnaire, the students' questionnaire was written in Spanish, which is usually the mother tongue of the participants, and special emphasis was put on using clear language adapted to students with a view to enabling them to understand the questions and express themselves with as much confidence as

possible. Unlike the teachers' questionnaires, which had quite a low return rate, the students' questionnaires were easier to collect and a total of 355 students answered it. This was mainly due to the fact that during the teachers' questionnaires and focus groups phase, previous contact had been already initiated with some high schools who agreed to distribute them in class both in its printed and online form, and which undoubtedly facilitated the process of data gathering.

Since most of the teachers work with both 'Program' and 'Section' groups, the students' questionnaire started with a preliminary question about the school's name and whether the students belonged to the Bilingual Program or the Bilingual Section so as to make sure respondents were filtered appropriately, and that the data gathered came exclusively from the Bilingual Section groups. Fortunately, 98.3% of the responses came from Bilingual Section groups. This identification question was followed by a set of questions related to the assessment tools teachers use, whether the focus of CLIL assessment is the English language, the content subject or both the language and the content, the aspects of the English language that are measured and penalized- fluency, grammatical accuracy and academic vocabulary-, who assesses- the language or the content teacher- the approximate weight of English in the final grade, and the treatment of language mistakes, i.e. whether students are informed about their language mistakes and whether these mistakes are penalized by content teachers. Finally, students were asked if they think that different teachers in the school use the same criteria or guidelines in order to assess students' work in CLIL subjects.

The assessment tools used by teachers and the role of the content teacher as the one assigning a grade in CLIL subjects were clear from both the teachers' questionnaires and the focus group interviews and consequently, no additional information was needed in that subject-matter. However, to double-check the findings, obtain information about the

content teachers' role as seen by students, and to confirm the prevalence of written tests over oral ones, the students were asked about those aspects too.

In addition, the following aspects were specifically included in the students' questionnaire so as to gather additional information as they were left unclear in the teachers' questionnaires and the focus group interviews: First, the weight given to oral presentations in class assessment as opposed to oral exams. Second, the way informal class observation is conducted. Third, the perception students have of the role of language in content subject assessment, and whether teachers follow the same guidelines in CLIL assessment. Special attention was given to other assessment instruments apart from written exams, as considered by students, and the aspects of the English language that are taken into account and/or penalized.

The questionnaire was piloted by a group of five teenage students coming from different backgrounds: two of them reported that the questions seemed quite predictable at first whilst the three others commented that they are sometimes not aware of the aspects being taken into consideration in assessment issues. Thus, taking the opinions from the piloting stage into account, I decided to open up the variables and give the students the possibility to answer in a broader way. As students are teenagers and it is acknowledged that they do not necessarily need to know about assessment guidelines, questions regarding the aspects taken into account to measure language proficiency and the weight of English in CLIL assessment had the option to answer: '*I don't know*'.

Students' interviews (SI) (Appendix 10) were conducted after having completed the teachers' and students' questionnaires and the teachers' focus group interviews. It is important to note that for the research to be as reliable as possible, the participants in the students' interviews were selected from high schools that did not take part in the first phase of data collection process and, consequently, had not answered the previous

students' questionnaires. The interviews were conducted in October 2015, during the English class, in the bilingual coordinator's office in the case of the first group and in one of the common halls in the second group. Although students might be shy and/or reluctant to be recorded by an adult in a formal scenario, and this was thought as a menace to naturalness in the conversations, all the interviews were tape-recorded, as opposed to note-taking, in order to facilitate the preservation of language and the analysis of the data afterward. In total, 24 students participated and answered the questions: eight in the first group, from 3<sup>rd</sup> grade of Bilingual Section in a High school in Getafe, a town in Southern Madrid and 16 students in 2<sup>nd</sup> grade of Bilingual Section in a High School in central Madrid, one of the first schools taking part in the CAM Bilingual Program. For privacy reasons, students' names were not revealed or recorded at any moment, and parental permission was required before accessing the students. Thus, codes have been assigned to refer to speakers' comments. The interviews were conducted in pairs so that students could benefit from their partner's help in case they did not understand the question and to express themselves in greater detail. Furthermore, the interviews were done in Spanish, which is usually the mother tongue of the students to facilitate comprehension and create a relaxed environment. The main purposes of the interviews were the following: In the first place, although fluency and academic vocabulary had been rated as important aspects in CLIL assessment (60.8% and 56.9% respectively), I intended to confirm the prevalent tendency for teachers to penalize language aspects in CLIL subjects. Secondly, since this strong trend is usually focused on grammatical accuracy - as had been pointed out by a substantial number of students (75.8%) - it was necessary to verify that, as 44.6% of teachers asserted that they take language aspects into account in CLIL assessment. Likewise, further information was needed about the type of language mistakes which are taken into account: spelling, grammatical accuracy, vocabulary, whether the ideas are

well linked and expressed, pronunciation, summarizing skills, etc., whether these mistakes are the same which teachers take into consideration in other non-CLIL subjects, and which teachers- content teachers or language teachers- are likely to be more flexible with these language mistakes. Other questions made reference to the kind of objectives- content-related objectives, language-related objectives or a combination of the two- that content teachers present at the beginning of each unit, whether teachers inform students about the linguistic aspects they need to improve to communicate content successfully, to what extent language proficiency can help content knowledge delivery/expression and finally, whether students in the Bilingual section are allowed to use Spanish as the vehicle for transmitting knowledge and skills in CLIL assessment.

#### 4.4.4. The analysis of CLIL assessment tools

The analysis of assessment instruments was the last step in the gathering of data in this research. Initially, this phase of the study was conceived to take place during the months of February and March 2016. However, due to the low rate of response on the part of the teachers to collect and show assessment tools, it was finally completed during the months of May and June in the same academic year. After having checked that according to the questionnaires, interviews and focus groups, written exams and class notebooks were the main assessment tools used by content teachers in the CAM Bilingual Project, it was time for assessment tools to be analyzed in detail. However, for the research to be as reliable and representative as possible, no information was given to the teachers sharing their assessment tools about the results gathered in the preceding stages. At this stage of the research, and due to the continuous visits to some centers, I had already managed to make some contacts with bilingual coordinators and other teachers who had helped in the data gathering.

A total of 15 schools were requested to compile and present some written and oral assessment samples from their subjects such as tests, essays, projects, portfolios, recording of debates and/or oral presentations, classroom registers and checklists along with any other assessment tool they use in their daily practice. Most of the assessment tools were gathered during informal school visits with the teachers in question. Teachers were also given the possibility to compile and send the assessment tools at some later point in time. At first, the participants were quite open to share materials and answer questions about their assessment procedure, but finally many of them were reluctant to show their assessment tools in spite of several reminders via email, and only four schools plus an external informant (an interim teacher who was on a leave but offered to compile some tools from the previous academic year) submitted the requested data. Consequently, in an attempt not to bother teachers, and as time passed, a deliberate decision was made in order to restrict the sample to the materials available at the end of the academic year 2015-2016. The submitted assessment samples are described below in relation to the grade, subjects and type of tool:

Table 11. Analysis of assessment tools

INFORMANTS	YEAR(S)	SUBJECT(S)	TOOLS
SCHOOL A (2 teachers)	1° ESO	Geography and History	Written exams
SCHOOL B (4 teachers)	3° ESO  4° ESO	Geography and History  Physical Education	Written exams and notebook grades. Written Exams, essays and portfolio samples; Informal interview
SCHOOL C (2 teachers)	1°, 2° and 3° ESO	Geography and History	Written exams, essays and project work (interviews, posters, timelines, diaries and newspapers articles)
External informant (1 interim teacher)	1°, 3° and 4° ESO	Arts and Crafts	Project work samples and 2 oral presentations
SCHOOL E (5 teachers)	3° ESO  1° and 4° ESO  1° ESO  2° ESO	Technology and Robotics  Biology and Geology  Natural Science  Tutorship materials	Written exams  Written exams  Written exams and their equivalents in Spanish  Project work and class work



INFORMANTS	TOOLS
SCHOOL F	Language guidelines + assessment criteria
SCHOOL G	Language guidelines
SCHOOL H	Language guidelines + assessment criteria
SCHOOL I	Assessment criteria
SCHOOL J	Assessment criteria
SCHOOL K	Assessment criteria
SCHOOL L	Assessment criteria for both Bilingual Section and Program
SCHOOL M	Assessment criteria
SCHOOL N	Assessment criteria
SCHOOL O	Assessment criteria
SCHOOL P	Assessment criteria

In total, samples pertaining to seven subjects were collected including the materials used in tutorship sessions, which can be delivered in English provided the center has teachers available to do so. The data collected consists of written exams, essays, interviews, posters, timelines, diaries and newspapers articles, assessment of class notebooks, a couple of oral samples and one example of portfolio work. Besides, three language guidelines, and the assessment criteria from ten schools were also gathered.

For space and time restrictions, not all the samples are commented on and described in detail in this study. A choice has been made in order to comment on common features, and gather general conclusions mainly about the type of assessment tools being used, and the presence of language in CLIL assessment. For this purpose, samples containing both exams and students' corrected work were selected randomly, and comparisons were made between those being granted with a full mark and those scored with a pass or a fail.

Likewise, a small group of teachers from three High Schools agreed to conduct a short interview about the main tools they use for assessment, any difficulties they find in the process, and the role and weight of the vehicular language in their subjects. During the interviews, as will be commented later on, some other interesting issues related to CLIL in general and CLIL assessment, in particular, arose, namely the differences between assessment in CLIL subjects and non-CLIL subjects and the lack of appropriate CLIL materials for some subjects.

The criteria used in the analysis of the sample was decided after consulting literature on assessment design (Brown & Hudson, 2002) and previous research in the field of CLIL (e.g. Hönig, 2009; Wewer, 2014). Likewise, criteria for the assessment of content disciplines (Odenstad, 2010; UNC, 2006) and national assessment criteria, referring to both the LOE and the LOMCE we also taken into account. In the case of the first and third years of secondary education (1° and 3° ESO) in which the LOMCE was already implemented during the academic year 2015-2016, special attention was paid to analyze them by taking into account the new regulations in terms of assessment such as the use of the portfolio and peer and self-assessment. Besides, since written comments about students' work are considered an important part of assessment practice and more specifically of the type of desired formative assessment suggested in the LOMCE (2013), special attention was given to whether the teachers add this type of comments, and whether they focus on content or language-related aspects. In this sense, in the case of teachers' comments and feedback dealing with language mistakes or language-related issues, particular attention was paid to the influence that language has on the final grade. Additionally, the analysis of the treatment of English as a vehicular language in CLIL assessment also aims to check whether language is aligned with content objectives and considered from a functional point of view- i.e. whether language is considered as the

instrument which allows expressing content knowledge in terms of using the appropriate forms to convey meaning in functional contexts (Mohan & Huan, 2002) or, on the contrary, if it is related to more traditional approaches, and thus it focuses on grammatical correctness. As with the other instruments such as teachers' and students' interviews and focus groups, the aim of the analysis of the assessment instruments was to observe if the CLIL approach has any effect on the assessment conducted by CLIL content teachers. For this reason, the emphasis was placed on analyzing the types of tools teachers use and more specifically, the presence of language corrections and any specifications about language grading in the exam instructions, the corrections made by teachers. In relation to the presence of language in CLIL assessment, which constitutes one of the cornerstones in this study, other essential instruments which are hoped to contribute to the observation of the treatment of language mistakes in CLIL subjects in the absence of a larger data sample were the analysis of the "Language Guidelines" and the grading or scoring criteria teachers use for each subject. The language guidelines (also called the "Action Plan" in one of the schools) consist of three documents from three High Schools each compiling information about the criteria to grade language in CLIL subjects both in the Bilingual Program and the Bilingual section, and guidelines to assess language-related issues. On the other hand, the Grading Criteria constitutes another source of information which includes assessment criteria in a variety of CLIL subjects and some useful indications on the weight of the English language in CLIL assessment from 10 High Schools in the CAM Bilingual Project.

Due to the absence of specific guidelines regarding assessment in CLIL by Madrid's regional government, samples could not be analyzed in terms of bilingual national or regional course goals or written assessment features. Instead, I looked mainly for exam-type, the types of items included in the exams, and the treatment of language in the

samples including any specifications on the weight and role of language, and the type of corrections and feedback made by teachers. Other criteria used to analyze the samples refer to the use of accommodation strategies (Bentley, 2010: 140) in order to facilitate scaffolding, such as the use of questions to elicit answers, glossaries, models and pictures or diagrams to offer visual support.

#### 4.5. Limitations of the study

The present study is based on a limited sample of teachers (TQ=57 + TFG= 27) and students (SQ= 353 + SI= 24) plus the data gathering from seven subjects, four assessment guidelines and ten assessment criteria from 16 schools. Since the opinions and materials might not be considered to be representative of the population in the target study, the study does not claim in any way to be exhaustive. Rather, it is hoped to serve as an indicator of the balance between the English language and the content in CLIL subjects, and the main assessment tools used in the CAM Bilingual Project, and to be able to shed some light on the actions and assessment practices which should be implemented in the future. In order to draw firmer conclusions, the study could be repeated with a more significant percentage of responses in the teachers' questionnaires and assessment tools, which would ideally require the cooperation of principals and bilingual coordinators in the CAM Bilingual Project, which it was not possible to gain in this study.

Furthermore, the data collection was conducted in three differentiated periods which cover the academic years 2014-2015 and 2015-2016. During that time, two educational laws were in force, namely the LOE (2006-2013) and the LOMCE (2013 to the present). In this sense, another significant limitation was in relation to Formative Assessment and Spanish legislation in terms of assessment. Although assessment in CLIL is conceived

from the premises of Formative Assessment, the first evidence of the formative character of assessment in Spanish education dates from 2015 when the LOMCE was implemented in secondary education. As educational changes usually take time to be implemented in real practice, assessment practices in the context of this study basically adjust to the LOE rather than the LOMCE. Likewise, it has been difficult to find innovations such as the portfolio, self and peer-assessment and rubrics in the data collected, and thus, it would be appropriate to replicate the study in the near future to observe to what extent Formative Assessment tools have found their way into the CAM Bilingual Project.



## **CHAPTER 5: Analysis of Results**

This section offers a description of the main results in the teachers' and students' questionnaire, the focus groups, the students' interviews, and the analysis of the assessment tools used by content teachers. Besides, at the end of each section, a summary of the main findings is provided, and some quotations voicing the teachers' and students' opinions are included in order to clarify data.

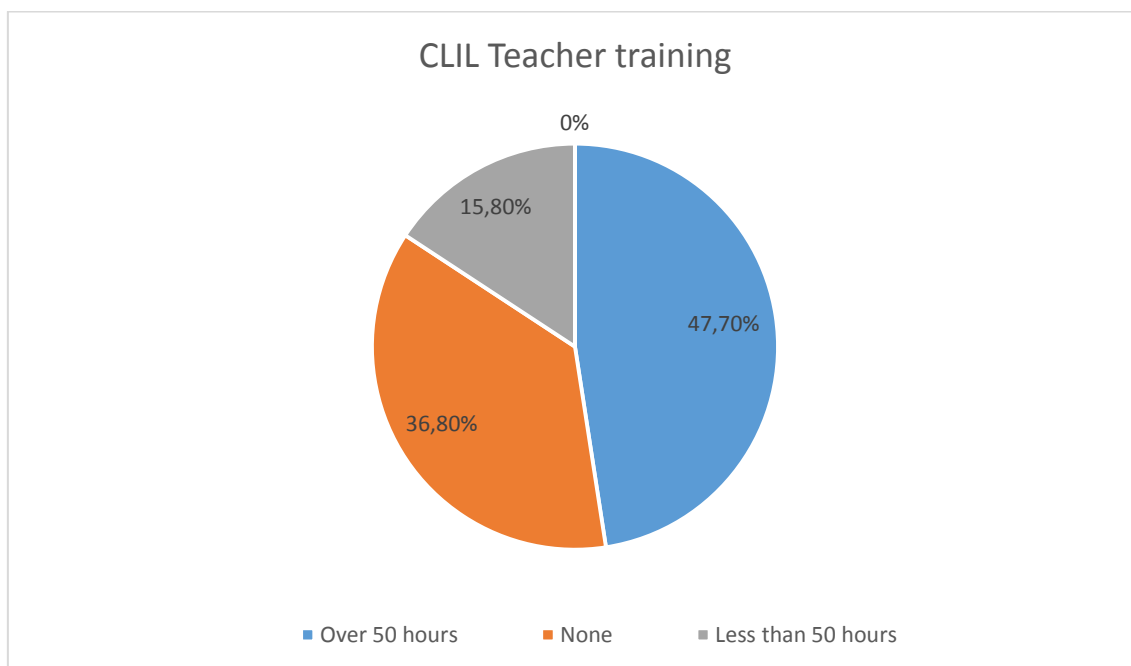
### **5. 1. The teachers' questionnaire**

The analysis of results in the teachers' questionnaire combines quantitative and qualitative data. Quantitative data is analyzed and expressed using statistics whilst the qualitative data analysis has been done on a question by question basis for both similarities and differences amongst participants' responses. Thus, key words and ideas were cross-referenced in order to make generalizations and highlight disparity among responses, and participants' additional comments from closed multiple choice questions were analyzed using the same pattern.

The questionnaire started with a preliminary section on the teachers' background. The teachers who completed the questionnaire are content specialists in diverse subjects such as History and Geography, Biology, Technology, Physical Education, Music and Arts and Crafts working in public high schools in Madrid. Most of teachers (65%) have over five years of experience in bilingual education, 20% have less than five years of experience, , 10% have between five and ten years of experience, and for only 5% of respondents it is their first year teaching content subjects in English. As for specific CLIL training, none of them has completed a Postgraduate course in CLIL methodology and/or bilingual education, nearly half of the teachers (46.4%) have over 50 hours of undergraduate CLIL

training, a significant 37.5% point out they have received no training at all, and 16.1% have less than 50 hours' training.

Graph 12. CLIL teacher training



Those who have received formal training have usually done so through short courses offered by Madrid Regional Government in both ‘Centros de Formación del Profesorado’ and European and American Universities during summer holidays, and a few others have accessed other institutions for that purpose.

It is also interesting to point out that approximately 30% of the teachers are completing an “internship period” that is, they are covering a temporary position. To be an interim in a bilingual school, and until they get a permanent job in public education through the state competitive exam, teachers need to certify the required language level (C1) through “examen de habilitación”. The presence of interim teachers generally has a direct impact



on teaching staff and consequently, on assessment procedures. This temporary staff is normally new to teaching, and sometimes they need to adapt to different schools for undetermined periods of time. This provisional condition may mean that they have no opportunity to develop any innovative or different approaches to teaching and assessing, or that they have to adapt to previously established conditions and regulations in the department. Furthermore, even if they complete the whole academic year, these teachers have no contract for the summer period and consequently, do not always assess students during September exams, a fact that could have a direct influence on students' assessment, as they expressed during informal interviews (Personal communication with students and teachers, academic year 2014-2015).

### **Frequency and assessment tools**

Regarding the frequency of assessment, most teachers express they adhere to continuous assessment as they point out they are constantly assessing students. They also state that they provide formal tests three or four times a trimester, depending on the subject they teach and their personal preferences: "I assess my students on a regular basis, everything adds to the final mark, from exams to their class behavior, every day"<sup>15</sup> (TQ-1.1)

However, the frequency of assessment can vary among teachers as eight of the respondents assert they assess every didactic unit independently: "I assess them continuously although I also grade the exams we do after each didactic unit; besides, there is the global mark in which everything is taken into account, once in the academic trimester" (TQ-1.1). What seems to be common practice for many teachers is the need to take class observation into account as is observed by 91.1% of respondents. Through this

---

<sup>15</sup> All the quotes are the author's translation.

class observation, teachers measure several aspects such as effort, level of engagement, day-to-day work and communication, partner and group work, oral presentations, projects and class debates. This is usually done through regular observation, when students have completed a unit, and through files of students' work although just 5.3% of respondents use checklists, which might mean that the rest generally take notes on students' performance: "I assess class participation every single day, and I register it. I mean, then I also take that info into account to decide on the final score, even if it's not expressed quantitatively" (TQ-1.1).

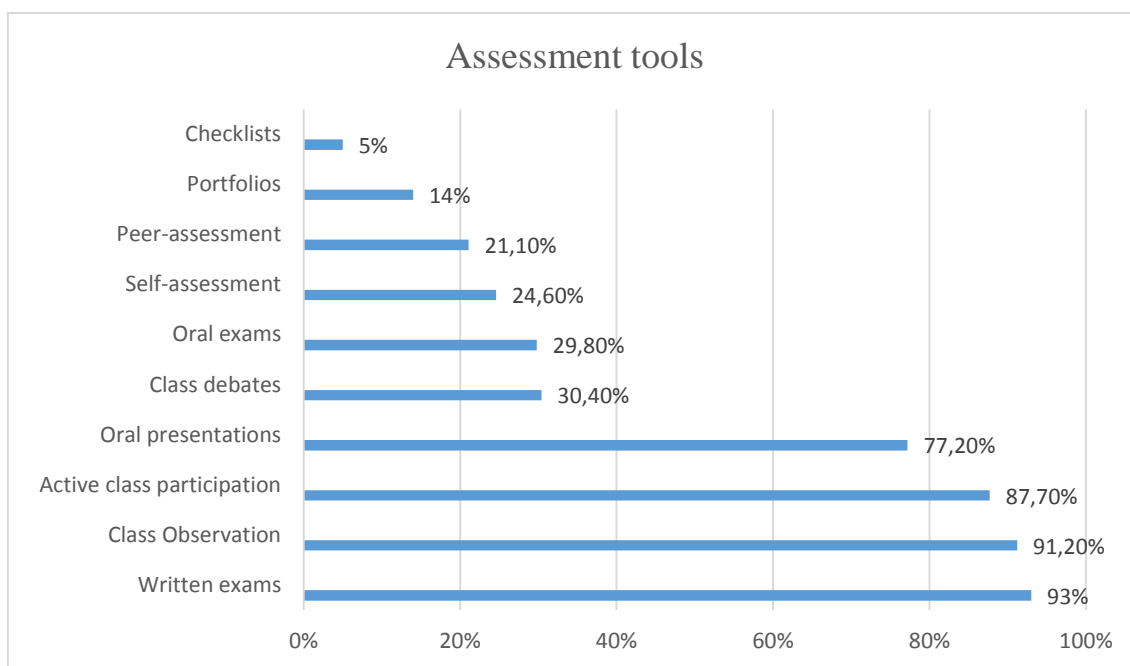
Regarding assessment tools, the data show that the most noticeable trend is the use of written exams over oral assessment procedures. As can be seen in the graph below, the vast majority (92.9%) use written exams, and only 30.4% also use oral exams. This figure, however, shows the growing tendency to use oral assessment tools which undoubtedly, has risen significantly in comparison with the predominance that written tests have traditionally had in Spanish education (OECD 2015, Lukas, Santiago, Joaristi, & Lizasoain, 2006). The systematic prevalence of written exams over oral ones might be due to the washback effect of the 'PAU/EvAU' - the university entrance exam at the end of secondary education. However, if we have a look at the *Teaching and Learning International Survey* (TALIS 2013), we observe that standard written tests are still quite predominant when compared to the use of tests in other countries (38%).

Another way to take oral skills into consideration is the use of oral presentations where students need to show their ability to express content in English individually or in groups with the help of a power point or some other IT tool. If we have a look at the number of oral presentations that are currently part of teaching practice in Madrid, we find 76.8% who use oral class presentations regularly. Looking at other forms of oral assessment, there is also a significant 30.4% who engage students into class debates or discussions

about a certain topic previously dealt with in class, a practice that seems to be gaining ground today.

Moving onto alternative forms of assessment such as self-assessment, peer-assessment or the use of portfolios, which are considered effective for CLIL instruction (Mehisto, Marsh & Frigols 2008, Coyle, Mehisto & Marsh 2010, and Massler 2011) just a quarter (25%) use self-assessment techniques along with a small number of respondents (19.6%) who use peer-assessment techniques, while several (14.3%) use portfolios, and 5.4% use checklists to measure classroom activities and homework. The limited number of teachers opting for alternative ways of assessment in the line of Formative Assessment for CLIL might be explained on the one hand because of the lack of teacher training, and on the other hand, to the fact that the use of self and peer-assessment and the portfolio in mainstream education was not recommended by Spanish educational authorities until the arrival of the LOMCE in secondary education in 2015, one year after the questionnaires were completed.

Graph 13. Assessment format according to teachers' questionnaires



It is interesting to note that although the use of portfolios is not a widespread practice in the Madrid area, “cuadernos de clase”- the compilation of students’ work through time- are still considered an excellent tool to check students’ regular work and progress for a significant number of teachers in some high schools: “We assess didactic units through the class notebook. This way we can check how the student worked on the objectives, and whether they have finished all the activities in the didactic unit during the academic year”. (TQ- 1.2). The difference between portfolios and “cuadernos de clase” is that the latter rely on a systematic account of students’ work throughout time rather than the selection of classroom materials and the students’ reflection about their progress in the portfolio.

Teachers were also encouraged to comment on other alternative assessment tools they might use. In this sense, it is important to clarify that variety in the answers stems from the different subjects they teach. For instance, two Physical Education teachers mentioned the need of focusing on more practical and “physical” activities whilst those teaching

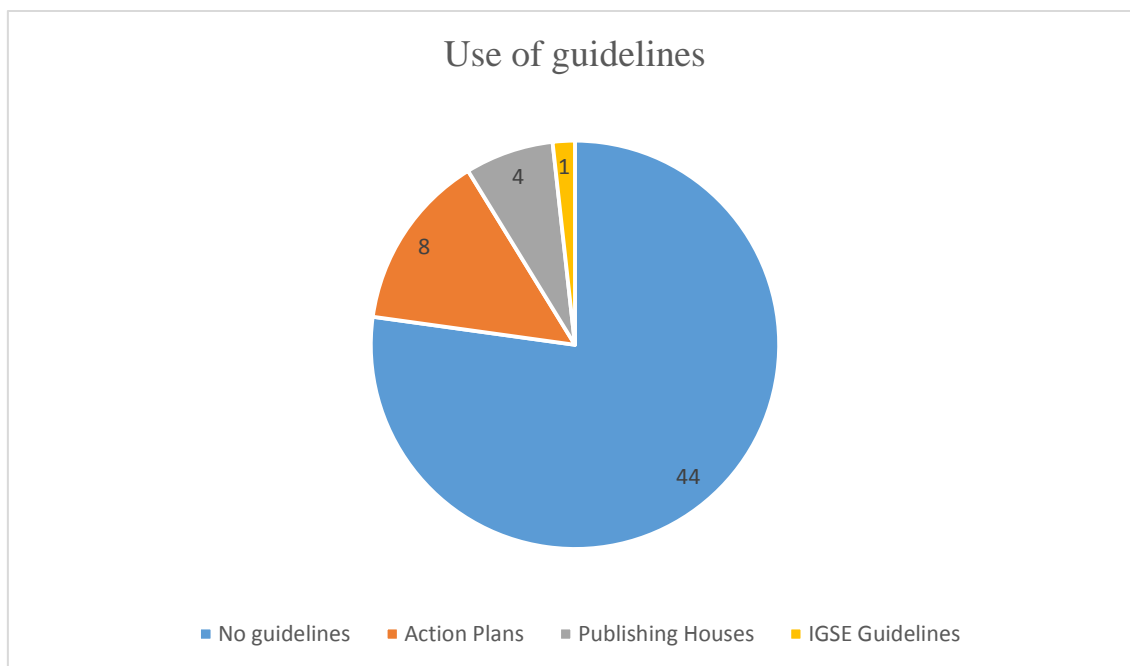
music referred to musical interpretation in terms of musical and vocal performance as valid approaches to assessment in that students do not need the aid of the English language in order to express content knowledge and objectives. Likewise, Biology and Physics and Chemistry teachers turn to laboratory work when they need to check their students' competences that is, whether they have fulfilled and covered the requirements for the subject in terms of can-do descriptors or the LOMCE's "assessable learning standards" rather than content or attitudinal aspects.

The rest of the survey respondents express their reliance on a variety of methods and tools such as project and group work (nine teachers), practical activities including homework and class work (eight teachers), fill-in-the gap activities (six teachers), power point presentations (five teachers), maps (three Geography teachers), computer room practice (two teachers), topic quizzes and work on the wiki (two teachers), worksheets students must complete and hand-in both in class and for homework (one teacher), and musical compositions (one teacher). As for students' active participation in class activities, once again, a vast majority (87.5%) confirm that they take active class participation and interest in the subject into account even if it is not always scored quantitatively.

As regards the use of guidelines for CLIL assessment from educational authorities or any other institutions, twelve teachers declare they do not use any specific guidelines. In fact, the most common answer is that they have their own and/or use the assessment criteria of the department, although teachers state they miss some agreement or set guidelines from educational authorities in Madrid Regional Government: "The CAM has not provided any guidelines in this respect. I use my own, the ones I learnt in training courses." (TQ-1.3). In general, teachers assert that they use their own ideas or have compiled some "practical tips" from the training they have received during teacher training sessions. Five teachers also declare that they lack this information about assessment, and in case of

doubt, they ask other teachers in the department for advice or, as three teachers comment, they use the criteria for assessing which is common in their didactic department regardless they teach a bilingual or a non-bilingual group: “The truth is that I don’t have any set guidelines. I have my own criteria, and I sometimes ask the department for help” (TQ-1.3). Eight teachers coming from two different schools express they have agreed on some basic rules in the Department or in the Bilingual Coordination meetings, and that their schools have developed an “Action Plan”. This so-called ‘Action Plan’ is the document collecting joint rules for assessment, created by the Bilingual Team in order to deal with students’ systematic mistakes in English in content subjects, and it will be referred to in the section devoted to the focus group interviews and in the analysis of assessment tools. As one teacher points out: “The High School has given us some guidelines we can rely on, and these are the ones we follow and adjust to. We have an ‘Action Plan’ (TQ-1.3). Also, four teachers declare they use guidelines and advice from publishing houses, and one of them asserts that although he misses set criteria from educational institutions, he takes IGSE guidelines into account.

Graph 14. Use of guidelines according to teachers' questionnaires



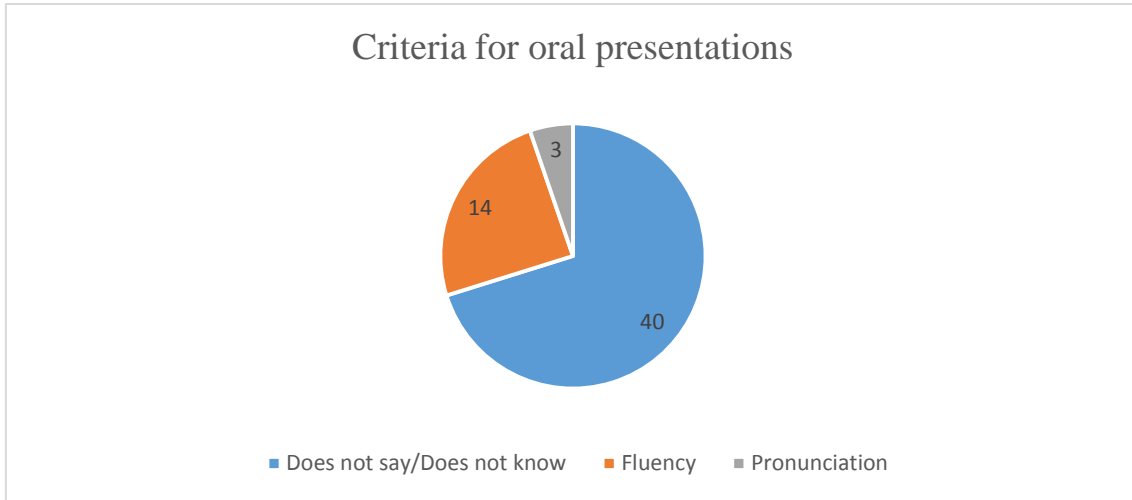
The next question was addressed to the teachers using oral presentations as a method of assessment. More specifically, the question was posed in order to disclose the criteria they use for that purpose. Answers greatly vary from one another but in general, teachers turn to content, vocabulary, delivery or the way concepts/ideas are presented, grammatical accuracy and whether the audience was successfully engaged during presentation time. The use of the Power Point and other technical tools is also valued by four teachers.

One common feature is the importance content teachers give to fluency, which is regarded as the most important factor for oral assessment, and one of the main indicators of students' proficiency in the foreign language by 14 teachers. Apparently, it seems that for some teachers, the concept of fluency is sometimes intertwined with other factors to be taken into account such as grammar accuracy and general correctness: "Fluency: Students being able to express themselves in correct and fluent English, using appropriate

vocabulary, being coherent”. (TQ- 1.4) “Content development being expressed correctly in English” (TQ- 1.4) “Oral fluency: They need to be able to use verb tenses in a correct way, and also coherence in the use of the articles, pronouns and nouns” (TQ- 1.4). Perhaps the confusion between grammatical accuracy and fluency in some teachers’ comments might be due to the fact that content teachers are not generally as used to paying attention to language-related objectives as language teachers are, and consequently, they might find it difficult to describe them. Thus, when trying to specify the concept of fluency they focus on the clarity of expression, and the students’ ability to create meaningful messages, and forget about the ability to employ formulaic language and to speak without hesitations. Likewise, it is interesting to note that the terms “correct” and “correctly” are common in teachers’ descriptions of fluency, a clear indication of the importance teachers give to grammar mistakes since both terms are more related to accuracy than fluency. Conversely, some teachers seem to have a clear idea of what fluency is but the concept of fluency also appears intertwined or in relation to other categories such as the use of academic vocabulary: “The message has to be clear enough for the content be transmitted in an effective way; you have to understand what the student means. They have to use specific vocabulary (academic vocabulary from the subject)” (TQ- 1.4). It can be concluded then that teachers might be unsure of what fluency means or that the concept of fluency might differ depending on the teacher, an aspect which confirms the lack of language awareness among content teachers. Finally, three teachers out of 57 state they take pronunciation into account and two teachers assert they use a rubric to grade oral presentations, but no further information is provided on that point.

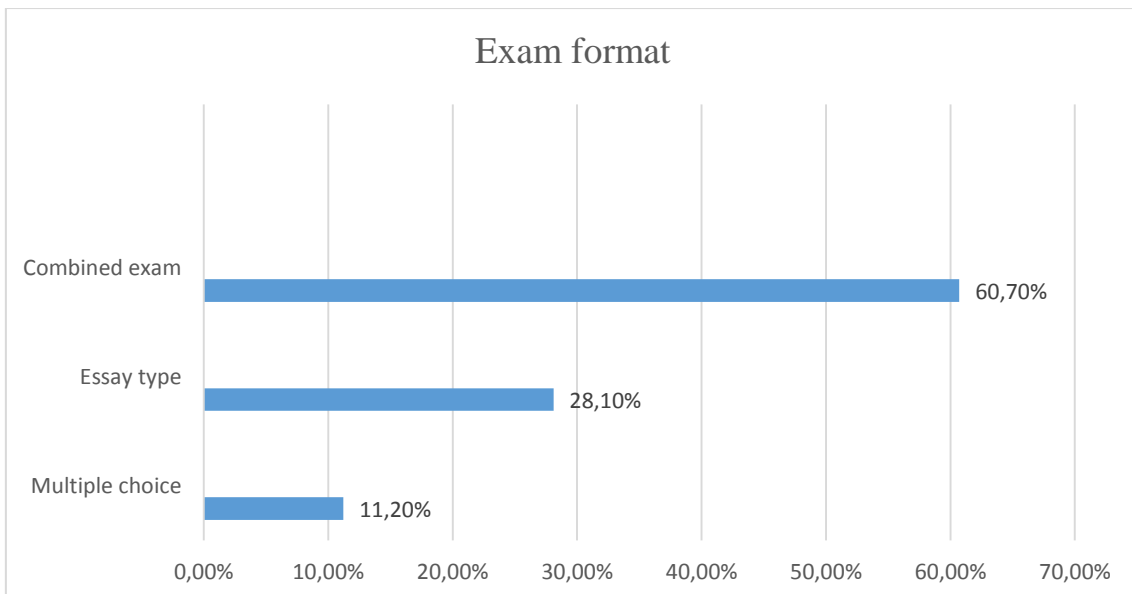


Graph 15. Criteria for oral presentations



About the typical format teachers use for written exams, over half of the respondents (60.7%) prefer a combination of multiple choice questions plus some questions students need to explain or an answer they have to justify, over a quarter (28.10 %) choose essay-type questions, and 11.20 % prefer a Multiple Choice.

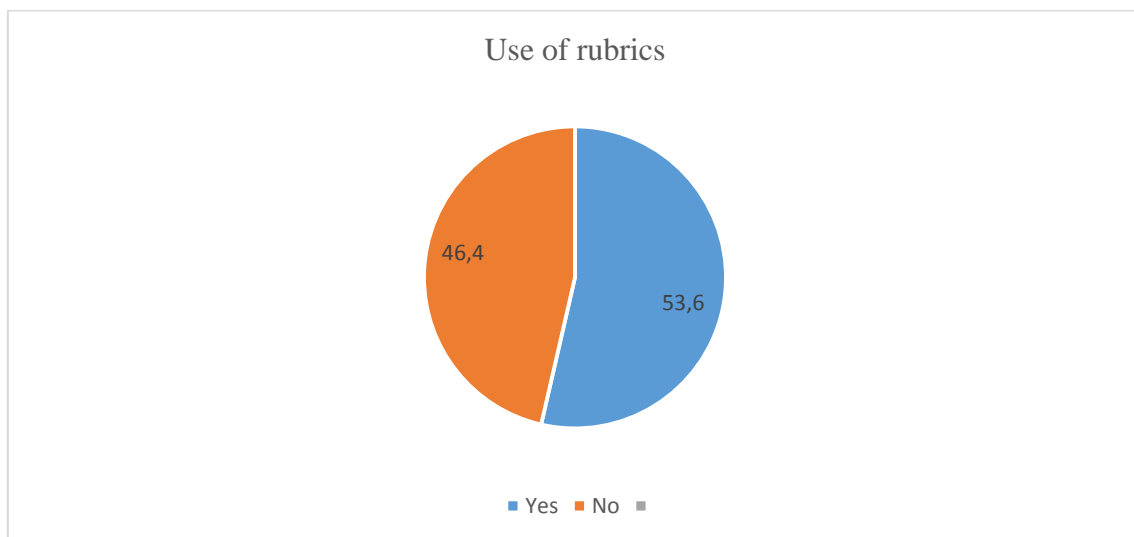
Graph 16. Exam format according to teachers' questionnaires



Some teachers comment that they choose a combination of both because they wish to favor those students who have limited English skills as they are conscious that being able to express concepts through a foreign language is not always an easy task for all students, which at the same time, is a common concern for researchers, and one of the recommendations for successful and fairer CLIL assessment (Bentley, 2010).

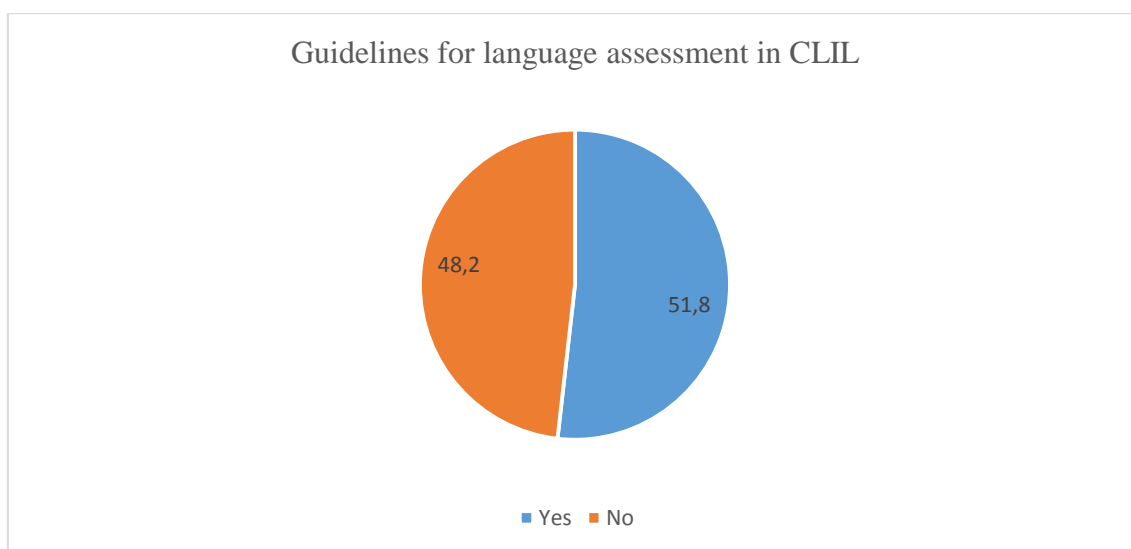
Regarding the use of rubrics for both written and oral production, over half of the teachers (53.6%) say they do use rubrics whilst 46.4% say they do not do so. As the last phase of the research consists of the compilation of assessment tools, no further information was collected regarding the type of rubric - analytic or holistic - in this preliminary stage - nor whether teachers use them for oral or written skills, if they measure linguistic aspects, and which ones (if any).

Graph 17. Use of rubrics by teachers



As for the existence of guidelines teachers use to assess the English language in CLIL subjects, answers diverge radically as 51.8 % state they do exist and 48.2 % state it is the teacher's choice whether to assess it or not.

Graph 18. Guidelines for language assessment in CLIL



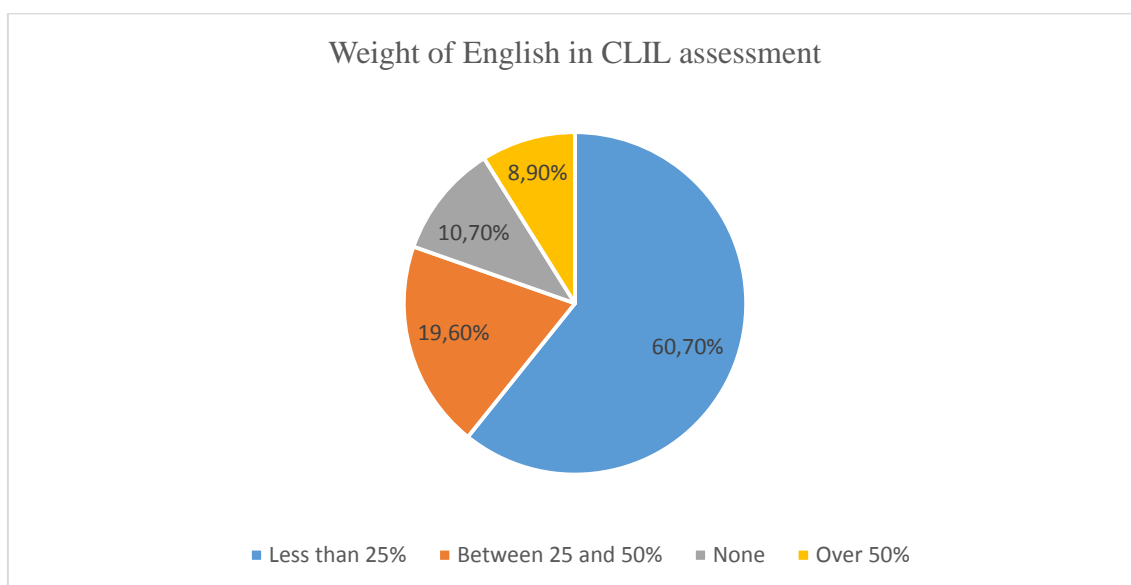
In contrast, teachers agree on the person/people deciding the assessment criteria (content teacher, subject teacher or both) showing that unlike in primary education, roles in secondary education are clearly specified due to the teachers' profiles as specialists in subject(s) and fields. The vast majority of teachers (92.9 %) claim that they take decisions on the assessment criteria themselves since they are the subject specialists as opposed to those who assess in collaboration with the language teacher. Here, it is interesting to consider the teachers' comments about the role of the department in defining the assessment criteria, and how these criteria match the agreements taken for non-bilingual groups. In this sense, a large number of teachers highlight the significant role of the department in this type of agreements, and the fact that assessment in CLIL subjects might

not adjust to Formative Assessment at all but to the criteria common to the rest of the curricular subjects in mainstream education: “Criteria are the same as the ones we use in non-bilingual groups. If we penalize the incorrect use of Spanish, we need to do the same in English”. (TQ- 37); “Criteria are decided in the department meetings, and they are exactly the same criteria we use for “Spanish groups” (TQ- 22). Another teacher explains: “Assessment criteria are decided in each department. Criteria are the same but using a different (CLIL) methodology, and using the English language” (TQ-37), which means that the assessment criteria are the same used in non-bilingual groups but adding the language component which is intrinsic to CLIL. On the other hand, some participants also state that although criteria are usually discussed in the department, the bilingual coordinator is a relevant person in that s/he can help establish criteria regarding the use of the English language, in such a way that these criteria are added to the normal criteria for non-bilingual groups in mainstream education. Bilingual coordination meetings seem to be a relevant opportunity for all members of the bilingual teaching staff to communicate with each other and deal with common matters, one of them being assessment. Finally, those teachers who benefit from the use of the so-called “Action Plan” or criteria for assessing English, express that guidelines come from this document which has generally been agreed upon in a formal meeting. In this sense, some teachers mention the fact that they have received feedback about the type of errors they should penalize, although there is not always a follow-up to see the extent to which these measures are implemented in the school: “In our High School, we received some information about mistakes we should penalize in English depending on the students’ level. But then it’s up to the teacher to follow that or not” (TQ-26). Conversely, two teachers mention there is no real coordination among teachers in the department. This suggests that they might feel something should be done in this line as collaboration among

language teachers and content teachers is paramount in CLIL (Kelly 2014, Pavón & Ellison, 2013).

As regards to the weight that the English language is given in CLIL assessment, as can be seen in graph 19, over half of the teachers (60.7 %) state that it is less than 25 % out of the total mark, under a quarter (19.6 %) between 25% and 50%, several (10.7%) claim they do not assign it a mark, and 8.9% state they assign it over 50% out of the total mark.

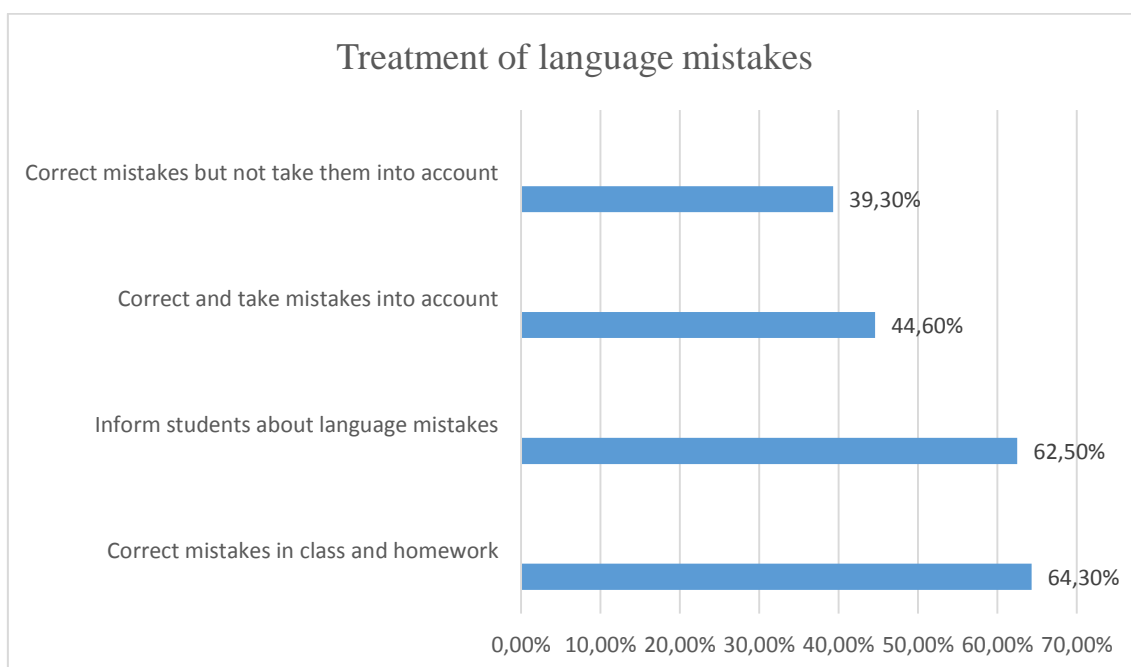
Graph 19. Weight of English in CLIL assessment according to teachers



Following the previous question, the two next ones were aimed at providing further information firstly, on the way content teachers deal with language mistakes and secondly, on the aspects of the English language (if any) which are taken into account. In this question, teachers were given the possibility to select more than one option. Most teachers (64.3 %) admit they correct basic mistakes in class and/or writing assignments, 39.3% state they check basic mistakes, but they do not take them into consideration for

the final mark; nearly half of the teachers (44.6 %) assert they check mistakes and take them into account for the final mark, and a significant 62.5 % confirm that they inform students about the language mistakes they make but do not penalize them or take them into account for the final mark.

Graph 20. Treatment of language mistakes according to teachers

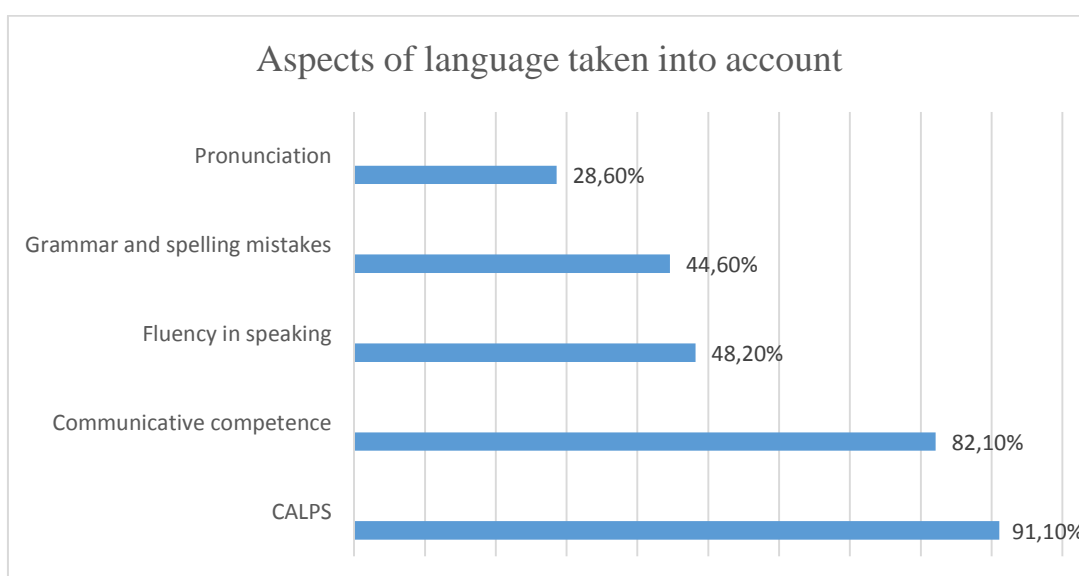


According to the data, it seems that in general, most teachers differentiate between mistakes in the class or homework and mistakes made in exams, the latter being regarded more seriously, which clearly shows that language assessment is far from being fair in this respect. Besides, teachers are aware of the importance of mastering the vehicular language for the expression of content and skills in CLIL subjects, and thus, they usually inform students about their most common mistakes. Regarding the mistakes students make in exams, the biggest trend is to assess them, and to offer students feedback about

them. However, the number of teachers who advocate for not assessing students' language mistakes nearly equals the number of those who do. It might be the case that content teachers are unsure of their role, and feel that even if they are not expected to teach the language as such, some information about language-related issues is expected from them.

As for the aspects being taken into consideration when correcting and penalizing language mistakes in English, and as can be observed from the graph below, a vast majority (91.1%) claim they consider academic or "basic" vocabulary from the subject followed by 82.1 % consider communicative competence as the ability to produce clear messages which manage to communicate content knowledge. Fluency in speaking assignments is the next category to be selected by nearly half of the teachers (48.2%). Similarly, a very significant 44.6% consider grammar mistakes and spelling in written tasks whilst over a quarter (28.6%) consider pronunciation as relevant in CLIL assessment.

Graph 21. Aspects of language taken into account according to teachers



The next questions: “What happens if a student uses a very basic vocabulary or fails to inflect the third person singular correctly? What about those students making spelling mistakes or using incorrect pronunciation?” were included in the questionnaire to check whether, in the absence of official guidelines, any agreements exist which were actually used in assessing students’ work. The majority of teachers point out that these type of mistakes are usually corrected especially when students show poor vocabulary, as they consider students need to master specific vocabulary from the subject: “They need to know the basic academic vocabulary to pass” (TQ-23); “I try to correct the mistakes in English but what I find really important, what I emphasize the most, is that they need to master the academic vocabulary in the subject” (TQ-37); “As long as the message is transmitted in a clear way, if it is easy to understand, and the academic vocabulary is ok, it’s fine for me and I don’t correct the mistakes in English” (TQ-46). As most teachers state, comprehensibility of the message is paramount for a task to be considered successful: “I always try to check the mistakes, and if they are systematic, and students are unable to communicate the message (they impede comprehension), they need to correct them, say that again, and they could be penalized if the message is not clear” (TQ-6). As for the rest of mistakes (spelling, pronunciation and the incorrect use of the third person singular), answers vary greatly. In general, most teachers (30 out of 57) assert they insist on the importance of accurate writing and that they take spelling mistakes into consideration and, consequently, penalize them: “It’s so important for them to try to write in a correct way” (TQ-17); “I correct mistakes all the time and I insist on correct spelling, and pronunciation” (TQ-12). Conversely, eight teachers claim they try to give feedback when mistakes are systematic but do not penalize them at all:



I don't take mistakes into account. In our subject, we focus mainly on students' comprehension and being able to communicate in English, so we let them express themselves regardless of the mistakes they could make. When mistakes are highlighted, we check them later, but we let students to speak freely in English, and that's the way they learn from mistakes. I jot them down, and I focus on them but I don't go further than that (TQ-27).

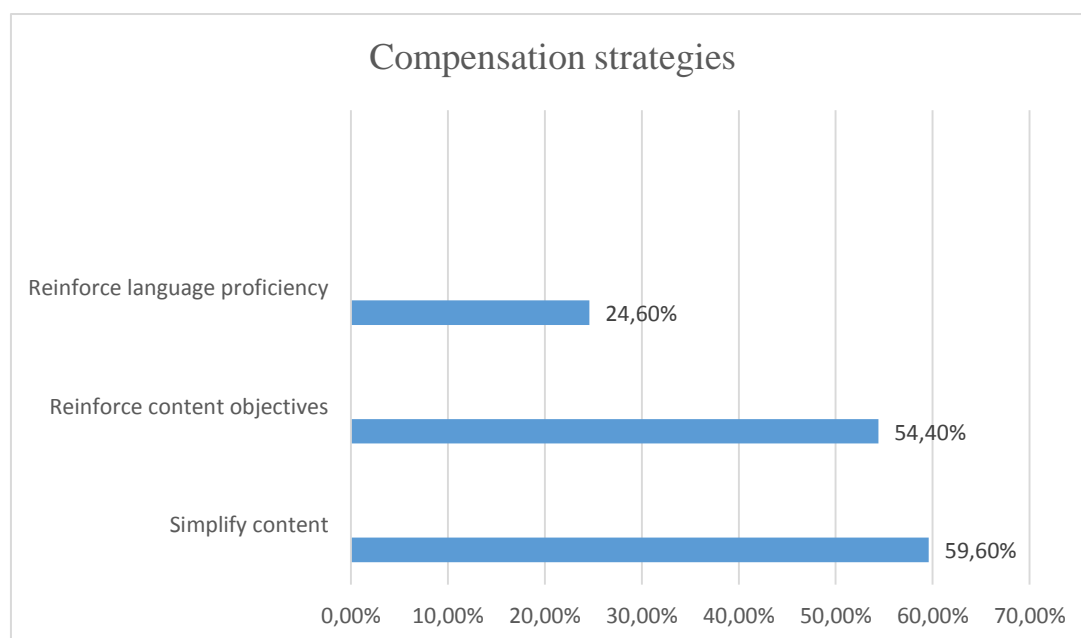
Once again, in relation with the role of content teachers, many of the teachers refer to the language assistant as a helper to work with students' language mistakes: "Mistakes (in English) are corrected in class with the help of the language assistant" (TQ-16); "The language assistant is paramount to help students with English mistakes" (TQ-3); "If I am alone in class, I just check content whilst if I am with the LA, I also take into account these criteria for language correction" (TQ-32). Some others point out to the help of the English Department or even the Bilingual Coordination since as one teacher comments, they are the ones who later deal with those serious mistakes which need remedial work. What seems to be common practice for seven teachers out of 57 is the process of taking notes about their students' mistakes and comment them afterward so that students can properly fix them. Finally, five teachers mention the "Action Plan" as the agreement on the type of mistakes that should be penalized.

The last question in this first part dealt with whether teachers inform students about the content and language goals of each unit of work, 84.2% say they do as opposed to 17.5% who say they do not. However, taking into account teachers' comments about the responsibility they have for correcting language mistakes, it remains unclear whether they mean content goals and language goals too.

The second part of the questionnaire focused on the main problems teachers face in relation to CLIL assessment. The first question centered on the strategies which are

commonly used with those students who need extra help. As can be seen from the graph below, responses vary amongst teachers but over half of the respondents (59.6%) confirm that they tend to simplify content or lower the level of requirement, 54.4% use strategies in order to reinforce content objectives and 24.6% state they count on extra support or specific strategies in order to reinforce language proficiency. No additional information has been gathered on the nature of these strategies though.

Graph 22. Compensation strategies



Again, we can conclude that content teachers tend to rely on strategies dealing with their area of expertise rather than on those related to language skills. This might be due to a variety of factors, namely the lack of teachers' training in language pedagogies on the one hand, and the fact that teachers might feel insecure about their own language

level on the other hand, or it is even possible that teachers might not be at ease devoting time to language-related issues.

The use of students' L1 as a vehicle of expression in CLIL subjects also deserves close attention, as nearly half of the respondents (47.4%) state they sometimes allow students to use their mother tongue when they seem to be unable to express their thought using the vehicular language, and to prevent students keeping quiet in class in the fear of making mistakes in English. As one of the teachers point out: "I let them speak in Spanish when they need to ask a question, for instance, after they have tried in English, and they can't ask the question properly" (TQ-27).

As for the strategies content teachers use in order to favor students with special needs, most of the teachers point out that this is not a real problem in Bilingual Sections. Students accessing the program generally show a better level of English compared to students in the Bilingual Program and consequently, the students' profile is a rather uniform one. Apart from that, teachers complain about the lack of time they have to prepare lessons and find appropriate materials which can suit their students' needs, and the big number of students they have per group, which of course puts extra pressure on the teaching process as it menaces quality. It is important to notice that in this section of the questionnaire, some teachers comment on the problems they find in both the Bilingual Section and the Bilingual Program, and the different levels students have in the Bilingual Program and thus, some of the answers have been discarded as the latter is not the focus of this research. Five teachers out 57 also state they would like to offer extra support to weaker students but unfortunately, this is not always possible due to external factors such as the large number of students per class and the limited class time they have. As one teacher comments, there is always the possibility to set additional

homework, prepare reinforcement activities, and ask voluntary students to help those who need it, and even answer questions in their mother tongue: “Sometimes I can solve doubts, even in Spanish, after the class, if they need so. Sometimes we don’t have that extra time. There’s not always time to do so...” (TQ-37)

Lastly, as for the main obstacles teachers find when it comes to CLIL assessment, many of them refer to the fact that the lack of special training in CLIL is a constant drawback in bilingual education. Other common answers are the lack of coordination in the school when it comes to assessment criteria, and the lack of clear guidelines for assessment in CLIL (five teachers), the different levels of language command we can have in the same class, overcrowded classes (up to 33 students in class) which makes it difficult to use communicative activities appropriate for CLIL such as class debates: “It is so difficult trying to do a student-centered approach with so many people, having so many different levels in the class, sometimes I am tempted to go back to traditional methodologies” (TQ-2). Other difficulties content teachers mention is the lack of good materials in some subjects - which inevitably forces teachers to prepare their own, with the consequent investment in terms of time -, the uncertainty about the real role and weight of English in a CLIL subject as has been already commented, the lack of clear objectives in the CLIL units- which might lead teachers to focus specifically on content-related goals and set aside language-related ones-, the opposition to Bilingualism from other teachers in the school and the department- which might impede content teachers to design specific content and language goals and assessment criteria suited to the CLIL purpose- and the limited amount of CLIL hours.

From the data gathered in the teachers’ questionnaire, we can conclude that a vast majority of teachers (92.9 %) use written exams as the main tool for assessment, along

with informal class observation to measure students' participation and attitude. Oral exams are rated by over a quarter of respondents (30%), which contrasts with the scarce use of alternative assessment tools, namely checklists (5 %), portfolios (14 %), and self (24 %) and peer-assessment (21 %). About the use of guidelines to grade linguistic issues, the majority of teachers (44 out of 57) declare that no common guidelines are used except for those created in the department or school. Likewise, no common criteria exist for the assessment of oral presentations although respondents assert they consider fluency and pronunciation above all. Regarding the exam format, over half of the teachers (60.7 %) prefer combined exams as opposed to 28.10 % who choose essay format, and only 11.20 % opt for multiple-choice questions. In relation with the use of rubrics, opinions seem to be divided since over half of respondents (53.6 %) who use them as compared to just under half of the respondents who do not (46.4 %). The same accounts for the use of guidelines for grading language as slightly over half of the teachers (51.8 %) state they use them, and 48.2 % state they do not. What seems to be clear, however, is that the responsibility for assessment in the content subjects lies solely with the content teachers rather than language teachers.

The weight of English is another interesting aspect in which teachers' opinions seem to be divided: over half of teachers (60.7 %) give language a value less than 25 %, under a quarter (19.6 %) give it a value between 25 and 50 %, and only 8.9 % assign it over 50 % of the final grade. As for language assessment, the biggest trend (62.5 %) is to inform students about their language mistakes, and nearly half of the teachers (44.6 %) assert that language is corrected and penalized in exams. In this assessment of language, academic language or CALP, is regarded as the most essential aspect by teachers (91.1 %), followed by the communicative competence - as the ability to get

the meaning across - (82.10 %), fluency in speaking tasks (48.2 %), grammar and spelling (44.6 %) and pronunciation (28.6 %). Finally, as regards the compensation strategies teachers use, and the main obstacles in conducting assessment, the data reveal that teachers prefer to simplify content (59.6 %) and reinforce content objectives (54.4 %) rather than reinforcing linguistic aspects. About the main obstacles they find, these are related to the lack of training, common agreements and clear guidelines along with other minor aspects such as the student ratio in the CLIL classes and the lack of good materials for assessment.

## 5. 2. Focus groups

In the focus groups, interaction among participants was based on a list of topics in relation to the main obstacles teachers find in CLIL assessment, the instruments they commonly use, the way informal class observation is used as a tool to measure students' progress, and whether language competence has a direct influence on the grade they assign to students' output. Attention was also given to the way teachers deal with the absence of specific guidelines for CLIL assessment - a common complaint according to data obtained from teachers' questionnaires and informal conversations. In this sense, the main goal was to find out whether teachers communicate with colleagues in their department and/or at school to know how to deal with assessment issues, and whether they have coordinated on that matter or have reached any agreements so far on topics such as the weight of the foreign language in CLIL assessment or the aspects that should be marked down (if any) in assessing the language. Other relevant questions

covered topics such as the treatment they give to language mistakes, and the use of rubrics among other assessment tools.

### **First focus group interview**

The first focus group interview took place in March 2015 in the library of the High School, in Rivas-Vaciamadrid. It involved 12 teachers- permanent and temporary staff- along with the Bilingual Coordinator. In this first group, the discussion focused mainly on the weight of the English language in CLIL subjects along with the criteria teachers have to correct language aspects and the teachers' roles. Teachers' views revealed that they found it extremely difficult to assess content knowledge without taking language proficiency into account. In fact, as they pointed out, the difficulties which bilingual education can entail in terms of students' production in the foreign language has always stood out as a controversial topic in the school, which attracted most teachers' interest. Consequently, this issue had been previously discussed on many occasions during school meetings since the implementation of the bilingual program three years earlier. Finally, in the academic year 2013-14, the teaching staff agreed on an improvement plan for writing skills or "plan de mejora de la expresión escrita" to be used by all content teachers in both non-bilingual and bilingual groups. The plan was aimed at improving writing skills in English and Spanish and for that purpose, it was initially devoted to agree on joint rules for the presentation and organization of students' class notebook and academic work, as well as for the outline of exams and project work. Those actions led teachers to agree on the assessment criteria regarding writing skills and grammar mistakes in exams and students' work. After having analyzed typical mistakes and having created a framework for written proficiency, both assessment and

grading criteria were modified accordingly in all the subjects, and families were informed about these guidelines through the students' school diary.

Regarding the role of English in CLIL assessment, teachers overtly showed their concerns about the topic and immediately started asking about the existence of general guidelines as they complained about the lack of information and teacher training in CLIL issues. "We don't have much idea about it, to be honest. What are we supposed to do about assessment?" (TFG-M). They also emphasized that their main goal as content teachers in relation with language is that students are successful in acquiring academic vocabulary or what they term as "CALP, the specific vocabulary from their subjects". In this sense, it is interesting to notice that although CALP is more than just academic vocabulary, teachers tend to simplify the concept to refer to the specific language of the subject.

We always emphasize the vocabulary of the subject. Students have to learn it and know how to use it to express content. In Music, for instance, it is essential to know ordinal and cardinal numbers, they learnt that in primary education. As for the new concepts, or definitions, etc. above all, they are names in Italian. Well, I suppose I can overlook some spelling mistakes. (TFG-A)

When asked about error treatment in CLIL subjects, all the members of the focus group seemed to be clearly concerned about how to deal with language errors as they commented on the most typical grammar mistakes- the -s in the third person singular, starting a sentence using "that" which is obviously Spanish-like word order: "Some mistakes need to be fixed immediately. Otherwise, they go viral..." (TFG-J).

Nevertheless, although teachers recognize the need to correct students while speaking, most of the teachers tend to favor intelligibility over accuracy. In this regard, it is interesting to see the tendency they show to contrast accuracy and fluency as if the first



did not help the latter in the process of content expression, as can be seen in the following comment: “I usually focus on whether the writing is easy to understand. I go for comprehensibility because CLIL is a communicative approach” (TFG-C). In fact, accuracy in writing had also been a controversial issue they had been discussing for years. As the different departments were not in agreement on the best ways to deal with language mistakes in CLIL subjects i.e. whether they should just be highlighted or also marked down, they asked the English language department for advice. Apparently, although the English teachers had not agreed on a taxonomy of errors themselves, this request proved useful for them so as to identify common mistakes which were later used to design the improvement plan for written skills. However, despite these agreements, it might be the case that in current practice, each teacher corrects what s/he finds appropriate depending on the level, the subject and the group with a focus on fluency over accuracy: “I sometimes come across sentences with no -s in the third person singular but they express so much content knowledge that for me it’s fine, it is enough” (TFG-R); “I know there were some agreements about the way we correct but we also need to look at other aspects which have not been considered, and which are also necessary” (TFG-E). In this sense, and regarding the joint rules they agreed on the improvement plan for written skills, it is interesting to notice that although the plan was globally perceived as positive, some teachers complain that there is more flexibility in CLIL subjects than in Spanish:

This is like when a student goes and starts a definition using “when”. We don’t accept that in non-bilingual groups. Students can’t start a definition using “when” in Spanish. But then we allow them to do that in English. You can even find a definition like that in a textbook! So of course, I believe we take comprehensibility rather than accuracy or grammar mistakes into account. (TFG-M)

The discussion also raised issues about the role of the content teacher as opposed to that of the language teacher, and it revealed the fact that content teachers seem to be uncomfortable when correcting and grading language mistakes: “I am afraid if I devote too much time to check and fix English mistakes, I will end up being a teacher of English. However, my students sometimes don’t know how to express content in my subject...” (TFG-P)

Finally, an additional difficulty that teachers have dealing with the weight of English in CLIL is that they are also afraid that in some situations their language level might not be good enough, and they might make mistakes that students could repeat, as one the teachers state: “Sometimes, I also need to have a grammar or a dictionary around when I am grading exams. Yes, that happens sometimes, to make sure this guy is writing this and that the correct way. How am I supposed to do that if I am not sure to have that proficiency level in English? I am a Science teacher, not an English teacher” (TFG-C)

As in some of the comments from the teachers’ questionnaire, the participants also expressed their concern about the difficulties they find when selecting appropriate assessment tools for CLIL contexts. Despite the presence of the improvement plan for written skills in this school, the general procedure for assessment criteria in Spanish secondary education is set by the didactic department which usually comprises bilingual and non-bilingual groups. Thus, exam formats and assessment tools are usually designed for non-bilingual groups, namely, tests including essay questions. These essay parts might be problematic for bilingual groups even in the case of Bilingual Sections where students need to express content knowledge through productive skills- being writing the preferred mode- which is challenging since the

language level in English is lower than in Spanish: “The main problem is that regardless of whether you have bad, good or excellent materials, when it comes to assessment tools, I mean the way exams and tests are designed, it’s completely different. I don’t know about you, but I can’t expect my students will be able to write in English the way they would write in Spanish” (TFG-A).

As for the absence of clear guidelines for CLIL assessment, comments showed that teachers agree that the Ministry of Education or Regional Government of Madrid should offer specific guidelines regarding assessment regulations for bilingual schools in the CAM Bilingual Program. As respondents put it, the assessment tools designed for non-bilingual groups are not in line with bilingual education, and a great deal of effort needs to be made to create specific CLIL materials which are not mere translations from Spanish. Apart from that, in the absence of guidelines, more freedom should be given to bilingual schools so that assessment tools, methods and criteria can be set apart from those recommended by the didactic department which are common for both bilingual and non-bilingual schools. In fact, a common complaint by parents, they assert, is that bilingual students can have ‘easier’ exams than their non-bilingual partners, which some people think can devalue bilingual education:

Besides, we have that pressure from the parents. When families come, they tell us non-bilingual students have much more difficult exams, essay-type exams whilst bilingual groups sometimes do that, but not always, they have these matching activities, more visual support...But we are aware we can’t expect the same linguistic level in the other groups, the Spanish groups, that’s a fact. (TFG-M)

Regarding class observations, which according to the findings in the questionnaires is an essential component of the assessment process, no significant information was

gathered. Teachers assert they use informal assessment techniques in the form of class observation, checklists and revision of homework, workbooks or “cuadernos de trabajo” on a regular basis, as is common practice among secondary teachers. However, the frequency and weight they assign to this informal assessment is far from being systematic but rather depends on individual teachers, and the use of portfolios and peer and/or self-assessment is also rare among teachers: “We correct the activities at the end of the term, we assess the didactic units. This is the best way to check they were working on a regular basis. No, we don’t really use the portfolio” (TFG-A); “I don’t know about the rest of the teachers in the department, but I don’t use self or peer-assessment. The students do know about their progress because the activities are corrected in class. Activities are always corrected here” (TFG-O).

The same can be said about rubrics, which according to respondents have been designed by the bilingual coordinator and a language assistant to measure oral skills in English lessons but they are not present in content assessment yet. Practicalities regarding the urgent need to agree on basic guidelines among departments and share rubrics and model exams were also raised at this point, along with the compromise by all members to continue in this line of improvement they started with the implementation of the plan for written skills. However, regardless of the joint agreements or guidelines in the school, there seems to be a lack of real coordination among teachers as can be observed by some teachers’ comments (“I don’t know about the rest of the teachers in the department”/ “I am not sure what other people do”).

### **Second focus group interview**

The second focus group interview took place in May 2015 in the meeting room of the High School in Getafe, a town in the South of Madrid. It included 15 teachers, five

language assistants and the bilingual coordinator who expressed her wish to include all the members in the bilingual team in the meeting. First of all, it is important to point out that this high school has extensive experience in bilingual education since it was one of the first MEC-British Council Project centers back in 2006 until they became part of the CAM Bilingual Program in 2010. This has given the teaching staff a deeper understanding of CLIL methodology, materials and the functioning of a bilingual school and above all, a strong commitment by all members in the bilingual group to work in collaboration with each other as will be shown later on.

Although the questions were the same as in the first focus group, before discussing the weight of English in CLIL assessment, the conversation started with the main assessment tools they use for CLIL subjects, and the assessment and grading criteria. In this regard, all the teachers indicate they use both open-ended and closed questions: fill in the gaps, multiple choice questions, short questions and answers and essay type questions:

I usually combine the two: short and essay-type questions. The multiple-choice type and longer questions. And I add images so that they can complete the task with the help of visual support. I do it that way because I know there are also visual students, and they learn this way, I don't want the final grade to be so influenced by the CLIL methodology (TFG-N).

As can be observed from the quote above, teachers are conscious that the lack of proficiency in the foreign language might hinder the expression of content, and thus apart from traditional essay-type questions, they try to offer some matching or multiple-choice questions in which students can demonstrate content knowledge and skills without being burdened by linguistic issues. Also, in more practical subjects such as Technology or Arts and Crafts, students are asked to solve problems or demonstrate

skills. Again, the main goal for teachers seems to be vocabulary knowledge since students are required to master the specific academic vocabulary from a subject: “There are some questions in which they have to write a definition so that I can see they master the concept, they have understood the subject” (TFG-MO)

Other assessment tools which respondents use in order to give prominence to language in content subjects are oral presentations. This is a regular requirement in most subjects since students need to prepare them on a monthly basis whilst some others ask for group expositions once a week. When asked about the criteria to assess oral expositions, teachers agree that the focus lies on content knowledge, presentation skills such as the ability to create a good Power Point presentation, and to address the audience appropriately. Besides, they recognize they assess fluency over accuracy, i.e. they expect students to be able to express themselves with acceptable fluency according to their level although they might make some mistakes or inaccuracies: “I guess the most important thing is whether they know how to express content knowledge in English. Rather than reading from their cue notes, they have to be able to speak fluently and confidently, and of course, to know the vocabulary” (TFG-S). Oral presentations are important because they allow students to show understanding of the subject and express it. In relation with content expression, and in order to abandon memorization in favor of fluency in oral presentations, some teachers also expressed their concern about the students’ need to develop critical thinking and skills as is noticed in Bloom’s taxonomy where students can move from LOTS- remembering and understanding knowledge- to upper-level HOTS, in which they are able to apply, analyze, evaluate and create from the knowledge they have acquired: “Then I can see if they understood a historical fact. I check they were able to understand not just memorize concepts and facts, to

understand that a historical fact comes as the result of other direct previous factors. This is the type of knowledge that people in our department acknowledge is difficult to measure by means of a multiple-choice test” (TFG-R). Another teacher points out: “The most important thing is the message. The message should be transmitted in a clear way. In this sense, I’d say it is important to demonstrate they understood the main contents, that important information was assimilated. They also have to be able to reflect critically, in terms of cognition” (TFG- E).

In Arts and Crafts, teachers state that portfolios are used to measure students’ progress, but no additional information was offered on the topic. On the other hand, teachers reveal that the use of self and peer-assessment techniques are not current tools yet.

In relation with the selection of assessment tools, no difficulties were highlighted. Nevertheless, teachers noted that they sometimes miss good materials for exams and tests in their textbooks. Although the quality of materials has improved over the past years, some teachers complain that most CLIL materials are translations from Spanish textbooks and consequently, the assessment tools do not serve Spanish CLIL contexts very well.

As regards informal assessment, class notebooks or “cuadernos de clase” are of high importance for teachers in order to check students’ daily work. This process of gathering students’ pieces of work is rather systematic among teachers in the school. The weight of these assessment tools is set by the department and it is also made public and sent to first and second graders’ families at the beginning of the academic year so that both students and parents know about the school’ assessment and grading criteria

in advance. These notebooks are measured using quantitative marks along with some qualitative comments which students can read and learn from.

Informal assessment, teachers assert, is complemented with other tools such as class observation, checklists, students' behavior and active class participation and interest-known as "attitudinal contents" in Spanish secondary education. Criteria for informal assessment is also set by the department- not the bilingual team- as is common for both bilingual and non-bilingual groups, and it can amount to approximately 20% of the final mark. According to the data from the teachers' questionnaires, the rest can be obtained by one or more written tests, which shows a big prevalence of written tasks over oral tasks and other forms of assessment.

Moving on to the weight content teachers assign to English in CLIL assessment, as in oral presentations, teachers (overtly) focus on fluency over accuracy but they insist that in production activities, the students' level is taken into consideration: "In assessment, language is part of the final grade, but the most important aspect is always content, and as such it is considered over the English language" (TFG-L).

Apparently, students with a good command of English do not have difficulties in expressing content knowledge. The problem arises with those students who are less proficient in English and whose final grade can be affected by their English level. It might be the case - they point out - that these students find that the foreign language represents an additional challenge and they could (possibly) obtain better results in non-bilingual programs.



In both oral and written productive skills, some actions and agreements have been made. Contrarily to the criteria in some other schools<sup>16</sup>, where the weight of English in content subjects is clearly specified by each department, some general joint rules have been agreed from the introduction of the so called 'Action Plan'. This Plan was implemented in the academic year 2014-15 as a strategy to prevent the fossilized errors which teachers observed had started to be rather common among 3rd and 4th graders. The teachers worried that students' language proficiency might be compromised by the overt focus on fluency, and consequently, a group of English teachers supported by the bilingual coordinator met to agree on criteria to grade language mistakes in both English as a foreign language and CLIL subjects so that they could subtract from two to four points in the exam or final mark. Although typical mistakes are the same for all subjects, they are penalized differently depending on whether they occur in content subjects or in English as a foreign language, English teachers being stricter regarding language accuracy. Nevertheless, apart from the criteria in the 'Action Plan', teachers point out that some additional factors regarding students' level, effort and attitude are also taken into account. The language mistakes in this plan are the ones which teachers supposedly consider for assessing and marking down students' written output in essays and exams (See appendix 11).

Finally, another problematic issue for teachers in this focus group was how to deal with these language mistakes especially during students' oral participation in class and oral presentations. At this point, they asked about European guidelines on this subject matter, at the same time that they insisted on the importance of accuracy, and they pointed out that some errors cannot be overlooked and need to be corrected

---

<sup>16</sup> Personal communication with students and teachers, academic year 2015-2016

immediately: “I have this group, they are the best group in the 4rd grade (4º ESO). And then there are these two boys who are so confident, self-assured, they have very fluent English but they make mistakes all the time, so I also need to stop them at times. Otherwise, they would think they are doing it fine and they aren’t...” (TFG-F).

About the duality between fluency and accuracy, some teachers clarify it is still fluency over accuracy the criterion that prevails among them, and that they tend to let students talk without correcting unless it is a very serious mistake. One teacher exemplifies her teaching procedure when she describes the way these mistakes can be later retrieved in class and come under scrutiny as in the ‘Language Clinic’ (Coyle, Hood & Marsh, 2010) which, she points out, is very common practice in this high school. As for the type of mistakes which have been typified in the Action Plan, evidence shows that the focus is on grammatical accuracy, namely correct verb tenses, the obligation to include the subject at the beginning of declarative sentences- a typical mistake among Spanish students- and correct comparative and superlative forms, to name just a few.

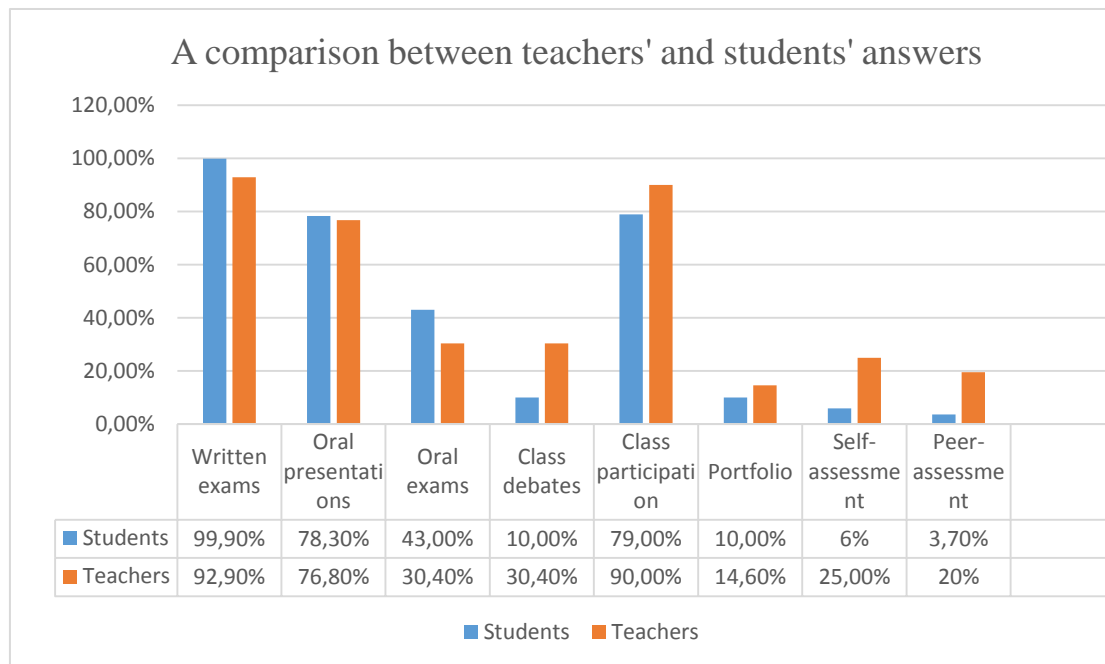
The focus group interviews helped me in the first place to compare findings with the results from teachers’ questionnaires, secondly, draw conclusions, and finally, frame the structure of the next step in research - the students’ questionnaires and interviews. It also offered me an in-depth view and understanding of the topic of CLIL assessment in Madrid, which clearly has the challenge of following the same guidelines that in non-bilingual schools even if the bilingual program deals with a different reality. According to the data collected, the most frequent assessment tools are exams combining multiple choice and essay type questions, and offering visual support. Class notebooks are also highly regarded as the best way to check students’ work over time, and some oral presentations are used to help students develop their language skills. Teachers

expressed their concerns about the lack of guidelines, resulting in the creation of their own agreements, and the difficulties they find in the implementation of alternative assessment such as portfolios and self and peer-assessment due to the presence of standardized exams, especially the university entrance exam. Another significant issue was raised in relation with their role as content teachers since teachers do not consider themselves as language experts, and thus feel they might not be in a position to deal with language-related aspects, although they recognize language is paramount in the expression of content and skills. In this sense, they insist they focus on academic vocabulary along with grammar, and do not penalize language mistakes unless the message is not clear. Finally, they point out to the need of better assessment materials which are suited to CLIL and not just secondary education.

### 5. 3. Students' questionnaire

As in the teachers' questionnaire, the analysis of results in the students' questionnaire combines quantitative and qualitative approaches. Besides, key ideas from the teachers' questionnaires and the focus groups were cross-referenced in order to draw general conclusions and highlight disparity amongst responses. Last but not least, the participants' additional comments from closed multiple choice questions were analyzed using the same pattern. The following bar chart compares students' answers with those by the teachers in the questionnaires:

Graph 23. A comparison between teachers' and students' answers in questionnaires



Regarding the assessment tools teachers' use to assess students' performance, it is evident that written exams are still the most common tool for assessing students' work according to 99.9% of students and 92.9% of teachers. As was commented in the analysis of teachers' results, standardized exams are still a very widespread practice among lower and upper secondary teachers whilst oral presentations seem to be growing in popularity (78.3% and 76.8% respectively) in Bilingual Schools. Also, almost half of the students (43%) report they take oral exams as compared with 30.4% of teachers declaring they do so. As for class debates, the findings seem to be contradictory as only 10% of students declare they do them in class opposed to 30.4% of teachers who asserted they are a common practice. Although standardized or traditional exams are not representative of the Formative Assessment and alternative tools which are needed in CLIL contexts (Coyle, Mehisto and Marsh 2010, Kiely 2011),

we can observe the growing tendency of using oral presentations and exams as a clear attempt to focus on students' oral skills, which have been traditionally left aside in Spanish education.

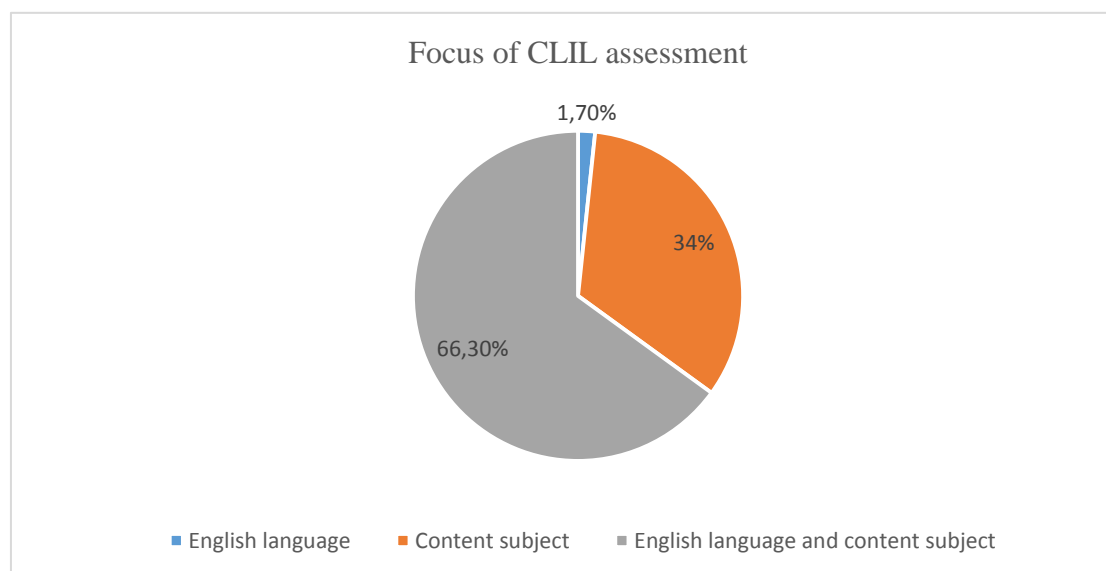
Students' class notebooks are also considered highly important by a vast majority of respondents (79.4%), which coincides with teachers' responses in the form of comments about their daily assessment practice. This work compilation is regularly supervised by teachers mainly as a tool to check homework completion, and in some cases to correct some practical activities due to the relevance traditionally given to "procedural contents" in the Spanish secondary education curriculum along with concepts and attitudinal aspects (Madrid and Hughes, 2011). Class participation is also observed by students as essential for assessment as is reported by 79% of respondents, which confirms the teachers' answers (90%) in the questionnaires. In general, students assert that class behavior is considered by teachers as a relevant indicator of good class performance along with active class participation (78.9%). This class behavior conforms to criteria to assess attitudinal aspects although apparently there are no common guidelines regulating them, but it is left to schools to interpret criteria. When asked about other assessment tools, seven students refer to homework, and three to written assignments and voluntary activities. Less attention is given to the use of the portfolio, it being only 10% of students who report using it, a figure which is close to teachers' opinions (14.6%).

Finally, as for self and peer-assessment, which are also indicators of formative and innovative assessment practice, these are rather infrequent according to students' responses. In fact, there is no statistically significant correlation between teachers' and students' opinions on these two techniques. Thus, only 5.9 % of students assert they

assess their own work as opposed to 25% of teachers, and only 3.7 % of students state they do pair-assessment activities as opposed to 19.6% of teachers.

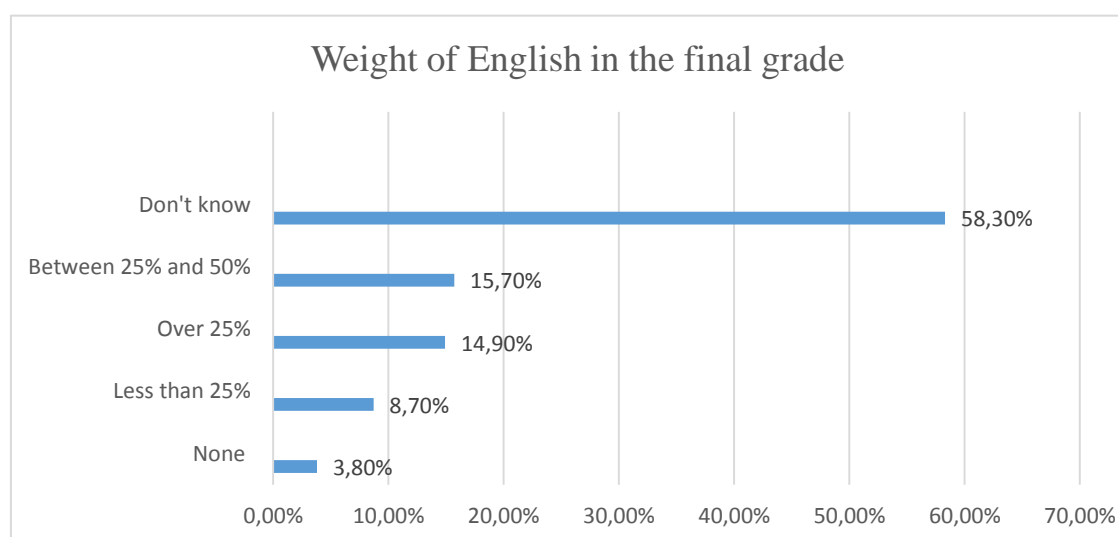
The next question dealt with the main focus of CLIL assessment- the content, the language or both content and language-related aspects. As can be seen in the chart below, according to most students (66.3%), both language and content aspects are taken into account by teachers in assessment. Conversely, 34% of respondents report that content is the only focus and 1.7% consider the focus lies in language-related aspects. It seems thus that students perceive language as an essential component in assessment, even though only 44.6% of teachers declared they check language mistakes and take them into account for the final mark.

Graph 24. Students' questionnaires: Treatment of language mistakes



About the weight of English in the final grade, over half of the students (58.3%) report that they have no idea about the weight of English in the final grade as opposed to 15.7% students who think it is measured with between 25% and 50%, 14.9% students who consider it is measured over 25%, even if they are not sure of the exact weight it might have, and only 8.7% who state that it is measured with less than 25% of the final grade. In this sense, it is noteworthy that most students have no knowledge about the approximate weight of English in CLIL subject assessment.

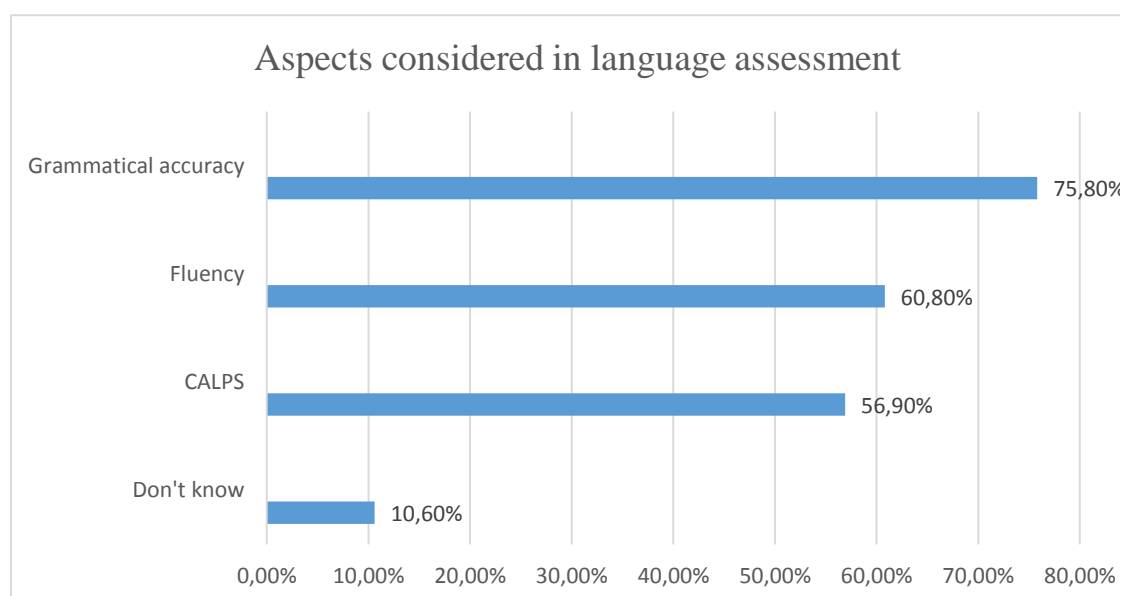
Graph 25. Students' Questionnaires: Weight of English in the final grade



After having confirmed the relevant role of language in the CAM Bilingual Program, another aspect which it was necessary to deal with - as responses among participants in both the teachers' questionnaires and focus group interviews vary - referred to the aspects considered in language assessment. In terms of fluency, which is rated 60.8% by students, and the use of academic vocabulary (56.9%), the findings coincide with those of the teachers' questionnaires and comments in the focus group interviews.

However, grammatical accuracy is perceived by the vast majority of students (75.8%) as the most important aspect content teachers take into account when it comes to language assessment in CLIL. Grammatical knowledge is considered an important component of language knowledge as is frequently viewed as the sine qua non condition of task accomplishment (Bachman, 1990; Weigle 2014), and it seems that teachers tend to base their language assessment on grammatical ability related to error counts rather on vocabulary richness, text structure or clarity of ideas in the text.

Graph 26. Students' questionnaires: Aspects considered in language assessment



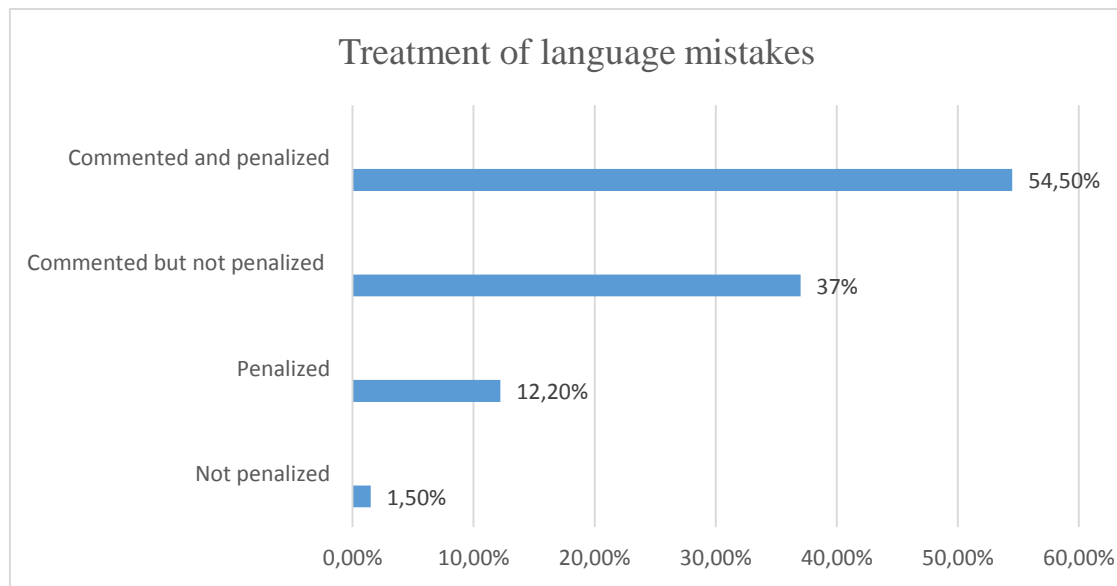
In this question, students were given the possibility to comment on other aspects considered by teachers in assessing the language but only a small number of students provided feedback and consequently, it was difficult to draw any inferences. Two students refer to fluency and the ability to communicate the message successfully, one student reports pronunciation is important in speaking skills, another one points to the



importance of mastering the vocabulary of the academic subject, and the last student states that all aspects mentioned in the question are important to obtain a good grade but maybe being able to communicate avoiding grammar mistakes is paramount.

As for the treatment of language mistakes, over half the students (54.5%) state that English mistakes are commented on by teachers in class so that students can correct them, and they are also penalized in written work and exams. On the other hand, 37% of students report that language mistakes are commented on in class so that students can correct them but that they are not marked down on those mistakes. Finally, several (12.20 %) students point out that they do not receive any kind of linguistic feedback by their content teachers but they are penalized if they make language mistakes, and 1.50 % report that they do not receive any kind of linguistic feedback by their content teachers nor they are marked down for them in exams. Here it is also interesting to remark the fact that although content teachers do not necessarily feel responsible for language-related objectives, they tend to check and take them into account while they also inform students about those mistakes so that they can be properly corrected. This last aspect is significant as it reveals the teachers' implication in the bilingual program.

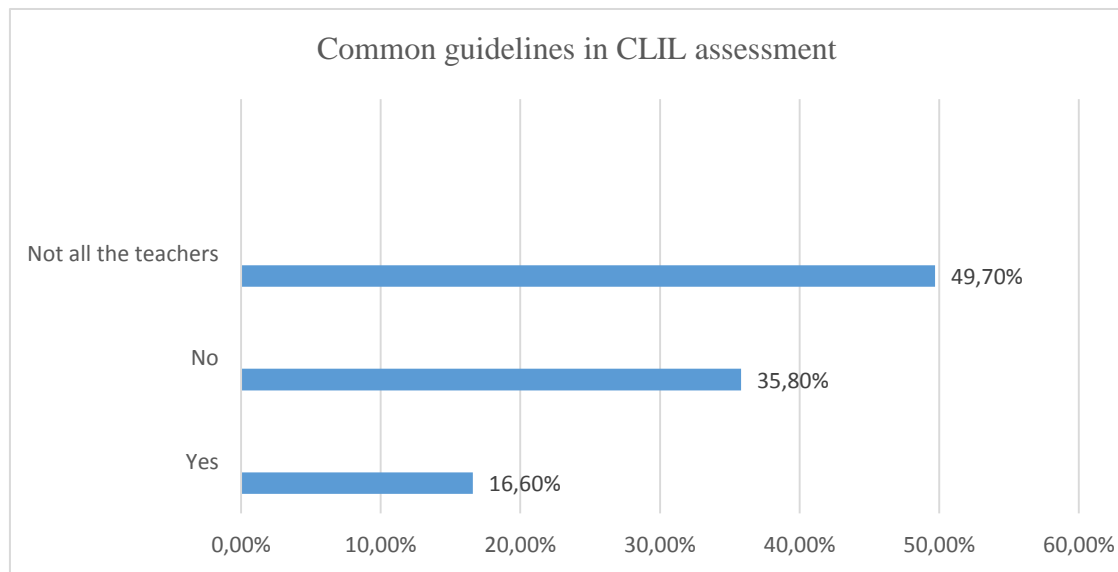
Graph 27. Students' questionnaires: Treatment of Language Mistakes



With regards to the person in charge of assessing content subjects, it is clear from both students' questionnaires (98%) and teachers' questionnaires (99%) that this role corresponds exclusively to content teachers. However, language teachers can be consulted at times for advice on language mistakes, and language assistants can check students' language production in class.

Finally, as regards the presence of common guidelines shared by teachers in the high school, roughly half of the students report that not all the teachers use the same guidelines (49.7%), 35.8% say they do not, and a small number of respondents (16.6%) declare they do. Students' opinions in this regard closely resemble those in the teachers' questionnaires in which 51.8% assert they use them whilst 48.2% say they do not.

Graph 28. Students' Questionnaires: Common guidelines in CLIL assessment



Thus, from the data collected in the students' questionnaire, we can conclude that the most frequent assessment tool is the written exam, followed by class participation, oral presentations and oral exams. Several students also report the use of portfolios but self and peer-assessment is rated as very rare. As for the main focus of assessment, over half of students seem to be clear about the prevalence of both content and language-related issues, and over one quarter state that content is the only focus as opposed to a very limited number of respondents who state language is considered over content. However, over half of respondents state they are unsure about the real weight of English in the CLIL subjects whilst several students think it is measured with over 25 % of the final grade, and a very limited number of respondents state it is weighed less than 25 %. About the aspects being considered in this language assessment, a vast majority assert grammatical accuracy plays a significant role followed by fluency, and CALP. These language mistakes, they consider, tend to be commented on, and marked down by slightly over half of the teachers in contrast with some others who consider they are

not marked down. Finally, it is also significant that although students know about the existence of guidelines in the school, nearly half of the respondents consider it is up to the teacher whether to follow those guidelines or not.

#### 5. 4. Students' interviews

To start with, and in relation with the type of mistakes which are penalized, students state that language mistakes, and more specifically spelling and grammar mistakes, are penalized by teachers especially in exams and written homework, which confirms the findings in the students' questionnaires. In fact, all the students report that grammar mistakes are always under close scrutiny by teachers, and 22 students out of 24 state that spelling is also essential for a good grade. However, it seems that there are no common guidelines, and as was previously pointed out in the section devoted to the focus group interviews, on some occasions criteria depend on the subject and the teacher. Even in those cases where, apparently, teachers have agreed joint rules, students comment that not all the teachers finally use the criteria or that the criteria are used differently depending on the teacher. In fact, three students point out that these spelling and grammar mistakes are not so important in some subjects in which they do not need to write essay-type questions such as Physics and Chemistry or Arts and Crafts. It seems also that the higher the level, the higher the linguistic standards that are set by teachers. As one student points out, the level required becomes more demanding in the second grade of the Bilingual Section as compared with the first grade: "Last year (1st grade-1º ESO) we were told that grammar, I mean English, what English is, was not being taken into account, but it's different this year; they will take spelling mistakes into account, and all that" (SI- A1). Another student points out that in the third

grade, teachers are more demanding about the language component in CLIL: “Because we are supposed to have a higher level, and (grammar) mistakes do matter” (SI-A3). In this sense, it is important to bear in mind that students pursuing the Bilingual Section itinerary are the ones with the highest command of English in the high school and, consequently, teachers tend to set the highest standards for these groups as compared with the Bilingual Program groups. It is also frequent that for these groups standards are raised and common guidelines regarding assessment criteria are also designed as in the case of “The Action Plan” or “Plan de mejora de competencia lingüística” which were described in the focus group interviews. Usual grammar mistakes, according to students’ opinion, relate to verb tenses, the lack of the subject in declarative sentences, subject-verb agreement - especially in the third person singular - and wrong comparatives and superlatives. In this sense, as can be observed from the following two comments by students, grammar can greatly influence the way the message is delivered and how the teachers correct the task: “If you fail to write something well in English, it will mean a lot of deduction from the final mark. Sometimes, you may know the contents in Spanish very well, and then you make a mistake in one grammar part and then it does count” (SI-A5). “In verb tenses, where we sometimes make mistakes, they can also lower your grades” (SI-B2).

Nevertheless, there is no clear evidence about whether these comments refer to the grammatical inaccuracies which do not obscure sentence meaning or on the contrary, to those mistakes which impede or obscure meaning. As Mohan and Huang (2002) state, it is difficult to see whether the focus is on correcting language forms as grammatically correct or whether grammar is used appropriately to convey a meaning in functional contexts. On the contrary, some other students point out that some

grammar mistakes are considered in terms of correctness or incorrectness rather than in functional terms. This is the case, for instance, of the subject omission in declarative sentences, which is reported by students as systematic and which teachers insisted on in the previous year with the subsequent penalization of that aspect: “Yes, everything related to English is corrected by teachers; for example, last year when we didn’t write the subject, that lowered your grade, it deducted points, and spelling mistakes too!” (SI-B1). Some students, however, stress that language mistakes are penalized solely when they are very systematic: “If there are a lot of mistakes, like in a row, yes, it deducts points from the final score” (SI-A6).

As for spelling, apparently, it is regarded as paramount by students especially with those words which have been dealt with previously in class, i.e. academic vocabulary. Some students compare mistakes in English with mistakes they can have in Spanish, which shows the relevance that this aspect has, and the tendency to focus on accuracy amongst both content and language teachers: “Sometimes in the exam, having spelling mistakes does deduct points, as if it was the Spanish language exam” (SI-B6).

However, five students state they are sure about how spelling mistakes are rated in Spanish (0.25 each mistake) whilst apparently, they have no knowledge about the way mistakes are rated in CLIL subjects: “This year the teachers already take it (spelling mistakes) into account. I can’t remember how much it was... Was it 0.5? 0.05? I don’t know” (SI-B11); “In Spanish we know that each spelling mistake can deduct 0.25. In other subjects we don’t really know” (SI-B4). These spelling mistakes are also taken into account in the subject of English as a Foreign Language but according to students’ comments, they seem to be graded differently, the English language teachers being more demanding about mistakes when compared to content teachers. Contrarily,

pronunciation is not rated by students as relevant in CLIL assessment, the focus being primarily on written tasks.

The next question deals with the objectives content teachers present at the beginning of each unit - content-related objectives, language-related objectives or both. This question was introduced to check the initial finding from the questionnaires that content specialists might lack language awareness in that they do not introduce language objectives in the lesson, maybe because these language-related objectives are not always clearly articulated and/or identified (Hönig, 2012). All the students confirm that content-related objectives are the only ones presented in the CLIL lessons, and that language-related ones are regarded as the exclusive domain of language teachers. However, with regards to the feedback on the language aspects students need to improve, six students agree that some teachers comment on those typical mistakes after exams and once written work has been graded, and two students state that teachers know they often forget to write the subject or how to write properly in English: “This year, the Social Sciences teacher (History and Geography) tells us what we have to write at the beginning, when we start writing” (SI-B9). Two students also state that this linguistic information is also written in the corrected exams, so that they can know the aspects they need to improve in the future and the language mistakes which have been penalized by the teacher. The rest of the respondents (16 students) assert that this linguistic feedback is not common in the CLIL subject. Taking into consideration that students supposedly have the same teachers, this might confirm the data gathered both through the teachers’ and students’ questionnaires about language assistants providing linguistic feedback in content lessons.

In the next question, about which students are more likely to obtain better grades, the findings reveal that regardless of the penalization of language mistakes in CLIL assessment, English is not perceived by students as the most important aspect in the final grade: “Your English level does not affect as much as the time you devote to study. If you don’t know much English but you study, you’ll have a good mark, that’s for sure” (SI-A6); “The one getting the best grade is the one who studies harder, not the one with the best English level” (SI-B5). In fact, only two students out of 24 report that those students with a better command of English are the most likely ones to obtain better grades: “English can be the difference between a B and an A but that doesn’t mean that if you have some grammar mistakes you can’t have an A” (SI-A8).

Finally, students were asked about whether they are allowed to answer in Spanish in CLIL assessment. The vast majority of the students (23 out of 24) confirm that although students can occasionally use their mother tongue in class - for instance when they have tried to formulate a question in English without success - English is the only permitted vehicle of expression in assessment. In fact, only one student says he was allowed to answer in Spanish once during the first grade in the Bilingual Section, but of course, as it was an exception, that question was graded differently: “Everything you have to write in the exam has to be in English; otherwise, they will mark it as poor. (SI-B)

The data from the students’ interviews confirm that spelling and grammar mistakes along with academic vocabulary are regarded as the most important aspects in language assessment. As for the existence of guidelines, it seems that not all the teachers adjust to them. In relation with language mistakes, students agree that they deduct points from the final score, especially when mistakes are very systematic and refer to the academic vocabulary which is specific to the subject in question. However, students are not sure



about the exact deduction for each mistake. Students also stress that English is the only vehicular language in CLIL assessment, and that teachers inform them about the linguistic aspects they need to improve in order to get a better grade, and to express content successfully. Finally, in some cases, scaffolding is used to highlight the type of language students need to use in the expression of content and skills.

#### 5.5. Analysis of the assessment tools

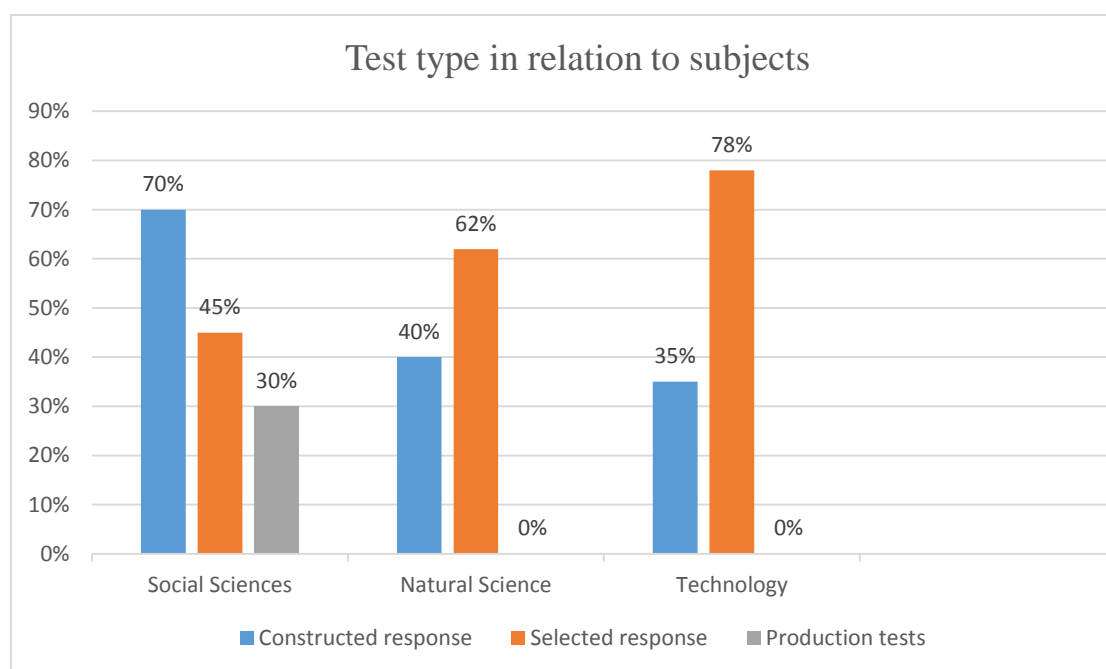
The analysis of the assessment tools was conducted to cover exam type, scaffolding in exams, guidelines for assessing language-related aspects, and teachers' feedback on both exams and homework. According to the sample, written exams are still the prevalent form of assessing students' development in CLIL subjects. Although other types of assessment tools apart from exams were requested - mainly due to the fact that in the questionnaires over three-fourths of the teachers and students mentioned the use of oral presentations as an assessment tool - only one of the informants offered samples regarding oral tasks. The written samples were thus analyzed in terms of their type.

Most of the tests from the sample are constructed response type including fill-in-the-blank/completion questions, open-ended questions, short-answer questions and essay questions especially in the subjects of History and Geography, in which students are usually asked to engage in more demanding cognitive tasks. In subjects like Technology, Robotics, and Biology, the most common type of tests are selected response tests with a combination of Binary choice/True False, matching and multiple-choice activities, as in the examples in Appendix 12, which demand the least possible language on the part of the learners.

Essays, papers, reports and other forms of production tests are not common according to the collected data, as no evidence has been found except for very short essay question type used in combination with constructed and selected response type in History. However, essays and project work are a natural part of everyday tasks and activities in some schools, and another way to assess students' work in a continuous way. For example, the following samples were considered by one of the teachers as essential part of the subject in History and Geography and thus, included in the class notebook: Essays, timelines, newspaper articles, opinion articles and diary entries.

The following chart shows how different test types differ depending on the subject:

Graph 29. Analysis of assessment tools: Test type in relation to subjects



To start with, and as can be observed in the chart, although teachers tend to combine different types of questions, the Social Sciences display a majority of constructed response exams including open-ended and short answer questions over selected

response. In this subject, it is also very significant that two-fifths of the respondents still use production tests or the same traditional essay-type tests students typically find in non-bilingual education. This is probably due to the fact that in some schools it is the didactic department who sets guidelines for both bilingual and non-bilingual groups as was stated in the chapter devoted to the focus groups. Second, in Biology and Chemistry, teachers tend to prefer selected response - matching, true/false and multiple-choice activities over constructed-response activities. This could be due to the attempt to facilitate tasks due to complexity in terms of vocabulary and cognition, and to check vocabulary and definitions in a more visual way. Finally, written samples in Technology show a vast majority of selected-response tasks as this subject usually complements exams with laboratory work and other types of project work using production tasks which mirror daily instruction in non-bilingual groups.

As for exam types in Arts and Crafts, no exams have been collected. This subject is very practical in nature, and teachers tends to assess students using traditional compilation of homework, class work, and other works presented by students, which are sometimes collected in the form of class notebooks or what some teachers call portfolio work. Regarding Physical Education, exams also tend to be in the form of physical performance, and sometimes there is a small percentage of contents which can be assessed in written constructed response exams which measure approximately up to 30% of the final grade. In this sense, it is quite significant that even though the portfolio is recommended by the law in force (LOMCE, 2013) no samples were collected. However, one of the teachers offered a compilation of activities under the name of portfolio<sup>17</sup>, containing several files such as the teacher's grading criteria, personal

---

<sup>17</sup> It is interesting to notice that some teachers refer to portfolios as the simple compilation of materials typical done in the form of students' class notebooks.

fitness records, and an endurance diary amongst others. In this sense, although this “portfolio” shows a compilation of materials and techniques such as self-assessment or personal reflection so that students can assess their own progress in areas such as physical strength, and the distance they are able to run, no selection of materials was observed so it cannot be considered a real portfolio.

### **Scaffolding or accommodation strategies**

The samples were also analyzed to see whether they provide any type of guidance or scaffolding techniques in terms of visual support, glossaries and language simplification or adaptation. Regarding visual support, the samples show a large presence of visual support (68 %) or pictorials, which are commonly associated with matching activities as in the case of Biology, and maps and timelines in Geography and History (Appendix 13).

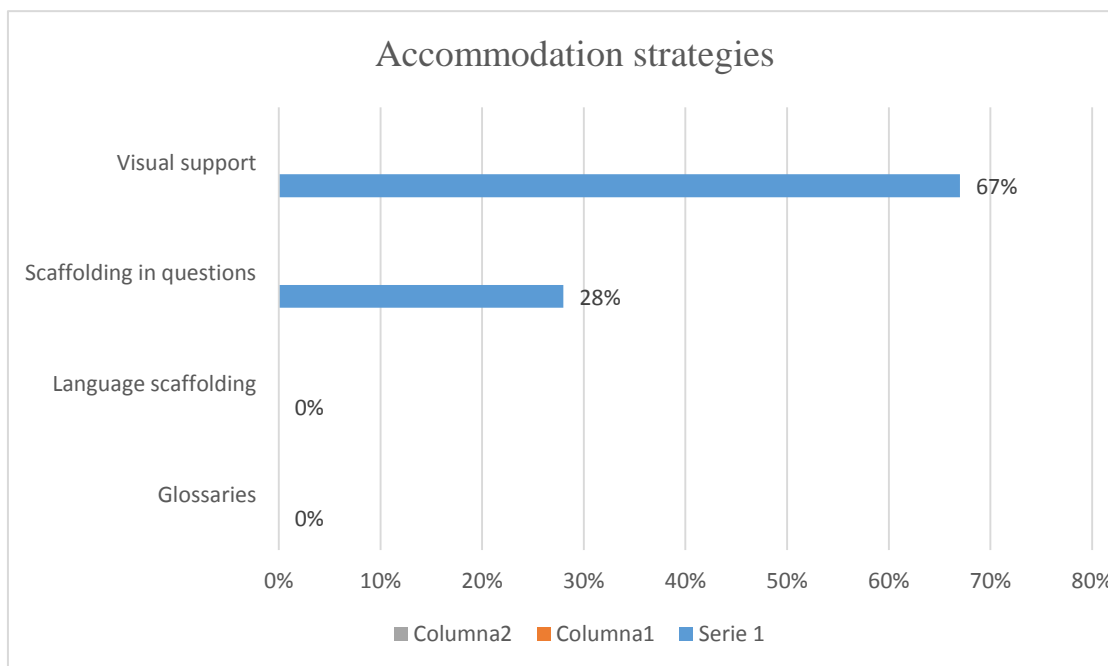
Although glossaries are a frequent tool in CLIL subjects<sup>18</sup>, no glossaries associated with the tests or tasks have been found in the sample, and no teachers assert to use them in order to scaffold content and help students in the production stage mainly because students are supposed to know the academic vocabulary. Likewise, probably due to the importance of academic language which is perceived by teachers as essential in the CLIL subject, no efforts or attempts have been found to simplify language or provide any examples of it. However, over a fourth (28 %) of the exams include some scaffolding to make content production easier for students, as can be seen in Appendix 14, in which students are guided in the process of writing with the help of prompts so

---

<sup>18</sup> Informal conversations with teachers, and observation of CLIL materials.

that it is easier for them to retrieve content knowledge, and some others offer very brief questions maybe in the attempt to simplify production tasks.

The following chart shows the presence of scaffolding or accommodation strategies in assessment tools:



Regarding the use of rubrics, although half of teachers (51.8 %) assert they use their own rubrics especially in order to assess oral presentations, no rubrics have been collected in any of the subjects for oral presentations or written assignments. This might mean that rubrics are not as common as it might seem, or even that the term rubric is misunderstood, or perceived as the teachers' own criteria to assess students' work.

### **Self and peer-assessment, and other alternative assessment tools**

The only samples in relation with self and peer-assessment refer to the self-assessment in the portfolio, which is part of Physical Education in one of the centers. This confirms the data from the questionnaires, and the fact that as in the case of the limited use of portfolios, the educational changes brought about by the implementation of the LOMCE in secondary education in 2015 have not been fully implemented so far.

On the other hand, regarding students' work and other assessment tools, namely essays, oral presentations and portfolio work, the following are the main conclusions: Students' work can offer lots of samples for assessment, but it is usually considered as a complement to exams or to round off the final grade, i.e. its use has not been generalized. As was pointed out at the beginning of this chapter, the majority of the data consist of written exams and thus, it is difficult to draw conclusions about the real weight and effect that project work might have on the students' final score. If we base our analysis on the data from the questionnaires, interviews and the focus group work, it seems clear that students' daily work is also measured in a continuous way through or class notebooks. The data collected includes essays, summaries, diary entries (Appendix 15), newspaper articles (Appendix 16), timelines and maps. They may serve as an example of the activities students complete as part of their regular class work or homework and which, according to the questionnaires have a big impact on CLIL assessment.

### **Assessment of skills**

Regarding the assessment of skills, the samples were analyzed to check whether cognitive, language and learning skills were taken into account in the assessment practice. To start with, concerning cognitive skills, most of the activities which are typically included in exams are of the type of non-cognitive demanding tasks (e.g. match, define, explain, etc.), i.e. the type of activities which require lots of context but very little cognitive challenge. Although cognitive demanding tasks are more difficult for students, they also trigger various levels of thinking as more linguistic and cognitive effort is needed on the part of the learner.

As for language skills, writing and reading largely dominate the assessment practice as opposed to speaking and listening which are not so common in the context of the study. Even in multiple choice questions, the information students are expected to retrieve is in terms of content-related aspects and not linguistic ones. Regarding speaking, just one oral sample was collected in the data, so no sound conclusions can be drawn in this sense. Finally, regarding attitudinal and learning skills, only one sample of a classroom checklist was gathered, and thus no sound conclusions can be drawn on this subject matter either.

### **Language correction and feedback**

In this section, the main samples of corrected work will be analyzed to comment on the teachers' criteria for language correction and feedback. For that purpose, attention will be paid to the most common language mistakes, the weight that English has in the assessment practice, and the feedback provided by teachers in CLIL exams, essays and

written projects. Furthermore, guidelines created by centers in order to offer instructions on how to deal with language in CLIL and the grading criteria some schools display in their websites will also be analyzed. Oral samples are very limited in number - two oral presentations pertaining to the same task, and no debates, discussions or rubrics associated with them - and they are just presented as pieces of students' work without including any grade, correction or feedback on the part of the teacher. Consequently, no conclusions can be drawn about the corrections of oral English, so the data collected refers exclusively to written pieces of work.

As for the language students use to express content knowledge, it is important to point out that all the sampling shows that English is the only vehicular language allowed in written tasks in the CAM Bilingual Project. This validates the assumption in the questionnaires and students' interviews about the use of the students' mother tongue not being allowed in the Bilingual Sections at least in the written form.

Regarding the weight that English as a vehicular language has in CLIL subjects, it seems clear firstly that generally up to 10 % of the final grade is deducted as a result of English mistakes, a fact that will be later validated by a closer look at the centers' grading criteria. Usually, in Social Sciences, more attention is given to language issues whilst in subjects such as Physical Education and Arts and Crafts in which more practical contents are required on the part of the students, the foreign language is not usually graded. However, one of the ten schools analyzed includes specifications about the value given to language in Music which are similar to the rest of the centers: 10%

About the consistency in the criteria regarding the weight of English in different subjects, it is necessary to point out that although no guidelines to assess CLIL are established by Madrid's Regional Government, most schools show similar percentages



(between 10 and 20 %). This could be explained firstly because of the lack of these guidelines, teachers tend to communicate with each other and/or create working groups to share criteria and ideas. Secondly, because interim teachers and those with a temporary destination in the region increase teachers' mobility, which undoubtedly favors communication and the dissemination of teaching practices.

As for the type of mistakes, and in order to define the construct, the treatment of language mistakes in the CLIL subjects has been observed in the students' corrected work, Language Guidelines and Grading Criteria, taking into account whether errors impede or impair communication (Ernst, 2005). Language corrections and feedback in exams refer, on the one hand, to academic vocabulary, and on the other hand, to the kind of language correction dealing with form which is typical of Foreign Language Assessment (Bigelow, 2001; Bigelow & Ranney, 2005), and which does not impede understanding. In this sense, it is easy to observe that the most common mistakes which are penalized by teachers are common to those in the Guidelines and Grading criteria described in the next section, namely subject omission in declarative sentences, subject-verb agreement, correct use of auxiliary verbs, verb tenses and adjectives along with issues relative to word order and in relation to academic vocabulary (spelling, wrong word formation and the creation of made-up words, which could obscure sentence meaning. Thus, teachers' comments and feedback about language use in exams focuses generally on the lack of completeness from students' answers as in the examples below:

*IN The metal year (ERA?), THE society was undeveloped; Settlements WERE surrounded by walls to defend WHAT?? (History)*

Besides, most significant corrections relate to form rather than content i.e. grammatical correctness as in the following examples:

*Ferdinand of Aragon was born ~~in~~ ON the 10<sup>th</sup> of May 1452 and ~~dead-in~~ DIED ON the 23 January 1516. (History)*

Another interesting aspect to observe is the presence of specifications on the weight of English as vehicular language in the written exams of two of the schools. However, as these specifications are not commonplace, they can serve to confirm the data from the questionnaires and students' interviews regarding the lack of guidelines to measure the weight of English as a vehicular language. In the first school, the general instructions at the beginning of the exam state that language is scored along with other "formal aspects" of the exam such as content deviation, grammar problems or slim margins as is shown below:

Serious spelling mistakes will be punished with 0.25. Besides the contents, other features of the exam will be considered to get the mark. Those are grammar problems, handwriting, confusing or wrong writing, slim margins, crossing outs, messy or dirty aspect in general. The exam mark can decrease or increase up to 1 point (Exam specifications, school C).

In the second school, language can also deduct up to 1 point from the final mark although a more detailed analysis is given as it also refers to style in the written form apart from grammar and lexis:

Language Code: Up to 1 point may be taken from your mark as a result of misspelling, grammar or style mistakes (Exam specifications, school E).

## **Language Guidelines and Grading Criteria**

Language Guidelines or guidelines for language in CLIL subjects are measures taken in some schools in order to deal with language mistakes. The data about guidelines has been collected from three sets of guidelines in three schools, in the Madrid area, and has been complemented with informal conversations with teachers, bilingual coordinators and talks during seminars on the practice of CLIL assessment in four Bilingual High Schools. In creating guidelines, the schools usually differentiate between the Bilingual Program and Bilingual Section, and the levels, mistakes in the upper courses of the Bilingual Section being regarded as more serious.

### **Types of mistakes taken into account in the Language Guidelines**

In general, according to the criteria established by the schools or the Language Guidelines, the mistakes that teachers take into account in CLIL subjects relate exclusively to those errors of form or “grammatical mistakes” which do not impede or impair understanding. The following table shows a compilation of the most typical mistakes penalized in CLIL subjects.

Table 12. Types of mistakes in CLIL subjects

TYPES OF MISTAKES PENALIZED IN CLIL SUBJECTS
Subject omission in declarative sentences.
Lack of subject-verb agreement
Incorrect use of “there is/there are”
Incorrect use of auxiliary verbs
Incorrect use of verb tenses, particularly of irregular verbs
Incorrect use of adjectives (especially comparatives and superlatives)
Incorrect use of question formation (especially in WH- questions)
Incorrect use of demonstratives
Incorrect use of relative pronouns
Incorrect use of possessive pronouns
Incorrect use of modal verbs
Incorrect use of linking words as related to the context

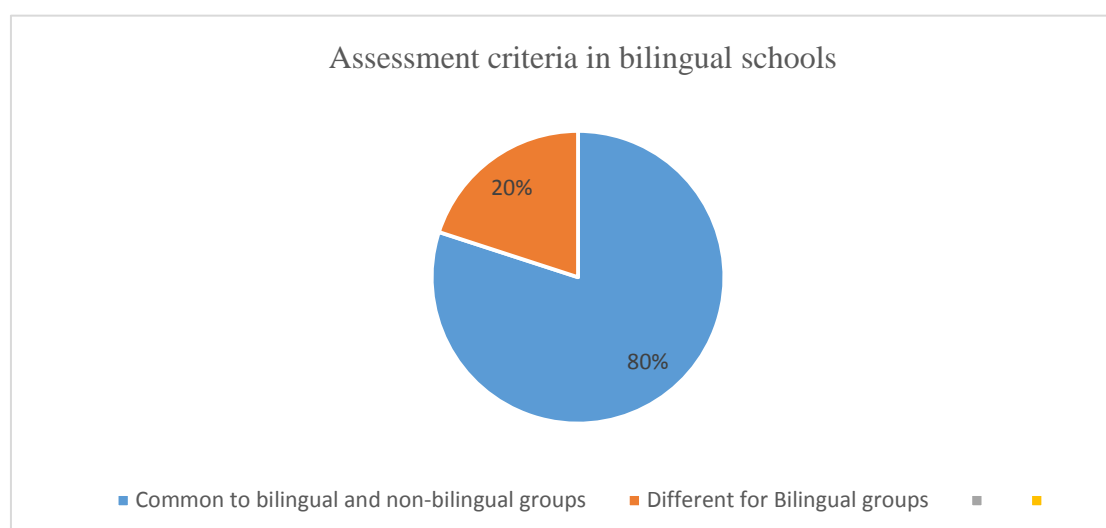
As can be seen from the table, the criteria refer to grammatical issues except for the use of linking words in relation with the type of context and text in question. On the other hand, mistakes related with academic vocabulary are termed as spelling mistakes. As for the deduction of points in the exam/task, in general, for each wrong structure in an exam, up to 0.10 will be subtracted, up to one or two points in the total score, depending on the school.

### **Grading criteria in CLIL subjects**

The following criteria have been taken from the grading criteria 10 schools created for the assessment of their CLIL subjects. The analysis was conducted as follows: First, it is necessary to observe whether assessment and grading criteria are the same for CLIL

subjects and non-CLIL subjects, and if they differ in the way the foreign language has been accommodated in the criteria for CLIL groups. Second, the weight of content, procedural and attitudinal aspects will also be analyzed, and third, special attention will be given to observe if assessment of the foreign language is included in the criteria. In this regard, assessment tends to follow the same criteria for non-bilingual schools since only 20% of the schools have different criteria for CLIL subjects, as can be observed in the graph below:

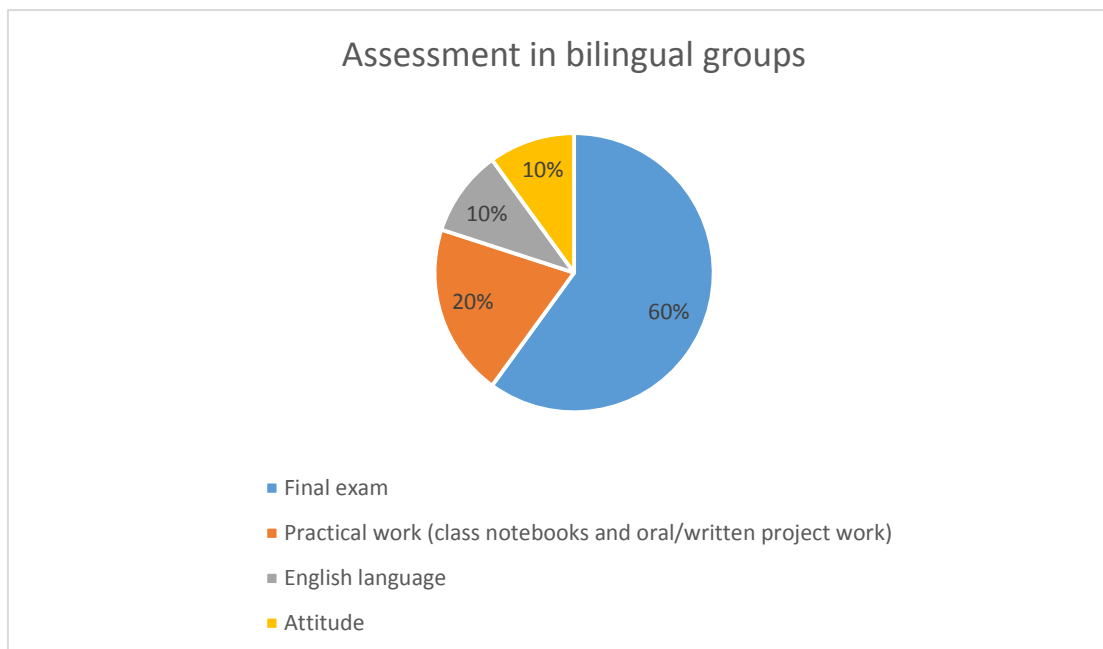
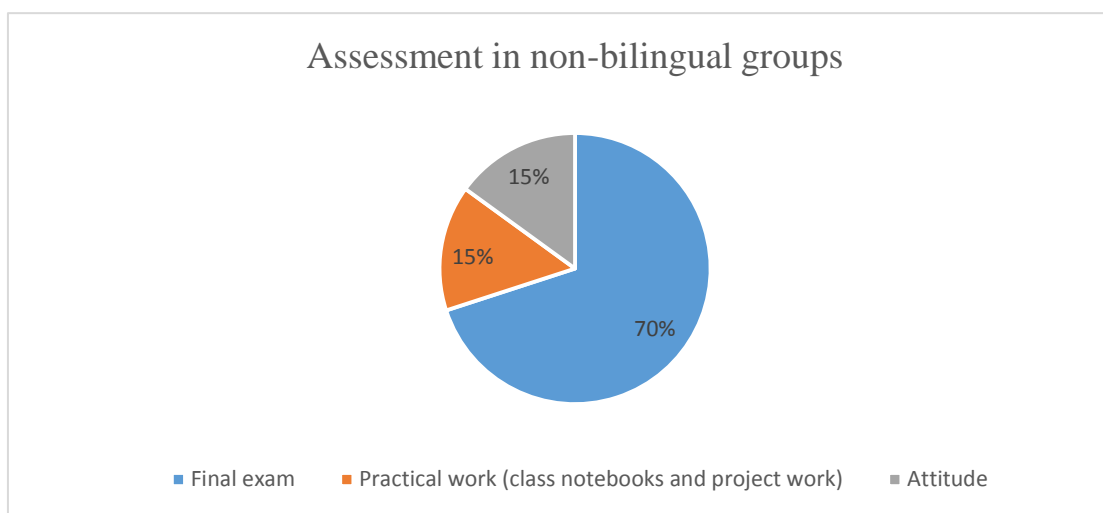
Graph 30. Analysis of assessment tools: Assessment criteria in bilingual schools



Regarding language-related issues, those schools showing different criteria for bilingual and non-bilingual groups have established criteria to assess and mark down the foreign language in CLIL assessment practice. The procedure is thus, to adjust the percentages so as to accommodate the foreign language into the final grade and differentiate criteria for the Bilingual Section from criteria for the non-bilingual groups or the groups in the Bilingual Program. The following graph exemplifies the grading

in Social Science and History in one of the schools. As seen in the graph, the exam constitutes 70% of the final grade whilst the rest is divided between students' practical work and attitude:

Graph 31. Assessment criteria in non-bilingual and bilingual groups



As the last graph shows, in an attempt to accommodate the foreign language in the subject curriculum, the written exam, which is usually focused on conceptual knowledge, is also adjusted and has a smaller weight than in non-bilingual groups. Also, it is quite significant that in this school, practical work comprises both oral and written activities whilst in non-bilingual groups it only refers to written work. In this sense, due to the limited linguistic competence of students in the foreign language, additional efforts have been made to find the most appropriate assessment tools for CLIL subjects. Consequently, in the first years of secondary education, the students' language will be less demanding in terms of discursive development. This language demand will gradually increase in the four years of Compulsory Secondary Education until linguistic competence is paralleled to that of non-bilingual groups, but ensuring at the same time the degree of achievement of the objectives or outcome indicators (School Website, 2016).

Summing up, written exams are the prevalent form of assessment over under a fourth of samples consisting of students' work (essays, timelines, entry diaries, etc.), and only two oral samples. As for exam type, constructed-response type exams are the most frequent in the Social Sciences whilst in Natural Science, Robotics, Biology and Technology teachers prefer selected response type questions. Regarding alternative assessment tools, no rubrics or self/ peer-assessment, and only one portfolio and one checklist have been collected. Regarding scaffolding strategies, most of the data (68 %) present visual support, questions are guided in 28 % of the sample, but no glossaries or language support have been found to support language in exams, maybe due to the low level of cognitive demand in questions, which facilitates the expression of content. Furthermore, it seems that in the absence of general guidelines, schools tend to create

their own, and in this sense, the language in CLIL subjects is usually counted with up to 20 % of the total mark although the general trend is 10%. Criteria for the assessment of CLIL subjects is the same used for non-bilingual groups in 80% of the cases, and in the rest of schools, the foreign language is added to the common criteria. Finally, as for the most typical mistakes teachers penalize along with academic vocabulary, they are the same mistakes which tend to be penalized in EFL, i.e. dealing with form but not impeding communication, namely spelling, concord, and mistakes in the use of verb tenses.



## **CHAPTER 6. Discussion**

This research aimed to investigate current assessment practice in the Bilingual Sections of the CAM Bilingual Project, and more specifically, whether and to what extent linguistic proficiency in the foreign language influences the grade students obtain. Due to the absence of language guidelines for CLIL assessment, a neglected area in the field, and of previous research in both primary and secondary contexts apart from Serra (2007), Serragiotto (2007), Hönig (2010), Wewer (2014), and Reierstam (2015), there is still need for discussion on actual assessment practices, and the best ways to assess CLIL subjects effectively so that language objectives are aligned with content ones. This study is hoped to contribute to clarifying to what extent theoretical considerations about CLIL assessment have been translated into practice in the CAM Bilingual Project. Furthermore, in the Spanish context, in which assessment legislation is common for bilingual and mainstream education, this research can help to identify teachers' needs and to suggest good practices.

In this section, the results from the questionnaires, focus groups, interviews and assessment tools are considered to provide a discussion of the results of the entire study, which will be the basis for the conclusions in the last chapter. For this purpose, the research questions are revisited. Furthermore, this section offers a discussion of the validity of assessment tools and the role and weight of English in the present study.

### **6.1. Language-related issues in CLIL**

In the first research question, the aim was to investigate to what extent teachers are conscious of the importance of the language in CLIL, and whether these linguistic aspects are taken into account in assessment. Language is an ongoing concern among secondary

education teachers in the CAM Bilingual Project. According to the data obtained from the teachers' and students' comments, and the work with the focus groups, teachers recognize the relevance of language-related issues for students' efficient expression of content and skills, although they have difficulties in assessing CLIL subjects through the vehicle of the foreign language. Among these difficulties, they stress the lack of guidance on the best ways to deal with language mistakes, and the best assessment tools to suit both the CLIL context and the secondary context they are immersed in, the latter being a growing concern among teachers. In the attempt to adapt assessment tools to the CLIL reality, and thus facilitate students' production in the foreign language, exam activities and tasks in content subjects usually combine essay type questions with short answers and matching activities. The problem, teachers assert, is that these activities do not parallel PAU/EvAU exams, and differ from the ones in non-bilingual groups, in which students are usually required to write longer texts so that bilingual students' academic level is sometimes questioned by other teachers and also families. In this regard, content teachers also try to deal with the restrictions imposed by their didactic departments and Spanish legislation as regards the need to use the same assessment tools which are common for mainstream education, as also happens in Swedish upper-secondary schools (Reierstam, 2015).

The teachers' awareness of language can also be seen in the relevance language has in assessment practices, as will be analyzed in detail in research question three. On the other hand, this language awareness is observed in the creation of school guidelines for correction and weighing of language due to the absence of official recommendations. However, assessing the language does not necessarily entail that language-related aspects are present in daily teaching practice. In fact, apart from commenting on students' language mistakes in exams from time to time, language is not visible in class as happens

in other European CLIL contexts where teachers recognize the relevance of language in daily teaching practice as a preparation for content expression in exams (Reierstam, 2015), and a tool for learning in general. Thus, in the context of our study, even if errors are treated by means of the “language clinic”, the objectives teachers present refer exclusively to content and not language, and in the need to compensate for students’ deficiencies, teachers opt for simplifying or reinforcing content objectives. This invisibility of language (Llinares et al, 2012) in the class contrasts with the prominence it has in exams, and it shows the lack of alignment between teaching practice and assessment. Therefore, students perceive that language is essential in CLIL subjects, but they are not trained in linguistic aspects which are later present and taken into account in exams and other assessment tools.

The lack of focus on language may be attributed to several factors. Firstly, language objectives and tasks are still absent in some CLIL models (Hönig, 2010), and scarce in most CLIL textbooks and materials (López Medina, 2016; Martín del Pozo & Rascón Estébanez, 2015; Kelly, 2010). Secondly, listening and speaking skills still receive little attention in secondary education assessment in Spain (García Laborda & Fernández Alvarez, 2011). Thirdly, teachers are usually reluctant to be made responsible for the language in CLIL, a role they think suits the language teacher best. This is also common in other countries such as Slovakia (Gondová, 2012 b), probably due to their background as content specialists, which usually implies a lack of training in language pedagogies, and because of their lack of confidence in their own language skills (Clegg, 2012). This tendency to overlook language issues and take them for granted can be explained because of the teachers’ lack of language awareness (Andrews, 2007; Pavón, 2010). In fact, although content teachers master the topic and the academic registers, they see language as a natural part of the text, are already trained to using academic literacy, which prevents

them to notice the difficulties students might encounter in dealing with academic texts (Personal communication with students and teachers, academic year 2015-2016). Besides, another factor impeding language visibility is that, as teachers point out, students have a limited vision of subjects (Personal communication with students and teachers, academic year 2015-2016), and when content teachers highlight language-related issues, students tend to see them as adopting the English teachers' roles. It also seems that students are not used to seeing teachers collaborating with each other, and thus they consider content teachers as the only ones responsible for the subject, which contrasts with the recommendations of subject integration by recent Spanish regulations, and the cross-curricular approach necessary in bilingual education. Teacher collaboration and coordination are, in fact, commonplace in other countries (TALIS, 2013) such as Italy and Austria, where content teachers and language teachers can co-assess the subjects (Serragiotto, 2007; Hönl, 2010).

## 6. 2. Integration of content and language in CLIL assessment

In research question two, the aim was to identify whether teachers separate content and language in assessment, and the strategies and tools they use to do so (if any). The data shows that teachers in the CAM Bilingual Sections opt for differentiated assessment, in which content and language are considered through the same tools and tasks. There is a clear focus on content-related aspects, and the final grade is a general one compiling both content and language issues. However, language is assessed separately as language mistakes are taken into account in the final grade, an aspect which is often criticized in CLIL assessment (Frigols, in Megías Rosa, 2012). This type of assessment is mainly motivated by the particular context in which CLIL is framed. Contrarily to other contexts such as Italy and Austria in which content teachers collaborate in teaching and scoring

with foreign language teachers (Nakanishi, 2015; Hönig, 2010; Serragiotto, 2007), in the CAM Bilingual Project, the responsibility for grading their subjects lies solely on content teachers. These teachers feel the pressure of having to assess language aspects, and deduct scores out of the final grade, as is typical in non-bilingual groups taught in the students' mother tongue. Third, although teachers' coordination is still subject to improvement, language teachers in general and Bilingual Coordinators, in particular, exert a big influence on content teachers as experts on language in the school, who are usually asked for advice in Spanish Bilingual Programs (Pavón, 2016: 78). In this respect, it is important to remember that in Spain, language teaching and learning has been traditionally associated with the study of grammar, vocabulary and writing skills (Romero Lacal, 2011; Rubio & Tamayo, 2012). As a result, an excessive focus on form (FoF) is quite generalized, which might lead teachers to assess and mark down language mistakes. On the other hand, in the absence of assessment criteria from educational authorities, language is assessed following the teachers' criteria, recommendations by the Bilingual Team, and the guidelines some schools create for that purpose focused on marking down language mistakes- spelling and grammatical inaccuracies over sociolinguistic, discourse and strategic ones.

6. 3. The balance between content and language skills in assessment. Teachers' learning intentions as communicated to students

The third research question focused on the balance between content and language objectives, and whether teachers' learning intentions are communicated to students. According to the data, and following CLIL recommendations (Coyle et al, 2010), the main focus of assessment is on content, and most teachers reveal that language is weighed less than 25% of the final grade. The predominance of content over language skills can

also be noticed in the schools' language guidelines, used by over half the teachers, and the teachers' assessment criteria consulted in the analysis of assessment tools. Even in those schools using no guidelines for CLIL assessment, it seems that there is an "unspoken rule" of grading language issues up to 20%. So, to facilitate the process of grading, teachers have anchored the proportion of target language in CLIL to approximately 20% as recommended by Cushing Weigle & Jensen (1997).

Regardless of the treatment of mistakes, this reflection on the part of the teachers about the role and weight of language already shows a turning point in the way different members act in collaboration with each other. An obvious example of this mind shift is the creation of guidelines to be used in the school to have some common criteria to deal with language-related issues. In two of the sampled schools, these guidelines are shared with students and families, and they are either published on the school website or sent to families via the school diary, to make assessment as transparent and objective as possible. In other schools in which no guidelines are used, students are aware of the importance of writing using correct English (Personal Conversations with Students and Teachers, Academic Year, 2014-2015), but they seem to be uncertain about the aspects this language correction entails. In this sense, it is interesting to note that although over half the teachers reveal that they inform students about their language mistakes, half the students report that language is paramount, but that they are unsure about the real weight of language in assessment. (Students' questionnaires and interviews). This might imply either that the teachers inform students about the importance of accurate writing and speaking, but the weight language mistakes have in the final grade is not explicitly specified due to the lack of guidelines or even that this weight really depends on the teachers themselves. However, due to the limited number of assessment tools, and the lack of rubrics in the sample, it is hard to observe to what extent agreements about the

weight of language are translated into current practice or if, on the contrary, the weight depends on the teachers in question.

#### 6 .4. The aspects of language taken into account in CLIL assessment

The next research question aimed to spot the aspects of language which are taken into account in CLIL assessment. To start with, and as is common in other contexts (Dalton-Puffer, 2007; Wewer, 2014; Reierstam, 2015), the ability to use the correct academic vocabulary for the topic and academic discipline in question is perceived by both teachers and students as paramount in CLIL. Mistakes being marked down in relation with academic vocabulary deal with correct spelling, the use of target vocabulary in English, and the makeup and misuse of words. According to teachers, other aspects which are taken into account apart from academic vocabulary are “communicative competence” or the ability to get the meaning across, fluency in speaking, grammar, spelling, and pronunciation.

However, although students also regard academic vocabulary as essential in the expression of content knowledge, a vast majority of them consider aspects such as grammar and spelling mistakes as the most important aspect taken into account by teachers in CLIL assessment (Students’ questionnaires and interviews). This inconsistency between teachers’ and students’ opinions can be explained on the one hand, by students’ perception of their limited language proficiency regarding grammatical accuracy, an aspect about which students consistently receive feedback, and which they thus perceive as essential for content expression. On the other hand, it can also be due to the teachers’ emphasis on accurate writing, which contrasts with assessment practice in other European contexts (Reierstam, 2015), and in even in Spain (Fuentes Arjona, 2013).

Perhaps in an attempt to compensate for the students' lack of awareness of grammatical features of the language, teachers turn to CALP and use traditional corrective feedback instead of focusing on the students' ability to express themselves using functions i.e. appropriate language structures and forms dependent on the academic discipline and topic in question.

Regarding the emphasis on academic vocabulary, it might have been influenced by several factors. First, the legislation in the Madrid region emphasizes academic vocabulary in both English and Spanish in CLIL primary schools, which sometimes inevitably results in the creation of lists where students memorize words and expressions out of context. Second, content teachers' lack of language pedagogies might impede them considering language aspects from a holistic viewpoint, as performing functions and not solely from the focus on grammatical form.

As for the corrective feedback, a closer look at the guidelines shows that teachers tend to treat language mistakes in the same way as in English as a foreign language, that is deducting score from the final grade when mistakes are very systematic such as subject-verb agreement, errors in verb tenses, singular and plural nouns to name just a few. The severity of mistakes depends on the level, those in the last years of Compulsory Education being considered more serious if compared to the first year in Compulsory Secondary Education (ESO).

Only one school seems to treat language mistakes differently since the assessment criteria are adapted to the assessment of content subjects, and both meaning and function are taken into account rather than form. An example of the latter is the way teachers deal with the correct use of quantifiers "some/any" in this school. In the sentence: "There were any children in the playground", the use of "any" is considered as a mistake which is marked down in the subject of English as a foreign language, but this is not applicable in CLIL



subjects. Therefore, although it is difficult to draw conclusions, most evidence points to the “unspoken rule” that language can deduct up to 20% of the final score although most teachers tend to deduct up to 10%. As for language mistakes, they are considered regarding grammatical accuracy, concord and spelling, these aspects being the same taken into consideration while correcting the students’ mother tongue production in non-bilingual subjects.

Consequently, as was stated in research question two, although the students’ capacity of expressing knowledge of the subjects, and getting the meaning across is the main goal of teachers, a clear focus on form is also observed in the correction of mistakes. The grammatical approach or focus on form (FoF) prevails over the functional approach. So, the focus is on corrective feedback<sup>19</sup> typical of traditional EFL lessons i.e. directing students to write the correct grammatical form, and avoiding spelling mistakes in academic vocabulary, and not so much in terms of language structure and functions, which is recommended in CLIL assessment (Coyle, 2010; Maggi, 2009; Frigols, 2012; Mohan & Huan, 2002).

Apart from leaving the item blank, providing an inadequate response, and providing an incomplete response which are common for assessment in mainstream education, the most typical mistakes resulting in a deducted score are language mistakes. On the one hand, those dealing with academic language (misspellings in academic vocabulary, using the L1 alternative, made up words). On the other hand, grammatical inaccuracies which do not impede or impair understanding (Ernst, 1995) such as word order, concord, incorrect use of adjectives and adverbs, question formation, relative sentences, and errors

---

<sup>19</sup> The term “corrective feedback” (Ellis, 2009) as the information a learner receives about an error s/he has made, is used here to refer to the written feedback by teachers in the context of the study.

in verb tenses. This focus on form is not recommended in CLIL contexts as it highlights the duality between content and language, and it menaces the validity of assessment by separating language-related objectives from content-related ones. As language objectives are not aligned with content ones, it is difficult to verify whether the student was not successful in completing the task due to lack of knowledge of the topic or because of her/his limited language proficiency. This focus on form clearly contrasts on the one hand, with the CEFR recommendations to treat language as an instrument to get the meaning across, and never as an end in itself (Council of Europe, 2001: 112-113). On the other hand, it is also in contrast with the focus on function in CLIL lessons in the Spanish context, in which language is simply used as the medium of instruction and not as an end in itself (Pérez-Vidal, 2007: 49). For instance, in the MEC-British Council Bilingual Program, the focus of form is combined with the focus on meaning and function (Dobson, Pérez Murillo & Johnston, 2010: 41). This corrective feedback can be beneficial for students' linguistic production in language lessons (Lightbown & Spada, 1994; Pica, 2002; Ellis, 2009). However, in CLIL, it is better to concentrate on meaning, consider CLIL students as users of the language (Dalton-Puffer, 2015), and leave high expectations of native like competence behind, as they could influence the way the language is assessed favoring accuracy over communication (Lasagabaster, 2010). Besides, an excessive emphasis on grammar correction does not necessarily favor students' interlanguage but rather results in pseudo-learning of the target language, it has a negative impact on students' ability to write accurately (Truscott, 1996, 2007), and it can have a direct influence on the final grade. In fact, this specific focus on written expression is avoided in other CLIL contexts (Dalton-Puffer, 2007), where students are allowed to express content knowledge using their mother tongue (Hönig, 2010), and teachers feel they should not correct language mistakes because they are not native speakers themselves

(Reierstam, 2015: 147). The same criteria on formal correction and use of academic vocabulary seem to prevail for the spoken language, where pronunciation, vocabulary and grammar play a dominant role according to students' and teachers' opinions. However, more oral samples would be needed to verify the data, and draw conclusions on that subject matter.

#### 6 .5. Most common assessment tools and main reasons to use them

In research question five, the focus was on the preferred assessment tools, the type of tests, and the reasons to use them over others. From the data gathered through the teachers' and students' questionnaires, it is clear that written exams are the most frequent assessment tool, followed by students' work through class notebooks and informal class observation. This emphasis on written exams is not common in pre-primary and primary education contexts in other European countries (Serra, 2007; Hönig, 2010) where oral tasks prevail, and specifically avoided in others such as in the German state of Baden-Württemberg, where students are assessed through oral tasks and activities. However, they are frequent in upper secondary education in Sweden (Reierstam, 2015) because they are easier to grade, and in the Spanish context, mainly due to the predominance of standardized exams in education as compared to other countries (TALIS, 2013). Unlike assessment in some primary CLIL contexts where the testing methods are adapted to the students' level of language development (Zangl, 2000), the testing methods in the context of this study are the same for all type of learners. This is probably because the students in the Bilingual Sections have an advanced level if compared to the students in the Bilingual Program (a minimum B1 level in the two first academic years, and B2 in the two last academic years), and because Spanish mainstream education tends to assess students uniformly regardless of students' characteristics. On the other hand, class notebooks,

consisting mainly of written homework (essays, reflections on experiments, timelines, projects, etc.) are very highly considered among secondary teachers to check students' skills or practical knowledge over time. Nonetheless, in the absence of a significant number of samples regarding class notebooks, no further conclusions can be drawn about their real weight on the final grade, and the range of activities they include or the way they are used. Furthermore, debates and oral presentations which are commonplace in other contexts (Hönig, 2010) have increased dramatically in the CAM Bilingual Project because of the need to practice oral skills. However, their weight in the final grade is still far from being systematic, and they tend to be graded along with class notebooks, that is as part of additional criteria to measure students' progress. In this regard, it is interesting to highlight that although teachers and students value them as positive assessment tools, only two oral samples were gathered in this study, which clearly indicates their small impact on the final grade, and the fact that teachers tend to associate assessment with written exams.

Likewise, regarding alternative assessment tools, namely self and peer-assessment and portfolios, which are recommended for CLIL contexts as well as by the law in force (LOMCE), timid movements are being made to implement them in content subjects. Nevertheless, their use is still very limited or even inexistent in some schools as is also common in other countries (Hönig, 2010). The reasons for not using self and peer-assessment are often relative to the lack of consistency these tools seem to have for teachers, and the students' lack of training in their use. The same can be said about the portfolio, which in contrast with the mere compilation of activities presented in class notebooks typical of the Spanish context, should involve reflection on the part of the students. For the practical implementation of these tools, apart from specific training, the teachers need to accept them as valid assessment tools, and therefore include them in the

final grade so that students develop reflection skills, and see their purpose in the subject (Personal communication with students and teachers, academic year 2015-2016). Since educational changes and tools are slowly implemented, it is hoped that to compensate for the supremacy of written exams, and to conduct assessment in a formative way, more efforts will be made to include alternative assessment tools in the near future.

As for the type of questions included in written exams, exams combining multiple choice and essay-type questions prevail in the Social Sciences while in Music, Robotics, Biology and Technology teachers prefer selected response type ones. It seems that in the Social Sciences, students are expected to write longer and more coherent texts contrarily to other subjects where attention is focused on the writing of definitions, and/or providing short answers. Lastly, in Physical Education and Plastic and Visual Arts, students are expected to show their progress through projects and practical work although written exams can also be used to check students' knowledge of theoretical concepts.

In CLIL assessment, attention also has to be given to the scaffolding strategies or the assessment methods used to minimize the impact of students' limited proficiency in English, and the way teachers assess skills and processes. In this regard, it is important to highlight the effort made by teachers to adapt the exam format to students' language level if compared to the type of exams in non-bilingual schools, and the presence of visual support in most of the samples. Also, a significant number of teachers are aware of the need to combine several open-ended questions or essay type questions with multiple-choice/matching ones, and the need to use visual support for cognitive reasons. However, recommendations by departments are not always in line with the purpose of CLIL as students' limited proficiency in the foreign language is not taken into account, a fact which makes the tasks less appropriate for those students with a more limited command of English. For instance, scaffolding in exam questions, i.e. prompts or guided questions

to facilitate students' written output is present in just over one fifth of the sample, and no glossaries or language support have been found in exams either. This lack of scaffolding contrasts with assessment conducted in Swedish upper secondary schools, where teachers make some adjustments such as allowing the use of students' L1, providing dictionaries, giving more time for tasks, and "the opportunity to make clarifications afterwards if ambiguities appear" (Reierstam, 2015: 138). In this sense, it is essential to highlight that legislation concerning the vehicular language in Bilingual Sections in Madrid does not specify that English is the only language students can use to answer questions and tasks, but contrarily, there seems to be an unspoken rule about the use of English in assessment among teachers. Due to the difficulties associated with content expression especially in the first grades, some timely assistance in the form of dictionaries or guided questions, or even allowing the expression of content in the students' L1 would facilitate assessment tasks.

Finally, skills and processes are necessary for CLIL as the vehicles to enable students to express their knowledge through tasks in real-life situations, and thus make learning meaningful. In relation with language skills, the analysis of the sample of assessment tools shows that writing stands out as the language skill used most frequently. Language skills are essential for students in content subjects, but this development is perceived to be the responsibility of language teachers only. Although oral presentations have found their way into current assessment practice, speaking tasks in this context seem to be relegated to class activities. Besides, language skills are assessed differently in students' work and exams. In the first, the language is not usually graded as the focus is on content correction whereas in the latter, the language is not assessed in a systematic way, which does not allow students to reflect on their language mistakes regularly. Consequently,

more efforts are still needed to integrate the spoken language in current assessment practice, and score language aspects systematically, using rubrics for instance.

Content skills are measured through written exams and class notebooks. However, as language and content skills are inextricably integrated with each other, it is sometimes difficult to ascertain whether the task was not successful because of the lack of language ability or rather because of the limited knowledge of the topic. Thus, for effective assessment in CLIL, and to prevent “muddied assessment” (Weir, 1990), skills need to be linked with learning outcomes (Coyle, 2005) or LOMCE’s assessable learning standards. Similarly, teachers should explicitly specify which language aspects they take into account in assessment. Furthermore, as written tests and tasks require proficient levels in the foreign language to enable students to express themselves efficiently, if language is to be graded, teachers should offer the type of scaffolding necessary to facilitate students’ output in the foreign language in the first grades of secondary education.

As far as cognitive skills are concerned, teachers ask for Lower Order Thinking Skills (LOTS) in detriment of more challenging tasks for students. This choice is probably justified by the attempt to facilitate content expression by offering cognitively accessible activities which also demand less language proficiency, and which is present in other CLIL contexts (Reierstam, 2015). The choice of LOTS might also be due to the traditional emphasis laid on memorization, and the lack of task-based work aimed at promoting thinking skills in mainstream education.

Finally, as regards learning and attitudinal skills, the first are not really assessed in the context of the study due to the lack of alternative assessment tools such as self and peer-assessment, and the latter are usually evaluated through classroom observation checklists. In this sense, it is important to point out that even if students’ participation and attitude

were highly considered by both teachers and students in the questionnaire, it seems their performance sometimes serves to round-up the final score, or it is not assessed systematically. Thus, although it is difficult to draw any conclusions because of the under-representation of this kind of assessment tools in the sample, more systematic assessment procedure need to be agreed upon teachers. This would have a direct impact on the success of assessment as it could change the way students perceive assessment of learning as punitive, and occurring at the end of instruction.

#### 6. 6. Recommended CLIL assessment tools

The last research question was aimed at identifying the best assessment tools for CLIL, and finding out how current assessment methods might be changed to suit bilingual education.

In chapter five, the best assessment tools for CLIL were described taking into account the following criteria: First, assessment tools should be designed as pedagogical tools, that is to serve CLIL instruction in line with Formative Assessment. Second, they should mainly be used to judge students' progress in relation with content knowledge and skills in the subject. For that purpose, and given the prevalence that content should have over language, assessment tools also need to balance cognitive and language demands, choose the best assessment method in relation with the subject in question, and consider the linguistic aspects as an aid to expression i.e. emphasizing communication and language use in getting the meaning across rather than seeing formal correction as an end in itself. Third, assessment should be conducted in such a way that tasks mirror real-life activities i.e. using alternative and authentic performance-based tasks such as portfolios, journals



and projects measured as objectively as possible, for example, with the help of rubrics (Coyle et al, 2010; Short, 1993).

As was pointed out in chapter four, the type of assessment has a direct influence on the approach adopted- the product or the process-, the assessment tools, and most importantly, on daily teaching practice, and the way students perceive the subject. Although in the Spanish context, the first references to assessment as formative and continuous are present in the current legislation (LOMCE, 2013), actual assessment still relies on summative procedures rather than on pedagogical tools to improve students' performance. The final score is obtained from the sum of different scores in exams, mainly based on memorization as the main referent in learning, along with class work and homework. Continuous assessment is thus more connected with classroom assessment, usually registered in class checklists, and measured through qualitative rather than quantitative data, namely effort, interest, participation and attendance, usually associated with positive and negative scores, and having a minor weight on the final score. Likewise, students' marks, especially those coming from exams, are not formative i.e. they inform students about their typical mistakes, and sometimes the best ways to improve them through remedial work, but they do not grant students the possibility to re-write their piece of work. As a consequence, assessment and learning processes are not dynamic, the final product and not the process of learning is prioritized, and assessment does not help teachers or students to reflect on the teaching and learning practice, and to improve instruction. In this sense, more attention should be devoted to process writing as opposed to product writing in language lessons, as the former allows students to engage in the formative assessment processes critical to students' progression.

The prevalence of written exams over oral ones can be attributed to the weight of tradition on school practices in general, and the lack of oral tasks in particular, as in the PAU/EvAU

exam. Traditional written exams are regarded as more objective, and easier to grade. However, they favor norm-referenced assessment rather than criterion-referenced assessment, and might not be appropriate for students with limited language proficiency. In this regard, it is necessary to find those tools which ask for less demanding discursive development to compensate for students' limited foreign language competence without compromising content knowledge. Besides, written exams tend to focus on reproductive learning (memorization) in detriment of constructive learning (understanding and applying). This inevitably lessens the cognitive challenge of questions, maybe due to the limited time students have for their completion on the one hand, and to students' limited language proficiency on the other. A possible solution to balance the cognitive challenge might be to leave behind the predominance of the written exam over performance-based tasks where HOTS and real-life activities can be easily accommodated. By using authentic and alternative assessment tools such as portfolios, oral presentations, journals and projects, students are allowed to show their competences in both content and language, and reflect on both the process and the final product. However, for alternative tools to be effective, students need to be clearly informed and familiarized with the assessment criteria, for instance, using rubrics.

Furthermore, to combine content and language skills successfully, assessable learning standards need to be aligned with the language competences that are essential to be able to express content knowledge in the different content subjects, and these language competences need to be highlighted. On the one hand, a solid framework combining content knowledge, thinking skills and the language required to perform those functions using at least two levels of difficulty could be helpful for content teachers. This framework would enable "to increase the understanding of language demands of test items, to address specific strategies to attend students' needs, and afford foreign language

learning beyond incidental language gains” (Leal, 2016: 310). On the other hand, if the language is to be graded, more attention needs to be given to language issues from the beginning of instruction, by including more linguistic skills in daily teaching practice, and making language objectives explicit. Finally, given that in written exams all learners’ understanding involves language use and production to a certain extent, as with alternative assessment tools, the criteria to grade language-related issues should be clarified, and some scaffolding might be offered at least in the first levels.

#### 6. 7. Assessment quality in the context of the study

In chapter five, some characteristics of effective assessment were highlighted in relation with validity, reliability and fairness. Validity in CLIL refers to whether the assessment tools measure what they are supposed to measure, i.e. students’ knowledge and skills in the academic subject. Likewise, the assessment of language aspects is considered valid when it focuses on sociolinguistic, discourse and strategic competence or language functions, and not solely on form, i.e. linguistic competence understood as grammatical and spelling inaccuracies. Introducing only the linguistic components in the assessment of the content subject might contaminate the grade since teachers tend to mix academic and language competences (Marzano, 2010: 17). So, the validity of the assessment process as described in this study is compromised by the overt focus on form used to assess language, and also by the lack of alternative assessment tools measuring students’ work in a formative way.

Reliability deals with the use of assessment methods which produce stable and consistent results thus avoiding subjectivity. In this sense, guidelines by the educational authorities or those created in the schools are paramount to rely on common criteria concerning the

treatment of language mistakes. Besides, rubrics are also essential as they offer teachers a series of descriptors or can-do statements that students can also use for assessing their work and their peers'. In this study, the creation of guidelines for language mistakes is seen as positive by half of the respondents, and contributes to the reliability of the assessment process. However, due to the absence of rubrics in the sample, it seems that assessment in productive tasks is dependent on teachers' experience, flexibility or personal preferences among other possible factors, or that the guidelines created by some schools are used as general rubrics for all the content subjects.

Finally, fairness in CLIL assessment relates to giving all students the possibility to demonstrate knowledge of the subject regardless of their language level. However, as the demonstration of content and skills is tied up with traditional standard exams following the PAU/EvAU models, in some subjects such as Social Sciences students are constrained by their limited language proficiency, and the impossibility to express themselves in their L1, which impedes them showing the depth of their knowledge. Thus, to ensure fairness in the process of assessment, more receptive tasks and scaffolding in productive tasks should be offered to students in the future.

## **CHAPTER 7: Conclusions**

The purpose of this study has been to shed some light on one of the most contested issues in CLIL, assessment, and how it is conducted in practice in the context of Bilingual Sections of the CAM Bilingual Project. This chapter is divided into three different parts: First, some conclusions are drawn from the results of this research. The conclusions have been contrasted with best practice suggestions from other CLIL contexts, and the informal conversations with teachers and students about the difficulties they face in their daily assessment practices. Second, some recommendations are included concerning, on the one hand, assessment practice and the treatment of language issues, and on the other hand, some guidelines for both content and language teachers. Finally, some proposals for further research concerning CLIL assessment, and several surrounding issues such as teachers' cooperation are also suggested.

### **7.1. Main Conclusions**

As was pointed out in the discussion, despite recommendations about the implementation of Formative Assessment in CLIL, practices in the context of this study demonstrate that assessment is conducted in a summative way. Assessment does not serve to inform instruction, and the main tools being used to assess students in content subjects still conform to traditional assessment patterns mostly in the form of written tests, leaving communicative language competence behind. Thus, although the impact of CLIL can be observed in aspects such as the increase in the number of oral activities in daily teaching practice, and the implementation of accommodation strategies catering for students with limited foreign language proficiency, this impact is not as evident in relation to assessment practices. Assessment in this study does not exclusively depend on issues

suiting to bilingual education but also on assessment legislation for secondary education, which undoubtedly exerts a significant influence on current assessment practices. In fact, the PAU/EvAU exam - the entry exam to access Higher Education - has a big impact on secondary education, and it shapes assessment practices (Rodríguez-Muñiz, Díaz, Mier & Alonso, 2016; Smith, 1991; Zakharov, Carnoy & Loyalka, 2014). Due to this washback effect, CLIL assessment tends to follow the same patterns typical of non-bilingual groups as regards the main assessment tools and exam format. To start with, the EvAU exam in Madrid is conducted in Spanish, a fact that commonly worries teachers, students and families because of the effect that bilingual education might have on content learning, and students' expression in their L1. Second, although attempts have been made to introduce listening tasks in English as a Foreign Language, this entry exam consists predominantly of written tests. Even though bilingual education is already well established in the Madrid Region after ten years' experience, these secondary education standardized exams are common for both bilingual and non-bilingual groups, a fact which might lead teachers to adopt more traditional approaches suited to the entry exam format to train students accordingly in the long term.

As for the agents taking part in the assessment process and its main goals, assessment is conducted by content teachers, centered on the final product and not on the process, students' results being presented using a score. Differences in assessment practices are related first, to teachers' perceptions of the role and weight that language should play in assessment; second, to different subjects and the level of language production they require; and third, to teachers' preferences about assessment tools.

Since, apparently, content teachers do not present language objectives in class, it might seem that language plays an invisible role (Linares et al., 2012) in daily teaching practice. However, this study has evidenced that the foreign language is assessed as separate from

content issues, it is not necessarily linked to the achievement of content-based learning objectives (Mohan & Huang, 2002), and thus, it plays a major role in assessment. Regarding correction guidelines, English mistakes are weighed up to 20% of the final grade, and marked down following traditional corrective feedback (CF) focused on form, i.e., taking into account the linguistic competence over discourse, strategic and sociolinguistic competences. This emphasis on form or grammatical accuracy shows that writing is still tied to traditional styles of teaching and assessing, which emphasize more superficial aspects of the text, and leave behind aspects such as cohesion, vocabulary richness (Fernández, Lucero & Montanero, 2016), and naturalness of expression in relation to the particular academic genre. Therefore, we can conclude that although the impact of CLIL in secondary education can be observed in the increase of oral activities in daily teaching practice, a similar effect is not observed in the assessment practice of the Bilingual Sections of the CAM Bilingual Project.

## 7.2. Recommendations for improving assessment in CLIL

Given the lack of research on CLIL assessment, the different CLIL realities among countries, regions and even schools, and the fact that the type of Formative Assessment recommended for bilingual education has not been translated into real practice in some educational contexts, there is an urgent need to create some guidelines for CLIL assessment. What follows is a series of recommendations for improving assessment in CLIL in general, and to deal with linguistic aspects in content subjects in particular so that the language can be made visible along with content knowledge and skills.

Previous research on CLIL has concluded first, that assessment should be conducted in a formative way, by means of carefully selected assessment tools depending on the learning

goals. Second, that regardless of the treatment given to the language in CLIL, linguistic elements are paramount in the expression of content and skills and as such, they cannot be separated from content. The present study agrees with previous findings in all these regards. However, as CLIL is an "umbrella" term covering a broad range of scenarios, for adequate assessment in CLIL, the particular context in question should also be taken into account. Thus, the following guidelines are suited to the context of the study: The Bilingual Sections in the CAM Bilingual Project:

#### 7.2.1. Recommendations for the program organization

1. Specific guidelines and policies for bilingual education are urgently needed given the fact that the general ones from the Ministry of Education and the Madrid Regional Government refer to mainstream education and as such, they are insufficient for the reality of assessment in Bilingual Secondary Education. These guidelines might come from the educational administration or in their absence, the secondary schools in the CAM Bilingual Project could agree on a model and basic CLIL guidelines to deal with assessment in general, and the role and weight of the vehicular language in particular.
2. Assessment should mirror daily practice. The type of exams (if any) and the questions in them should be similar to the ones students deal with on a daily basis in that they are rooted in real life. In this regard, more innovative assessment tasks in line with Formative Assessment are needed for a variety of reasons: First, to abandon the prevalence of the traditional exam, which does not always allow the integration of competences in real-life, in favor of more task-based learning using for instance portfolios and journals. Second, to allow the students to show content knowledge and skills in a meaningful way, focusing not just on the final product



but also on the process. Third, to assess language “for a real purpose in a real context” (Coyle et al., 2010: 131). Likewise, although oral tasks are already implemented in the CLIL lessons, more efforts should be made to include them in assessment practice and thus, to give them more weight in the final grade.

3. Concerning both content and language support, and given the students’ limited language proficiency as compared to their L1, more scaffolding techniques apart from visual support are needed in formal and informal assessment situations. Contrarily to what can be assumed, having language support such as guiding questions or using language prompts does not necessarily menace the quality of assessment, but makes it possible for more cognitive challenging questions to be used in both daily teaching practice and assessment.
4. To measure language as integrated with content, and to avoid inhibiting students’ willingness to communicate in the foreign language, language mistakes in CLIL should be assessed differently than in FL teaching. Thus, apart from assessing the ability to recall subject-specific vocabulary, and show awareness of the grammatical features of language, the assessment tasks should be designed to measure the students’ ability to operate using functions which are appropriate for academic discourse. This focus on language functions, which is the responsibility of both content and language teachers, is not necessarily opposed to the focus on form but can complement it.
5. Since students need to be able to demonstrate thinking/reasoning in CLIL subjects, more cognitively demanding tasks should be included both in textbooks and in assessment tools. However, it is important to consider how different subjects use language to express content knowledge, and to choose receptive or productive tasks accordingly. This reflection on productive and receptive

language use could help relieve the pressure of using the foreign language in some contexts. It can also alleviate the problems of those students having limited language proficiency, and diminish the effect that language aspects have on their grades. Similarly, as CLIL is supposed to enhance and deepen the awareness on both the students' L1 and L2, letting students use their mother tongue in those tasks requiring more cognitive effort could be extremely beneficial to help them reach the level they need to establish the connections necessary for successful learning, at least in the initial stages or the first grades of secondary education.

6. If language production is still so present in CLIL assessment tools, as is the case in Social Sciences, maybe more writing components can be included as part of the curriculum planning (Boscardin et al., 2008: 7), and more attention should be given to them in the class, with the help of language teachers or language assistants.
7. As content teachers' practices reveal the lack of language and CLIL pedagogies typical of content teachers' background (Dalton-Puffer, 2013), more teacher training is needed in the context of the study to give the language aspects the importance they deserve.
8. Due to the lack of Formative Assessment practices in the context of our study, an urgent change is needed to suit assessment to both CLIL and the recommendations in the current law, and to abandon the testing culture in favor of the learning culture (Taras, 2005; Stobart, 2008). These changes should be welcome in bilingual schools since they are currently encouraged to introduce innovations regarding quality-oriented actions. Thus, on the one hand, achievement indicators reflecting clear assessment criteria in relation to both content and language goals should be shared with students. On the other hand, more Formative Assessment

tools promoting students' reflection on their learning such as self- and peer-assessment and portfolio use need to be implemented.

9. To ensure validity in the assessment process, and if teachers want to assess language-related objectives, language objectives should be made visible to students, worked with in the class, and its weight along with the most typical mistakes should be made explicit in the assessment practice.
10. To ensure reliability, and avoid subjectivity in the assessment practice, more rubrics should be created for productive tasks. Likewise, students should become familiar with these rubrics.
11. To ensure fairness, that is to give all students the possibility to express content knowledge, more scaffolding is needed at the initial stages to express complex or cognitively demanding ideas. It would even be possible to let students use their mother tongue occasionally. In any case, language errors in exams should never be considered as more serious than the ones students make in class activities and homework.
12. As for the standardized university entry exam students take at the end of compulsory education, which is thought to shape assessment practices, we need to highlight that a particular standard exam suited for bilingual education seems to be difficult to implement for several reasons: On the one hand, because unlike the bilingual model by the MEC-British Council, the CAM Bilingual Project lacks a British-Spanish specific curriculum. On the other hand, because taking the EvAU exam in the students' L1 allows educators to check whether students' level of content knowledge and the expression of skills has not been compromised by the effect of bilingual education (Personal conversation with Ismael Sanz, 29 th June 2017). Nevertheless, even in the absence of that specific entry exam in the

vehicular language, students might have the possibility to obtain a double certificate as is the case of the Bachibac in the French Linguistic Sections, or IGCSE (International General Certificate of Secondary Education). This could favor internationalization and student mobility, and reduce the feeling that students' efforts to learn in the foreign language may not pay off.

### 7.2.2. Recommendations for content teachers

In this scenario of traditional standard exams, and the lack of CLIL curricular guidelines for real integration of content, language and skills, more efforts are clearly needed so that content and language teachers work in collaboration with each other. Collaboration among teachers is recommended in the current educational law (LOMCE, 2013) as one of the signs of an effectively integrated and integrative curriculum, and by CLIL research (Pavón & Ellison, 2012; Kelly, 2014; Otto, 2017a). Teachers' willingness to collaborate with each other, and to discuss and agree on the most effective ways to deal with bilingual education in general, and language issues in particular, has been shown in this study. In fact, a common concern in bilingual programs today is how to adapt current assessment tools used in mainstream education, and whether it is possible to use the same type of assessment as in non-bilingual groups. What teachers might lack, however, is the time to gather, the awareness of the most urgent actions to be undertaken, lack of training in language pedagogies and some clarifications on the specific roles of both content and language teachers in CLIL instruction. Thus, more coordination time is required especially at the beginning of the academic year so that content and language teachers can plan their subjects in a cross-curricular way, and identify linguistic objectives in content subjects. Furthermore, the whole bilingual team should plan to join efforts, design common guidelines regarding the grading of language aspects, create a holistic

curriculum in which different areas could complement each other, and students holistically perceive the curriculum. The following are recommendations on how this collaboration can contribute to content teachers' assessment practice:

1. Every content teacher is to be considered a language teacher and as such, s/he is responsible for the language-related aspects in her/his subject (s). This teachers' language awareness, however, does not mean that content teachers are expected to teach language as in the foreign language class. Instead, as was pointed out on several occasions in this research, the content teacher should make these language objectives visible along with the content ones, and make sure that students notice the language patterns which shape the academic discourse.
2. To assess language aspects adequately, academic discourse deserves closer attention as it entails more than the use of CALP, but also to be able to operate using language functions. Thus, to move from academic language to academic discourse, attention should be devoted to training students to use the functional language which makes content expression possible (Schleppegrell, Achugar & Oteíza, 2004). In this regard, to identify language objectives in the Spanish bilingual education scenario, teachers can analyze “assessable learning outcomes” (LOMCE, 2013), and try to find the language-related objectives in the content subjects.
3. Once language goals have been analyzed, for language issues be treated effectively, the CLIL teacher should ideally work in constant collaboration with the language department. Collaborative work can be helpful on the one hand, to check whether those linguistic aspects are present in the curriculum of English as a Foreign Language, or if students studied them in the past, and spot the areas which might need reinforcement (if any). On the other hand, the content teacher

might need to know about the students' English level according to the CEFR, and the most common challenges they encounter as regards expression and comprehension in the foreign language. This will help teachers to adapt the activities and assessment tasks to suit the learners' needs especially at the initial stages. Finally, the content teacher can also ask the language specialists for advice about techniques to deal with language learning strategies, for instance, to ask students to underline the stressed syllable when they encounter new or difficult words (Otto, 2017 b: 6).

### 7.2.3. Recommendations for language teachers

As for language teachers, the suggestions below from Otto (2017 b: 6-7) are hoped to help language teachers work in close collaboration with content teachers:

1. In a bilingual school, the language department needs to be at the content teachers' disposal for help and advice throughout the academic year. Collaboration is paramount especially at the beginning of the school year, when the first meetings can be held to plan subjects in a cross-curricular way as is recommended by the CLIL approach and in the LOMCE (2013). In case CLIL teachers need it, language awareness seminars (Schlepppegrell, Achugar. & Oteíza, 2004) can also be scheduled to share ideas and good practice among all teachers. These seminars can help to train content teachers deal with the different aspects of the CEFR model of language use which are most appropriate to the CLIL context, and the type and level of language performance that needs to be assessed (Council of Europe, 2001: 27) depending on the subject in question. Similarly, coordination time can be used to train content teachers to assess communicative competence in

CLIL subjects by considering linguistic competence along with discourse, strategic and sociolinguistic competences. This focus on all types of communicative competence could help to avoid the prevalence of corrective feedback on form, and to preserve content-related issues as the main focus of CLIL assessment.

2. Students sometimes need to be trained in academic language since they find it challenging in certain contexts. Consequently, the language teacher might consider the inclusion of some language learning strategies, which can be shared with the content teachers, so that the students can holistically perceive the curriculum. For instance, students can be encouraged to learn vocabulary through chunks and collocations, or to learn in it in real contexts of use. Also, concerning learning how to learn, the language teacher can train students in skimming and scanning for reading and listening, and encourage them to use these techniques while tackling challenging texts in English regardless of the subject. Similarly, the language teachers can help the bilingual team with the design of rubrics which students can use for oral tasks, presentations and reports, to name just a few class activities. These rubrics might be shared with other colleagues in the team, so that they can be adapted and used in other CLIL subjects.
3. Promoting grammar learning from a notional/functional viewpoint can also help students to notice language functions in both foreign language contexts and CLIL subjects efficiently, and to move gradually from academic language to academic discourse. In noticing language structures and functions, as in content subjects, the language teacher could implement reflection on text structure: analyze different texts with students so that they see how ideas are linked and progression is made and then transfer those learning skills to the CLIL subjects. In this sense,

language teachers can also be encouraged to introduce language functions in the EFL curriculum (Lorenzo, Trujillo & Vez, 2011), and use academic texts in the foreign language class so that students perceive the curriculum in a holistic way.

4. Finally, bilingual coordinators are also paramount for the successful assessment practice of a bilingual school. First, they need to encourage joint planning, make sure that practices are agreed on during periodical bilingual coordination meetings, and promote discussion on standard criteria for assessing language-related aspects. They might also be in charge of coordinating collaborative work with the language assistants so that they can help content teachers with the language aspects which are challenging for students, and which impede them express content knowledge.

### 7.3. Suggestions for future research

Given the enormous impact that assessment has on teaching and learning, the lack of official guidelines, and the scarcity of research on CLIL assessment, further research on the following areas is still needed to ensure the development and success of bilingual programs. To start with, due to the limited sample in this study, replication of the present study with a larger sample to see if results are confirmed would be necessary. Second, given that this study was conducted under two educational laws, namely the LOE and the LOMCE, a new study in the context of the implementation of the LOMCE would help to observe whether Formative Assessment practices concerning, for instance, alternative assessment tools and self and peer-assessment, have already been implemented on a larger scale. Third, although study of the washback effect of different assessment formats on CLIL students' learning has timidly started (Lofft Basse, 2016; Pascual, 2017), more research is still needed on this subject matter. Fourth, since current recommendations to



improve assessment in CLIL refer specifically to the need for teacher training, it would be interesting to observe the effect on assessment procedures of training content subject teachers in language awareness, and more specifically on the weight given to language. Finally, as for teachers' cooperation in bilingual contexts, there is a need to evaluate experiences of collaboration between language and content teachers and their impact on assessment.



## REFERENCES

- A Teachers' Guide to TALIS 2013. Teaching and Learning International Survey* (2014). Paris, France: OECD Publishing. Retrieved from <http://www.oecd.org/edu/school/TALIS-Teachers-Guide.pdf>
- Activities that Take Place between Twinned Schools* (2014). Madrid, Spain: Subdirección General de Programas de Innovación. Dirección General de Innovación, Becas y Ayudas a la Educación. Retrieved from [http://www.educa2.madrid.org/web/educamadrid/principal/files/8c5dec3a-4665-4b9d-a373-46ab12fbef6b/Documentos/Curso%202016-2017/Triptico\\_madrid\\_bilingual\\_community\\_2016\\_v5.pdf?t=1478078255053](http://www.educa2.madrid.org/web/educamadrid/principal/files/8c5dec3a-4665-4b9d-a373-46ab12fbef6b/Documentos/Curso%202016-2017/Triptico_madrid_bilingual_community_2016_v5.pdf?t=1478078255053)
- Aguirre-Muñoz, Z., Park, J.E., Amabisca, A. & Boscardin, C. K. (2009). Developing Teacher Capacity for Serving ELL's Writing Instructional Needs: A Case for Systemic Functional Linguistics. In *Bilingual Research Journal*, 31, pp. 295-322.
- Airey, J. (2012). I don't teach language. The linguistic attitudes of physics lecturers in Sweden. In *AILA Review. Integrating Content and Language in Higher Education*, vol. 25, pp. 64-79.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? In *Applied Linguistics*, 14 (2), pp. 115-129.
- Alderson, J. C., & Wall, D. (1996). Editorial in *Language Testing*, 13 (3), pp. 239-240.
- Alternative assessment (2003)*. In Washington, USA: National Capital Language Resource Center (NCLRC, The George Washington University). Retrieved from <http://www.nclrc.org/essentials/assessing/alternative.htm>

- Alternative Assessment Strategies*. University of Minnesota Center for Educational Innovation (2015). Retrieved from <https://cei.umn.edu/node/856>
- Álvarez Valdivia, I. M. (2005). Evaluación como situación de aprendizaje: evaluación auténtica. In *Perspectiva Educacional*, 35, pp. 45-67.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete edition). New York, USA: Longman.
- Anderson, L. W. (2002). Curricular alignment: A re-examination. In *Theory into Practice*, 41 (4), pp. 255-260.
- Andrews, S. (2003). Teacher language awareness and the professional knowledge base of the L2 teacher. In *Language Awareness*, 12, pp. 81-95.
- Andrews, S. (2007). *Teacher language awareness*. Cambridge, UK: Cambridge University Press.
- Apple, D. K. & Krumsieg, K. (1998). *Process education teaching institute handbook*. New Hampshire, USA: Pacific Crest.
- Arkoudis, J. (2006). *Teaching International Students: Strategies to Enhance Learning*. Victoria Centre for Study of Higher Education, The University of Melbourne.
- Assessment for Learning Formative Assessment (2008). In *OECD/CERI International Conference Learning in the 21<sup>st</sup> Century: Research, Innovation and Policy*. Retrieved from <http://www.oecd.org/site/educeri21st/40600533.pdf>
- Assessment Reform Group, The (2002). *Assessment for Learning: Beyond the Black Box*. Retrieved from [http://www.nuffieldfoundation.org/sites/default/files/files/beyond\\_blackbox.pdf](http://www.nuffieldfoundation.org/sites/default/files/files/beyond_blackbox.pdf)
- Assessment Reform Group, The (2012). *Assessment for Learning: Ten Principles*.

Retrieved from

[http://www.hkeaa.edu.hk/DocLibrary/SBA/HKDSE/Eng\\_DVD/doc/Afl\\_principles.pdf](http://www.hkeaa.edu.hk/DocLibrary/SBA/HKDSE/Eng_DVD/doc/Afl_principles.pdf)

*Australia - Editable WILF WALT Display Signs* (2016). Sheffield, UK: Twinkl

Educational Publishing. Retrieved from <http://www.twinkl.co.uk/resource/au-m-031-australia---editable-wilf-walt-display-signs>

Ausubel, D. P. (1968). *Educational Psychology: A Cognitive View*. New York, US: Holt, Rinehart, & Winston.

Bachman, L. F. (1990). *Fundamental considerations in Language Testing*. Oxford, UK: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language Assessment in Practice: Developing Language Assessments and Justifying their Use in the Real World*. Oxford, UK: Oxford University Press.

Baeher, L. Funsworth, T. & Ediger, A. (2014). The Challenges of Planning Language Objectives in Content-Based ESL Instruction. In *Language Teaching Research 2014*, vol. 18 (I), pp. 118-136.

Baehr, M. (2010). Distinctions between Assessment and Evaluation. In S. Beyerlein, C. Holmes, & D. Apple, *Program Assessment Handbook*, pp. 441-444. Retrieved from [http://www.pcrest2.com/institute\\_resources/PAI/index.html](http://www.pcrest2.com/institute_resources/PAI/index.html)

Bailey, K. M. (1996). Working for Washback: A Review of the washback concept in language testing. In *Language Testing*, 13 (3), pp. 257-279.

Baker, C. (2002). Bilingual education. In Kaplan, R. B. (ed.) *The Oxford Handbook of Applied Linguistics*. Oxford: Oxford University Press.

Baker, C. (2011). *Foundations of Bilingual Education and Bilingualism*. Bristol, UK: Multilingual Matters.

- Baker, C., Lewis, G. & Jones, B. (2013). 100 Bilingual Lessons: Distributing Two Languages in Classrooms. In R. C. Abelló-Contesse et al (eds.), *Bilingual and Multilingual Education in the Twenty-First Century: Building on Experience*. Bristol, UK: Multilingual Matters.
- Ball, P. (2009). Does CLIL work? In D. Hill & P. Alan (eds.), *The Best of Both Worlds? International Perspectives on CLIL*, pp. 32-43. Norwich, UK: Norwich Institute for Language Education.
- Ball, P., Kelly, K. & Clegg, J. (2015). *Putting CLIL into Practice*. Oxford, UK: Oxford University Press.
- Banegas, D. (2012). *The challenges of CLIL Implementation in EFL Contexts*. In BETA 2012, Bulgaria: BETA-IATEFL Annual International Conference.
- Barbero, T., & Clegg, J. (2005). *Programmare percorsi CLIL*, Rome, Italy: Carocci
- Bentley, K. (2010). *The TKT Course: CLIL Module*. Cambridge, UK: Cambridge University Press.
- Bertaux, P., Coonan, C.M., Frigols, M.J., Mehisto, P. (2010). *The CLIL Teacher's Competence Grid*. Retrieved from <http://lendtrento.eu/convegno/files/mehisto.pdf>
- Bigelow, M. (2010). Learning to plan for a focus on form in CBI: The role of teacher knowledge and teaching context. In J. Davis (ed.), *World language teacher education: Transitions and challenges in the twenty-first century* (pp. 35-56). Greenwich, UK: Information Age Publishing.
- Bigelow, M., & Ranney, S. (2005). Pre-service ESL teachers' knowledge about language and its transfer to lesson planning. In N. Bartels (ed.), *Applied linguistics and language teacher education*, vol. 4 (pp. 179-200). New York, USA: Springer Science/ Business Media.

- Black, P. & William, D. (1998). Assessment and Classroom Learning. In *Assessment in Education: Principles, Policy & Practice*, vol. 5:1, pp. 7-74.
- Black, P. & William, D. (2001). Inside the Black Box: Raising Standards through Classroom Assessment. Retrieved from <http://www.rdc.udel.edu/wp-content/uploads/2015/04/InsideBlackBox.pdf>
- Block, D. (2003). *The Social Turn in Second Language Acquisition*. Edinburgh, Scotland: Edinburgh University Press.
- Boscardin, C. K., Jones, B., Nishimura, C., Madsen, S., Park, J. (2008). *Assessment of Content Understanding Through Science Explanation Tasks, CRESST Report 745*. Los Angeles, USA: University of California Graduate School of Education and Information Studies, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Brindusa, A., Cabrales, A., Carro, J. M. (2012). *Evaluating a bilingual education program in Spain: the impact beyond foreign language learning*. Retrieved from <https://e-archivo.uc3m.es/bitstream/id/65176/we1214.pdf/>
- Brinton, D. M, Snow, M. A. & Wesche, M. B. (1990). *Content-based Language Instruction*. New York, USA: Newbury House.
- Brown. H. D. (2004). *Language assessment: Principles and classroom practices*. New York, USA: Longman.
- Brown, H.D. & Hudson, T. (2002). *Criterion-referenced Language Testing*. Cambridge Applied Linguistic Series. Cambridge, UK: Cambridge University Press.
- Bruton, A. (2011). Is CLIL so beneficial or just selective? Re-evaluating some of the research. In *System* 39, pp. 523-532.
- Buckingham, L. (2015). Top ten tips for teachers with LAs. In *Revista digital EducaMadrid*. Retrieved from

[https://comunidadbilingue.educa2.madrid.org/web/revista-digital/inicio/-visor/lyndsay-r-buckingham-language-assistants2?visor=WAR&cms\\_tools\\_bookmarks=true](https://comunidadbilingue.educa2.madrid.org/web/revista-digital/inicio/-visor/lyndsay-r-buckingham-language-assistants2?visor=WAR&cms_tools_bookmarks=true)

- Byrner, H. (2008). Assessing Content and Language, in N. H. Hornberger, (ed.) *Encyclopedia of Language and Education*, pp. 2182-97. Boston, MA: Springer.
- Cabezuelo Gutiérrez, P. & Fernández Fernández, R. (2014). A case study on teacher training needs in the Madrid bilingual project. In *Latin American Journal of Content and Language Integrated Learning*, 7(2), pp. 50-70.
- Canale, M. & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. In *Applied Linguistics*, vol. 1 (1/1), pp 1-47.
- Casanova, M. A. (1998) *Evaluación: Concepto, tipología y objetivos*. Retrieved from [http://148.208.122.79/mcpd/descargas/Materiales\\_de\\_apoyo\\_3/evaluacion-tipologia\\_casanova.pdf](http://148.208.122.79/mcpd/descargas/Materiales_de_apoyo_3/evaluacion-tipologia_casanova.pdf)
- Cenoz, J. & Etxague, X. (2013). From Bilingualism to Multilingualism: Basque, Spanish and English in Higher Education. In R. C. Abelló-Contesse et al (eds.), *Bilingual and Multilingual Education in the Twenty-First Century: Building on Experience*. Bristol, England: Multilingual Matters.
- Chapelle, C. & Douglas, D. (1993). Foundations and directions for a new decade of language testing. In D. Douglas & C. Chapelle (eds.), *A new decade of language testing research*, pp. 1-22. Alexandria, VA: TESOL.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. In *Language and Education*, 11 (1), pp. 8-54.
- Cheng, L., Watanabe, Y. & Curtis, A. (eds.), (2004). *Washback in Language Testing: Research Contexts and Methods*. NJ, USA: Lawrence Erlbaum and Associates.



Clegg, J. (2007). Analysis the language demands of lessons taught in a second language.

In *Revista española de lingüística aplicada*, vol. extra 1, pp-113-128. Retrieved from

[http://ateneu.xtec.cat/wiki/form/wikiexport/media/cmd/lle/clsi/modul\\_2/dialnet-analysingthelanguagedemandsoflessonstaughtinasecon-2575499.pdf](http://ateneu.xtec.cat/wiki/form/wikiexport/media/cmd/lle/clsi/modul_2/dialnet-analysingthelanguagedemandsoflessonstaughtinasecon-2575499.pdf)

Clegg, J. (2012). Preface. In Quartapelle (ed.), *Assessment and Evaluation in CLIL*.

CLIL Assessment (2011). *Template for Analytic Rubrics. Assessment Grid for a Written*

*Test. CLIL/EMILE Project*. Retrieved from <http://aeclil.altervista.org/Sito/wp-content/uploads/2012/11/Grid-for-written-tests.pdf>

*CLIL Magazine*. Retrieved from:

[https://www.researchgate.net/publication/262414356\\_Developing\\_Literacy\\_through\\_CLIL](https://www.researchgate.net/publication/262414356_Developing_Literacy_through_CLIL)

CLIL/EMILE European Project for CLIL assessment. Lifelong Learning Program

(2011). *CLIL Assessment. Template for Analytic Rubrics. Assessment Grid for a*

*Written Test*. Retrieved from <http://aeclil.altervista.org/Sito/wp-content/uploads/2012/11/Grid-for-written-tests.pdf>

Chadwick, T. (2012). *Language Awareness in Teaching: A Toolkit for Content and Language Teachers*. Cambridge, UK: Cambridge University Press.

Communication from the Commission to the Council, the European Parliament, the

Economic and Social Committee and the Committee of the Regions. *Promoting language learning and linguistic diversity: an action plan 2004-2006*. Council

of Europe, 24<sup>th</sup> July 2003. Retrieved from <http://eur-lex.europa.eu/legal-content/LV/TXT/?uri=LEGISSUM:c11068>

- Content and Language Integrated Learning (CLIL) at School in Europe* (2006).  
 Bruselas, Bélgica: Eurydice. Retrieved from  
[http://www.indire.it/lucabas/lkmw\\_file/eurydice/CLIL\\_EN.pdf](http://www.indire.it/lucabas/lkmw_file/eurydice/CLIL_EN.pdf)
- Coonan, C. M. (2007). Insider Views of the CLIL Class through Teacher Self-observation-Introspection. In *International Journal of Bilingual Education and Bilingualism*, n.º 10, pp. 625-646.
- Coonan, C.M. (2012). The Foreign Language Curriculum and CLIL. In *Synergies Italy*, n.º 8, pp. 117-128.
- Coonan, C.M. (2013). *The role of CLIL teachers*. Paper presented at CLO CLIL 2, Bari. Retrieved from <https://www.youtube.com/watch?v=doN8oLApaSU>
- Cope, B. & Kalantzis, M. (eds.) (1993). *The powers of literacy: A genre-based approach to teaching literacy*. London, UK: Falmer Press.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.
- Coyle, D. et al (2003). *Developing CLIL: Towards a Theory of Practice*. Gerona, España: APAC.
- Coyle, D. (2015). Pluriliteracies, pedagogies, principles and practices: what research can tell us about integrated learning. Talk delivered in the CLIL COLLOQUIUM. *Integrating content and language for teacher development in bilingual/multilingual educational settings: From research to practice*. Madrid, España: UAM.

Coyle, D. (2005). Planning and Monitoring CLIL. University of Nottingham: UK.

Retrieved from [https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/coyle\\_clil\\_planningtool\\_kit.pdf](https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/coyle_clil_planningtool_kit.pdf)

Coyle, D., Holmes, B. & King, L. (2009). *Towards and Integrated Curriculum. CLIL National Statements and Guidelines*. London, UK: The Languages Company.

Retrieved from

<http://www.rachelhawkes.com/PandT/CLIL/CLILnationalstatementandguidelines.pdf>

Coyle, D., Hood, P. & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge, UK: Cambridge University Press.

Cross, R. & Gearon, M. (2013) *Research and Evaluation of the Content and Language Integrated Learning (CLIL) Approach to Teaching and Learning Languages in Victorian Schools*. Retrieved from

<http://www.education.vic.gov.au/Documents/school/teachers/teachingresources/discipline/languages/CLILtrialresearchrpt.pdf>

Cummins, J. (1991). *Language Development and Academic Learning*. In L. Malave & G. Duquette, *Language, Culture and Cognition*. Bristol, UK: Multilingual Matters.

Cummins, J. (2001). *Negotiating identities: Education for empowerment in a diverse society*. 2<sup>nd</sup> Edition. Los Angeles, USA: Association for Bilingual Education.

Cummins, J. (2008). BICS and CALP: Empirical and Theoretical Status of the Distinction. In Cummins, Street, B. & Hornberger, N. H. (eds.). *Encyclopedia of Language and Education*, 2<sup>nd</sup> Edition, vol. 2: Literacy, pp. 71-83. New York, USA: Springer Science + Business Media LLC.

- Cushing Weigle, S. & Jensen, L. (1997). Issues in Assessment for Content Based Instruction. In M. A. Snow & D. Brinton, (eds.) *The Content-Based Classroom: Perspectives on Integrating Language and Content*, pp. 200-220. New York, USA: Longman.
- Dale, L. and Tanner, R. (2012). *CLIL Activities: A Resource for Subject and Language Teachers*. Cambridge, UK: Cambridge University Press.
- Dalton-Puffer, C. (2007). Discourse in Content and Language Integrated Learning (CLIL) Classrooms. In *Applied Linguistics*, Volume 31, Issue 3, 1 July 2010, pp. 475–478.
- Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In W. Delanoy and L. Volkman (eds.), *Future Perspectives for English Language Teaching*, pp. 139-157. Heidelberg, Germany: Universitätsverlag Winter GmbH.
- Dalton-Puffer, C. (2013). A Construct of Cognitive Discourse Functions for Conceptualising Content-language Integration in CLIL and Multilingual Education. In *European Journal of Applied Linguistics*, 1(2), pp. 216-253.
- Dalton-Puffer, C. (2015). *Policy and Practice of CLIL in Europe and Beyond*. Retrieved from <http://www.cliljapan.org/wp-content/uploads/2015/02/Dalton-Puffer-CLIL-Sophia-University-2015.pdf>
- Datos y cifras. Curso escolar 2016-2017* (2016). Madrid, Spain: Secretaría General Técnica. Subdirección General de Documentación y Publicaciones. Ministerio de Educación, Cultura y Deporte. Retrieved from <http://www.mecd.gob.es/servicios-al-ciudadano-mecd/dms/mecd/servicios-al-ciudadano-mecd/estadisticas/educacion/indicadores-publicaciones-sintesis/datos-cifras/Datosycifras1617esp.pdf>

- Davison, C. M. (2006). Collaboration Between ESL and Content Teachers: How Do We Know When We Are Doing It Right? In *International Journal of Bilingual Education and Bilingualism*, v. 9 n. 4, pp. 454-475. Retrieved from [https://www.researchgate.net/profile/Chris\\_Davison3/publication/254242346\\_Collaboration\\_Between\\_ESL\\_and\\_Content\\_Teachers\\_How\\_Do\\_We\\_Know\\_When\\_We\\_Are\\_Doing\\_It\\_Right/links/55aeb008aed614b09a7855.pdf](https://www.researchgate.net/profile/Chris_Davison3/publication/254242346_Collaboration_Between_ESL_and_Content_Teachers_How_Do_We_Know_When_We_Are_Doing_It_Right/links/55aeb008aed614b09a7855.pdf)
- Del Pozo, E. (2009). Portfolio: Un paso más allá en CLIL. In *Prácticas en educación bi/plurilingüe*, 1, pp. 35-41. Retrieved from [http://www.letra25.com/ediciones/bilingue\\_01/index.html](http://www.letra25.com/ediciones/bilingue_01/index.html)
- Diario Oficial de las Comunidades Europeas. N.º C 207/1. *Resolución del Consejo de 31 de marzo de 1995 relativa a la mejora de la calidad y la diversificación del aprendizaje y de la enseñanza de las lenguas en los sistemas educativos de la Unión Europea*. 12 de agosto de 1995. Retrieved from [http://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:31995Y0812\(01\)&from=ES](http://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:31995Y0812(01)&from=ES)
- Díaz Cobo, A. (2009). Assessment instruments for CLIL written production tasks. In D. Marsh et al (eds.), *CLIL Practice: Perspectives from the Field*, pp. 139-148. Jyväskylä, Finland: University of Jyväskylä.
- Dobson, A., Pérez Murillo, M. D. & Johnstone, R. (2010). *Bilingual Education Project Spain Evaluation Report*. Ministerio de Educación. Madrid, Spain: Instituto de Formación del Profesorado, Investigación e Innovación Educativa (IFIIE) & British Council, Spain.
- Doiz, A., Lasagabaster, D. & Sierra, J. M. (2013). *English-Medium Instruction at Universities. Global Challenges*. Bristol, UK: Multilingual Matters.

- Echevarría, J, Vogt, M. E. & Short, D. (2013). *Making Content Comprehensible for English Learners: The SIOP Model*. New York, USA: Pearson.
- Educamadrid P. (2012). Senise, N., Taller: Organización del Programa Bilingüe en los colegios Públicos de la Comunidad de Madrid. In *I Congreso Internacional sobre Bilingüismo en Centros Educativos el 14 de junio de 2010 dirigido a profesores de primaria, secundaria y universidades, a investigadores y responsables políticos interesados en la educación bilingüe y en metodología AICOLE (Aprendizaje Integrado de Contenidos y Lengua)*. Retrieved from <https://mediateca.educa.madrid.org/video/rsb7cefsk2swfz6c>
- Education at a Glance 2015: OECD Indicators* (2015). Paris, France: OECD Publishing. Retrieved from [http://download.ei-ie.org/Docs/WebDepot/EaG2015\\_EN.pdf](http://download.ei-ie.org/Docs/WebDepot/EaG2015_EN.pdf)
- European Commission/EACEA/Eurydice. (2012). *Developing Key Competences at School in Europe: Challenges and Opportunities for Policy*. Eurydice Report. Luxembourg: Publications Office of the European Union.
- European Commission/EACEA/Eurydice. (2017). *Key Data on Teaching Languages at School in Europe – 2017 Edition*. Eurydice Report. Luxembourg: Publications Office of the European Union.
- Ellis, R. (2009). Corrective Feedback and Teacher Development. In *L2 Journal*, vol. 1, pp. 3-18. Retrieved from <http://repositories.cdlib.org/uccllt/12/vol1/iss1/art2/>
- Ernst, M. (1995). Fehlerkorrektur und Leistungsbewertung im bilingualen Sachfachunterricht. In *Praxis des neusprachlichen Unterrichts*, 42 /3, pp. 258-264.
- Escobar Urmeneta, C. (2006). Una propuesta PEL de evaluación criterial para la educación secundaria. In D. Cassany, (ed.), *Portfolio Europeo de las Lenguas de*

- secundaria y su aplicación en el aula*, pp. 77-108. Madrid, Spain: Instituto Superior de Formación del Profesorado del MEC de España.
- Ferjan Ramírez, N. & Kuhl, P. (2017). Bilingual Baby: Foreign Language Intervention in Madrid's Infant Education Centers. In *Mind, Brain, and Education*, 11, pp. 133–143.
- Fernández, M. J., Lucero, M. & Montanero, M. (2016). Red on Black: How Do Teachers Assess Their Students' Narrative Texts? In *Revista de Educación*, 372, pp. 63-86. Secretaría General Técnica. Retrieved from:  
[https://sede.educacion.gob.es/publiventa/descarga.action?f\\_codigo\\_agc=17517](https://sede.educacion.gob.es/publiventa/descarga.action?f_codigo_agc=17517)
- Fernández, R. & Halbach, A. (2011). Analysing the Situation of Teachers in the Madrid Bilingual Project after Four Years of Implementation. In Y. Ruiz de Zarobe, J. Manuel Sierra and F. Gallardo del Puerto (eds.), *Content and Foreign Language Integrated Learning. Contributions to Multilingualism in European Contexts*, pp. 41-70. Bern, Switzerland: Peter Lang.
- First European Survey on Language Competences. Final Report* (2011). Retrieved from  
[http://ec.europa.eu/dgs/education\\_culture/repository/languages/policy/strategic-framework/documents/language-survey-final-report\\_en.pdf](http://ec.europa.eu/dgs/education_culture/repository/languages/policy/strategic-framework/documents/language-survey-final-report_en.pdf)
- Firth, A. & Wagner, J. (1997). On discourse, communication and (some) fundamental concepts in SLA research. In *The Modern Language Journal*, 81, 3, pp. 285-300. New Jersey, USA: Wiley-Blackwell.
- Flowerdew, J. & Peacock, M. (2001). Issues in E. A. P.: A Preliminary Perspective. In Flowerdew & Peacock, (Eds.), *Research Perspectives on English for Academic Purposes*. Cambridge, UK: Cambridge University Press.
- Foran, D. (2012). *Rubrics for Formative Assessment*, talk at the CIEB Congress. Unpublished manuscript, Madrid.

*Formative Assessment that Truly Informs Instruction* (2013). Illinois, USA: NCTE.

Retrieved from

[http://www.ncte.org/library/NCTEFiles/Resources/Positions/formative-assessment\\_single.pdf](http://www.ncte.org/library/NCTEFiles/Resources/Positions/formative-assessment_single.pdf)

Fortanet-Gómez, I. (2013). *CLIL in higher education: Towards a multilingual language policy*. Bristol, UK: Multilingual Matters.

Fortune, T. & Teddick, D. (2015). *Some features of language objectives for the CBI-classroom. Adapted after Fortune and Tedick*. From the CoBaLTT Project website: <http://carla.umn.edu/cobalitt/modules/curriculum/features.pdf>

Frigols, M. J. (2008). CLIL implementation in Spain: an approach to different models. In Coonan, C. M (ed.). *CLIL e l'apprendimento delle lingue. Le sfide del nuovo ambiente di apprendimento*. Venezia, Italy: Librería Editrice Cafoscarina.

Retrieved from <http://arcaold.unive.it/bitstream/10278/1013/1/13Frigols.pdf>

*From Linguistic Diversity to Plurilingual Education: Guide for the Development of Language Education Policies in Europe*. Council of Europe, 2007. Retrieved from [www.coe.int/t/dg4/linguistic/Source/Guide\\_Main\\_Beacco2007\\_EN.doc](http://www.coe.int/t/dg4/linguistic/Source/Guide_Main_Beacco2007_EN.doc)

Fuentes Arjona, M. A. (2013). Which score is adequate: approximation to the assessment rationale used in Science through English CLIL written text. In *Bellaterra Journal of Teaching & Learning Language & Literature*, vol. 6 (4), pp 54-73. Retrieved from

<http://www.raco.cat/index.php/Bellaterra/article/viewFile/297851/386834>

Fullana, J. y Pallisera, M. (1995). Aplicación de las pruebas criterios en la evaluación de los procedimientos. In *Bordón*, 47 (1), pp. 41-51.



- Gablasova, D. (2014). Learning and Retaining Specialized Vocabulary from Textbook Reading: Comparison of Learning Outcomes Through L1 and L2. In *The Modern Language Journal*, 98, pp. 976–991.
- García Laborda, J. y Fernández Álvarez, M. (2011). Teachers' opinions towards the integration of oral tasks in the Spanish University Examination. In *International Journal of language studies*, 5 (3), pp. 1-12.
- García López, M. & Bruton, A. (2013). Potential Drawbacks and Actual Benefits of CLIL Initiatives in Public Secondary Schools, In C. Abello-Contesse et al., *Bilingual and Multilingual Education in the 21st Century*. Bristol, UK: Multilingual Matters.
- García, O. & Lin, A. (2014) *Translanguaging in Bilingual Education* (PDF Download Available). Retrieved from:  
[https://www.researchgate.net/publication/312053254\\_Translanguaging\\_in\\_Bilingual\\_Education](https://www.researchgate.net/publication/312053254_Translanguaging_in_Bilingual_Education)
- Gardner, J., Harlen, W., Hayward, L. & Stobart, G. (2008). *Changing Assessment Practice: Process, Principles and Standards*. London, UK: Assessment Reform Group. Retrieved from <http://www.aiaa.org.uk/content/uploads/2010/06/ARIA-Changing-Assessment-Practice-Pamphlet-Final.pdf>
- Gibbons, P. (2002). *Scaffolding Language, Scaffolding Learning: Teaching Second Language Learners in the Mainstream Classroom*. New Hampshire, USA: Heinemann.
- Gómez Arceo, B. M. (1990). *Evaluación criterial (una metodología útil para diagnosticar el nivel de aprendizaje de los alumnos)*. Madrid, Spain: Narcea.
- Gondová, D. (2012). Identifying the Needs of CLIL Teachers at Lower and Upper-Secondary Schools. In *Journal of Interdisciplinary Philology*, 3 (3), pp. 5-22.

- González, J. A., & Barbero, J. (2013). Building bridges between different levels of education: Methodological proposals for CLIL at university. In *Language Value*, 5(1), pp. 1-23.
- Gottlieb, M. (2006). *Assessing English language learners: Bridges from language proficiency to academic achievement*. Thousand Oaks, Calif: Corwin Press.
- Graddol, D. (2006). *English next: Why global English may mean the end of 'English as a foreign language'*. London, UK: British Council Publications.
- Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, USA: Harvard University Press.
- Halbach, A. (2008). Una metodología para la enseñanza bilingüe en la etapa de primaria. In *Revista de educación*, 346, pp. 455-466. Retrieved from [http://www.ince.mec.es/revistaeducacion/re346/re346\\_17.pdf](http://www.ince.mec.es/revistaeducacion/re346/re346_17.pdf)
- Halbach, A. (2010). From the Classroom to University and Back: Teacher Training for CLIL in Spain at the University de Alcalá. In Lasagabaster & Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 243-256.
- Halbach, A. (2012). Same but different: La enseñanza de la lengua extranjera en el Proyecto bilingüe. In A. Halbach & T. Hernández González (eds.), *Retos de la Educación Bilingüe*, pp. 24-37. Madrid, España: Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación.
- Halbach, A., Lázaro Lafuente, A. & Pérez Guerra, J. (2013). La lengua inglesa en la nueva universidad española del EEES. In *Revista de Educación*, 362. pp. 105-132. Retrieved from [http://www.revistaeducacion.mec.es/doi/362\\_154.pdf](http://www.revistaeducacion.mec.es/doi/362_154.pdf)
- Halliday, M. A. K. (1994a). The Construction of Knowledge and Value in the Grammar of Scientific Discourse, with Reference to Charles Darwin's *The Origin of*

- Species*. In M. Coulthard (ed.), *Advances in Written Text Analysis*, pp. 136-156.  
New York. USA: Routledge.
- Halliday, M. A. K. (1994b). On language in relation to the evolution of human consciousness. In S. Allen (ed.), *Of thoughts and words (Proceedings of Nobel Symposium 92: The relation between language and mind)*. London, UK: Imperial College Press.
- Halliday, M. A. K. & Martin, J. R. (eds.), (1993). *Writing science: Literacy and discursive power*. London, UK: The Falmer Press.
- Halliday, M. & Matthiesen, C. (2004). *An Introduction to Functional Grammar*. London, UK: Edward Arnold.
- Hargett, G. R. (1998). *Assessment in ESL and Bilingual Education: A Hot Topics Paper*. Retrieved from <http://files.eric.ed.gov/fulltext/ED425645.pdf>
- Hargreaves, A. (1994). Collaboration and contrived collegiality: Cup of comfort or poisoned chalice? In A. Hargreaves (ed.), *Changing Teachers, Changing Times*, pp. 186-211. London, UK: Cassell.
- Hargreaves, A. & Macmillan, B. (1994). The balkanisation of teaching: Collaboration that divides. In A. Hargreaves (ed.), *Changing Teachers, Changing Times*, pp. 212-240. London, UK: Cassell.
- Harlen, W. (2012). The Role of assessment in developing motivation for learning, in J. Gardner (ed.), *Assessment and Learning*, pp. 61-80. London, UK: SAGE
- Harlen, W. (2012). On the Relationship between Assessment for Formative and Summative Purposes. In J. Gardner (ed.), *Assessment and Learning*, pp. 87-102 London, UK: SAGE.

- Harris, L. & Brown, G. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. In *Assessment in education: principles, policy & practice*, vol. 16. London, UK: Routledge.
- Hasan, R. & Williams, G. (eds.), (1996). *Literacy in society*. London, UK: Longman.
- Heather, B. (2012). The amazing effects of WALT and WILF: assessment for learning versus assessment of learning. In A. Halbach & T. Hernández González, *Retos de la educación bilingüe*, pp. 160-174. Madrid, Spain: Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación.
- Hellerman, J. (2008). *Social Actions for Classroom Language Learning*. Clevedon, UK: Multilingual Matters.
- Heredia Manrique, A. (2009). Two Differentiated Forms of Didactic Evaluation: Normative Evaluation in order to Select Students and Criterial Evaluation to Master the Basic Content. In *Bordón*, 61 (4), pp. 39-48. Retrieved from <https://recyt.fecyt.es/index.php/BORDON/article/viewFile/28812/15367>
- Heritage, M. (2008). *Learning Progressions: Supporting Instruction and Formative Assessment*. Washington, USA: Council of Chief State School Officers.
- Hertz, M. B. (2013). *Using e-portfolios in the Classroom*. Retrieved from: <http://www.edutopia.org/blog/e-portfolios-in-the-classroom-mary-beth-hertz>
- Hönig, I. (2010). *Assessment in CLIL: Theoretical and Empirical Research*. Saarbrücken, Germany: VDM Verlag Dr. Müller.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge, UK: Cambridge University Press.
- Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished manuscript. Berkshire, UK: University of Reading.

- Hulstijn, J. H. (2015). *Language proficiency in native and non-native speakers: Theory and research*. Amsterdam, The Netherlands: John Benjamins.
- INEE (2012). *Estudio Europeo en Competencia Lingüística. Informe Español*. Secretaría Madrid, Spain: General Técnica. Subdirección General de Documentación y Publicaciones. Retrieved from <http://www.mecd.gob.es/dctm/ievaluacion/internacional/eeclvolumeni.pdf?documentId=0901e72b813ac515>
- Ioannau-Georgiou S. & Pavlou, P. (2011). *Guidelines for CLIL Implementation in Pre-Primary and Primary Education*. Nicosia, Chipre: PROCLIL. Retrieved from [http://www.schools.ac.cy/klimakio/Themata/Anglika/teaching\\_material/clil/guidelinesforclilimplementation1.pdf](http://www.schools.ac.cy/klimakio/Themata/Anglika/teaching_material/clil/guidelinesforclilimplementation1.pdf)
- Jacobs, M., & Farrell, T. (2001). *Paradigm Shift: Understanding and Implementing Change in Second Language Education*. Retrieved from: <http://tesol.org/ej17/a1.html.archive>
- Järvinen, H. M. (ed.), (2009). *Handbook: Language in Content Instruction*. Turku, Finland: University of Turku. Retrieved from [https://broadyesl.files.wordpress.com/2012/09/language-in-content\\_handbook\\_en.pdf](https://broadyesl.files.wordpress.com/2012/09/language-in-content_handbook_en.pdf)
- Johnston, R. & Watson, J. (2005). *The Effects of Synthetic Phonics Teaching on Reading and Spelling Attainment*. Retrieved from <http://www.gov.scot/Resource/Doc/36496/0023582.pdf>
- Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A. (2007). Toward a Definition of Mixed Method Research. In *Journal of Mixed Method Research*, Volume 1, Issue 2, pp. 112-133.

- Johnstone, R. & Dobson, A. (2012). Educating Pupils at School through the Medium of More than Language: Some Reflections on Spain in the International Context. In A. Halbach, (ed.) & T. Hernández González. *Retos de la Educación Bilingüe*, pp. 8-23. Madrid, Spain: Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación.
- Jones, C. (2009). Interdisciplinary Approach-Advantages, Disadvantages, and the Future Benefits of Interdisciplinary Studies. In *ESSAI*, vol. 7 (26). Retrieved from <http://dc.cod.edu/essai/vol7/iss1/26/>
- Kasper, G. (2009). Locating Cognition in second language interaction and learning. Inside the skull or in public view? In *IRAL-International Review of Applied Linguistics in Language Teaching*, 47, 1, pp. 11-36.
- Kelly, K. (2010). CLIL in natural science subjects: Language and task. In *Ikastaria*, 17, pp. 91-108.
- Kelly, K. (2014). *COOP E-CLIL*. Retrieved from <https://www.teachingenglish.org.uk/sites/teacheng/files/Language%20teachers%20working%20with%20subject%20teachers%20-%20TE.pdf>
- Kiely, R. (2009). CLIL: The Question of Assessment. Retrieved from [https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/kiely\\_r.\\_clil\\_assessment.pdf](https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/kiely_r._clil_assessment.pdf)
- Kizlik, B. (2017). *Measurement, Assessment and Evaluation in Education*. Retrieved from <http://www.adprima.com/measurement.htm>
- Krashen, S.D. (1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.
- Krueger, R. A. (1994). *Focus Groups: A Practical Guide for Applied Research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

- Krueger, R.A., & Casey, M.A. (2009). *Focus Groups: A Practical Guide for Applied Research* (4th Ed.) Thousand Oaks, CA: Sage Publications.
- Lasagabaster, D. & García, O. (2014). Translanguaging: Towards a dynamic model of bilingualism at school /Translanguaging: hacia un modelo dinámico de bilingüismo en la escuela, In *Cultura y Educación*, vol. 26:3, pp. 557-572.
- Retrieved from:  
[https://www.researchgate.net/publication/280171216\\_Translanguaging\\_towards\\_a\\_dynamic\\_model\\_of\\_bilingualism\\_at\\_school\\_Translanguaging\\_hacia\\_un\\_modelo\\_dinamico\\_de\\_bilinguismo\\_en\\_la\\_escuela](https://www.researchgate.net/publication/280171216_Translanguaging_towards_a_dynamic_model_of_bilingualism_at_school_Translanguaging_hacia_un_modelo_dinamico_de_bilinguismo_en_la_escuela)
- Lasagabaster, D. & Sierra, J. M. (2010). Immersion and CLIL in English: More Differences than Similarities. In *ELT Journal*, vol. 64/4, pp. 367-375. Oxford, UK: Oxford University Press.
- Lasagabaster, D., & Ruiz, Z. Y. (2010). *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle, UK: Cambridge Scholars.
- Laorden Gutiérrez, C. & Peñafiel Pedrosa, E. (2010). Proyectos bilingües en los centros de la Comunidad de Madrid: Percepción de los equipos directivos. In *Revista de Investigación Educativa*, 28 (2), pp. 325-344. Alcalá de Henares, Spain: Escuela Universitaria Cardenal Cisneros. Retrieved from  
<https://revistas.um.es/rie/article/viewFile/98101/121621>
- Ley orgánica para la mejora de la calidad educativa (LOMCE)*, (Ley Orgánica 8/2013, 9 de diciembre). Boletín Oficial del Estado, n.º 295, 10 diciembre de 2013.
- Lightbown, P. M., & Spada, N. (1994). An innovative program for primary ESL students in Quebec. In *TESOL Quarterly*, 28 (3), pp. 563-579.
- Llinares, A. (2012). *The challenge of integrating language and content in CLIL*. II Foro Nebrija en Enseñanza Bilingüe. Retrieved from:

[http://www.nebrija.com/la\\_universidad/facultades/facultad-artes-letras/actividades/foro-nebrija-bilingue/pdf/the-challenge-of-integrating-in-clil-linares.pdf](http://www.nebrija.com/la_universidad/facultades/facultad-artes-letras/actividades/foro-nebrija-bilingue/pdf/the-challenge-of-integrating-in-clil-linares.pdf)

- Llinares, A. and Whittaker, R. (2006). Linguistic analysis of secondary school students' oral and written production in CLIL Contexts: Studying Social Science in English. In Dalton-Puffer et al., *Current Research on CLIL*, vol. 15, 3, pp. 28-32. Retrieved from [http://anglistik.univie.ac.at/fileadmin/user\\_upload/dep\\_anglist/weitere\\_Uploads/Views/views15\\_3\\_clil\\_special.pdf#page=28](http://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/views15_3_clil_special.pdf#page=28)
- Llinares A. and Dafouz, E. (2010). Content and Language Integrated Programmes in the Madrid Region: Overview and Research Findings. In Lasagabaster & Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 95-114.
- Llinares, A., Morton, T. & Whittaker, R. (2012). *The Roles of Language in CLIL*. Cambridge, UK: Cambridge University Press.
- Lo, Y. & Lin, A. (2014). Designing Assessment Tasks with Language Awareness: Balancing Cognitive and Linguistic Demands. In *Assessment and Learning*, 3, pp. 97-119. Retrieved from [http://wlts.edb.hkedcity.net/filemanager/file/A&L3\(11\)\\_Lo&Lin.pdf](http://wlts.edb.hkedcity.net/filemanager/file/A&L3(11)_Lo&Lin.pdf)
- Lodico, M., Spaulding, D. & Voetgle, K. (2006). *Methods in Educational Research: From Theory to Practice*. San Francisco, USA: Jossey-Bass.
- Lofft Basse, R. (2016). Assessment for learning in the CLIL classroom: A corpus based study of teacher motivational L2 strategies and student motivation and metacognitive abilities. (Doctoral Dissertation). Madrid, Spain: Universidad



- Autónoma de Madrid. Facultad de Filosofía y Letras/Departamento de Filología Inglesa. Retrieved from <https://repositorio.uam.es/handle/10486/671766>
- López-Medina, B. (2016). Developing a CLIL Textbook Evaluation Checklist. In *Latin American Journal of Content & Language Integrated Learning*. Vol. 9 (1), pp. 159-173.
- Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. In *Applied Linguistics*, 31 (3), pp. 418-442.
- Lorenzo, F., Trujillo, F. & Vez, J.M. (2011). *Educación bilingüe: integración de contenidos y segundas lenguas*. Madrid: Síntesis.
- Lukas, J. F., Santiago, K., Joaristi, L. & Lizasoain L. (2006). Usos y formas de la evaluación por parte del profesorado de la ESO. Un modelo multinivel. In *Revista de Educación*, 340, pp. 667-693. Retrieved from: [http://www.revistaeducacion.mec.es/re340/re340\\_24.pdf](http://www.revistaeducacion.mec.es/re340/re340_24.pdf)
- Lynch, B. K. (2001). Rethinking Assessment from a Critical Perspective. In *Language Testing*, 18 (4), pp. 351-372.
- Lyster, R. (2007). *Learning and Teaching Languages through Content: A Counterbalanced Approach*. Amsterdam and Philadelphia, PA: Benjamins
- Maggi, F. (2012). Towards and Integrated Evaluation. In F. Quartapelle, *Assessment and Evaluation in CLIL*. Ibis, Pavia: Italia. Retrieved from <http://riviste.unimi.it/index.php/promoitals/article/viewFile/2827/3030>
- Marsh, D. (2000). Using Languages to Learn and Learning to Use Languages. An introduction to CLIL for parents and young people. In *TIE-CLIL*. Retrieved from [www.tieclil.org/html/products/pdf/%201%20UK.pdf](http://www.tieclil.org/html/products/pdf/%201%20UK.pdf)

- Marsh, D. (2007). Language Awareness & CLIL. In J. Cenoz, J & N. Hornberger, (eds.), *Encyclopedia of Language and Education: Knowledge about Language*, vol. 6, 2<sup>nd</sup> edition, pp. 233-246. New York & Berlin: Springer Science and Business Media.
- Marsh, D. & Frigols Martín, M. J. (2012). Content and Language Integrated Learning. In *The Encyclopedia of Applied Linguistics*. New Jersey, USA: Wiley-Blackwell.
- Marsh, D., Marshland, B., & Stenberg, K. (2001). *CLIL compendium. Profiling European CLIL classrooms*. Jyväskylä, Finland: University of Jyväskylä.
- Marsh, D., Mehisto, P., Wolff, D., Frigols Martin, M. J. (2010). *European Framework for CLIL Teacher Education: A framework for the professional development of CLIL teachers*. Graz, Austria: European Centre for Modern Languages.  
Retrieved from [https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/european\\_framework\\_for\\_clil\\_teacher\\_education.pdf](https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/european_framework_for_clil_teacher_education.pdf)
- Martin, J. R. & Veel, R. (eds.), (1998). *Reading science: Critical and functional perspectives on discourses of science*. London, UK: Routledge.
- Martín del Pozo, M. A. & Rascón Estébanez, D. (2015). Textbooks for Content and Language Integrated Learning: policy, market and appropriate didactic. In *Foro de Educación*, 13 (18), pp. 123-141. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/5153348.pdf>
- Martínez López, J. & Cantero García, V. (2014). In Search of an Effective Model for Assessing Learning in Bilingual Education: The Authentic Assessment. In *Porta Linguarum*, 21, pp. 137-150. Retrieved from [http://www.ugr.es/~portalin/articulos/PL\\_numero21/9%20%20Juan%20A.pdf](http://www.ugr.es/~portalin/articulos/PL_numero21/9%20%20Juan%20A.pdf)

- Martínez Rizo, F. (2013). Los sistemas nacionales de indicadores como instrumentos de transparencia. Limitaciones y posibilidades. In *Participación Educativa*, vol. II, 3, pp. 67-74. Retrieved from [https://sede.educacion.gob.es/publiventa/descarga.action?f\\_codigo\\_agc=16171](https://sede.educacion.gob.es/publiventa/descarga.action?f_codigo_agc=16171)
- Marzano, R. (2010). *Formative Assessment and Standard-Based Grading*. Indiana, USA: Solution Tree.
- Marzano, R. (2012). *Providing Clear Learning Goals and Scales (Rubrics)*. Single School Culture for Academics. Palm Beach Country. Retrieved from <http://staff.camias.wednet.edu/blogs/tpep/files/2012/06/DQElement1PacketforTeachers.pdf>
- Massler, U. (2011). Assessment in CLIL learning. In Ioannau-Georgiou S. & Pavlou, P. (2011). Guidelines for CLIL Implementation in Pre-Primary and Primary Education, pp.114-136. Nicosia, Chipre: PROCLIL. Retrieved from [http://www.schools.ac.cy/klimakio/Themata/Anglika/teaching\\_material/clil/guidelinesforclilimplementation1.pdf](http://www.schools.ac.cy/klimakio/Themata/Anglika/teaching_material/clil/guidelinesforclilimplementation1.pdf)
- Mateo, J. A. (2000). *La evaluación educativa, su práctica y otras metáforas*. Barcelona, Spain: ICE, Universidad de Barcelona. Retrieved from <https://web.ua.es/es/ice/documentos/recursos/materiales/la-evaluacion-educativa.pdf>
- McKay, P., Davies, A., Devlin, B., Clayton, J., Oliver, R., & Zammit, S. (1997). *The bilingual interface project report*. Canberra, Australia: Department of Employment, Education, Training & Youth Affairs. Retrieved from <http://www.iier.org.au/qjer/qjer15/mckay.html>
- McNamara, T. (2000). *Language Testing*. Oxford University Press.

- Megías Rosa, M. (2012). Formación, Integración y Colaboración: palabras claves de CLIL. Una charla con María Jesús Frigols. In *Encuentro: revista de investigación e innovación en la clase de idiomas*, n.º 21, pp. 3-14. Retrieved from [http://dspace.uah.es/dspace/bitstream/handle/10017/14561/Megias\\_Entrevista.pdf?sequence=1&isAllowed=y](http://dspace.uah.es/dspace/bitstream/handle/10017/14561/Megias_Entrevista.pdf?sequence=1&isAllowed=y)
- Mehisto, P., Marsh, D., & Frigols, M. (2008). *Uncovering CLIL: Content and Language Integrated Learning in Bilingual and Multilingual Education*. Oxford, UK: Macmillan Education.
- Mertler, Craig A. (2001). Designing scoring rubrics for your classroom. In *Practical Assessment, Research & Evaluation*, 7 (25). Retrieved from <http://PAREonline.net/getvn.asp?v=7&n=25>
- Meyer, C. A. (1992). What's the difference between Authentic Assessment and Performance Assessment?. In *Educational Leadership*, pp. 39-41. Retrieved from [http://www.ascd.org/ASCD/pdf/journals/ed\\_lead/el\\_199205\\_meyer.pdf](http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_199205_meyer.pdf)
- Meyer, O. (2010). Introducing the CLIL-Pyramid: Key Strategies and Principles for Quality CLIL Planning and Teaching. In M. Eisenmann and T. Summer (eds.), *Basic Issues in EFL-Teaching and Learning*. Heidelberg, Germany: Universitätsverlag Winter.
- Mohan, B. A. (1986). *Language and content*. Reading, MA: Addison-Wesley
- Mohan, B. & Huang, J. (2002). Assessing the Integration of Language and Content in a Mandarin as a Foreign Language Classroom. In *Linguistics and Education* 13(3), pp. 407–35.
- Montalbán, J. (forthcoming). *Evaluating bilingual high schools in the Region of Madrid / Evaluación del bilingüismo en los Institutos de la Comunidad de Madrid*.

- Versión inicial en [http://www.parisschoolofeconomics.eu/docs/montalban-castilla-jose/evaluacion-de-institutos-bilingues-en-madrid\(1\).pdf](http://www.parisschoolofeconomics.eu/docs/montalban-castilla-jose/evaluacion-de-institutos-bilingues-en-madrid(1).pdf)
- Morton, T. (2016). *Integrating Content and Language by Design: Principles for Designing Teaching Units for Content and Language Integrated Learning (CLIL)*. Talk at Nebrija University. Conferencias inaugurales del Departamento de Lenguas Aplicadas.
- Mueller, J. (2016). What is Authentic Assessment? In *Authentic assessment toolbox*. Retrieved from:  
<http://jonathan.mueller.faculty.noctrl.edu/toolbox/whatisit.htm#authentic>
- Muñoz, C. (2007). CLIL: Some Thoughts on its Psycholinguistic Principles. In F. Lorenzo, S. Casal Madinabeitia, V. de Alba Quiñones and P. Moore (eds.), *Revista Española de Lingüística Aplicada*, pp. 17-26. Retrieved from  
[https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/lorenzo\\_et\\_al.\\_models\\_and\\_practice\\_in\\_clil.pdf](https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/lorenzo_et_al._models_and_practice_in_clil.pdf)
- Nakanishi, C. & Nakanishi, H. (2015). Assessment of Scoring Content and English in the first CLIL Project at a High School in Sardinia. In *Proceedings of the International Conference on EEE2015*, pp. 92-97. Retrieved from  
<http://worldcomp-proceedings.com/proc/p2015/EEE3130.pdf>
- Novak, J. D. (1994). A view on the current status of Ausubel's assimilation theory of learning. In *CADMO: Giornale Italiano di Pedagogia, Sperimentale, Didattica, Docimologia, Tecnologia dell'instruzione*, 2(4), pp. 7–23.
- Odenstad, C. (2010). *Prov och bedömning i samhällskunskap - En analys av gymnasielärares skriftliga prov*. Karlstad, Sweden: Karlstad University Press. Retrieved from <http://www.diva-portal.org/smash/get/diva2:379242/FULLTEXT01.pdf>

Official Journal L 394. Lifelong learning - key competences. In *Recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning*, pp. 10-18. 30<sup>th</sup> december 2006. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=URISERV:c11090&from=EN>

Olivares Leyva, M. & Pena Díaz, C. (2013). How do we train our CLIL Teachers. A case Study form Alcalá University. In *Porta Linguarum*, 19, pp. 87-99.

Olson, J. & Platt, J. (2000). *The Instructional Cycle. Teaching Children and Adolescents with Special Needs*, pp. 170-197. New Jersey, USA: Prentice-Hall, Inc.

*ORDEN 1275/2014, de 11 de abril, de la Consejería de Educación, Juventud y Deporte, por la que se regula el procedimiento de obtención de la habilitación lingüística en lengua extranjera para el desempeño de puestos bilingües en centros docentes públicos y en centros privados sostenidos con fondos públicos, de Educación Infantil, Educación Primaria y Educación Secundaria, de la Comunidad de Madrid. Área de Coordinación Legislativa y Relaciones Institucionales, Subdirección General de Régimen Jurídico y Desarrollo Normativo S.G.T. de la Consejería de Presidencia, Justicia y Portavocía del Gobierno. Comunidad de Madrid. 11 de abril de 2014. Retrieved from: [http://www.madrid.org/wleg\\_pub/secure/normativas/listadoNormativas.jsf#no-back-button](http://www.madrid.org/wleg_pub/secure/normativas/listadoNormativas.jsf#no-back-button)*

*Orden ECD/65/2015, de 21 de enero, por la que se describen las relaciones entre las competencias, los contenidos y los criterios de evaluación de la educación primaria, la educación secundaria obligatoria y el bachillerato. Boletín Oficial del Estado, N.º 25, disposición n.º 738, pp. 6986-7003, 29 de enero de 2015.*

- Ortega Martín, J. L. (2012). The Primary and Secondary School Curriculum in Spain. In D. Madrid D., & S. Hughes (2011) (eds.), *Studies in Bilingual Education*, pp. 51-75. Bern, Switzerland: Peter Lang.
- Otto, A. (2017a). Assessment Issues in CLIL: What You've Been Wondering but Were Afraid to Ask. In *RDIM*, n.º 1, pp. 1-13. Retrieved from <https://app.rdim.es/archivos/publicacion/e8a4d38d3b45a4831eebde1c2dcbe47e.pdf>
- Otto, A. (2017b). Colaboración efectiva entre profesores CLIL y profesores de lengua extranjera. In *Revista CDL (Encarte lenguas extranjeras)*, n.º 268, pp. 17-19. Retrieved from <https://www.cdlmadrid.org/wp-content/uploads/2017/01/042017.pdf>
- Pascual, I. (2017). Assessment for learning in primary CLIL classrooms and its co-construction in classroom discourse (Unpublished Doctoral Dissertation). Madrid, Spain: Universidad Autónoma de Madrid. Facultad de Filosofía y Letras/Departamento de Filología Inglesa.
- Pavón, V. (2010). The introduction of multilingual teaching in Andalusia: heading towards a newly proposed methodology. In *Journal of Border Educational Research*, 8 (1), pp. 31-42.
- Pavón, V. & Ellison, M. (2013). Examining Teachers' Roles and Competences in Content and Language Integrated Learning (CLIL). In *Linguarum Arena*, vol. 4, pp. 65-78.
- Pavón, V. & Rubio, F. (2010). Teachers' concerns about the introduction of CLIL programmes. In *Porta Linguarum*, 14, pp. 45-58.
- Pavón, V. (2016). *Informe de diagnóstico de centros BEDA de excelencia y referencia*. Madrid, Spain: Escuelas Católicas de Madrid.

- Pawlak, M. (2014). *Error Correction in the Foreign Language Classroom. Reconsidering the Issues*. Heidelberg–New York: Springer.
- Pekarek Doehler, S. (2010). Conceptual Changes and Methodological Challenges: On language, learning and documenting learning in conversation analytic SLA research. In P. Seedhouse, S. Walsh, & C. Jenks (eds.), *Conceptualizing Learning in Applied Linguistics*, pp. 105-26. Basingstoke, UK: Palgrave MacMillan.
- Pérez Vidal, C. (2007). The need of focus on form (FoR) in content and language integrated approaches: an exploratory study. In *Revista española de lingüística aplicada*, pp. 39-54. Retrieved from [hispadoc.es/download/articulo/2575491.pdf](http://hispadoc.es/download/articulo/2575491.pdf)
- Phelan, C. & Wren, J. (2005) Exploring Reliability in Academic Assessment. Retrieved from <https://chfasoa.uni.edu/reliabilityandvalidity.htm>
- Pica, T. (2002). *Subject-Matter Content: How does it assist the Interactional and Linguistic Needs of Classroom Language Learners?* Retrieved from [http://repository.upenn.edu/cgi/viewcontent.cgi?article=1039&context=gse\\_publications](http://repository.upenn.edu/cgi/viewcontent.cgi?article=1039&context=gse_publications)
- Plan de formación en lenguas extranjeras. Dirección General de Innovación, Becas y Ayudas a la Educación* (2016). Consejería de Educación, Juventud y Deporte. Comunidad de Madrid. 2016. Retrieved from <http://gestiondgmejora.educa.madrid.org/plan2016/>
- Pokrivčáková, S. et al. (2015). *CLIL in Foreign Language Education: e-textbook for foreign language teachers*. Nitra, Slovakia: Constantine the Philosopher University.
- Polias, J. (2006). Assessing Learning: A Language-Based Approach. In M. Olofsson (ed.) *Symposium 2006*. Stockholm, Sweden: Nationellt Centrum för SFI, HLS.



- Popham, W. (2008). *Transformative Assessment*. Virginia, USA: Association for Supervision and Curriculum Development.
- Primera fase de la evaluación del Programa Bilingüe de la Comunidad de Madrid*. Consejería de Educación, diciembre de 2016. Retrieved from [http://www.madrid.org/cs/Satellite? blobcol=urldata&blobheader=application %2Fpdf&blobheadername 1=Content- Disposition&blobheader value1=filename %3DEvaluaciones+Bilinguismo+Diciembre+2016+VE RSION+FINAL.pdf&blobkey=id&blobtable=MungoBl obs&blobwhere=1352921 373721&ssbinary=true](http://www.madrid.org/cs/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=filename%3DEvaluaciones+Bilinguismo+Diciembre+2016+VERSION+FINAL.pdf&blobkey=id&blobtable=MungoBlobs&blobwhere=1352921_373721&ssbinary=true)
- Proyecto Bilingüe Madrid: Inmersión en Inglés en Educación Infantil*. Consejería de educación, Comunidad de Madrid and Institute for Learning and Brain Sciences Washington University, 28 de octubre de 2016.
- Rea-Dickins, P. (2000). Classroom Assessment. In T. Hedge, *Teaching and Learning in the Language Classroom*, pp. 375-401. Oxford, UK: Oxford University Press.
- Rea-Dickins, P. (2001). Mirror, mirror on the Wall: Identifying Processes of Classroom Assessment. In *Language Testing*, 18 (4), pp. 429-462.
- Rea-Dickins, P. (2008). Classroom-based Language Assessment. In N. H. Hornberger (ed.), *Encyclopedia of Language and Education*, pp. 239-405. New York, USA: Springer.
- Reierstam, H. (2015). *Assessing Language or Content? A comparative study of the assessment practices in three Swedish upper secondary CLIL schools*. (Licentiate thesis in Subject Matter Education at the Department of Education and Special Education). Gothenburg, Sweden: Faculty of Education, University of Gothenburg. Retrieved from [https://gupea.ub.gu.se/bitstream/2077/40701/4/gupea\\_2077\\_40701\\_4.pdf](https://gupea.ub.gu.se/bitstream/2077/40701/4/gupea_2077_40701_4.pdf)

- Reiss, J. (2005). *Teaching Content to English Language Learners: Strategies for Secondary School Success*. New York, USA: Pearson Education.
- Resnick, L. B., Rothman R., Slattery, J. B, & Vranek, J. L. (2003). Benchmarking and alignment of standards and testing. In *Educational Assessment*, 9 (1-2), pp. 1–27.
- Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching*. Cambridge, UK: Cambridge University Press.
- Rodríguez-Muñiz, L. J., Díaz, P, Mier, V., Alonso, P. (2016). Washback Effect of University Entrance exams in Applied Mathematics to Social Sciences. In *PLoS ONE* 11(12): e0167544.
- Romero Lacal, J. R. (2011). ¿Por qué nuestros alumnos no hablan bien inglés? In *Revista digital innovación y experiencias educativas CSIF*. Retrieved from [http://www.csif.es/andalucia/modules/mod\\_ense/revista/pdf/Numero\\_41/JOSE\\_LUIS\\_ROMERO%20LACAL\\_1.pdf](http://www.csif.es/andalucia/modules/mod_ense/revista/pdf/Numero_41/JOSE_LUIS_ROMERO%20LACAL_1.pdf)
- Rubio, F. D. & Tamayo Rodríguez, L. (2012). Estudio sobre prácticas docentes en evaluación de la lengua inglesa en la ESO. In *Profesorado. Revista de Currículum y Formación del Profesorado*, 16 (1), pp. 295-316. Retrieved from <http://www.ugr.es/~recfpro/rev161COL5.pdf>
- Ruiz, M. (2017). *Bilingual Education: experience from Madrid*. Madrid, Spain: CEMFI. Retrieved from <https://www.cemfi.es/ftp/mt/t1701.pdf>
- Sadler, D. R. (1989). Formative assessment and the design of instructional strategies. In *Instructional Science*, 18, pp. 119-144. Dordrecht, The Netherlands: Kluwer Academic Publishers. Retrieved from

<http://michiganassessmentconsortium.org/sites/default/files/Formative%20Assessment%20and%20Design%20of%20Instructional%20Systems.pdf>

Salaberri Ramiro, S. (2010). Teacher Training Programmes for CLIL in Andalusia. In Lasagabaster & Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 140-161.

Savage, J. (2011). *Cross-Curricular Teaching and Learning in Secondary Education*. London, UK: Routledge.

Semenova, Y.L. & Bukharova, G.D. (2012). Interdisciplinary and Metadisciplinary Integration as a Means of Developing Students' Communicative Competence. In *Образование и наука*. 2012. № 7 (96).

Serra, C. (2007). Assessing CLIL at Primary School: A Longitudinal Study. In *International Journal of Bilingual Education and Bilingualism*, 10:5, pp. 582-602.

Serragiotto, G. E (2007). Assessment and Evaluation in CLIL. In D. Marsh & D. Wolff (eds.), *Diverse Contexts- Converging Goals: CLIL in Europe*, pp. 271-286. Frankfurt, Germany: Peter Lang.

Scanlan, D. (2012). *Assessment, Evaluation, Testing and Grading*. Retrieved from [https://resource.mccneb.edu/edutut/Online%20Pets%20Course/Assessment\\_evaluation\\_grading.htm](https://resource.mccneb.edu/edutut/Online%20Pets%20Course/Assessment_evaluation_grading.htm)

Schleicher, A. & Zoido Lobatón, P. (2013). La transparencia y la mejora de los sistemas educativos. El caso de PISA. In *Participación Educativa*, vol. II, 3, pp. 45-51.

Schleppegrell, M. J., Achugar, M. & Oteíza, T. (2004). The Grammar of History: Enhancing Content-Based Instruction Through a Functional Focus on Language.

- In *TESOL Quarterly*, vol. 38, 1, pp. 67-93. Retrieved from:  
[https://www.jstor.org/stable/3588259?seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/3588259?seq=1#page_scan_tab_contents)
- Swain M. (2006). Languaging, Agency and Collaboration in Advanced Second Language Proficiency. In Byrnes H. (Ed.), *Advanced Language Learning: The Contribution of Halliday and Vygotsky*, 95-108. London-New York: Continuum.
- Shepard, L. (2000). The Role of Assessment in a Learning Culture. In *Educational Researcher*, 29, 7, pp. 4-14.
- Shepherd, E. & Ainsworth, V. (2017). *English Impact: An Evaluation of English Language Capability*. Retrieved from  
[https://www.britishcouncil.org/sites/default/files/g303\\_03\\_english\\_impact\\_report\\_madrid\\_web1.pdf](https://www.britishcouncil.org/sites/default/files/g303_03_english_impact_report_madrid_web1.pdf)
- Shohamy, E. (1993). *The power of tests: The impact of language tests on teaching and learning*. *NFCL Occasional Papers*. Washington, USA: NFLC. Retrieved from  
<http://files.eric.ed.gov/fulltext/ED362040.pdf>
- Short, D. (1993). Assessing Integrated Language and Content Instruction. In *TESOL Quarterly*, vol. 27, 4, pp. 627-656.
- Smit, U. & Dafouz, E. (eds.), (2012). Integrating Content and Language in Higher Education. Gaining Insights into English-Medium Instruction at European Universities. In *Special Issue of AILA Review*, vol. 25.
- Smith M. L. (1991). The effects of external testing on teachers. In *Educational Researcher*, 20 (5), pp. 8-11.
- Sotoca, E. & Muñoz Hueso, A. (2015). The impact of bilingual education in academic achievement of students enrolled in public schools in the Autonomous Community of Madrid. In *Journal of Education Research*, vol. 9, 1, pp. 27-40.

- Retrieved from  
[https://www.novapublishers.com/catalog/product\\_info.php?products\\_id=56184](https://www.novapublishers.com/catalog/product_info.php?products_id=56184)
- Stobart, G. (2008) *Testing Times: The Uses and Abuses of Assessment*. Abindon, Oxon: Routledge.
- Tamariz, M., & Blasi, D. E. (2016). *Consequences of Bilingual Education in Primary and Secondary Schools in the Madrid Region (Comunidad de Madrid)*. Madrid, Spain: Consejería de Educación, Juventud y Deporte.
- Taras, M. (2005). Assessment –summative and formative– some theoretical reflections. In *British Journal of Educational Studies*, vol. 53, 4, pp. 466-478.
- Teaching Geography Through English. A CLIL Approach*. (2014). University of Cambridge. Retrieved from  
[http://www.unifg.it/sites/default/files/allegatiparagrafo/21-01-2014/teaching\\_geography\\_through\\_clil.pdf](http://www.unifg.it/sites/default/files/allegatiparagrafo/21-01-2014/teaching_geography_through_clil.pdf)
- Thornbury, S. (1997). *About language: Tasks for English Teachers*. Cambridge, UK: Cambridge University Press.
- Ting, T. (2011). Review of CLIL in Spain: Implementation, Results and Teacher Training. In *ELT Journal*, Volume 65, Issue 4, 1 October 2011, pp. 495–498,.
- Truscott, J. (1996). The Case against Grammar Correction in L2 Writing Classes. In *Language Learning*, 46, 2, pp. 327-369.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. In *Journal of second language writing*, 16, pp. 255-272.
- UNC Serve Center, (2006). *How to Assess Student Performance in History: Going Beyond Multiple Choice Tests*. University of North Carolina at Greensboro.
- Unsworth, L. (ed.) (2005). *Researching language in schools and communities: Functional linguistic perspectives*. London, UK: Bloomsbury Academic.

- Vaughn, S., Shay Schumm, J. & Sinagub, J. (1996). *Focus Group Interviews in Education and Psychology*. California, USA: SAGE.
- Vygotsky, L. (1978). Interaction between learning and development. In *Mind and Society*, pp. 79-91. Cambridge, USA: Harvard University Press. Retrieved from [http://www.colorado.edu/physics/phys4810/phys4810\\_fa08/4810\\_readings/vygotsky\\_chap6.pdf](http://www.colorado.edu/physics/phys4810/phys4810_fa08/4810_readings/vygotsky_chap6.pdf)
- Vygotsky, L. (1986). *Thought and Language*. Cambridge, USA: The MIT Press.
- Walqui, A. (2006). Scaffolding Instruction for English Language Learners: A Conceptual Framework. In *The International Journal of Bilingual Education and Bilingualism*, vol. 9, 2, pp. 159-180. Retrieved from <https://pdfs.semanticscholar.org/eeaf/a59d9a254ac843f9d8fb94aebd017dae6b76.pdf>
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge, UK: Cambridge University Press.
- Weir, C. (1990). *Communicative Language Testing*, 2nd edition. New York: Prentice Hall.
- Westwood, P. (2006). *Teaching and Learning Difficulties: Cross-curricular Perspectives*. Camberwell, Australia: Australian Council Educational Research (ACER).
- Wewer, T. (2014). *Assesment of young learners' English proficiency in bilingual content instruction CLIL* (Doctoral Dissertation). Turku, Finland: University of Turku. Faculty of Education/Department of Teacher Education. Retrieved from <https://www.doria.fi/bitstream/handle/10024/96838/AnnalesB385Wewer.pdf?sequence=2>
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development.

- Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Edward Arnold.
- Whittaker et al. (2011). Written Discourse Development in CLIL at Secondary School.  
In *Language Teaching Research*, 15 (3), pp. 343-362.
- Whittaker, R., O'Donnell, M., & McCabe, A. (2006). *Language and literacy: Functional approaches*. London, UK: Continuum.
- Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. In *Journal of Child Psychology and Child Psychiatry*, 17, 89–100.
- Zakharov, A., Carnoy, M., Loyalka, P. (2014). Which teaching practices improve student performance on high-stakes exams? Evidence from Russia. In *International Journal of Educational Development*, 36, pp. 13-21.
- Zangl, R. (2000). Monitoring Language Skills in Austrian Primary (Elementary) Schools: A Case Study. In *Language Testing*, 17 (2), pp. 250-260.

## APPENDICES





**APPENDIX 1: Rubric Model in relation with CEFR levels (from Díaz Cobo, 2009)**

NIVEL B2					
1. Organización interna de la información y argumentos del texto		A	B	C	D
La información y los argumentos se organizan de forma lineal y coherente con la tarea.					
<i>La información y los argumentos se presentan de forma lineal y sencilla</i>					
<i>La información y los argumentos son coherentes con la tarea y dan sentido global al texto</i>					
2. Grado de extensión de la información (ejemplos, detalles..)		A	B	C	D
La información sobre el tema que se aporta en la tarea es extensa y detallada					
<i>La información se ajusta en extensión a los requisitos de una tarea compleja</i>					
<i>La información se extiende a lo más relevante tratado en clase sobre el tema de estudio.</i>					
<i>Se dan detalles sobre la información que se aporta</i>					
<i>Se dan ejemplos puntuales sobre las ideas que se exponen.</i>					
3. Tratamiento de la información del texto (síntesis, argumentación..)		A	B	C	D
El alumno trata la información de forma efectiva para mostrar su nivel de dominio de la asignatura.					
<i>Se aportan informaciones y argumentos procedentes de las fuentes básicas de la asignatura.</i>					
<i>Las informaciones y argumentos del texto se tratan de forma efectiva (se resumen / se expanden),(se sintetizan /se analizan);(se comparan / se contrastan) (se evalúan)</i>					
4. Complejidad estructural del texto		A	B	C	D
La estructura del texto se ajusta a la exposición de temas con cierta complejidad .					
<i>Se usan con precisión los conectores oracionales para tratar temas de cierta complejidad</i>					
<i>Se usan con precisión las preposiciones (lugar, tiempo, modo, causa, finalidad, compañía, instrumento....) para tratar temas de cierta complejidad.</i>					
<i>Se usan con precisión las preposiciones en frases convencionales (frases hechas, palabras que rigen preposición, etc) para tratar temas de cierta complejidad</i>					
5. Grado de conexión / cohesión entre las ideas		A	B	C	D
El texto tiene cohesión porque sus ideas se relacionan entre sí y con el entorno del mensaje de forma apropiada y consistente..					
<i>Las concordancias (gén /núm./persona) entre sus elementos son apropiadas y consistentes</i>					
<i>Las referencias a elementos dentro y fuera del texto (pronombres, uso de tiempos, espacio, etc..) son apropiadas y consistentes</i>					
<i>El texto presenta cohesión entre sus ideas</i>					
6. Familiarización del alumno con los contenidos de la asignatura		A	B	C	D
El alumno demuestra conocer adecuadamente los contenidos de la materia					
<i>Aporta información adecuada y veraz sobre los contenidos de aprendizaje de la tarea.</i>					
<i>Aporta información veraz sobre contenidos de la asignatura no incluidos en el programa.</i>					
7. Grado de interacción entre el alumno y el texto		A	B	C	D
El alumno muestra cierto grado de implicación personal en la producción del texto para transmitir al lector un mensaje pragmáticamente relevante para el lector.					
<i>Se implica en la selección y tratamiento de la información sobre el tema</i>					
<i>Usa un registro (neutro/formal/ informal) adecuado en el ámbito de la tarea</i>					
<i>Usa marcadores de dialecto y acento coherentes con el registro utilizado.</i>					
8. Efecto que el texto produce en el lector (grado de recepción e inteligibilidad)		A	B	C	D
El texto es claro, y, aun con limitaciones de expresión no ostensibles, tiene recursos formales apropiados para una recepción suficiente del mensaje					
<i>El orden de la información facilita una comprensión prácticamente completa del texto</i>					
<i>El uso de las palabras es suficiente para una recepción satisfactoria del mensaje.</i>					
<i>La ortografía y la puntuación son apropiadas para una recepción suficiente del mensaje.</i>					
<i>El registro es el que espera el receptor para recibir satisfactoriamente el mensaje.</i>					
Calificación global de la tarea (Valoración de la competencia funcional)		A	B	C	D
A = 100 B = 80 C = 60 D = 40					



<b>SUBJECT:</b> History.	<b>UNIT AND TOPIC:</b> U5 (Ancient Rome).	<b>LEVEL:</b> 1ºESO	<b>CEFR:</b> A2
<b>SPECIALIST TEACHER:</b> Juan Pérez.	<b>LANGUAGE TEACHER:</b> Noelia Fernández.	<b>ACADEMIC YEAR:</b> 2015 / 2016	

<b>Areas in which students must improve</b> (weak points of students, presentations)	<b>Strong points of students' presentation</b>	<b>Grammar structures or vocabulary to revise in English Lesson</b>
<b>My Personal reflection</b>	<b>Students' suggestions, a peer assessment, self-assessment</b>	

\*Please note: speaking tasks like presentations will be recorded in case of anyone makes a complaint to the History Department about their marks.

### APPENDIX 3: Rubric for 4 C's (external informant)

	OUTSTANDING	VERY GOOD	ADEQUATE	INSUFFICIENT
<p><b>COGNITION</b></p> <p>The students have been able to:</p> <ul style="list-style-type: none"> <li>- Make a presentation of a given economic sector.</li> <li>- Collect, organize and represent data in their presentation.</li> <li>- Customize their diorama showing creativity and originality.</li> <li>- Prepare an oral presentation to introduce their presentation.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Demonstrates ability to plan and conduct a research that results in a creative and attractive presentation.</li> <li>- Shows self-confidence, personal initiative, curiosity, interest and creativity in the learning process.</li> <li>- Shows a comprehensive use of presentation tools.</li> </ul> <p><b>The presentation:</b></p> <ul style="list-style-type: none"> <li>- Shows an innovative format by the use of tools such as Piktochart, Knovio and Clearslide.</li> <li>- Has been consciously planned and shows original thought and ideas.</li> <li>- Flows well and logically.</li> <li>- Includes many relevant tables, graphs, images and summaries.</li> <li>- Shows wide variety of data from different sources.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Manages sufficient planning and research, making possible the collection, organization and representation of data in his/her presentation.</li> <li>- Shows interest and curiosity to investigate intellectually and creatively.</li> <li>- Shows general understanding of presentation tools.</li> </ul> <p><b>The presentation:</b></p> <ul style="list-style-type: none"> <li>- Shows an innovative format by the use of tools such as Piktochart, Knovio and Clearslide.</li> <li>- Has been suitable planned.</li> <li>- Flows appropriately.</li> <li>- Includes some relevant table, graphs, images and summaries.</li> <li>- Shows data from relevant sources.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Is able to plan and conduct research that leads to collection, organization and representation of data in a presentation, but with frequent teacher help.</li> <li>- Develops ideas that are sometimes meaningful or important.</li> <li>- Shows an acceptable understanding of presentation tools.</li> </ul> <p><b>The presentation:</b></p> <ul style="list-style-type: none"> <li>- Has been done through a Powerpoint, a collage or a poster.</li> <li>- Has been appropriately planned.</li> <li>- Flows appropriately.</li> <li>- Includes tables, graphs, images and summaries, but not all of them are relevant to the topic.</li> <li>- Shows sufficient data from different sources.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Has not participated in the learning process.</li> <li>- Does not show neither interest nor enthusiasm in the presentation.</li> <li>- Shows little or no understanding of technology.</li> </ul> <p><b>The presentation:</b></p> <ul style="list-style-type: none"> <li>- Could have been more carefully planned.</li> <li>- Does not include any image, table, graph and summary.</li> <li>- Shows data that is not relevant to the topic.</li> <li>- Is difficult to follow.</li> </ul>

	OUTSTANDING	VERY GOOD	ADEQUATE	INSUFFICIENT
<p><b>COMMUNICATION</b></p> <p>The student's presentation shows language accuracy and coherence. The student's presentation shows the student's capacity for oral communication (introducing the topic, showing organization of ideas, concepts and information...)</p>	<ul style="list-style-type: none"> <li>- The student speaks with clear voice and the volume is appropriate.</li> <li>- Language use is more than satisfactory.</li> <li>- The student uses elaborated sentences and logically connected.</li> <li>- Information is well communicated and the presentation is endowed with a coherent and cohesive structure.</li> <li>- Ideas are very well delivered through the use of connectors.</li> <li>- Absence of serious grammar mistakes.</li> </ul>	<ul style="list-style-type: none"> <li>- The student speaks clearly most of the time and the volume is adequate.</li> <li>- Language use is appropriate.</li> <li>- The student uses more elaborated language through the use of longer sentences.</li> <li>- Information is reasonable well communicated and the presentation's organization is quite consistent.</li> <li>- Pauses were momentary and did not interrupt fluency speech.</li> <li>- Few grammar mistakes.</li> </ul>	<ul style="list-style-type: none"> <li>- The student needs to speak a little bit louder/softer.</li> <li>- Language use is sufficient.</li> <li>- The student uses longer sentences, but the vocabulary used is too simple.</li> <li>- Information is correctly communicated and the presentation's organization is consistent enough.</li> <li>- Some serious grammar mistakes, but the student is able to correct them.</li> </ul>	<ul style="list-style-type: none"> <li>- It is difficult to understand when the student speaks.</li> <li>- The student uses simple sentences and his/her speech lacks in elaboration.</li> <li>- Information is poorly communicated due to the absence of connectors and linkers.</li> <li>- Long pauses interrupt flow of speech.</li> <li>- Some serious grammar mistakes.</li> </ul>
<p><b>CULTURE</b></p> <p>Student's presentation shows he/she is able to make connections between the content and how it is projected in the real life.</p>	<ul style="list-style-type: none"> <li>- The students' presentation shows many connections between the professions in a given economic sector and the activities that are carried out.</li> <li>- The student relates the studied professions with his/her ambitions for future employment.</li> </ul>	<ul style="list-style-type: none"> <li>- The students' presentation shows some connections between the professions in a given economic sector and the activities that are carried out.</li> <li>- The student relates the studied professions with his/her ambitions for future employment.</li> </ul>	<ul style="list-style-type: none"> <li>- The students' presentation shows only few connections between the professions in a given economic sector and the activities that are carried out.</li> <li>- The student is not able to relate to the studied professions with his/her ambitions for future employment.</li> </ul>	<ul style="list-style-type: none"> <li>- The presentation shows no connection between the professions in a given economic sector and the activities that are carried out.</li> <li>- The student is not able to relate to the studied professions with his/her ambitions for future employment.</li> </ul>

	OUTSTANDING	VERY GOOD	ADEQUATE	INSUFFICIENT
<p><b>CONTENT</b></p> <p>Student's presentation shows the student have learned CLIL contents:</p> <ul style="list-style-type: none"> <li>- The differences amongst the primary, secondary and tertiary sectors</li> <li>- Identifies the main types of activities within the service sector and the people who carry them out</li> <li>- Research on a particular economic sector.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Demonstrates full knowledge of a given economic sector, including many examples, facts and/or statistics that support the previous research.</li> <li>- Names 10-15 professions related to a given economic sector.</li> <li>- Provides an abundance of material clearly related to the research that is presented.</li> <li>- Shows a logical progression of ideas</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Demonstrates full knowledge of a given economic sector, including many examples, facts and/or statistics that support the previous research.</li> <li>- Names 5-9 professions related to a given economic sector.</li> <li>- Provides sufficient information with many good points made.</li> <li>- Shows a logical progression of ideas.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Demonstrates some knowledge of a given economic sector, including some examples, facts and/or statistics that support the previous research.</li> <li>- Names 3-5 professions related to a given economic sector.</li> <li>- Includes essential information about the topic but there are 1 or 2 factual errors.</li> </ul>	<p><b>The student:</b></p> <ul style="list-style-type: none"> <li>- Demonstrates little knowledge of the economic sector, including very few examples, facts and/or statistics that support the previous research.</li> <li>- Provides information that it is not clearly integrated or connected to the research.</li> <li>- Provides information that is not logically organized.</li> </ul>

## **APPENDIX 4: Cover letter to introduce a portfolio (external informant)**

### **COVER LETTER**

*Hello all, and welcome to our e-portfolio creator in eduClipper.*

*Why are we doing this portfolio? This is a tool that is going to help me to assess your learning, but it is also a great tool that will help you organize your work, analyze the results obtained, and reflect on the carrying out and the possible ways to improve your performance.*

*We work hard in every lesson, we learn in most of them but here, what we are going to show are our best works; best in performance and best because of the goals we followed by it.*

*We will also be able to share it with our teacher, our peers and our parents and show them work at a click!*

**APPENDIX 5: Descriptors used in portfolio work for Geography (del Pozo, 2009: 37)**

What can you do in Geography? Colour the squares using a green crayon.

☐☐ I can't do this  
 ☐ I need to work more on this  
 ✓ I can do this well  
 ✓✓ I can do this very well

SKILLS		YOU				YOUR TEACHER			
		✓✓	✓	☐	☐☐	✓✓	✓	☐	☐☐
Writing	● I can write the definition of a geographic/historic term.								
	● I can use and complete vocabulary webs on geography/history								
	● I can group sentences into paragraphs speaking about geography/history facts								
Searching	● I can introduce, develop and conclude pieces of writing appropriately								
	● I can carry out geographical enquiry, including identifying questions and developing my own opinion								
	● I can use some problem-solving skills and resources, including different types of maps, atlases, ICT images etc.								
Spoken Production	● I can use appropriate geographical vocabulary.								
	● I can ask geographical questions.								
	● I can talk about a natural disaster.								
Reading	● I can use atlases/globes/maps								
	● I can understand a text about a student and answer detailed questions.								
	● I can analyse evidence and draw conclusions.								
Listening	● I can understand a native speaker giving a talk on geography/history								
	● I can listen for specific information on geography/history in a recording								
	● I can understand specific information on geography/history in the radio, tv, etc.								
Spoken Interaction	● I can have a conversation with a partner about a geographic/historic issue.								
	● I can have a conversation in a group about a geographic/historic issue.								
	● I can hold a conversation with a native speaker on the issue.								



## APPENDIX 6: Questions for portfolio reflection (external informant)

PORTFOLIO EXAMPLE: What have we learnt in this unit?

### PRESENTATION

- My language goals are...
- My subject goals are...
- The attitude I work to develop is...
- I want to learn...

### OBJECTIVES



In this unit we will learn...

VOCABULARY...

### CONCLUSIONS

- The attitude I worked to develop was...
- I achieved my goals because...
- I need to pay more attention to the following evaluation criteria: ...
- My portfolio is well organized and easy to follow...

## APPENDIX 7: Teachers' questionnaire

Este cuestionario se enmarca dentro de una investigación de Posgrado sobre la evaluación de materias no-lingüísticas en las **Secciones Bilingües** de Institutos bilingües de la Comunidad de Madrid. Pretende recoger información sobre los procesos de evaluación CLIL/AICOLE, con objeto de analizar críticamente y mejorar los mecanismos de formación.

Para ello, solicito tu valiosa colaboración al rellenar este cuestionario, cuyos datos me proporcionarán información relevante para este estudio.

Si estás interesado/a en obtener feedback sobre los resultados del cuestionario, estaré encantada de hacerlo llegar.

Instrucciones:

- 1) Lee atentamente las preguntas e intenta contestar con la mayor sinceridad posible. No existen respuestas correctas o incorrectas y cualquier aportación/opinión es válida.
- 2) Contesta a todas las preguntas en la medida de lo posible y explica tus respuestas si así se te indica.

El cuestionario es anónimo.

Agradezco enormemente tu colaboración.

Ana Otto

Profesora Asociada Universidad de Alcalá.

LA EVALUACIÓN DE MATERIAS NO-LINGÜÍSTICAS EN LAS SECCIONES  
BILINGÜES (INGLÉS) DE LA COMUNIDAD DE MADRID.

CUESTIONARIO PARA PROFESORES.

INFORMACIÓN GENERAL
Nombre del centro:
Localidad:
Año en que se implantó la sección bilingüe:
Asignatura (s) que impartes en la sección bilingüe:
Años de experiencia docente en centros bilingües:
Formación en CLIL: <input type="checkbox"/> Ninguna <input type="checkbox"/> Menos de 50 horas de formación <input type="checkbox"/> Más de 50 horas de formación <input type="checkbox"/> Posgrado en CLIL

## 1. Evaluación CLIL

1.1. ¿Con qué frecuencia evalúas a tus alumnos?

1.2. ¿Qué instrumentos usas para evaluar? (Márquese lo que proceda)

- Exámenes escritos
- Exámenes orales
- Presentaciones orales en clase
- Debates
- Observación en el aula
- Auto-evaluación (self-assessment)
- Evaluación por pares (peer-assessment)
- Portafolios
- Checklists
- Participación activa en clase
- Otros: (Indica cuáles)

1.3. ¿Usas alguna guía sobre CLIL, indicación de la CAM o pautas de tu centro o alguna otra institución a la hora de evaluar a tus alumnos? Indica cuál(es).

1.4. En caso de utilizar presentaciones orales como instrumento de evaluación, ¿qué criterio(s) usas?

1.5. Indica el formato o formatos que sueles utilizar en los exámenes escritos:

- Examen de opción múltiple
- Examen de desarrollo
- Examen mixto (opción múltiple + preguntas de desarrollo)

Explica tu respuesta:

1.6. ¿Usas alguna rúbrica de evaluación?

- Sí
- No

1.7. A la hora de evaluar tu asignatura (s): (Márquese lo que proceda)

- Existen pautas claras acerca de si se debe o no evaluar la lengua inglesa
- Cada profesor/a decide si tiene en cuenta la lengua inglesa en la evaluación y de qué manera la puntúa y/o penaliza

1.8. La evaluación de tu asignatura(s):

- La decides tú como profesor/a especialista de la materia (s) que impartes
- La decides conjuntamente junto con el profesor/a de lengua inglesa

Explica tu respuesta o añade cualquier comentario que creas oportuno:

1.9. ¿Qué peso le otorgas a la lengua inglesa con respecto al contenido en las pruebas de evaluación?

- Ninguno
- Menos del 25%
- Entre 25% y 50%
- Más del 50%

1.10. En caso de evaluar y/o corregir la lengua inglesa, ¿con cuál(es) de las siguientes afirmaciones te identificas?

- Corrijo errores básicos en intervenciones en clase o en redacciones de los alumnos
- Corrijo errores pero no los tengo en cuenta para la nota
- Corrijo errores y los tengo en cuenta para la nota
- Informo a los alumnos acerca de los errores lingüísticos que han cometido

1.11. En caso de medir la lengua inglesa, ¿qué aspectos de la misma tienes en cuenta? (Márquese lo que proceda)

- Competencia comunicativa (que el mensaje sea claro y logre comunicar el contenido)
- Vocabulario específico de la asignatura
- Errores gramaticales
- Fluidez en pruebas orales
- Pronunciación en pruebas orales
- “Spelling” en pruebas escritas

1.12. ¿Qué ocurre si un alumno utiliza un vocabulario muy básico o si no utiliza la tercera persona de singular correctamente? ¿Y si la pronunciación o “spelling” no son correctos?

1.13. ¿Informas a tus alumnos sobre los objetivos de cada unidad didáctica?

- Sí.
- No

Explica tu respuesta

## 2. Aspectos problemáticos en evaluación CLIL

### 2.1. ¿Qué estrategias utilizas para integrar al alumnado con dificultades?

- Apoyo o estrategias para reforzar la lengua extranjera
- Estrategias para reforzar el contenido de la asignatura
- Simplifico el contenido o rebajo el nivel de exigencia
- En ocasiones permito que el alumnado use la lengua materna
- Otros: Indica cuáles.

### 2.4. Con respecto a la evaluación CLIL, ¿cuáles son los principales obstáculos que encuentras en la práctica docente?

Muchas gracias por tu colaboración. Si deseas recibir feedback sobre los resultados de esta investigación, por favor indica tu dirección de correo electrónico a continuación:

Email:

## **APPENDIX 8: Focus group questions**

1. Do you find any difficulties when selecting CLIL assessment tools? Which one (s)?
2. Do you use any classroom observation techniques to measure students' participation and progress? How do you use "cuaderno de trabajo" or students' units of work?
3. Do you talk to colleagues in your department or at school about how to assess CLIL subjects? Are you/they concerned about it? Have you drawn any conclusions or reached any agreements?
4. Are there clear guidelines in your school about how to assess CLIL subjects? Do you use IGSE Guidelines or similar ones? Which one (s)? Any materials from publishing houses?
5. Do you negotiate assessment criteria with students? Do you comment it in class?
6. Would you say it is easy to assess content competence without taking language proficiency into account? Why (not)?
7. Do you find portfolios and peer and self-assessment are useful in CLIL assessment? Do you use them?
8. How do you deal with language mistakes? Do you correct and/or penalize them? Which ones?
9. Do you use any assessment rubric? Which one(s)?
10. What is the (approximate) weight of language in CLIL assessment?
11. Anything else you would like to point out or clarify: differences between levels, lack of teacher training, etc.

## APPENDIX 9: Students' questionnaire

### CUESTIONARIO ALUMNOS "ASSESSMENT IN CLIL"

El siguiente cuestionario forma parte de una investigación doctoral y pretende medir la percepción que tienen los alumnos sobre los procedimientos de evaluación en CLIL.

Instrucciones:

- 1) Todas las preguntas se refieren a las asignaturas que se imparten en lengua inglesa en un centro bilingüe. Lee atentamente las preguntas e intenta contestar con la mayor sinceridad posible. No existen respuestas correctas o incorrectas y cualquier aportación/opinión es válida.
- 2) Contesta a todas las preguntas en la medida de lo posible y explica tus respuestas si así se te indica.

El cuestionario es anónimo.

Agradezco enormemente tu colaboración.

Ana Otto

Profesora Asociada Universidad de Alcalá.

Indica la modalidad de educación bilingüe a la que perteneces: \*

Selecciona todos los que correspondan.

- Programa Bilingüe
- Sección Bilingüe

1-Indica a continuación qué usan los profesores para evaluar tu progreso en las asignaturas que se imparten en lengua inglesa, por ejemplo, "Natural Science".

Selecciona todos los que correspondan.

- Exámenes escritos
- Exámenes orales
- Presentaciones orales en clase
- Debates
- Auto-evaluación
- Evaluación por parejas
- Portafolios
- Cuaderno de clase o unidades didácticas
- Participación activa en clase
- Otros: (Indica cuáles)

2- A la hora de examinarte, ¿a qué aspecto(s) crees que le dan más importancia o más peso tus profesores?

Selecciona todos los que correspondan.

- A la lengua inglesa
- Al contenido de la asignatura
- A la lengua inglesa y al contenido de la asignatura

3- En caso de tener en cuenta la lengua inglesa para la nota, ¿qué se tiene en cuenta?

Selecciona todos los que correspondan.

- La fluidez: cómo te expresas en inglés, si eres capaz de expresar las ideas con soltura
- La corrección en el uso del inglés: si eres capaz de expresarte correctamente sin cometer muchos fallos de gramática
- El vocabulario de la asignatura
- No lo sé
- Otros aspectos. Indica cuáles.

4- ¿Quién decide la nota de la asignatura, por ejemplo, en "Social Science" o "Technology"?

Selecciona todos los que correspondan.

- El profesor de la asignatura
- El profesor de inglés
- El profesor de inglés y el profesor de la asignatura

5- ¿Cuánto cuenta el inglés en la nota?

Selecciona todos los que correspondan.

- Nada
- Menos del 25%
- Entre el 25% y el 50%
- Más del 50%
- No lo sé

6- ¿Qué ocurre si cometes errores en la lengua inglesa?

Selecciona todos los que correspondan.

- Nada
- Te bajan la nota
- Te corrigen los errores y te los dicen para que mejores, pero no bajan nota
- Te corrigen los errores y te los dicen para que mejores y se baja la nota



7- ¿Crees que en las asignaturas que se imparten en inglés todos los profesores siguen las mismas pautas a la hora de evaluar la lengua inglesa?

Selecciona todos los que correspondan.

- Sí
- No
- No todos

Muchas gracias por tu colaboración. Si deseas recibir información sobre los resultados de esta investigación, por favor indica tu dirección de correo electrónico a continuación:

Email:

## APPENDIX 10- Students' interview questions

- 1- En las asignaturas CLIL, como Geography, Physical Education, etc., ¿en qué aspectos se centran los profesores: ¿en el contenido, la lengua inglesa, ambos? ¿Qué ocurre si cometéis errores en lengua inglesa? ¿Baja nota? ¿Pasa lo mismo en las asignaturas que se imparten en español?
- 2- En caso de que la lengua inglesa baje nota, ¿estos errores son los mismos que los errores que bajan en la asignatura de inglés (Lengua extranjera)? ¿Dónde hay más flexibilidad con respecto a errores, en la clase de inglés o en las asignaturas que se imparten en inglés?
- 3- ¿Qué aspectos de la lengua inglesa se tienen en cuenta? ¿*Spelling* u ortografía? ¿Gramática? ¿Vocabulario? ¿La capacidad de enlazar bien las ideas, saber transmitir la información o resumir? ¿La pronunciación?
- 4- ¿Los profesores de asignaturas en inglés os muestran los objetivos al principio de cada unidad? ¿Son objetivos de contenido o relacionados con la lengua inglesa? ¿Ambos? ¿Os dan información sobre los aspectos relativos al inglés que debéis mejorar en estas asignaturas?
- 5- ¿Pensáis que el alumno que sabe mucho inglés saca mejor nota que el que tiene un nivel más bajo en lengua inglesa? ¿Os dejan contestar en español en los exámenes y en clase?

**APPENDIX 11: Action plan for language mistakes** (anonymous High School in Madrid)

**ACTION PLAN FOR CORRECTING GRAMMATICAL ERRORS**

**BILINGUAL PROGRAM**

1° and 2° ESO
- Omission of the subject in a sentence
- Subject-verb agreement
- Correct use of adjectives
- Correct use of “there is/there are”
- Correct use of verb tenses, particularly irregular verbs
- Correct use of possessive forms
- Correct use of the auxiliary verbs “do/does/did” in interrogative and negative sentences
- Correct use of WH-questions
- Correct use of “some/any”
- Correct use of demonstratives (this-that-these-those)
3° and 4° ESO
- Comparatives and superlatives
- Verb tenses (present/past/perfect tenses)
- Modal verbs
- Relative pronouns or adverbs
- Linking words
N.B. For each mistake in an exam, 0,10 will be deducted up to 1 point

## BILINGUAL SECTION

1° and 2° ESO
- Omission of the subject in a sentence
- Subject-verb agreement
- Correct use of “there is/there are”
- Correct use of verb tenses, particularly of irregular verbs
3° and 4° ESO
- Omission of the subject in a sentence
- Subject-verb agreement
- Correct use of “there is/there are”
- Correct use of verb tenses, particularly of irregular verbs
- Correct use of the auxiliary verbs “do/does/did” in interrogative and negative sentences
- Correct use of WH-questions
- Correct use of demonstratives (this-that-these-those)
- Relative pronouns
N.B. For each mistake in an exam, 0-10 will be deducted up to 2 points

**APPENDIX 12: Constructed response test type** (data gathered for analysis of assessment tools)

**Natural Science**

1. Are the following sentences true (T) or false (F)? Correct the false ones:
  - a) A water molecule is composed of nitrogen and oxygen atoms
  - b) A considerable part of the hydrosphere is made up of pure water
  - c) Ice floats on liquid water because it has more density than water
  - d) Water can dissolve some solids, liquids and gases
  - e) The lowest amount of water is found in vapour state

**Biology:**

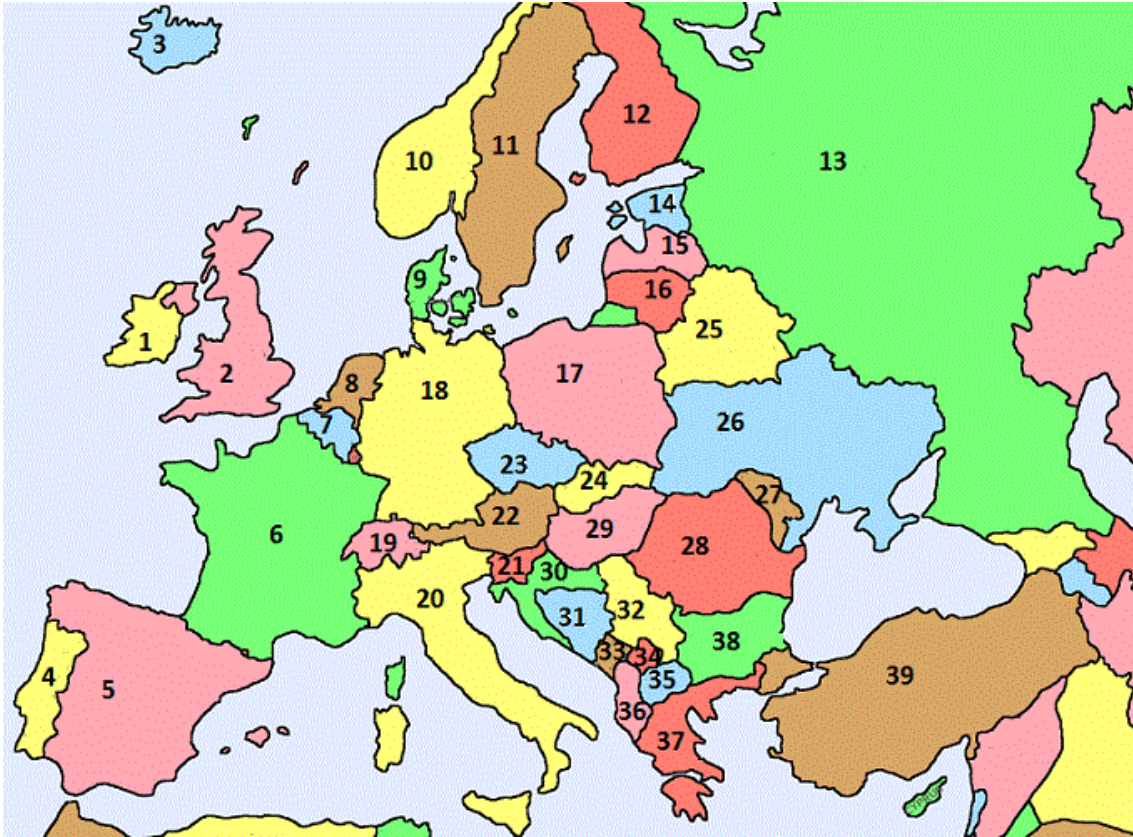
1. Identify which features belong to animal eukaryotic cells, plant eukaryotic cells or prokaryotic cells.

	Prokaryotic	Animal eukaryotic	Plant eukaryotic
It has cellular organelles			
It contains nucleic acid			
It has the genetic material in a membrane			
It is smaller than the others			
It has a cell wall			
It has chloroplasts			
It has the genetic material in the cytoplasm			
It has cytoplasm			
It has mitochondria			
It has plasmatic membrane			



Geography exam

1. Write on the list below the name of the numbered countries



1.	11.
2.	12.
3.	13.
4.	14.
5.	15.
6.	16.
7.	17.
8.	18.
9.	19.
10.	20.

# APPENDIX 14: Scaffolding in CLIL History exam

EXAM UNIT 1-8

NAME: \_\_\_\_\_

Date: \_\_\_\_\_

Class: \_\_\_\_\_

Excellen  
10

1. What are the two prehistoric ages (dates)? (0.5)
2. Characteristics, utensils, advances and location of *Homo habilis* (1 mark)
3. What was the natural landscape like in the Neolithic? (1 mark)
4. What was society like in the Metal Age? (1 mark)
5. Architecture and art in the Metal Age. (1.5 mark)
6. Write all you know about *Millares* and *Talotic* cultures. (1 mark)
7. Complete the chart. (3 marks)

	PALAEOLITHIC	NEOLITHIC
Beginning and end	Began → 5.000.000 B.C End → 7.000 B.C	Began → 7.000 B.C End → 3.500 B.C
Materials used	They made stone axes (bifaces), animal hides such as clothes.	They used vegetable fibres (baskets) but then they used cotton, wool, exports grass and linen to make cloth. They used stones and hammers of wood.
What did they eat?	They were predators, they lived from hunting, fishing and gathering (wild fruits)	They became producers, they discovered agriculture. They grow plants (cultivate) they domesticate animals.
Where did they live?	They live in caves or in shelters made of bones, stone or branches... they were nomads, so they move from place to place.	They live in permanent settlements; they became sedentary which means that they didn't move from the place. Their settlers were surrounded by palisades to defend it.
What was society like?	There was an hierarchy within each group, some were more important than others. Women treated hides and take care of children and men made tools, hunted...	There was a division of labour, everyone had a job to do (in livestock, agriculture, hunting...). Historians thought that was women who invented agricultural methods.
What art did they produce?	They made cave paintings (cave of Altamira, Lascaux). They made Venus statues (statues of women with fertility symbols) and also bones.	They made rock painting, but very schematic. They made funeral idols (with big eyes). Paintings represented hunting scenes, and also they painted animals as a vertical line.
What did they believe in?	They deified the forces of nature (spirits) and they had certain religious beliefs.	They believe in spirits, they buried people in cemeteries or in necropolis with their personal possessions (pots, jewellery...). They worshipped the dead.

8. Read the text and answer to the questions:

1. The two Prehistoric Ages are:
  - The Stone Age → the Palaeolithic Age (old stone), Neolithic (new stone)
    - ↓
    - (From 5.000.000 - 3.500 B.C)
    - Palaeolithic → from 5.000.000 to 7.000 B.C
    - Neolithic → from 7.000 to 3.500 B.C.
  - Metal age → the Copper Age, Bronze Age and Iron Age.
    - ↓
    - (From 3.500 to 300 B.C)



## APPENDIX 15: Diary entry

### DIARY OF FERNANDO II

8.5

Year 1492:

We are planning, my wife Isabel and I, how to conquer the southern part of the peninsula specifically the Nasrid Kingdom of Granada. Isabel proposed me to demand Boabdil the land of Granada and have some negotiations. My idea was to conquer them by force as always. But I don't know what to do yet.

Finally we've decided to do what Isabel proposed. I honestly preferred my option but she leads more than me so... Another reason we are doing her option is because we signed a treatment with Boabdil that we gave him a manor if he give us the Kingdom of Granada. This treatment was signed in 1486, around that year. *in or around??*

FINALLY! We've conquered Granada! We've finish the Reconquest! I hope everybody is proud of us because all the Peninsula is finally Christian. Now we are focusing on another thing. We signed with Cristobal Columbus the Capitulations of Santa Fe. Cristobal wanted to search a new route for the Silk Route and we said him that when the Reconquest finished we were going to financed it. So Cristobal has already talk<sup>d</sup> with us and he's going to start his trip. I hope what he does is remembered in the future...



I have chosen Fernando II because I think he's been important in the history of Spain. One book where I've search information at is:

FERNANDO EL CATÓLICO, CRÓNICA DE UN REINADO. FERNANDO MARTÍNEZ.

*Very good essay because you concentrated a lot of historical information in a diary  
Quote in the text?*

APPENDIX 16: Newspaper article

# EL PAÍS

www.elpais.com

EL PERIÓDICO GLOBAL EN ESPAÑOL

THURSDAY, FEBRUARY 11, 2016

ENGLISH EDITION WITH THE INTERNATIONAL HERALD TRIBUNE

## INDEPENDENCE OF CATALONIA

7.5

The problem in Catalonia is:

*Independence: yes or not?*

I'm going to talk about independence of Catalonia. First, I'm going to say my opinion, and then I will say the opinion of some Catalonian members of my family.



Personal opinion:

I am agree with independence. My

opinion is that if they want to be independent, they can, but they have to get prepare for the consequences, because if they leave us, they leave us for all: they must have its own economic system, its politics, its institutions and even its own football league.

Very good

A Catalanian cousin opinion (Neus):

She isn't agree. She thinks that it is a topic that has been invented by Artur Mas to distract the people attention. And she told me that all the money they have used for consults and campaigns, should have been used for other more important things. But she thinks that Spanish govern should know of all that Catalonia brings, because it is the community that more taxes pays.

Very good reference!!

A Catalanian uncle opinion (Paco):

He isn't agree with Catalonia independence, but he isn't neither agree with remain united to Spain as usual. He thinks that Catalonia has different things that the rest of Spain and he prefers that Spain were a federal state. Together, we can be stronger, but maintaining cultural differences.

At least, other catalonian cousin, Sergi, thinks that, like people, Catalonia has come of age and they must walk alone.

What is the readers opinion?

← Really good ending!

Excellent writing, I only miss the El País 27/01/2016  
historical references: Spanish March, W. the Hair, Borrell II, its role during the reconquest