Trabajo Fin de Máster en Tecnologías de la Información y la Comunicación en la Enseñanza y Tratamiento de Lenguas

THE INFLUENCE OF ORAL CALL ACTIVITIES ON CLIL ASSESSMENT PROCESSES

LA INFLUENCIA DE ACTIVIDADES ELAO ORALES EN LOS PROCESOS DE EVALUACIÓN AICLE

Autor: Emilio Jesús Cifuentes Simón

Tutora: Dra. Dª Ana Ibáñez Moreno

Curso Académico: 2017/2018

Convocatoria de junio

Facultad de Filología

Universidad Nacional de Educación a Distancia

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AICLE	Aprendizaje Integrado de Contenidos y Lenguas Extranjeras
CAI	Computer Assisted Instruction
CALL	Computer-Assisted Language Learning
CALT	Computer-Assisted Language Testing
CASLA	Computer Applications in Second Language Acquisition
CELL	Computer Enhanced Language Learning
CLIL	Content and Language Integrated Learning
CLL	Community Language Learning
ELAO	Enseñanza de Lenguas Asistida por Ordenador
EPT	Empowerment and Participation Technologies
ICALL	Intelligent Computer Assisted Language Learning
ICT	Information and Communication Technologies
LKT	Learning and Knowledge Technologies
MALL	Mobile-Assisted Language Learning
TELL	Technology Enhanced Language Learning
WELL	Web Enhanced Language Learning

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation attempts to explore the ways in which Computer-Assisted Language Learning (CALL) tools can enhance the assessment processes in Content and Language Integrated Learning (CLIL) settings, thereby using oral skills as a means of assessment; thus, the purpose is not merely to assess oral skills, but to evaluate pupils through them. To that end, a case study is carried out on the basis of a pre-experimental design, in which principally qualitative data are collected from a non-probability and purposive sample through participant observation, self-completion questionnaires and semi-structured group interviews. After the analysis and triangulation of data, this research concludes that oral-based CALL activities are suitable for CLIL assessment purposes as they increase pupils´ motivation throughout the evaluation processes and elicit students´ both linguistic and content knowledge, as CLIL assessment requires.

Key words: CALL, CLIL, oral skills, assessment processes, case study.

RESUMEN

Esta investigación intenta explorar las formas en que las herramientas ELAO (Enseñanza de Lenguas Asistida por Ordenador) pueden mejorar los procesos de evaluación en entornos AICLE (Aprendizaje Integrado de Contenidos y Lenguas Extranjeras), utilizando las destrezas orales como medio de evaluación; así, el propósito no es únicamente evaluar las destrezas orales, sino evaluar a los alumnos por medio de éstas. Para ello, se lleva a cabo un estudio de caso partiendo de un diseño pre-experimental, en el que se extraen datos principalmente cualitativos de una muestra con carácter no probabilístico y propositivo, a través de una observación participativa, encuestas personales y entrevistas grupales semiestructuradas. Tras el análisis y la triangulación de datos, esta investigación concluye que las actividades orales ELAO son adecuadas para propósitos de evaluación AICLE, ya que aumentan la motivación de los alumnos durante los procesos de evaluación y recaban el conocimiento lingüístico y de contenidos por parte del alumnado, tal y como la evaluación AICLE requiere.

Palabras clave: ELAO, AICLE, destrezas orales, procesos de evaluación, estudio de caso.

1. INTRODUCTION

The teaching of foreign languages has been an utmost priority in virtually all education systems since the twentieth century onwards, and this issue has recently gained even more relevance as a consequence of the globalisation process. This way, different approaches and paradigms, mostly influenced by the fields of psychology and linguistics, have historically shed light on the area of foreign language teaching, stretching from the grammar-translation method, to the current communicative approach, in which the emphasis is placed on the development of pupils´ communicative competence through the integration of all the linguistic skills; namely listening, speaking, reading and writing (Gilbert, 2018). Taking the communicative approach as the basis, different methodologies have come to the fore, and especial mention deserves the so-called Content and Language Integrated Learning (henceforth CLIL), which presents noteworthy benefits such as the improvement of the linguistic competence, oral communication and intercultural skills; the increase of pupils´ motivation and confidence in both the language and the content; as well as the development of multilingual attitudes and cognitive strategies, as pointed out by Nieto-Moreno-de-Diezmas (2018).

Besides the appearance of communicative approaches, the evolvement of new technologies has also had a tremendous impact on the field of education in general, and the matter of foreign language teaching in particular. As such, throughout the past decades the role of Information and Communication Technologies (henceforth ICT) has gradually gained more and more relevance to the field of teaching, to the extent that a more specific concept has raised in order to refer exclusively to the issue of pedagogy; Learning and Knowledge Technologies (henceforth LKT) (Reig, 2012). In turn, the appearance of the Web 2.0. has widened the possibilities to interact and take a more active role when using the Internet and the new technologies, resulting in a more recent term; Empowerment and Participation Technologies (EPT), as coined by Reig (2012). More concretely to the teaching of languages with the support of new technologies, it is worth highlighting the significance of Computer-Assisted Language Learning (henceforth CALL). The use of computers as a language teaching instrument dates back to the 60s (Araujo, 2013) notwithstanding that the concept of CALL was first coined by 1983, being understood as all the applications that computers offer in language teaching and learning (Pérez Torres, 2004).

Currently, both CLIL and CALL are of such magnitude that they are almost bound to relate to one another. In this regard, these two fields are linked in the present dissertation in order to delve into an issue of paramount importance; assessment. In this sense, this study attempts to explore the ways in which CALL activities can influence CLIL assessment. More to the point, the emphasis is placed on the role that oral skills can play as the main means for assessing in CLIL settings with the support of CALL instruments and activities.

This way, this paper follows a definite structure; in the present section, the focus of interest is identified, and the justification of the research is exposed together with the state of the question, the research questions, the objectives and the limitations. In the second section, the theoretical framework is developed, placing especial emphasis on the characteristics of CALL and CLIL, as well as on the key information regarding the assessment processes related to both concepts. The third section deals with the methodology followed to collect the data, which is analysed and discussed in the fourth and fifth sections, respectively. Lastly, a conclusion of the research is presented, followed by the bibliography and some appendixes that are relevant to complement the information of this dissertation.

1.1. Justification of the investigation

CLIL approaches have gradually gained acceptance in Europe, and they are expected to keep being deployed and implemented at the European schools in order to raise European citizens´ plurilingual competence, which is a key objective as recently pointed out by the Common European Framework of Reference for Languages (CEFR, 2018[1]) in its publication of February 2018. Meanwhile, the involvement of new technologies and CALL techniques within foreign language teaching contexts is also expected to keep gaining more and more relevance, which is brought into relief, amongst others, in the educative legislation; for instance, The Spanish Royal Decree 126/2014, February 28th, which establishes the basic curriculum for Primary Education in Spain, highlights in its article 2.2. the significance of the digital competence as one of the key competences that children are expected to develop.

Consequently, the relevance of CLIL and CALL seems unquestionable. Nonetheless, as detailed in section 1.2., the current literature of these fields fail to address a key issue that can link both disciplines; the improvement of CLIL assessment through oral CALL

[1] Although its first publication dates from November 2001, subsequent CEFR papers have been published, the most recent one in February 2018.

activities. The purpose of this research is to shed some light on this issue, and for this reason the present dissertation may be interesting for these fields; hence the justification of this investigation.

1.2. State of the question

Regarding the field of CALL, significant research has been found related to the applicability of LKT and CALL to develop the oral skills in foreign languages (Sharples, Taylor & Vavoula , 2005; Farhan, 2007; Blasco Mayor, 2009; Talaván, 2010; Slaouti, Onat-Stelma & Motteram., 2013; Stanley, 2013; Vázquez Cano & Martín Monje, 2014), as well as research addressing the assessment of foreign language oral skills (Levy & Gertler, n.d.; Sharples et al., 2005; Wagner, 2010; Bahrani, 2011; Beaven & Neuhoff, 2012; Slaouti et al., 2013; Stannard & Basiel, 2013; Wu, 2013; Caruso, Colombi & Tebbit, 2017).

As for the field of CLIL, the literature presents plenty of research related to its nature and pedagogic implications (Marsh, n .d.; Attard Montalto, Walter, Theodorou & Chrysanthou., n. d; Coyle, 1999; Dodge, 2001; Cummins, 2008; Pérez Torres, 2009; Coyle, Hood, & Marsh, 2010; Dalton-Puffer, 2011; Quartapelle & Schameitat, 2012; Casal, 2016; CEFR, 2018; Nieto-Moreno-de-Diezmas, 2018).

However, as noted by Maggi (2012), astonishingly little has been researched and published regarding CLIL assessment and evaluation. Therefore, the issue of assessment in CLIL contexts is scarcely found in the literature, although some authors have shed some light on it (Barbero, 2012; Maggi, 2012; Casal, 2016). This is a paradoxical fact given that:

> *"Assessment is not something that comes after instruction, but is an indispensable part of instruction. It is by thinking about assessment that we really start to sharpen up our idea of what CLIL is about and the role of language within it"* (Llinares et al., 2012, p. 280; in Barbero, 2012).

More closely to the nature of this investigation, no previous research has been found addressing how CALL instruments can enhance CLIL assessment processes and, more specifically, how the oral skills (listening and speaking) can be used as the main medium in CLIL assessments, with the support of CALL tools. That is way this dissertation gains significance for the field.

1.3. Research questions

Bearing in mind the state of the question outlined above, in order to address this research three main questions arise:

(1)

a. Can CALL activities enhance the assessment processes in CLIL contexts?
b. Are oral CALL activities suitable to assess learners in CLIL settings?
c. Can CALL tools address the dual focus assessment (language and content) that CLIL requires?

The responses for these questions are formulated as the research develops, and in section 5.2. they are given on the basis of the obtained findings.

1.4. Objectives of the investigation

The general objective of this investigation is to explore the ways in which CALL strategies and instruments can enhance the assessment processes within CLIL contexts through oral skills. It is important to clarify that the purpose is not to assess merely the oral skills with the support of CALL instruments, but to rely on these skills as a means to carry out assessment activities in CLIL contexts, where both language and contents ought to be evaluated.

In turn, in order to delve into this main purpose, two specific objectives have been set in this research:

(2)

a. To examine the applicability of CALL within CLIL evaluation processes.
b. To explore the feasibility of oral-form assessment activities in CLIL settings.

The accomplishment of these specific objectives is significant insomuch as it will determine the achievement of the general purpose.

1.5. Limitations

Apart from the shortage of time and resources that typically affect this type of research, the limitations of this study stem from two main causes; bureaucratic hindrances and lack of appropriate materials. On the one hand, it was quite complicated to find a school willing to have a CLIL group participate in the field work and, when it was found, plenty of paperwork had to be carried out and followed. This paperwork included; finding a teacher who would let his/her classroom participate; arranging different appointments with the head teacher; writing a blueprint of the project that the school direction needed to approve of; sending a formal request to the school inspector; and waiting for the participants´ parents´ and legal guardians´ consent to go ahead. Following all these steps supposed a significant amount of time, yet eventually a sample group could be reached, and the school participation was certainly much appreciated.

On the other hand, once the sample was available and the field work had been designed, some extra limitations arose in terms of materials difficulties. For example, some CALL tools and activities that were initially included in the design had to be cancelled or replaced by others as they did not work on the laptops that the participants had available. For instance, initially a session was designed with the tool Voicethread, but it had to be discarded given that it could not be displayed on the laptops. Another example of hindrance was the session with the tool Voki, in which some of the laptops microphones did not work properly, so participants had to share the laptops whose microphones could be used. This meant a slight time-delay, although this sessions could be deployed without further complications.

Notwithstanding these limitations, the field work phase could eventually be carried out after some slight delays and modifications, so it can be claimed that the study objectives and purposes were not hindered by these issues.

2. THEORETICAL FRAMEWORK

Three main sections can be distinguished in this theoretical framework. The first one deals with the issue of CALL, making especial mention of how the oral skills can be treated in LKT and CALL environments. The second section copes with CLIL, in which its main features and pedagogic implications are discussed. Lastly, the third section, which is bound up with the two previous ones, intends to explore the ways in which the assessment of CLIL can be enhanced by CALL methods and instruments.

2.1. Computer Assisted Language Learning (CALL)

This section begins with a brief overview of CALL. Then, the concept of Computer Assisted Language Testing (henceforth CALT) is outlined, exposing the reasons why this field does not fit for the purposes of the present thesis, despite the fact that it is related to the issue of assessment. Later, the main advantages of CALL are discussed and, last but not least, it is analysed how CALL can be linked to oral skills.

2.1.1. History and new paradigms of CALL

The use of computers as a language teaching instrument dates back to the 60s (Araujo, 2013). Nevertheless, the concept of CALL was first coined by 1983, being understood as all the applications that computers offer in language teaching and learning (Pérez Torres, 2004). Currently, plenty of acronyms can be found referring to the same field, depending on the approach in which the computer is involved (Pérez Torres, 2004):

(3)

a. CAI - Computer Assisted Instruction
b. CASLA - Computer Applications in Second Language Acquisition
c. CELL - Computer Enhanced Language Learning
d. TELL - Technology Enhanced Language Learning
e. WELL - Web Enhanced Language Learning
f. CMC - Computer Mediated Communication
g. CLL - Community Language Learning

According to Araujo (2013), and regardless of the acronyms, in all of them the computer has a clear educative purpose, notwithstanding that CALL tends to be criticised for being frequently more related to the area of informatics than the area of pedagogy. Either way, CALL has got an interdisciplinary nature, as it has traditionally benefitted from different fields such as programmed learning, computational linguistics, automatic translation, educative technology, or Human Computer Interaction (HCI). Consequently, CALL tends to be renewed and updated on a regular basis, although it cannot always take fully advantage of the innovations from these fields (Araujo, 2013).

On a different note, Chaka (2009) notes that a change of role on the part of teachers and learners has stemmed from CALL, which has historically gone through three main pedagogical phases, shown in (4):

(4)

a. *Behaviourist CALL*; where the computer served as tutor and focus was on drilled and repetitive activities.
b. *Communicative phase*; in which there was more interaction of the students. It was based on cognitivist and constructivist theories.
c. *Integrative CALL*; whereby the integration of the four language skills was promoted and a socio-cognitive view took place.

In view of these phases, seemingly CALL has reinforced the communicative approach to language teaching, and constructivism is the paradigm expected to be developed in the coming years. This fact is expected because, although technology cannot improve learners´ learning experiences by itself, it can provide new contexts and tools which will do (Palalas & Hoven, 2016).

In this regard, one of the most recent CALL expressions is Mobile-Assisted Language Learning (MALL), which enhances the flexibility of learning due to its ubiquitous nature, through which learning can take place anytime and anywhere (Palalas & Hoven, 2016). Lastly, another acronym that has burst upon the scene of CALL is Intelligent Computer Assisted Language Learning (ICALL), which refers to the way in which artificial intelligence techniques can be involved within CALL strategies (Araujo, 2013).

2.1.2. Computer-Assisted Language Testing (CALT)

It can be argued that another sub-branch or natural consequence of CALL is Computer-Assisted Language Testing (CALT), which needs to be outlined in a specific section. CALT makes use of computer applications for evaluating test takers' performance in a second language. According to Chapelle & Douglas (2006), within CALT different matters can be involved, such as the use of multimedia in language test tasks, computer-adaptive testing (CAT), or automatic response analysis. In this vein, Chapelle (2010)[2] notes that technology is worth using for testing due to its efficiency, equivalence, and innovation.

Suvorov & Hegelheimer (2014) point out that CALT has traditionally been regarded as a means for traditional paper and pencil tests. Nonetheless, this field has been subject of significant developments since the 1990s, as new ideas have come to the fore, for instance; computer-adaptive testing, integrated skills assessment, or automated evaluation. (Suvorov & Hegelheimer, 2014).

However, although the present dissertation has to do with evaluation processes in language and learning environments, no more references will be made to CALT in this paper owing to two main reasons:

On the one hand, CALT is still too linked to tests, instead of to the wider nature of assessment and evaluation. And more to the point, CALT lacks evidence of being useful to assess one of the oral skills; speaking (Suvorov & Hegelheimer, 2014). Therefore, given that the objective of this thesis is to explore new ways of assessing oral skills through ICT, as well as exploring how oral skills can be employed to assess learners in CLIL contexts with the help of new technologies, it seems clear that CALT does not fit for these outcomes.

On the other hand, CALT is exclusively related to the testing of languages. This fact automatically excludes the assessment of non-linguistic area contents and, consequently, CALT is against the nature of the "dual focus" that must be taken into consideration in CLIL evaluation (as discussed in subsequent sections), where both language and content are simultaneously assessed. As a matter of fact, assessment in CLIL ought to be more similar to the models offered by the disciplines than those commonly used in language lessons (Barbero, 2012).

[2] In Suvorov & Hegelheimer (2014)

2.1.3. Characteristics and advantages of CALL

Admittedly, CALL can involve some disadvantages. For example, is has been claimed that it may hinder socialisation processes; overwhelm students who are not accustomed to handling computers; or place too much emphasis on the learning of languages, instead of their acquisition, just to name a few (Araujo, 2013). In view of these shortfalls that CALL may present, seemingly they depend more on how CALL is utilised rather than the nature of CALL itself.

Another problem that can be included is the issue of affordability and affordances. According to Pegrum (2014), affordability is the extent to which individuals have access to new technologies and devices, whereas affordances have to do with the possibilities and abilities that these devices offer. Hence, the more affordability a population has, the more probabilities to benefit from a wider range of affordances. To put it another way, a significant problematic of CALL has to do with the lack of resources that some contexts may have in comparison with others, both at hardware and software levels.

Anyway, owing to its characteristics, the positive effects of CALL outnumber its drawbacks. In this sense, Pérez Torres (2004) classifies the advantages of CALL from the perspective of the teacher, the pupils and the computer, all of which are reinforced by the use of the Web. According to her, from the teacher´s perspective, the use of computers prompt them to have a role of collaborator, rather than instructor; carry out more regular and accurate assessments; and explore new ideas and methods, which increases their motivation. Besides, CALL lets teachers invest more time in pedagogic purposes as it helps them save time of classroom management and organisation (Araujo, 2013).

From the pupils´ perspective, computers increase learners´ motivation by giving them a more prominent role; permit self-evaluation and feedback immediately; and promote collaborative work and different learning strategies in keeping with pupils´ different needs (Pérez Torres, 2004). Moreover, Araujo (2013) adds that CALL provides learners with privacy and autonomous self-guidance, which is particularly beneficial for shy students.

Lastly, as for the materials used in CALL, Pérez Torres (2004) brings into relief that generally they are attractive as they present multimedia elements such as images, texts, audios or videos; imply more interaction on the part of learners and self-evaluation activities; and foster a significant learning by providing access to authentic materials. This features

contribute to the gamification of the foreign language learning, as they lead learners to regard the activities as games (Araujo, 2013).

2.1.4. The role of CALL in Oral Skills treatment

Admittedly, in language classrooms, oral skills have traditionally been underdeveloped and not as considered as written skills (Hughes, 2017). This fact is gradually changing and nowadays oral skills play a more desirable role within teaching contexts, and this change is greatly reinforced and enhanced by the use of CALL strategies on the part of both teachers and learners (Araujo, 2013). In addition, the appearance of the Web 2.0. has also contributed to this change of paradigm, prompting the oral skills to be an utmost priority for different fields such as Psycholinguistics, Cognitive Psychology, Bilingualism, and Applied Linguistics (Blasco Mayor, 2009).

According to Stanley (2013), the use of ICT tools is important because they prompt learners to negotiate meaning and put their oral skills into practice within significant communicative contexts where they want to have a say. In this regard, the synchronous web tools deserve especial mention given that they allow learners to have a real-time conversation with a real receiver or audience, promoting telecollaboration, which involves a huge cultural value as well (Stanley, 2013).

Another significant concept to be considered is "seamless learning" (Slaouti et al., 2013), referring to the possibility of keeping up the learning process beyond the classroom. This means that the oral skills can be worked on both inside and outside the educative contexts due to the ubiquitous nature of CALL, and more specifically MALL approaches (Sharples et al., 2005).

In this sense, Web 2.0. tools such as blogs, wikis and podcasts are worth pointing out as well, since they provide internet-based project work, boosting cooperative learning, interaction and creativity, as well as prompting learners to work at home, not only in class, which contribute to the development of the oral skills, as well as the improvement of students´ digital literacy (Vázquez Cano & Martín Monje, 2014).

Besides, other authors highlight the relevance of the digital video as a teaching resource in CALL, insofar as it encompasses plenty of advantages, such as authenticity; motivation, interest and confidence; the sociolinguistic and pragmatic level of language;

non-verbal features; active involvement and participation; and real vocabulary acquisition, accomplishing this way learners´ automatization of L2 input processing, (Blasco Mayor, 2009).

Lastly, the use of subtitles and subtitling tasks within foreign language learning contexts should be mentioned too, given that it has proved fruitful to enhance learners´ oral comprehension skills in a motivating and appealing way (Talavan , 2010), whereas the cooperative computer-mediated techniques are deemed beneficial for learning and teaching oral skills too (Farhan, 2007).

2.2. Content and Language Integrated Learning (CLIL)

This section begins with a revision of the origin of CLIL and its main legislation in Spain, concretely at the level of Primary Education. Following this revision, the definition of CLIL and some noteworthy aspects are outlined. And lastly, the characteristics and advantages of CLIL are mentioned.

2.2.1. Origin of CLIL and situation in Spain

Undoubtedly, the globalization process has had a tremendous effect in virtually all the aspects of our lives. Amongst others, this fact has exponentially increased the need to learn, at least, one foreign language, which is commonly the current lingua franca; English. Studying foreign languages inexorably involves advantages in terms of personal, academic and professional aspects. For instance, it develops a more accurate understanding of the first language; broadens our students' minds; and prompts them to value other cultures.

The Common European Framework of Reference for Languages (henceforth CEFR) was created in 2001 to describe foreign language proficiency at different levels, highlighting that the sooner a second language is learned, the better (CEFR, 2018). In this regard, Attard Montalto et al. (n. d.) outline that the term CLIL, for Content and Language Integrated Learning, was first launched by 1994 within a debate on how to bring language learning excellence, party promoted by the European Commission.

CLIL approaches have gradually gained acceptance in Europe, and they are expected to be used more and more at the European schools in order to raise European citizens´

plurilingual competence. In conformity with the CEFR (2018), the plurilingual competence is intended to lead learners to:

(5)

a. Switch from one language to another.
b. Express themselves in one language and understand a person speaking another.
c. Call upon the knowledge of a number of languages (or dialects, or varieties) to make sense of a text.
d. Recognise words from a common international store in a new guise.
e. Mediate between individuals with no common language.
f. Exploit paralinguistic strategies (mime, gesture, facial expression, etc.).

More concretely, in the case of Spain, Guadamillas & Alcaraz (2017) estate that bilingual programmes at public schools got started by 1996, when a collaboration agenda was agreed between the Spanish Ministry of Education and the British Council. This agreement was gradually included in the educative legislation and curriculums and concreted in the Order 5th April, 2000. Concurrently, the different autonomous communities have been regulating and developing their own legislation regarding bilingual education (Guadamillas & Alcaraz, 2017). Lastly, given that this thesis is focused on the stage of Primary Education, it is important to note that all bilingual legislation for this educative level must currently be in keeping with Royal Decree 126/2014, February 28th, which establishes the basic curriculum for Primary Education at a national level.

2.2.2. Definition of CLIL. Key aspects and concepts

Since its coinage in 1994, not a clear definition of CLIL has been easily agreed, although currently the definition provided by Quartapelle & Schameitat (2012) can be arguably accepted. These authors define CLIL as "the teaching of any non-language-subject through the medium of a language which is not the mother tongue" (p. 29).

Different key aspects concerning CLIL can be found in the literature, yet three basic ideas of the CLIL practice ought to be highlighted (Pérez Torres, 2009); firstly, the use of the target language is significant because it is regarded as a means, rather than as the main goal; secondly, the language put into practice depends on the area that is taught, not the other way around; and thirdly, fluency is more relevant than grammatical and linguistic accuracy,

yet these issues should be addressed as well. In this sense, the collaboration between L2 teachers and the area specialists is crucial for the CLIL practice (Pérez Torres, 2009).

Besides, a remarkable concept in the field of CLIL is scaffolding. Scaffolding has to do with the activities and strategies that learners are provided with in order to build up their knowledge and reach their Zone of Proximal Development (ZPD), coined by Vygotsky (1978). In other words, scaffolding is the sequence of steps pupils need to go through in order to accomplish tasks that they can carry out with some kind of support.

In this regard, Dodge (2001) classifies up to three types of scaffolding; 1) reception scaffolding, which are the strategies that help pupils comprehend and assimilate new relevant information; 2) transformation scaffolding, which helps learners transform the information received into new contents; and 3) production scaffolding, which prompts students to create new ideas and materials. These three types of strategies tend to follow the given order, but they can take place at any time of the learning process (Dodge, 2001).

Another remarkable aspect is the 4 Cs framework established by Coyle (1999), referring to content, communication, cognition and culture:

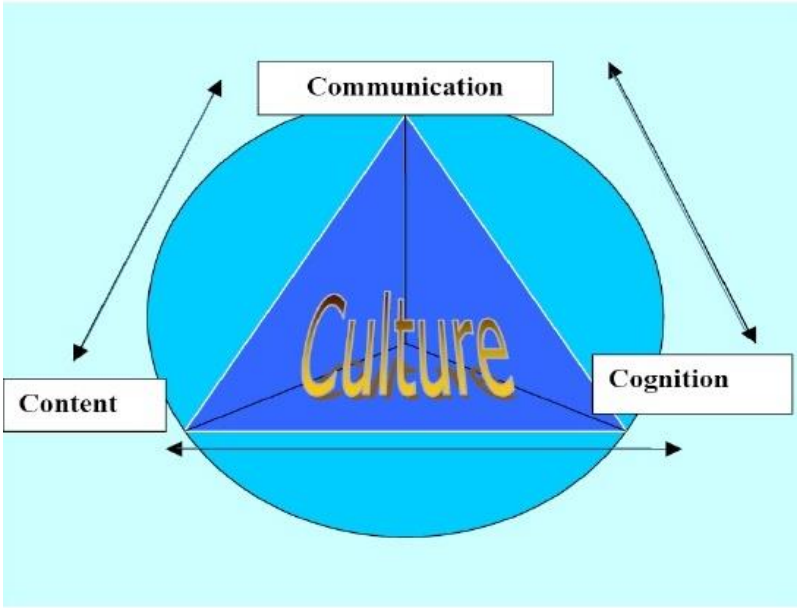

Figure 1. The 4Cs conceptual framework for CLIL

Source: Coyle (n. d.)

These elements have been widely accepted in the literature, and their knowledge is of paramount importance for the planning and deployment of CLIL programmes (Pérez

Torres, 2009). Therefore, due to their relevance, it is worth delving into these four components, which are given in (6):

(6)

a. Content.
b. Communication.
c. Cognition
d. Culture

With regards to (6.a), it makes reference to non-language-subject elements that are addressed within CLIL contexts, which would be attained in the L1 in non-bilingual programmes. Working on these elements is the main priority in CLIL, and more importance must be given to them than to language achievements (Maggi, 2012; Casal, 2016).

Regarding (6.b), it refers to the linguistic processes that take place throughout CLIL projects. Coyle et al. (2010) describe the language Triptych, distinguishing between; 1) language *of* learning, which are key words and phrases needed to access content; 2) language *for* learning, necessary to carry out class tasks and interact with classmates; and 3) language *through* learning, which sporadically and naturally results from the CLIL sessions. Besides, Cummins (2008) makes a distinction between Basic Interpersonal Communicative Skills (BICS), referring to the language of general use mostly studied in the L2 area; and Cognitive Academic Language Proficiency (CALP), which makes reference to technical and academic languages and is more related to CLIL programmes.

As for (6.c), it has to do with the mental and cognitive processes that are involved in CLIL environments. Coyle at al. (2010) make a distinction between; Higher-Order Thinking Skills (HOTS), such as hypothesizing, problem-solving, information exchange and interaction, summarizing, decision-making or self-assessing; and Lower-Order Thinking Skills (LOTS), such as remembering, identifying, reordering sentences, linking concepts to images, classifying items, listing words, and so on.

Finally, as to (6.d), it makes reference to the integration of cultural opportunities into the CLIL classroom. A noteworthy way to integrate culture is by getting and keeping in touch with foreign learners and groups. In this vein, this fact is nowadays a feasible practice due to the use of ICT, which promote telecollaboration (Pim, 2013; Slaouti at al., 2013; Stanley, 2013). In the literature, some authors also mention a fifth "C"; Community, referring to how CLIL can be linked to learners´ own culture and nearest context, although

this term can generally be included within the concept of culture as well (Attard Montalto et al., n. d.).

2.2.3. Characteristics and advantages of CLIL

Defining the most representative characteristics of CLIL is a hard task, given that there is no specific or universal methodology that best works with it (Coyle at al., 2010). This occurs because CLIL methods are flexible and dynamic, so they can adjust to any area, level or needs (Coyle, n.d.). As such, different approaches can take place in CLIL, and their effectiveness tend to depend more on the context, rather than the method itself (Casal, 2016). However, Attard Montalto et al. (n. d.) remark that whatever the methodology is, there are some common issues to be considered by any CLIL approach, for instance:

(7)

a. In CLIL, it is important to use audio-visual aids and multimedia in order to overcome difficulties caused by the use of a new language.
b. Language mistakes are permitted as long as communication takes place.
c. The use of scaffolding is crucial.
d. It is preferable to use the target language as much as possible, yet the use of the L1 is allowed if necessary.
e. Collaboration among the teaching staff is required, particularly between the L2 specialist and the area teacher.
f. Pupils´ language level, needs and interests should be considered when planning CLIL lessons.
g. At early stages such as Primary Education, plenty of language models and exposure must be given to children, and oral skills are more prominent than written ones in these early stages.

Considering these issues, seemingly CLIL models have provoked a change of methodology on the part of teachers, as the focus of the process is on the students. According to Dalton-Puffer (2011), teachers are now compelled to cater for learners´ future needs and features, who will be expected to solve problems, plan their own work and find out information for themselves using a range of sources, particularly the Internet (Pim, 2013).

In turn, these characteristics and pedagogic changes encompass a wide range of learning advantages. As exposed by Attard Montalto et al. (n. d.), CLIL presents remarkable benefits such as the improvement of the linguistic competence, oral communication and intercultural skills; the increase of pupils´ motivation and confidence in both the language and the content; as well as the development of multilingual attitudes and cognitive strategies. Moreover, Maggi (2012) notes that not only L2 skills are improved through CLIL, but also L1 abilities are. The reason why this occurs, according to Nieto-Moreno-de-Diezmas (2018), is that the learning strategies applied in both languages are similar, in such a way that improving L2 skills in CLIL contexts tend to automatically improve L1 skills as well.

Therefore, CLIL approaches encompass positive consequences for the pupils as they boost their cognitive development and improve the target language acquisition (Dalton-Puffer, 2011). Additionally, Casal (2016) points out that CLIL environments provide students with significant opportunities for them to work collaboratively and put into practice more sophisticated language structures and strategies than in ordinary language lessons. Because of its nature, cooperative learning completely fits in CLIL contexts, reinforcing the 4Cs framework (Casal, 2016).

Lastly, apart from the pedagogical benefits, Coyle (n. d.) notes that CLIL methodologies involve political, economic and social reasons as well, such as unifying a common language in territories where plenty of different languages coexist. In this regard, David Marsh (n .d.) points out that CLIL approaches are a way to overcome social inequalities as they provide all learners with opportunities to "pick up" the language, regardless of their socioeconomic backgrounds, due to the "naturalness" of the environment created.

2.3. The issue of assessment; linking CALL and CLIL

Regardless of the situation, when it comes to assessment, it is generally important to distinguish among three main questions; *what*, *how* and *when* to assess. The first question refers to the evaluation criteria; the second one makes reference to the instruments of evaluation; and the third question indicates three moments for assessing, as pointed out by Wu (2013); initial assessment (at the beginning of the process), formative assessment (throughout the learning process), and summative assessment (at the end of the learning experience).

In this section, the relationship between CALL and the assessment of oral skills is discussed, as well as the evaluation in CLIL contexts, in order to subsequently explore the ways in which CALL methods and instruments can enhance the assessment of CLIL pupils´ through their oral skills. Last, but not least, it is shown a selection of LKT tools that can be applied to such objective.

2.3.1. CALL and the assessment of the Oral Skills

According to Chapelle (2007), when evaluating CALL effectiveness, three different issues should be considered; the software used, the task deployed, and the students carrying it out. Assessment is gradually being deemed as part of the learning cycle, not merely the conclusive part. Therefore, teachers are giving more importance to formative assessment, rather than summative processes (Stannard & Basiel, 2013), and this tendency is in turn being reinforced by technology and the use of CALL (Levy & Gertler, n. d.). In this sense, in accordance with Bahrani (2011), technology-based assessment ought to be "authentic" (real life situations ought to be considered); "valid" (assessment is intended to check what students have learnt, not what they have not acquired yet); and "reliable" (learner´s final results should be representative of their actual performance).

Additionally, another noteworthy concept is "washback", which can be understood as the type and amount of feedback provided through technology-based assessment, and how it will impact on subsequent teaching and learning (Stannard & Basiel, 2013), as well as "transferability", which is the extent to which technology-based assessment encompasses language skills and strategies that can be taken out and used in other communicative situations (Caruso et al., 2017).

Keeping these features in mind, arguably traditional tests, such as fill in the gaps or multiple choice questions, cannot be considered authentic and valid, as they do not represent how language is typically used by speakers for communicative purposes in real life contexts (Caruso et al., 2017). Alternatively, technology-based assessment can provide activities that foster pupils´ creativity and motivation by encouraging group work and significant responses in which language skills are integrated, rather than assessed in isolation (Bahrani, 2011).

In view of this technology-based assessment aspects, some methodological implications can be highlighted with respect to how oral skills should be assessed and, by extension, what type of CALL tools are worth considering, given that, as noted by Stannard

& Basiel (2013); "One of the areas where the affordances of ICT is perhaps most pertinent is in the area of oral skills." (p. 166).

Concerning the most useful ways to assess the oral skills, Wu (2013) emphasizes the importance of relating language to its context, as well as culture (Beaven & Neuhoff, 2012). In this vein, when it comes to the listening skill, Caruso et al. (2017) note that, although there is still little research proving this fact, video technology is thought to offer more authentic language input and learning opportunities than audio-only texts, apart from providing students with direct contact with the culture represented on screen. This is due to the belief that including the non-verbal components of a spoken text will be useful for listeners in comprehending the aural messages (Wagner, 2010). In this sense, it is important to consider CALL methods, given that they are a reliable source of multimedia materials, such as videos (Pérez Torres, 2004).

Besides, Caruso et al. (2017) mention that listening must be deemed as a process, rather than as a product; if listeners are asked to recall information once the listening passage is over, students will be working on memory processes, not listening comprehension. This implies that learners should be evaluated during while-listening activities, not only during post-listening tasks. In this vein, CALL meets the requirements to enable assessment to take place throughout all the listening process (Araujo, 2013).

Apart from this, as for the assessment of speaking, Levy & Gertler (n. d.) point out that learners should be allowed to speak freely and redo their work as often as they wish whilst the assessor observes them. In this regard, according to Pihkala-Posti (2014), CALL tools permit learners to track their own oral productions and edit them as they see fit before sharing them with their classmates or teacher, who can then assess their work.

In this sense, Alonso (2012)[3] describes four stages to follow when dealing with oral production skills with technologies; planning, recording, edition and sharing. In relation with this, minor mistakes can be permitted as long as overall accuracy and fluency performance is acceptable. This way, learners will be prevented from going through stressful and frustrating processes which may impact negatively and hinder their spoken productions (Levy & Gertler, n. d.).

[3] In Vázquez Cano & Martín Monje (2014)

On a different note, other ways that are becoming more and more commonplace to deal with oral skills assessments are self and peer-assessment processes (Beaven & Neuhoff, 2012; Stannard & Basiel, 2013), although "the usefulness of self-assessment is not necessarily obvious to all students. In addition, it is important that learners can compare their self-assessment with other peer- or expert assessments of the same samples." (Beaven & Neuhoff, 2012, p. 4).

This change of paradigm is reflected in the instruments of evaluation that get involved when assessing the oral skills. For instance; the portfolio, which engages the learners to reflect on their progress throughout all the learning process, or its digital version; the e-portfolio, which may take the shape of, among others, a blog, wiki or even Virtual Learning Environment-VLE (Stannard & Basiel, 2013).

Some other examples of tools can be; chats, podcasts or mobile phones (Bahrani, 2011), which are being reinforced by the increase of online assessment in seamless learning contexts, where language teaching and learning expand beyond the traditional classroom (Slaouti et al., 2013), by taking advantage of the ubiquitous nature of CALL approaches, and more specifically MALL (Sharples et al., 2005).

2.3.2. Assessment in CLIL contexts

Roughly speaking, assessment in CLIL involves similar features than assessment in general (Barbero, 2012). In this regard, three different types of assessment are distinguished; assessment "of" learning, with a summative nature; assessment "for" learning, related to formative assessment; and assessment "as" learning, in which alternative forms of assessment take place, such as self- and peer assessment, as well as alternative tools such as portfolios or observation grids and rubrics (Maggi, 2012).

CLIL assessment must be valid, reliable and present feedback, as well as being authentic (Barbero, 2012). Authentic assessment occurs when it is linked to real life situations, rather than merely soliciting answers which only require simple and easy to assess responses. In other words, students should be assessed when performing tasks in which they behave as if they were in a real, meaningful and significant situation.

Given that learning in CLIL contexts is a social and interactive process, assessment ought to be deemed socially, thereby placing the emphasis on learners´ active participation

and interaction with others, even though two different aspects should be considered when assessing them; their contributions to the group and their individual performance (Casal, 2016). In this regard, it seems clear that "the more involved the student becomes in the group task, the better their performance at an individual level will be" (Casal, 2016, p. 145).

Besides, a dual focus should be taken into account, in which both content and language are assessed (Maggi, 2012; Barbero, 2012), notwithstanding that in CLIL more emphasis is placed on meaning, being the foreign language a means for learning, rather than a goal itself (Coyle et al., 2010). For that reason, Barbero (2012) considers that "assessment in CLIL should be more like the models offered by the disciplines than those commonly used in language lessons" (p. 54). Either way, the assessment of language and content must be performed simultaneously (Casal, 2016).

Apart from this, learning objectives and evaluation criteria must be very clear and, when possible, pupils should be aware of them (Maggi, 2012). In this sense, the use of rubrics can contribute to clarify learning outcomes, assessment criteria and feedback (Casal, 2016). It is important to note that rubrics can be applied to any kind of written or oral work in CLIL, and they can take two forms; holistic rubric; which assesses the work as a whole; and analytic rubric, including three main components; assessment criteria; an identified behaviour; and a score (Barbero, 2012).

Besides, self-assessment and peer assessment are of paramount importance as they have been proven to engage students in their own learning process, both individually and collaboratively (Maggi, 2012). In addition, peer and self-assessment involve positive effects on learners´ self-esteem and motivation, which are key factors affecting language learning, especially the latter (Moreno, 2010). Furthermore, it has been shown that combining rubrics with self-assessment processes increases the effectiveness of overall teaching practices (Barbero, 2012).

All these aspects considered, it can be argued that assessment is fruitful within CLIL settings when it is useful to adapt future teaching; understandable feedback is given to students on a regular basis; and pupils are aware of the learning outcomes and evaluation criteria or, to put it another way, it is crucial that students know what they are expected to achieve and how they should attain the objectives (Casal, 2016).

2.3.3. Linking CALL to CLIL assessment

After analysing in the previous two sections the role of CALL features in oral skills assessment, as well as the characteristics of CLIL assessment, a relationship can be established between how the nature of CALL may improve the quality of evaluation in CLIL contexts, making mention of the role of oral skills as well. In the table below the main connections are shown:

Characteristics of CLIL assessment	How they are enhanced by CALL
Formative assessment is more prominent than summative in CLIL.	CALL reinforces the role of formative assessment in the learning process.
Assessment in CLIL must be valid, reliable and authentic.	Proper technology-enhanced assessment increases validity, reliability and authenticity of assessment.
Pupils should be in significant and meaningful contexts when evaluation takes place.	CALL provides simulated significant and meaningful assessment contexts.
It is advisable to evaluate different language skills concurrently.	CALL promotes an integrated-skill approach and assessment.
CLIL assessment calls for social, interactive and communicative processes.	CALL methods can involve cultural and contextual aspects within assessment processes.
In CLIL, learners should be aware of what they are expected to do and how they are assessed.	CALL encourages students to keep track of their learning process and control the objectives to be achieved.

CLIL students should be assessed on a regular basis.	Technology-enhanced assessment is easy to be carried out regularly without significant difficulties.
In CLIL, peer and self-assessments can be of paramount importance.	CALL provides opportunities for peer and self-assessments to take place in the classroom.
CLIL calls for alternative oral and written instruments of assessment.	CALL provides alternative instruments to assess learners through written and oral means.

Table 1. Relation between CLIL assessment and CALL features

Source: Own elaboration

In view of these connections, it can be argued that CALL characteristics are suitable to assess pupils´ linguistic proficiency, but also their content knowledge. In other words, CALL features can be appropriate to fit the dual focus that CLIL evaluation processes require.

2.3.4. LKT tools to assess in CLIL through the Oral Skills

Bearing in mind the aforementioned CALL features that suit CLIL assessment, in this subsection some LKT instruments are exposed, which can be utilised in order to assess oral skills straightaway in the form of CALL activities, but that can also be used in order to assess CLIL students´ language and content performances by means of their oral skills. These tools have been classified into different headings, notwithstanding that some of them could be included in more than one category on account of their flexibility and applicability in language settings.

This way, in the table below (table 2) some of these instruments are selected and briefly described:

PODCASTS	
Vocaroo	With Vocaroo, learners can record a podcast and automatically share it online with anyone else.
Voki	This tool lets students create an avatar and give it a voice, enabling the teacher to assess their oral expression skill as well as the contents that pupils have their avatars reproduce.
Tellagami	This app is similar to Voki, but it lets students create a wider range of avatar features.
Songify	With this app, learners can record themselves to create a song with the podcast produced, as well as listening other learners´ contributions.
CHAT	
Whatsapp	Whatsapp, which is well known by students, can be used by pupils to communicate through recorded oral messages, which the teacher can store and assess.
Skype	Through Skype, students can maintain oral conversations and be unconsciously assessed by the teacher when discussing the contents of the subject.
MOBILE PHONE	
Movenote	With Movenote learners can make and record an oral presentation including pictures, videos or audios from their own phones and then share their work.
Chatterpix kids	Children take a picture or use a given one in order to edit it and give it a voice by recording themselves.

VIDEO TECHNOLOGY	
Voicethread	Either individually or collaboratively, learners work around audiovisual materials by recording themselves.
ThingLink	Similar to Voicethread. Students work with interactive videos or other materials, yet responses tend to be on written form.
Edpuzzle	Edpuzzle is an example of digital video use. This tool can edit videos so learners watch them and answer content questions meanwhile.
MailVU	Children tape themselves to send the result online straight away.
Elllo	On Elllo, learners watch and listen to a video. It includes pre, while and post-listening activities that teachers can look at to assess.
VIRTUAL LEARNING ENVIRONMENT	
Edmodo	Learners can watch videos or upload recordings on Edmodo, where teachers can keep track of learners´ progress.
Schoology	Similar to Edmodo, though Schoology is oriented to higher levels than in the previous case.
E-PORTFOLIO	
ClassDojo	Although it is more related to class control, through ClassDojo learners can upload their oral work and keep track of it.
Seesaw	With Seesaw, learners share with classmates, teacher and their parents what they do (e.g. presentations, oral descriptions, etc.).

Table 2. LKT instruments to assess CLIL pupils through the Oral Skills

Source: Own elaboration

In view of these tools, it seems clear that there is a wide range of possibilities offered by LKT and CALL instruments. For this research, some of them are used in the field work phase, such as Voki and Edpuzzle.

3. METHODOLOGY

According to Pino, Rodríguez & Mañana (n. d.), in linguistic research four[4] main phases can be distinguished; preparatory, field work, analytic, and informative phases. These phases are described in the figure below:

1. Preparatory phase	It is the first step of the investigation, and involves two sub-phases: 1. *Reflexive phase*; research of bibliographic material to determine the centre of interest for the investigation and the theoretical framework. 2. *Design phase*; step in which the research is designed.
2. Field work phase	Phase of the investigation in which the researcher collects data from the samples through different instruments such as surveys, interviews and so on.
3. Analytic phase	Phase of the investigation in which the analysis and appraisal of the obtained data take place.
4. Informative phase	Phase of the investigation intended to present and spread the research conclusions.

Table 3. Phases of linguistic research

Source: Own elaboration

Taking the aforementioned steps into consideration, the present part of this paper intends to delve into the design phase of this research. As such, in the following sections different design aspects are addressed and detailed, such as the hypotheses proposed, the research approach or the investigation method, as well as the instruments and techniques of data collection.

[4] Some authors also include a phase called *operative planning*, understood as the phase of the investigation in which the theoretical framework is built and developed.

3.1. Hypotheses of the study

Two main hypotheses are brought into relief in this research:

The first (and main) hypothesis is that introducing CALL strategies in a CLIL Primary Education classroom can enhance the evaluation processes and make them more appealing for learners by increasing their motivation.

Additionally, a second hypothesis is that the role of oral skills in assessment activities can be boosted by the use of CALL tools, which can prompt these skills to be not only the goal of the learning process, but also the means.

3.2. Research approach

This investigation is included within the so-called ideographic sciences, which cope with isolated facts that take place in singular and individual contexts (Nunan, 2012). In the field of education, ideographic studies are more significant than nomothetic sciences, which seek for universal principles.

Besides, the approach in this investigation is mainly qualitative, although quantitative data are also collected and considered. According to Gray (2004), qualitative research is a sort of study with subjective and inductive nature, highlighting the significance of the research process. Conversely, quantitative research is a sort of investigation with objective and deductive nature, which focuses on the research objective, rather than the process (Gray, 2004).

Either way, this piece of work is encompassed within the interpretative/hermeneutic paradigm, which generates scientific knowledge through the study of an isolated reality within its context, thereby relying principally on qualitative methods (Nunan, 2012). All in all, this investigation belongs to the experimental-qualitative-interpretative paradigm; a mixed approach that provides an interpretive and subjective analysis through experimental and predominantly qualitative methods (Nunan, 2012).

3.3. Applied method

The methodology applied has a deductive nature; in other words, the purpose of the investigation is to prove an initial theory or hypothesis through collected data, which in this case is gathered through a primary research[5]. Besides, a pre-experimental design is followed, because this research dispenses with a control group[6] and pre/pro tests, and the results obtained cannot be generalised. According to Nunan (2012), pre-experimental designs are valuable and useful to obtain rich and accurate information from specific situations, which eventually can usher in a deeper research or be the introductory phase to delve into more sophisticated investigations.

In addition, this study is a classroom-based research; to put it another way, the research takes into deep consideration the information collected in a genuine classroom designed for teaching and learning purposes, not research issues. Lastly, this investigation can be considered to be a process-product research; in accordance with Gray (2004), process-product research makes reference to a study intended to analyse causal relationships between classroom processes and learning outcomes.

3.4. Investigation technique

Given that it is the technique that best fits the proposed objectives, this investigation in developed through a case study. According to Cohen, Manion, y Morrison (2011), case studies can penetrate situations in ways that are not always susceptible to numerical analysis. In this regard, Robson (2002) points out that case studies have a flexible design; although traditionally they were oriented to single cases, they can also be done "on a group, on an institution, on a neighbourhood, on all innovation, on a decision, on a service, on a programme and on many other things" (p. 180-181).

Despite its criticisms, case study is deemed a scientific method insofar as it is the outcome of a process of enquiry carried out through critical norms and standards of

[5] Procedure in which primary sources are used principally (e.g. a group of students), and secondary sources are less significant (e.g. books or other researches).

[6] A sample that is analysed by the researcher in order to be compared to the experimental group. No independent variable is introduced in the control group.

rationality. Therefore, case studies present a wide range of advantages, yet some disadvantages should be pointed out as well. In the chart below, the main advantages and disadvantages of case studies are shown:

Strenghs	*Weaknesses*
1. The results are more easily understood by a wide audience (including non-academics) as they are frequently written in everyday, non-professional language.	1. The results may not be generalizable except where other readers/researchers see their application.
2. They can be undertaken by a single researcher without needing a full research team.	2. They are not easily open to cross-checking, hence they may be selective, biased, personal and subjective.
3. They are immediately intelligible; they speak for themselves.	3. They are prone to problems of observer bias, despite attempts made to address reflexivity.
4. They catch unique features that may otherwise be lost in larger scale data (e.g. surveys); these unique features might hold the key to understanding the situation.	
5. They can embrace and build in unanticipated events and uncontrolled variables.	
6. They are strong on reality.	
7. They provide insights into other, similar situations and cases, thereby assisting interpretation of other similar cases.	

Table 4. Strengths and weaknesses of case studies

Source: Nisbet and Watt (1984, in in Cohen et al. (2011)

Seemingly, the advantages of case studies outnumber its disadvantages. In addition, through this case study qualitative data are collected primarily, but also quantitative data are gathered, which results in data triangulation. Nunan (2012) defines triangulation as the combination of different methods to address the same subject. In this regard, Cohen et al. (2011) point out that triangulation processes enhance research quality as they provide them with more validity and reliability.

More concretely, two types of triangulation mentioned by Robson (2002) take place in this research; on the one hand, *data triangulation*, given that more than one method of data collection are involved (as exposed in subsequent sections); and on the other hand, *methodological triangulation*, since both quantitative and qualitative methods are combined. In this sense, the perspectives of the participants and the perspectives of the researcher will be compared and combined as well.

Lastly, in order to finish this section, it is important to clarify why a case study has been chosen instead of action research, as not distinguishing these concepts may lead to confusion because, as stated by Robson (2002), "there may be difficulties in defining and delimiting exactly what one means by the 'case' when the focus moves away from the individual person" (p.180).

According to Nunan (2012), action research is a sort of inquiry carried out by practitioners (e.g. a group of teachers) in order to deal with and figure out an issue affecting their daily work, whereas case study is the analysis and observation of how a phenomenon takes place and develops within a given situation, stretching from an isolated classroom to an entire school.

In this vein, the main reason why this investigation is a case study (and not action research) is that currently the researcher is not a practitioner. Either way, it is important to note that "the flexibility of case studies in design and approach, as well as in the use of method, encourages their use as a model for action research." (Robson, 2002, p. 217). In other words, it can be argued that case studies can be deemed as action research, and vice versa.

3.5. Instruments and techniques of data collection

In order to carry out a research it is indispensable to select the appropriate instruments and techniques of data collection (Robson, 2002; Nunan, 2012). In the following subsections, the emphasis is placed on the description of the chosen tools and techniques for this investigation.

3.5.1. Instruments

When it comes to social sciences research, three main types of instruments can be distinguished (Nunan, 2012);

(8)

a. *Physical instruments*; referring to any touchable tool that is used throughout the investigation process;
b. *Psychological instruments*; in relation with the psychological features that are involved in the research process, such as construct or languages;
c. *Social instruments*; which are tools related to social activity or involvement, such as samples and population.

The next subsections deal with some of these kind of instruments, focusing on the variables of the study and the sampling, respectively.

3.5.1.1. Variables of the investigation

In linguistic research, a variable can be defined as anything that does not remain constant, such as aptitude or motivation; it is a feature or characteristic that changes (Cohen et al., 2011). In this regard, there are several types of variables, yet for this research only two types are considered; dependent and independent variables. According to Nunan (2012), an independent variable is the variable expected to influence another one, whereas the dependent variable is a feature, behaviour or process expected to be influenced by the independent variable.

To put it another way, the independent variable remains unchanged and constant, and it affects the dependent variables, which can be influenced. In this sense, for this research the following variables have been set:

Independent variable	CALL instruments and activities.
Dependent variables	The improvement of oral evaluative activities. The increase of pupils´ motivation in CLIL evaluation processes.

Table 5. Variables of the study

Source: Own elaboration

Therefore, the purpose of this research is to explore if CALL activities can improve the evaluation activities based on the oral skills, as well as increasing pupils´ motivation when they are assessed. This way, the relation between the variables is linked to the objectives of the study.

3.5.1.2. Sample and population

According to Robson (2002), the sample is a smaller portion of a population used in an investigation, whilst the population is composed by all subjects or individuals that have one or more features in common. In this sense, the sample of this research is a group of pupils from fifth level of Primary Education belonging to a state school, who are immersed in a bilingual programme and make use of laptops on a regular basis, usually once or twice a week.

The school is called C.E.I.P. Príncipe Felipe, and it is located near the centre of the town of Albacete. Its pupils mainly belong to middle-class socioeconomic backgrounds, and different nationalities are found in it. Besides, the school caters for over 200 Infant Education children and around 450 Primary Education pupils, and it counts on different facilities such as playgrounds with basketball and football pitches, a library, a canteen, a gym and a language laboratory.

In the table below the characteristics of the sample and population of this investigation are specified:

Sample	Group of 24 pupils from 5th level of Primary Education	Girls	14
		Boys	10
Population	All 5th level learners from C.E.I.P. Príncipe Felipe		

Table 6. Sample and population

Source: Own elaboration

Therefore, this sample is composed of 24 subjects. This size may seem insufficient or scarce, but it is important to note that "sample size might also be constrained by cost – in terms of time, money, stress, administrative support, the number of researchers, and resources" (Cohen et al., 2011, p. 102). This way, in qualitative research, such as this case study, the size of the sample is likely to be small in comparison with other kinds of research. Nevertheless, experimental methodologies, such as pre-experimental designs, require a sample size of no fewer than fifteen cases (Cohen et al., 2011). Thus, arguably 24 subjects are appropriate for this research in terms of sample size.

Lastly, it is important to point out that this sample has a non-probability and purposive nature. This means that not all members of the population have chances to take part in the process (Cohen et al., 2011). In this research, this occurs because a group has been chosen on purpose since it involves the use of computers regularly, whereas the other groups compounding the entire population were automatically excluded because they do not take advantage of laptops on a regular basis. Therefore, it can be argued that this sample cannot be proven to be fully representative of the whole population.

3.5.2. Data collection

Since relying on multiple sources increases the validity and reliability of the study, in this research three techniques of data collection are involved within the field work phase, in order to address the proposed objectives, as given in (9):

(9)

a. *Participant observation*; in which the researcher participates in three sessions within the sampling group, in order to introduce the independent variable and observe how it influences the dependent variables.
b. *Self-completion questionnaire*; so that the participants can express their experiences and opinions after the sessions.
c. *Semi-structure interview*; in order to gather additional information that cannot be shown through the questionnaires, such as non-verbal communication, debates among the sample participants, and so on.

The following three subsections delve into these techniques of data collection, making mention of their main advantages and disadvantages.

3.5.2.1. Participant observation

Nunan (2012) defines observation as a way of collecting information through a planned and set strategy dependent on the previously established outcomes. Observation encompasses three stages; 1) Setting of goals; 2) Collection and categorisation of data; and 3) Interpretation of the results. In this research the first stage has already been addressed, the second phase takes place throughout the sessions with the sample group, and the third stage will take place once these sessions are over.

According to Robson (2002), observational methods, such as participant observation, are widely used in flexible designs, particularly in those which follow an ethnographic approach, and they are principally related to qualitative styles. Participant observation might be particularly suitable for projects with small groups that take a reasonably short time, although the preparation is time-consuming (Robson, 2002). Hence, this method seemingly fits for this research.

Observation raises criticisms, such as the thought that the researcher´s presence can influence the usual behaviour of participants, as well as difficulties in making sure that the investigator reflects what actually takes place. Nonetheless, observation is especially valuable when assessing the effectiveness and implications of innovative practices within a given context (Robson, 2002).

This way, data from direct observation contrasts with complementary information obtained by virtually any other techniques, such as interviews and questionnaires. In fact, direct observation is crucial to either underpin or contest the information provided by these other methods, given that it provides the researcher with the chance to observe and analyse how individuals behave in a given situation (Robson, 2002).

Additionally, Robson (2002) suggests that researchers should make a conscious effort to distribute the attention widely and evenly; keep an open mind throughout to avoid pre-judgements; put aside interpersonal factors and preferences; and write up field notes promptly before the behaviours gazed are forgotten.

For this case study, three sessions are designed and observed, in which CALL activities are proposed to analyse how CLIL learners can be assessed through them thereby using principally the oral skills. In this regard, three CALL tools have been selected to assess CLIL pupils´ linguistic and non-linguistic knowledge; Edpuzzle, Plickers and Voki. In order to establish the evaluation criteria and learning standards, it has been taken into account both Royal Decree 126/2014, February 28th, which establishes the basic curriculum of Primary Education nationwide; and Castilla-La Mancha Decree 54/2014, July 10th, which regulates the curriculum of Primary Education in this autonomous community (note that the sample selected belongs to a school from this community). Each tool and activity has been deployed in a different session, which are detailed in APPENDIX I.

Besides, two rubrics are involved in the process; the first one is used by the researcher to jot down what is observed in the sessions (placing especial emphasis on the variables), whereas the second rubric is intended to assess pupils´ performance both at linguistic and content level, in conformity with the dual assessment that CLIL calls for, and keeping in mind the corresponding legal frame for the stage of Primary Education.

The first rubric has been designed to be fulfilled as the three sessions develop. Conversely, the second one will be filled out mainly after the sessions, in order to assess pupils´ performance and work. Both rubrics are shown in APPENDIX IV and V, respectively.

3.5.2.2. Self-completion questionnaire

Questionnaire can be defined as a set of questions with a predetermined goal that are relevant for the topic being investigated (Nunan, 2012). According to Robson (2002), "the survey questions should be designed to help achieve the goals of the research and, in particular, to answer the research questions" (p. 241). For this reason, the questions of the questionnaire are designed keeping in mind the research objectives.

More to the point, for this case study a self-completion questionnaire takes place, where respondents fulfil the questions by themselves, yet they can be administered on a group basis (Robson, 2002). In this case, the learners from the class complete them with the supervision of the researcher, given that self-completion questionnaires can be subject to response bias.

For instance, people with reading or writing difficulties are less likely to respond, and this fact can be aggravated in the case of children (Robson, 2002). For that reason, in this investigation the Primary Education pupils completing the self-completion questionnaires are supervised by the researcher in order to make sure that all participants understand all the proposed questions. In this sense, the questionnaire, which can be seen in APPENDIX II, is carried out through the tool Plickers[7] in order to make the process more appealing for learners.

3.5.2.3. Semi-structured interview

According to Nunan (2012), an interview is a face to face or computer-based conversation in which the researcher interacts directly with the subjects of the sample to obtain accurate information provided by verbal and non-verbal communications. It is a bidirectional process and can be a group interview. This admits three types; 1) Structure interviewed, where questions are prepared beforehand; 2) Semi structured interview, in which there is a guide but participants can ask on their own; and 3) Non-structured interview, where the interaction is free.

[7] Plickers is a tool to create group quizzes and questionnaires in a motivating and appealing way for learners. It consists in handing out QR codes to pupils, which they will orient towards one of its four sides depending on the answer they intend to give, which are scanned though a mobile phone.

Interviews can be particularly useful when collecting data from children, as this method provides additional and rich information that cannot be visible in questionnaires, such as body language, facial expressions, gestures, or further explanations and justifications (Robson, 2002). For this research, a semi-structured interview is utilised, which has predetermined questions that can be modified as the conversation develops. This type of interview is widely used in flexible designs, such as case studies, either as the only method or combination with others.

Apart from being semi-structured, the interview used in this research is a group interview as well, where groups of four or five learners are interviewed concurrently. The purpose of this election is to give them the chance to exchange their ideas, share their impressions and discuss others´ opinions, which may lead the researcher to obtain valuable data as well. Therefore, a *group* and *semi-structured interview* takes place in this research, which can be found in APPENDIX III.

After discussing the chosen instruments of data collection, in the figure below a summary of the key information is shown:

Technique	***Type***	***Purpose***	***Timing***
Observation	Participant observation	Introduce and observe the variables. Assess participants´ performance.	During and after the sessions
Questionnaire	Supervised self-completion questionnaire	Know participants experiences and opinions.	After the sessions
Interview	Semi-structured interview Group interview	Elicit additional information (e.g. non-verbal language, participants´ debate, etc.).	After the sessions

Table 7. Techniques of data collection

Source: Own elaboration

With these techniques, it is expected that enough information can be collected to carry out a triangulation of data in subsequent sections.

4. ANALYSIS OF THE DATA COLLECTED

This section is aimed at analysing the information obtained in the field work phase, once the data have been compiled. Therefore, this section deals with the data analysis. In order to do so, this section is organised into four subsections; the first three subsections address the three techniques of data collection that have been followed; and the fourth subsection focuses on the data triangulation. It is important to note that, although the initial sample was of 24 learners, one of them missed two of the three sessions, so the analysis is made on the basis of the other 23 participants.

4.1. Participant observation

As pointed out in section 3.5.2.1., in order to carry out the observation throughout the three sessions, two rubrics were designed (see APPENDIX IV and V); the former focuses on the variables of the study, whilst the latter places the emphasis on assessing participants´ linguistic and content performance through the proposed oral CALL activities, which are shown in APPENDIX I. For this reason, this section analyses the data compiled from both rubrics separately, in two different subsections.

4.1.1. Observation of the variables

On the one hand, regarding the first dependent variable (the improvement of oral evaluative activities), it has been observed that almost all the participants perform appropriately the activities with the tools Edpuzzle and Plickers. In other words, they tend to show an acceptable knowledge when they are assessed through listening-based CALL activities. Nonetheless, this tendency seems to be slightly lessened in speaking-based CALL activities, insofar as in the activity with the tool Voki, approximately one out of three participants had difficulties in completing the task. Either way, it can be argued that, in this case study, the first dependent variable has been positively influenced by the independent variable (CALL instruments and activities).

On the other hand, concerning the second dependent variable (the increase of pupils´ motivation in CLIL evaluation processes), it has been seen that all pupils enjoy and like the

CALL activities of evaluation that they carried out, and they seem to show a receptive, participative and active attitude, yet shy learners were not as active in speaking-based activities as they were in listening-based ones. Even so, it can be argued that the second variable has been positively influenced as well, given that pupils´ motivation was high at all times.

4.1.2. Observation of participants´ performance

Firstly, regarding the oral linguistic performance on the part of participants, it has been observed that, roughly speaking, they deal with the listening skill better that with speaking skill; around 85% of them met the oral comprehension descriptors shown in APPENDIX V, whereas around 67% of the participants met the observed oral production descriptors throughout the activities. Nevertheless, in both cases learners required some kind of written support to handle the oral skills; for instance, in the activity with Plickers the questions were asked orally, yet the possible choices were presented in written form.

Besides this, when it comes to understanding and producing orally the contents of the assessed unit, learners seem to have more control of the vocabulary of the unit than of the grammar structures. In both cases the comprehension of messages appears to be more achievable for them than their production, yet the latter was generally positive as well. In this sense, the average linguistic performance in the evaluative activities has been successful, so it can be argued that the oral CALL activities have elicited a favourable linguistic performance from the learners.

And secondly, as for the evaluation criteria considered to assess pupils´ knowledge of the unit (shown in APPENDIX V), it has been observed that the majority of the learners meet the criteria. More concretely, the first evaluation criterion has been met by all the pupils, whereas the second, third and fourth criteria have been met by 19, 18 and 20 learners, respectively. In the activity with Voki, 15 pupils show a clear understanding on the question asked, whereas 8 learners show insufficient answers in terms of content knowledge (see APPENDIX VI to find some examples of participants´ oral production through the tool Voki).

As for the activities with Edpuzzle and Plickers, only 1 of the 23 students fails to reach at least 50% of right answers, whilst most of them get very positive results, as shown in the table below:

Percentage of right answers in Edpuzzle and Plickers	Number of participants achieving each percentage	Average percentage of right answers
38%	1	83%
50%	1	
63%	3	
75%	3	
88%	7	
100%	8	

Table 8. Percentages of pupils´ performance with Edpuzzle and Plickers

Source: Own elaboration

In view of this data, it can be pointed out that the oral CALL activities have been useful for learners to show their knowledge on the CLIL subject.

4.2. Self-completion questionnaire

The participants completed the questionnaire through the tool Plickers, which made the process more appealing for them. This way, when asked if they enjoyed the evaluative activities that they had carried out, 19 of the participants answered that they did; only 2 of them answered that they did not, and two others did not respond, as they were not sure. Therefore, most of the pupils showed a positive motivation towards the oral CALL activities deployed in the sessions.

In addition to this, when asked if they would like to be evaluated in a similar way more often in their CLIL subjects, again most of them agreed; 20 of the respondents think

they would like to be assessed more frequently through CALL activities, and solely 3 of them answered that they would not. Thus, it seems clear that the majority of the participants prefer being evaluated through oral CALL tasks on a regular basis, which confirms that their motivation throughout the sessions were rather high.

In light of the data collected from these two questions, arguably there is clear a relationship between the motivation elicited in the sessions and the participants´ willingness to be assessed though CALL activities more regularly, as shown in the chart below in percentages:

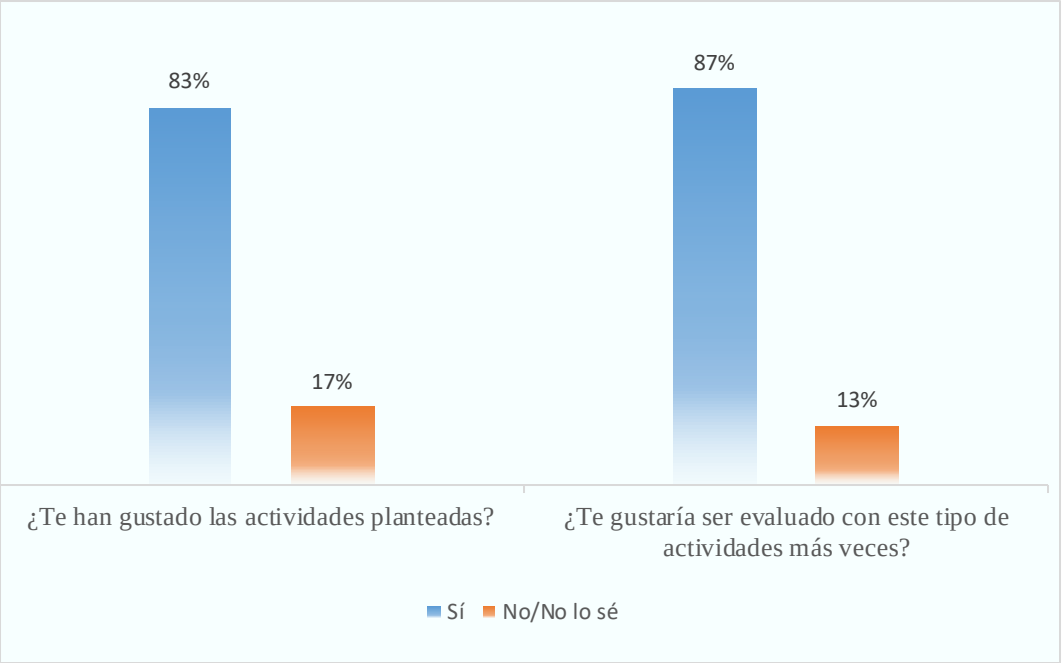

Figure 2. Pupils´ motivation and willingness to be assessed with CALL activities

Source: Own elaboration

Besides this, the respondents were asked about the use of oral skills in the assessment activities. Firstly, they were asked if having to listen in English in the CALL activities had been a problem for them to be assessed; 15 of them responded that it had not, 3 of the participants pointed out that it had been a problem, and 5 of them considered that it had been a handicap sometimes. Consequently, it appears that listening-based CALL evaluation activities prove fruitful to assess the learners in CLIL settings. To put it another way, CALL tools elicit the use of the listening skills as a means of evaluation.

Similarly, when asked if having to speak in English had been a difficulty to be assessed, 16 of the participants answered that it had not been an additional hindrance; 3 of them responded that it had; and 4 of the pupils held that speaking in English to be assessed had been a problem sometimes. Thus, seemingly most participants do not regard speaking-based CALL activities as a further limitation, so it can be argued that CALL tools prompt the speaking skill to be handy in CLIL assessment processes.

Bearing the information from these two questions in mind, it appears that both listening-based and speaking-based CALL activities are suitable for assessing CLIL students, notwithstanding that this fact does not apply to the overwhelming majority of the sample. In this regard, apparently for both oral skills the results from the questionnaire are in keeping, as the graph below shows:

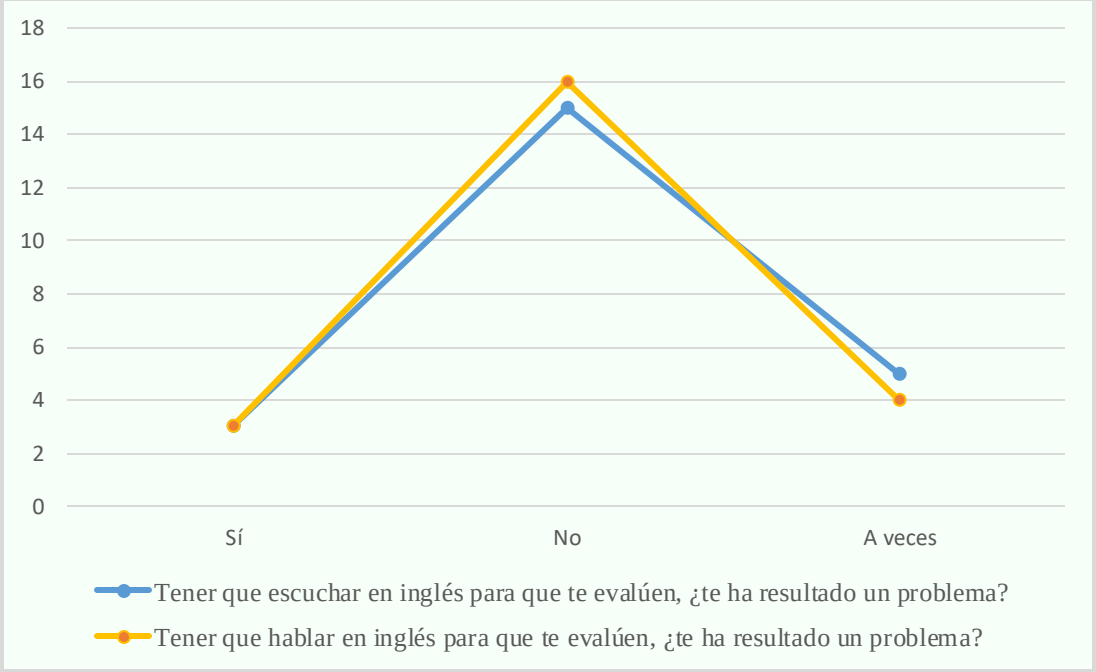

Figure 3. Pupils´ opinion on being assessed through the oral skills

Source: Own elaboration

Lastly, the respondents were asked about the knowledge that the activities had elicited, as well as the usefulness of the computers. As such, when asked if the activities were useful for them to show their knowledge about the CLIL topic, up to 17 of the participants answered that they were; solely 1 student responded that the activities were not useful to show their knowledge; and 5 learners considered that some activities were useful,

but others were not. Hence, the majority of participants think that the CALL activities deployed proved fruitful in order to show their CLIL area understanding.

Apart from this, students were asked if, in their opinion, the use of computers had improved the way in which they had been assessed. In this question, 14 respondents pointed out that computers had improved the assessment process; only 2 of them answered that computers did not mean an enhancement; and up to 7 of the learners were not sure about their usefulness. Therefore, in this question the results are evener than in previous ones, yet it can be argued that most pupils recognise the usefulness of laptops for their evaluation.

In these last questions, the results are similar, so it appears that there is a relation between the perception of the activities being useful to show knowledge, and the perception of the laptops being influencing in the assessment process. In other words, arguably students tend to perceive the activities useful because laptops are involved in them. Likewise, the number of participants who do not find the activities useful to elicit knowledge is akin to the number of participants who do not find the use of computers useful either:

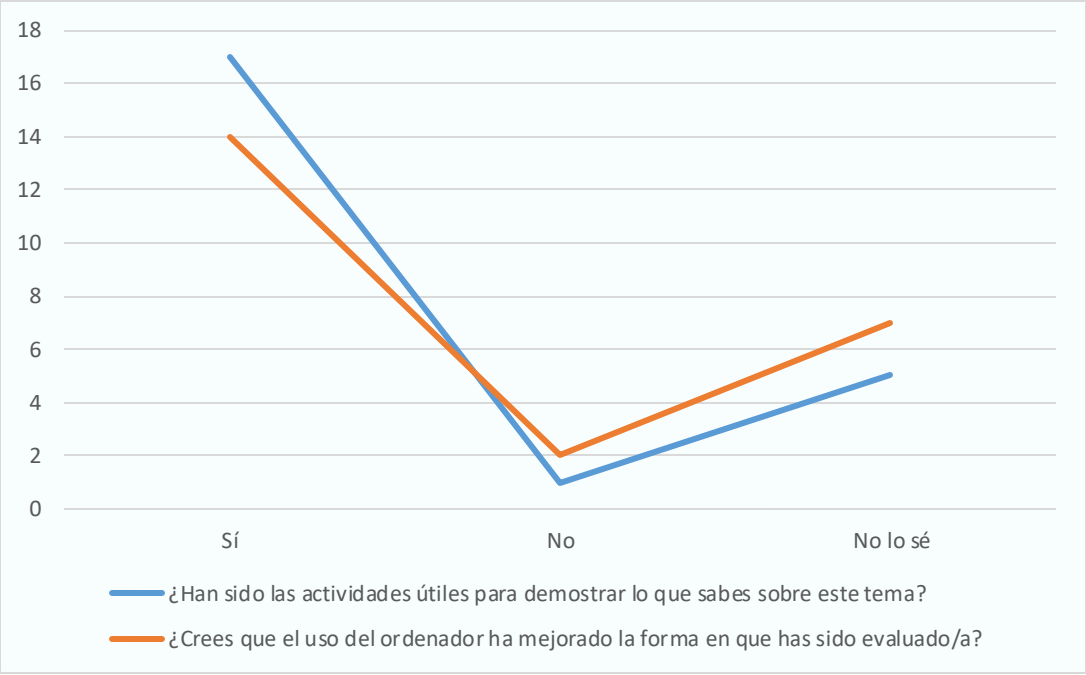

Figure 4. Pupils´ perception of activities and laptops usefulness

Source: Own elaboration

Once the data collected from the questionnaire have been analysed, the following section copes with the data collected in the semi-structured and group interviews.

4.3. Semi-structured interview

Although the interviews were designed to interview up to five respondents at once (as shown in APPENDIX III), they were filled out in groups of four learners, except one group that was composed by three pupils. Roughly speaking, the information gathered through the semi-structured interview reinforces the data collected through the questionnaire, yet some additional information elicited through open questions ought to be highlighted because of its relevance and significance. In APPENDIX VII a transcription of one of the interviews is shown.

The first question, given below, was:

(10)

a. *¿Os han gustado las actividades de evaluación que habéis hecho? ¿Por qué? (Did you like the assessment activities that you have done? Why?)*

This way, participants were asked if they liked the evaluation activities they had done, and more importantly; why. All of them agreed that they had enjoyed the activities, and the reasons they gave are rather similar. In this sense, the most common answer, given by around 80% of the participants, was that they had fun and learnt at the same time, which brings into relief the overall motivation of the students, as well as showing that pupils appreciate the feedback given in CALL activities, as they helped them learn apart from assessing them.

Similarly, 12 of the 23 pupils declared that the activities had helped them "remember things"; this may indicate that the activities proved fruitful in eliciting previous knowledge and, consequently, they were suitable for assessing purposes. As a matter of fact, a couple of pupils (interviewed in different groups) mentioned that the activities were "like taking an exam but easier", and they added that they found the activity with Plickers (listening-based) particularly easy because they were guided by the researcher as they listened to the questions orally; this shows the importance of assessing pupils in while-listening activities, rather than in post-listening tasks.

Another interesting answer that around half of the learners gave in the first question was that the activities was helpful to improve the language. Hence, it seems that they realize and are aware of their language progress in CLIL sessions and, according to their answers, the CALL activities they performed prompted them to reinforce this language. Therefore, in

view of these statements, the deployed CALL activities addressed the dual-focus evaluation that CLIL settings call for, insofar as both language and content were assessed and reinforced alike.

The second question, specified in (11), was:

(11)

a. *¿Habéis podido demostrar todo lo que sabíais del tema? ¿Hay algo que no?*
(Could you show all you knew about the unit? Is there anything that you could not show?

In the second question the respondents were asked if they could show all they knew about the topic through the activities, and if there was something that they could not demonstrate because of the sort of activities. Here most participants, over 85%, responded that they could show their knowledge of the topic, but three of the learners disagreed with this viewpoint. According to these, sometimes they did not have enough time to think the answers, so they were not able to demonstrate their actual understanding of the proposed questions; they referred to the session with Plickers, where all learners had to answer at the same time. With Edpuzzle and Voki, they could work in a more flexible way.

This comments bring into relief that the activities where pupils work individually with their own laptops tend to be more flexible than activities in which the computer is used by the teacher (o researcher), where all learners are expected to answer the questions at the same time. It is important to note, though, that the majority of participants had no significant problems in following the rhythm of the proposed activities.

As for the third question, it was:

(12)

a. *¿Ha sido difícil tener que escuchar y hablar en inglés durante estas actividades de evaluación? ¿Qué os ha costado más?*
(Was it difficult to listen and speak in English throughout these activities of evaluation? What was harder for you?)

Therefore, in the third question pupils were asked if having to speak and listen in English during assessment activities were hard for them. Only 3 of the respondents mentioned that relying on the oral skills as a means of assessment was hard for them, so for the majority of them it was not a meaningful problem. Additionally, when asked which skill had been harder

for them, most learners agreed that it was speaking, although a few pupils considered that listening-based activities had been more complicated for them. Therefore, it appears that oral CALL evaluation tasks do not represent a major problem for learners in CLIL contexts, yet speaking-based CALL activities tend to be more challenging than listening-based tasks.

Lastly, the fourth question was:

(13)

a. *En las asignaturas bilingües, ¿os gustaría que os evaluaran más veces a través de los ordenadores? ¿Por qué?*
(In your bilingual subjects, would you like to be assessed more often through the computers? Why?)

In the last question participants were asked if they would like to be assessed through similar activities more often in their CLIL subjects, and why. All participants answered that they would like, and the reasons why were similar to those given in the first questions. Principally, the main reason, given by 17 of the 23 students, was that they enjoyed the activities and learnt at the same time as they were assessed, and they found the tasks appealing and motivating. In this sense, up to 8 pupils also argued that they understood what they were expected to do at all times as they had more guidance provided by the computer activities.

The table below summarizes the main responses in each question, as well as specifying the number of participants who answered differently in each case.

Semi-structured interview questions	**Pupils´ close answers (SÍ / NO)**		**Pupils´ common open answers and comments**
¿Os han gustado las actividades de evaluación que habéis hecho? ¿Por qué? *Did you like the assessment activities that you have done? Why?*	23	0	Had fun and learnt at the same time. Easy way to take an exam Helpful to remember what they know. Helpful to improve the language.

¿Habéis podido demostrar todo lo que sabíais del tema? ¿Hay algo que no? *Could you show all you knew about the unit? Is there anything that you could not show?*	20	3	Activities were useful to show what they know. Not enough time sometimes.
¿Ha sido difícil tener que escuchar y hablar en inglés durante estas actividades de evaluación? ¿Qué os ha costado más? *Was it difficult to listen and speak in English throughout these activities of evaluation? What was harder for you?*	20	3	Speaking-based activities were more complicated that listening-based ones.
En las asignaturas bilingües, ¿os gustaría que os evaluaran más veces a través de los ordenadores? ¿Por qué? *In your bilingual subjects, would you like to be assessed more often through the computers? Why?*	23	0	Appealing way of being assessed. A motivating way of taking exams.

Table 9. Interview main answers and comments

Source: Own elaboration

As shown in the table, pupils´ motivation was high throughout the sessions, and they seem to realize the usefulness of the activities carried out to be assessed.

4.4. Data triangulation

After analysing individually the information compiled through the three techniques of data collection, in this subsection a data triangulation is carried out in order to compare and highlight the results that the different techniques show. In this sense, three general findings appear to be underpinned by the techniques followed.

Firstly, it has been proved that *pupils show a high motivation in oral CALL assessment activities*. In order to reach this result, some specific indicators have been examined; for instance, it has been observed that pupils showed a clear interest during the activities, as underpinned by the three techniques of data collection. Additionally, learners´ participation was highly active throughout the sessions, as retrieved from the observation and the interviews. And lastly, data from both the questionnaire and the interview indicate that participants were generally willing to be assessed through similar processes on a more regular basis. For these reasons, it can be claimed that participants´ motivation was high in oral CALL assessment sessions.

Secondly, it has been found that *it is feasible to rely on the oral skills as main means of CALL assessment activities*. This can be argued insofar as, roughly speaking, pupils´ performance was successful especially in listening-based CALL activities, as well as in speaking-based CALL activities in a slightly lower degree, as retrieved from the participant observation. Besides this, the three techniques of data collection indicate that neither listening nor speaking in English during CALL activities tended to be a hindrance for leaners. Consequently, it can be concluded that it is realistic to base the CALL assessment activities on the oral skills.

Lastly, the third general finding after the data triangulation is that *CALL activities are applicable to assess orally in CLIL settings*. This can be inferred after examining some specific indicators compiled in the analysis. For instance, it has been observed that pupils could generally show their linguistic level in the oral CALL activities deployed, as underpinned by the observation and the interview data. Likewise, participants could demonstrate their content knowledge in oral CALL activities as well, as all the techniques point out. Moreover, data from both the observation and the interview indicate that learners understood the purpose of oral CALL assessment activities, and they could benefit from the feedback provided by CALL tools. Lastly, it has been deduced that most pupils found laptops

useful in order to be assessed, which is retrieved from the questionnaire and interview data analysis. For these reasons, it can be argued that oral CALL activities can be suitable to assess students in CLIL settings.

The table below summarises the three general findings after the data triangulation, together with the specific indicators that lead to them and the techniques of data collection by which these are underpinned in each case:

General findings	Specific indicators	Techniques of data collection underpinning the indicators*		
		O	Q	I
Pupils show a high motivation in oral CALL assessment activities	Pupils show interest in the activities	**Yes**	**Yes**	**Yes**
	Pupils´ active participation in the sessions	**Yes**	**No**	**Yes**
	Willingness to be assessed similarly more frequently	**No**	**Yes**	**Yes**
It is feasible to rely on the oral skills as main means of CALL assessment activities	Generally, pupils´ performance is successful in listening-based CALL activities	**Yes**	**No**	**No**
	Generally, pupils´ performance is successful in speaking-based CALL activities	**Yes**	**No**	**No**
	Listening/speaking in English during CALL activities is not a hindrance for leaners	**Yes**	**Yes**	**Yes**

CALL activities are applicable to assess orally in CLIL settings	Pupils can generally show their linguistic level in oral CALL activities	**Yes**	**No**	**Yes**
	Pupils can generally demonstrate their content knowledge in oral CALL activities	**Yes**	**Yes**	**Yes**
	Pupils can benefit from the feedback provided by CALL activities	**Yes**	**No**	**Yes**
	Participants can clearly understand the purpose of oral CALL assessment activities	**Yes**	**No**	**Yes**
	Pupils find laptops useful to be assessed orally through CALL activities	**No**	**Yes**	**Yes**
* O = Observation	Q = Questionnaire	I = Interview		

Table 10. Summary of the data triangulation

Source: Own elaboration

It is important to note that, when the techniques of data collection in the table appear as "no" underpinning some indicators, it does not mean that they refuse the validity of the indicators, but it means that in these cases the techniques did not compile sufficient data to prove and underpin them; consequently, the general findings and their specific indicators have not been refuted by the techniques of data collection.

5. DISCUSSION OF THE RESULTS

Once the compiled data have been analysed and triangulated, this section is intended to discuss the results obtained. Firstly, the results are compared to the theoretical framework exposed in section 2 in order to examine the extent to which they coincide with the existing literature, as well as shedding some light on the field of this research. Secondly, emphasis is placed on the investigation hypotheses and questions, in order to discuss if they conform to the results and can be responded. And thirdly, in the last subsection some limitations of the results are pointed out, together with some proposals for further research.

5.1. Comparison of the results with the existing literature

First of all, regarding the first general finding (pupils show a high motivation in oral CALL assessment activities), it seems to agree with most of the literature, as plenty of authors point out the increase of pupils´ motivation that CALL techniques tend to provoke (Pérez Torres, 2004; Araujo, 2013). Nonetheless, these authors tend to refer to the use of CALL in general, whereas this research finds that when it comes to oral-based CALL activities, pupils´ motivation is high as well.

In addition to this, according to the reviewed literature, the reason why CALL activities are motivating is because they present multimedia elements such as images, texts, audios or videos (Pérez Torres, 2004); or because CALL activities contribute to the gamification of the foreign language learning, as they lead learners to regard the activities as "games" (Araujo, 2013). In this sense, this research adds that oral CALL activities are appealing to learners because these activities prompt pupils to regard assessment processes as an enjoyable task, not merely as an examination they are compelled to take in order to be graded.

Concerning the second main finding (it is feasible to rely on the oral skills as main means of CALL assessment activities), arguably no previous literature referring exactly to this conclusion has been encountered. Significant research has been found regarding the development of oral skills in foreign languages (Blasco Mayor, 2009; Talaván, 2010; Slaouti et al., 2013; Stanley, 2013; Vázquez Cano & Martín Monje, 2014); as well as concerning the assessment of the foreign language oral skills (Levy & Gertler, n.d.; Sharples et al., 2005;

Wagner, 2010; Bahrani, 2011; Beaven & Neuhoff, 2012; Slaouti et al., 2013; Stannard & Basiel, 2013; Wu, 2013; Caruso et al., 2017), but not dealing with the role of oral skills as means of evaluation with the support of computers. In relation with this, this research concludes that the oral skills can be treated as the main medium of evaluation through CALL activities.

And thirdly, as for the third main finding (CALL activities are applicable to assess orally in CLIL settings), the situation is similar to the previous finding; no significant research has been found addressing particularly this issue. The literature focuses on the nature and pedagogic implications of CLIL (David Marsh, n .d.; Attard Montalto et al., n. d; Coyle, 1999; Dodge, 2001; Cummins, 2008; Pérez Torres, 2009; Coyle at al., 2010; Dalton-Puffer, 2011; Quartapelle & Schameitat, 2012; Casal, 2016; Nieto-Moreno-de-Diezmas, 2018).

Nevertheless, as claimed by Maggi (2012), astonishingly little has been researched and published regarding CLIL assessment and evaluation, and this situation also applies to the influence of oral CALL activities on CLIL assessment, which is the main focus of this investigation. In relation with this, the present research brings into relief that CALL activities are appropriate to assess CLIL students through their oral skills. More concretely, this research points out that oral CALL activities are valuable in CLIL assessment because they elicit meaningful feedback for students and, more importantly, because they cater for the dual focus of CLIL evaluation, in which both language and content ought to be considered and assessed.

5.2. Revision of the investigation hypotheses and questions after the results

Once the research data and results have been exposed, this subsection discusses if the investigation hypotheses and questions have been met and responded. Firstly, as for the hypotheses (shown in section 3.1.), arguably both of them have been met; the first one because it has been shown that introducing oral CALL strategies in a CLIL classroom can enhance the evaluation processes and make them more appealing for pupils; and the second hypothesis is arguably met as well because it has been concluded that the role of oral skills in assessment activities can be boosted by the use of CALL tools.

Secondly, as to the research questions (shown in section 1. 3), it can be claimed that an answer can now be given to them, once the data have been analysed. Concerning research question 1, namely *Can CALL activities enhance the assessment processes in CLIL contexts?*, the answer is; yes, they can, because CALL activities elicit pupils´ knowledge, provide them with meaningful feedback and increase their motivation in CLIL assessment processes.

As for the research question 2, which is; *Are oral CALL activities suitable to assess learners in CLIL settings?*, the response is; yes, they are. Particularly listening-based CALL activities are suitable to assess CLIL pupils, yet speaking-based ones are too. Either way, some kind of written support is usually required by learners in order to have a visual guidance.

Lastly, regarding the research question 3, namely *Can CALL tools address the dual focus assessment (language and content) that CLIL requires?*, the answer is; yes, they can, since CALL tools have proven fruitful to elicit and assess both pupils´ linguistic competence and content knowledge.

The answers for the research questions have been formulated on the basis of the findings in this investigation. This means that the responses may differ depending on issues such as the approach of the study, the method applied, the investigation nature or the techniques of data collection.

5.3. Limitations of the results and future research

Admittedly, this research contradicts some issues suggested in the literature and fails to give an answer to some noteworthy questions. For example, during the sessions of the field work phase, some shy students had difficulties in completing some of the activities, particularly speaking-based ones. Thus, in this case study the CALL activities could not fully cater for inhibited students as some authors in the literature suggest, such as Araujo (2013), who notes that CALL provides learners with privacy and autonomous self-guidance, which is particularly beneficial for shy students. Maybe this hindrance does not stem from the CALL tools or tasks themselves, but due to the medium through which they were deployed; the oral skills. Either way, this study fails to delve into this issue and suggest a valid response.

Apart from this, in the field work phase it was also encountered that some of the listening-based activities were not flexible enough, as some pupils indicated in the interview. Presumably, individual CALL listening-based activities cater for pupils´ needs more accurately than whole group listening-based activities, given that in the activity with Edpuzzle (where pupils worked individually) no difficulties were reported, whereas in the activity with Plickers (in which learners worked in big group) is where concerns have been brought into relief. Once again, this research fails to address this issue.

For reasons like these, some questions for future research have arisen, such as; 1) To what extent can oral CALL activities suit inhibited learners?; 2) Can oral based CALL activities guarantee a valid assessment in big group processes?; 3) Can oral CALL tasks promote pupils´ self and peer-evaluation?, just to name a few. Hopefully, these questions and others can be delved into and answered in further research.

6. CONCLUSIONS

The main purpose of this research has been to explore the ways in which CALL strategies and instruments can enhance the assessment processes within CLIL contexts through oral skills. In other words, the ultimate outcome in this paper has been to examine the suitability and applicability of oral skills as a means to carry out assessment activities in CLIL contexts, where both language and contents need to be assessed. The reason why this general outcome was set stems from the revision of the existing literature, where no specific research is found concerning this issue; significant inquiries can be encountered regarding the development and assessment of foreign language oral skills through LKT and CALL, as well as concerning the nature of CLIL assessment, but not specifically addressing the use of oral-based CALL activities in CLIL assessment processes, which is the focus of this research.

In view of this lack of inquiry in the field, and having the main purpose established, three investigation questions arise; 1) Can CALL activities enhance the assessment processes in CLIL contexts?; 2) Are oral CALL activities suitable to assess learners in CLIL settings?; and 3) Can CALL tools address the dual focus assessment (language and content) that CLIL requires? Consequently, these questions ushered in two more specific objectives for the investigation; on the one hand, to examine the applicability of CALL within CLIL evaluation

processes; and on the other hand, to explore the feasibility of oral-form assessment activities in CLIL settings.

In turn, two hypotheses were formulated; the first hypothesis was that introducing CALL strategies in a CLIL Primary Education classroom can enhance the evaluation processes and make them more appealing for learners by increasing their motivation; and the second hypothesis was that the feasibility of treating oral skills as a means of CLIL assessment can be boosted by the use of CALL tools.

In order to address the investigation questions and corroborate the hypotheses, a methodology was designed for this research. In this sense, a pre-experimental design was followed, given that this research dispensed with a control group and pre/pro tests, and the results obtained cannot be generalised. In turn, the pre-experimental design was concreted through a case study, in which qualitative and quantitative data were retrieved from a sample of CLIL Primary Education students through three techniques of data collection; participant observation, self-completion questionnaire, and semi-structured group interview.

After analysing and triangulating the data retrieved from the techniques of data collection, three main findings can be highlighted. The first one is that *pupils show a high motivation in oral CALL assessment activities.* This fact can be deduced from more specific indicators such as pupils´ interest throughout the deployed sessions, their high level of interaction and participation, and their willingness to carry out similar assessment activities on a regular basis.

The second noteworthy finding is that *it is feasible to rely on the oral skills as main means of CALL assessment activities*. This finding is claimed on the basis of two specific indicators; the first one is that pupils´ performance is generally successful in listening and speaking-based CALL activities, and the second indicator is that listening and speaking in English is not deemed as a hindrance by learners when they carry out CALL tasks and activities.

The third and last remarkable discovery is that *CALL activities are applicable to assess orally in CLIL settings*. This can be argued because the proposed oral-based CALL activities proved fruitful to elicit learners´ linguistic and content knowledge, apart from providing them with significant feedback, which prompted students to value the usefulness of computers when they are assessed, as well as being aware of the purpose of computer-based assessment processes in CLIL areas.

Bearing these findings in mind, the three investigation questions can be answered affirmatively; 1) CALL activities can enhance the assessment processes in CLIL contexts because they elicit pupils´ knowledge, provide significant feedback and increase their motivation; 2) Oral CALL activities are suitable to assess learners in CLIL settings, notwithstanding that pupils require some kind of written support as well; and 3) CALL tools can address the dual focus assessment that CLIL calls for, as they elicit both learners´ linguistic competence and CLIL area knowledge.

In addition to this, the aforementioned hypotheses are corroborated by the findings of the study, and it can be argued that the specific research objectives have been accomplished; the applicability of CALL within CLIL evaluation processes has been examined, and the feasibility of oral-form assessment activities in CLIL settings has been analysed.

All these aspects considered, this research concludes that oral-based CALL activities can have a positive influence on CLIL assessment processes, given that they are appealing for learners and they adjust to the characteristics of CLIL assessment. Nonetheless, admittedly this study fails to give answer to other relevant questions, and further research is needed to delve into issues such as the extent to which oral CALL assessment activities can cater for inhibited learners; the validity of oral-based CALL assessments in big group activities; or the usefulness of oral CALL activities to promote self and peer-evaluation in CLIL settings.

Lastly, as a personal appraisal of this dissertation, undoubtedly it has been a valuable opportunity to realise and address such an important issue within the field of CALL from different perspectives; both by discussing the existing literature and by collecting meaningful data from a real CLIL classroom. Thus, the possibility to keep working on this issue in subsequent dissertations is welcomed. All these aspects considered, the overall appraisal of the development of this piece of work is positive, as it has clearly been beneficial in terms of academic, professional and even personal skills. For this reason, carrying out this dissertation has been a worthwhile experience, so all the knowledge acquired throughout its development will be taken into deep consideration as a Primary Education teacher.

7. REFERENCES

Araujo, J.C. (2013). Principales avances en el ámbito de la enseñanza de lenguas asistida por ordenador (ELAO). *Ikastorratza. e-Revista de Didáctica 11, retrieved from* *http://http://www.ehu.es/ikastorratza/11_alea/elao.pdf* ISSN: 1988-5911.

Attard Montalto, S., Walter, L., Theodorou, M. & Chrysanthou, K. (no date). The CLIL guidebook. *European Commission, Lifelong Learning Programme.* ISBN 978-88-7164-424-0

Bahrani, T. (2011). Technology as an Assessment Tool in Language Learning. *International Journal of English Linguistics* Vol. 1, No. 2.

Barbero, T. (2012). Assessment tools and practices in CLIL. In Quatarpelle, F. (2012). *Assessment and evaluation in CLIL* (pp. 38-56). European Commission: AECLIL-EACEA. ISBN 978-88-7164-424-0

Beaven, A. & Neuhoff, A. (2012). Assessing oral proficiency for intercultural professional communication: the CEFcult project. EUROCALL.

Blasco Mayor, M.J. (2009). CALL-enhanced L2 Listening Skills – Aiming for Automatization in a Multimedia Environment. *Indian journal of applied linguistics* Vol. 35, no. 1, jan-jun 2009.

Caruso, M., Colombi, A., & Tebbit, S. (2017). Teaching how to Listen. Blended Learning for the Development and Assessment of Listening Skills in a Second Language. *Journal of University Teaching & Learning Practice, 14(1).* Available at: http://ro.uow.edu.au/jutlp/vol14/iss1/14

Casal, S. (2016). Cooperative Assessment for Learning in CLIL Contexts. *Estudios sobre educación*. Vol. 31 / 2016 / 139-157.

Castilla-La Mancha Decree 54/2014, July 10th, which establishes the management and the curriculum in Primary Education in the autonomous Community of Castilla-La Mancha.

Chaka, C. (2009). Portable Handheld Language Learning: From CALL MALL to PALL. *Information Science Reference.*

Chapelle, C. & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge, England: Cambridge University Press.

Chapelle, C. (2007). Technology and Second Language Acquisition. *Annual Review of Applied Linguistics* 27 (2007): 98, doi: 10.1017/S0267190508070050.

Cohen, L., Manion, L. y Morrison, K. (2011). *Research Methods in Education*, 7th edition. Londres y Nueva York: Routledge.

Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with New Descriptors.* Language Policy Programme. Education Policy Division. Education Department.

Coyle, D. (1999). The next stage; is there a Future for the Present? The legacy of the communicative approach. *Francophonie19*, 13-16.

Coyle, D. (no date). Content and Language Integrated Learning; Motivating Learners and Teachers. *University of Nottingham*.

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

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In Street, B. & Hornberger, N. H. (Eds.). *Encyclopedia of language and education*: Vol. 2. Literacy (2nd ed.) (pp. 71-83). New York: Springer.

Dalton-Puffer, C. (2011). *Content-and-Language Integrated Learning: From Practice to Principles?* Cambridge University Press.

Dodge, B. (2001). Focus: Five rules for writing a great WebQuest. *Learning & Leading with Technology* 28(8), 6-9.

Farhan, A. (2007). Cooperative vs. individual learning of oral skills in a CALL environment. *Poznan Studies in Contemporary Linguistics*. DOI: 10.2478/v10010-007-0001-8

Gilbert, I. (2018). Developing Oral Communication Competence in a Second Language: The Communicative Approach. *World Academy of Science, Engineering and Technology, International Science Index, Cognitive and Language Sciences*, 12(8), 1172

Gray, D. E. (2004). *Doing Research in the Real World*. London: SAGE Publications.

Guadamillas, V. & Alcaraz, G. (2017). Bilingual Legislation. Analysing the Legal Framework for Primary Education in Spain. *Multiárea. Revista de didáctica*. Núm. 9 (2017): 82-103

Hughes, R. (2017). *Teaching and Researching Speaking*: Third Edition (Applied Linguistics in Action). Essex: Pearson Education.

Levy, T., & Gertler, H. (no date). Harnessing Technology to Assess Oral Communication in Business English. *Teaching English with Technology*, 15(4), pp. 52-59, http://www.tewtjournal.org

Maggi, F. (2012). Evaluation in CLIL. In Quatarpelle, F. (2012). *Assessment and evaluation in CLIL* (pp. 57-74). European Commission: AECLIL- EACEA.

Marsh, D. (no date). Using languages to learn and learning to use languages. An Introduction to CLIL for Parents and Young People. *TIE-CLIL*.

Moreno, C. (2010). Effective teachers –Professional and personal skills, en ENSAYOS, Revista de la Facultad de Educación de Albacete, Nº 24. Retrieved from http://www.uclm.es/ab/educacion/ensayos

Nieto-Moreno-de-Diezmas, E. (2018). Acquisition of reading comprehension in L1 in bilingual programmes of Primary Education. A comparative study. *Ocnos*, 17 (1), 43-54. doi: http://dx.doi.org/10.18239/ocnos_2018.17.1.1471

Nunan, D. (2012). *Research Methods in Language Learning*. Cambridge University Press.

Palalas, A., & Hoven, D. (2016). Emerging pedagogies for MALL. Chapter 2 in Palalas, A., & Ally, M. (2016). *The International Handbook of Mobile-Assisted Language Learning*. Beijing: China University Press.

Pegrum, M. (2014). *Mobile Learning; Languages, literacies and cultures.* Basingstoke. Hampshire: Palgrave MacMillan.

Pérez Torres, I. (2004). Diseño de WebQuests para la enseñanza-aprendizaje del inglés como lengua extranjera. Tesis doctoral. Universidad de Granada. Granada.

Pérez Torres, I. (2009). Apuntes sobre los principios y características de la metodología AICLE. In Pavón, V. & Ávila, J. (eds.), *Aplicaciones didácticas para la enseñanza integrada de lengua y contenidos.* Sevilla: Consejería de Educación de la Junta de Andalucía-Universidad de Córdoba.171-180.

Pihkala-Posti, L. (2014). Innovative training of oral communication: Berlin Kompass. In Jager, S., Bradley, L., Meima, E. J. & Thouësny, S. (Eds). CALL Design: Principles and Practice; Proceedings of the 2014. EUROCALL Conference, Groningen, The Netherlands (pp. 312-317). Dublin: Research-publishing.net. doi:10.14705/rpnet.2014.000237

Pim, C. (2013). Emerging technologies, emerging minds: digital innovations within the primary sector. Chapter 1 in Motteram, G. (2013). *Innovations in learning technologies for English language teaching*. Innovation Series; British Council.

Pino, M., Rodríguez, B. & Mañana, J. (no date). Métodos y Técnicas de Recogida y Análisis de Datos. Universidad Nacional de Educación a Distancia (UNED), Facultad de Filología.

Quartapelle, F. & Schameitat, B. (2012). Teaching and Learning with CLIL. In Quatarpelle. F. (2012). *Assessment and evaluation in CLIL* (pp. 29-37). European Commission: AECLIL- EACEA. ISBN 978-88-7164-424-0

Reig, D. (2012). Disonancia cognitiva y apropiación de las TIC. *Revista TELOS (Cuadernos de Comunicación e Innovación)* | ISSN: 0213-084X.

Robson, C. (2002). *Real World Research*. Oxford: Blackwell Publishing.

Royal Decree 126/2014, February 28th, which establishes the Basic Curriculum for Primary Education in Spain.

Sharples, M., Taylor, J. & Vavoula, G. (2005). Towards a Theory of Mobile Learning. *University of Birmingham. United Kingdom*.

Slaouti, D., Onat-Stelma, Z., & Motteram, G. (2013). Technology and adult language teaching. Chapter 3 in Motteran, G. (2013). *Innovations in learning technologies for English language teaching*. Innovation series. British Council.

Stanley, G. (2013). Integrating technology into secondary English language teaching. Chapter 2 in Motteran, G. (2013). *Innovations in learning technologies for English language teaching*. Innovation series. British Council.

Stannard, R., & Basiel, A. (2013). A practice-based exploration of technology enhanced assessment for English language teaching. Chapter 6 in Motteran, G. (2013).

Innovations in learning technologies for English language teaching. Innovation series. British Council.

Suvorov, R., & Hegelheimer, V. (2014). Computer-Assisted Language Testing. Chapter in Kunnan, A.J. (2014). *The Companion to Language Assessment*. First Edition. John Wiley & Sons, Inc. DOI: 10.1002/9781118411360.wbcla083

Talaván, N. (2010). Subtitling as a task and subtitles as support: pedagogical applications. Seen in Díaz, J., Matamala, A. & Neves, J. (2010). *New insights into audiovisual translation and media accessibility*. Amsterdam, New York; Rodopi.

Vázquez Cano, E. & Martín Monje, E. (2014). *Nuevas Tendencias en la Elaboración y Utilización de Materiales Digitales para la Enseñanza de Lenguas*. Mcgraw-HILL / Interamericana de España: Madrid.

Vygotsky, L. S. (1978). *Mind in Society. The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.

Wagner, E. (2010). Test-takers' interaction with an L2 video listening test. *Temple University, College of Education*. Available online at www.sciencedirect.com

Wu, Y, (2013). Conversation Analysis - A Discourse Approach to Teaching Oral English Skills. *International Education Studies*; Vol. 6, No. 5.

APPENDIX I. CALL instruments and activities used to assess CLIL pupils

FIRST SESSION: EDPUZZLE

Edpuzzle is an example of digital video use. This tool can edit videos so learners watch and listen to them as they answer content questions meanwhile.	
Groupings	Individual work
Materials	Laptops and headphones.
Description of the activity	Pupils watch and listen to this video on Edpuzzle, where they answer some questions after hearing the content; https://edpuzzle.com/media/5b07f16f37a2794052b2dd48.
Researcher´s role	The researcher explains the students what they are expected to do and checks their answers after the activity.
Skills involved	Mainly listening, but also speaking and reading.
Evaluation criteria	1. Conocer las propiedades de la materia y los estados de la misma, así como el paso de unos a otros.
Learning standards	1.1. Estudia y clasifica algunos materiales por sus propiedades (dureza, solubilidad, estado de agregación, conductividad térmica). 1.2. Identifica los cambios de estado y su reversibilidad.

Table 11. Activity to assess with Edpuzzle

Source: Own elaboration

SECOND SESSION: PLICKERS

Plickers is a tool to create group quizzes in a motivating and appealing way for learners. It consists in handing out QR codes to pupils, which they will orient towards one of its four sides depending on the answer they intend to give, which are scanned through a mobile phone.

Groupings	Whole group
Materials	A computer, a projector, a mobile phone and cards with QR codes.
Description of the activity	The researcher asks questions aloud, so that the pupils choose an answer from three possible choices, which they show by moving their QR codes towards one side or another.
Researcher´s role	In this activity, the researcher dictates the questions so that the learners show their answers, in such a way that they can read the possible answers, but they cannot read the questions, they only hear them from the researcher. Then, the researcher scans the QR codes with a mobile phone. The questions dictated and the right answers are the following: *Question 1:* What is the mass of an object? Right answer: It´s amount of matter. *Question 2:* What is the volume of an object? Right answer: The space it occupies. *Question 3*: What is the density of an object? Right answer: The relationship between mass and volume. *Question 4*: What are pure substances? Right answer: An only type of matter that can´t be separated. *Question 5*: In an alloy, there is always.... Right answer: Metal.

	Question 6: Which of the following is a chemical change? Right answer: Oxidation and Combustion. *Question 7:* What is solidification? Right answer: When a liquid becomes a solid *Question 8*: What is sublimation? Right answer: When a solid becomes a gas
Skills involved	Primarily listening
Evaluation criteria	2. Identificar la densidad como la magnitud que relaciona masa y volumen y relacionarla con la flotabilidad, así como reconocer y proponer distintos métodos de cálculo de la densidad de un cuerpo. 3. Diferenciar sustancias puras de mezclas y diseñar estrategias para separar distintos tipos de mezclas. 4. Identificar una combustión, oxidación y fermentación.
Learning standards	2.1. Conoce y utiliza diferentes procedimientos sencillos para la medida de la masa y el volumen de un cuerpo. 2.2. Planifica y realiza experiencias con el fin de averiguar la densidad de distintos cuerpos. 3.1. Identifica y diferencia sustancias puras de mezclas. 3.2. Realiza sencillas experiencias para separar los componentes de una mezcla mediante: destilación, filtración, evaporación o disolución, comunicando de forma oral y escrita el proceso seguido y el resultado obtenido. 4.1. Conoce las principales características de las reacciones químicas: combustión, oxidación y fermentación.

Table 12. Activity to assess with Plickers

Source: Own elaboration

THIRD SESSION: VOKI

This tool lets students create an avatar and give it a voice, enabling the teacher to assess their oral expression skill as well as the contents that pupils have their avatars reproduce.	
Groupings	Individual work
Materials	Laptops.
Description of the activity	Students read about the relationship between density and floatability of objects, in order to answer this question: *Why do ships not sink? Explain it*. Firstly they can jot down the answer in written form, and finally learners record their answers with the tool Voki.
Researcher´s role	The researcher explains the activity and observes how it develops, paying attention to the variables established. When the activity is finished, the researcher checks if the evaluation criteria and learning standards have been achieved by the participants.
Skills involved	Mainly speaking.
Evaluation criteria	2. Identificar la densidad como la magnitud que relaciona masa y volumen y relacionarla con la flotabilidad, así como reconocer y proponer distintos métodos de cálculo de la densidad de un cuerpo.
Learning standards	2.2. Planifica y realiza experiencias con el fin de averiguar la densidad de distintos cuerpos. 2.3. Identifica las principales características de la flotabilidad de determinados cuerpos en un medio líquido y la usa para explicar algún fenómeno físico observable en términos de diferencias de densidad.

Table 13. Activity to assess with Voki

Source: Own elaboration

APPENDIX II. Questionnaire

Questionnaire questions (Presented through Plickers)
¿Te han gustado las actividades planteadas?
A) Sí B) No C) Algunas sí, otras no
¿Han sido las actividades útiles para demostrar lo que sabes sobre este tema?
A) Sí B) No C) Algunas sí, otras no
Tener que escuchar en inglés para que te evalúen, ¿te ha resultado un problema?
A) Sí B) No C) A veces
Tener que hablar en inglés para que te evalúen, ¿te ha resultado un problema?
A) Sí B) No C) A veces
¿Crees que el uso del ordenador ha mejorado la forma en que has sido evaluado/a?
A) Sí B) No C) No lo sé
¿Te gustaría ser evaluado con este tipo de actividades más veces en tus asignaturas bilingües?
A) Sí B) No C) No lo sé

Table 14. Questions asked in the questionnaire

Source: Own elaboration

APPENDIX III. Semi-structured group interview

Semi-Structured Interview Questions

¿Os han gustado las actividades de evaluación que habéis hecho? ¿Por qué?

Participant 1		Participant 2		Participant 3		Participant 4		Participant 5	
Sí	No	Sí	No	Sí	No	Sí	No	Sí	No
Respuestas largas									

¿Habéis podido demostrar todo lo que sabíais del tema? ¿Hay algo que no?

Participant 1		Participant 2		Participant 3		Participant 4		Participant 5	
Sí	No	Sí	No	Sí	No	Sí	No	Sí	No
Respuestas largas									

¿Ha sido difícil tener que escuchar y hablar en inglés durante estas actividades de evaluación? ¿Qué os ha costado más?									
Participant 1		Participant 2		Participant 3		Participant 4		Participant 5	
Sí	No	Sí	No	Sí	No	Sí	No	Sí	No
Respuestas largas									

En las asignaturas bilingües, ¿os gustaría que os evaluaran más veces a través de los ordenadores? ¿Por qué?									
Participant 1		Participant 2		Participant 3		Participant 4		Participant 5	
Sí	No	Sí	No	Sí	No	Sí	No	Sí	No
Respuestas largas									

Table 15. Questions in the semi-structured interview

Source: Own elaboration

APPENDIX IV. Rubric used to observe the variables in the sessions

Independent variable	Dependent variables	Specific descriptors	Degrees		
			Most pupils	Some pupils	Few pupils
CALL instruments and activities.	The improvement of oral evaluative activities.	Pupils do well in listening-based activities			
		Pupils do well in speaking-based activities			
	The increase of pupils´ motivation in CLIL evaluation processes.	Pupils are motivated in the evaluation CALL activities			
		Pupils are active when being assessed			

Additional observations	

Table 16. Rubric to observe the variables in the sessions

Source: Own elaboration

APPENDIX V. Oral linguistic indicators and evaluation criteria observed

Oral Comprehension and Expression Indicators	Linguistic Indicators	Number of participants that meet the linguistic indicators (*to be marked by the researcher*)											
	Can understand/produce general oral messages throughout the sessions	**Understand**	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
		Produce	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	Can understand/produce the key vocabulary[8] of the unit orally	**Understand**	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
		Produce	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	Can understand/produce the key grammar structures[9] of the unit orally	**Understand**	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
		Produce	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	Additional observations													

Table 17. Oral linguistic indicators observed in the sessions

Source: Own elaboration

[8] The key vocabulary is: matter, atoms, mass, volumen, density, to sink, to float, pure substances, heterogeneous/homogeneous mixtures, alloy, oxidation, combustion, melting, solidification, vaporization, condensation, sublimation, reverse sublimation.

[9] The key grammar structures are: be made up of, general/specific properties are, be measured in, be the relationship between, there are, when (something) becomes a.

Evaluation criteria Castilla-La Mancha Decree 54/2014, July 10th (Annex I)	Evaluation criteria	Session in which they are assessed	Number of participants that meet the evaluation criteria (*to be marked by the researcher*)											
	1. Conocer las propiedades de la materia y los estados de la misma, así como el paso de unos a otros.	1st Session (Edpuzzle)	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	2. Identificar la densidad como la magnitud que relaciona masa y volumen y relacionarla con la flotabilidad.	3rd Session (Voki)	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	3. Diferenciar sustancias puras de mezclas y diseñar estrategias para separar distintos tipos de mezclas.	2nd Session (Plickers)	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	4. Identificar una combustión, oxidación y fermentación.	2nd Session (Plickers)	1	2	3	4	5	6	7	8	9	10	11	12
			13	14	15	16	17	18	19	20	21	22	23	24
	Additional observations													

Table 18. Evaluation criteria observed in the sessions

Source: Own elaboration

APPENDIX VI. Examples of participants´ oral production with the tool Voki

Example 1	Example 2
	
Link: http://tinyurl.com/y6v2yzfq	Link: http://tinyurl.com/yb745wpf

Figure 5. Examples of participants´ oral productions with Voki

Source: Own elaboration

APPENDIX VII. Transcription of one of the semi-structured interviews

	Original conversation		Translation into English
Entrevistador	¿Os han gustado las actividades de evaluación que habéis hecho?	Interviewer	Did you like the assessment activities that you have done?
Participante 1	Sí	Participant 1	Yes
Participante 2	Sí	Participant 2	Yes
Participante 3	Sí, mucho	Participant 3	Yes, a lot
Participante 4	A mí también	Participant 4	Me too
Entrevistador	¿Por qué?	Interviewer	Why?
Participante 3	Nos ayudan a recordar	Participant 3	They help us to remember
Entrevistador	Recordar, ¿el qué?	Interviewer	Remember, what?
Participante 3	Lo que sabíamos del tema	Participant 3	What we knew about the unit
Participante 1	Es verdad…	Participant 1	That´s right…

Participante 2	Y han sido muy divertidas, nos has facilitado mucho el trabajo	Participant 2	And they were so fun, they made our work much easier
Entrevistador	¿Todos pensáis eso?	Interviewer	Do you all think so?
(Todos)	Sí	(All)	Yes
Entrevistador	Vale, entonces… ¿habéis podido demostrar todo lo que sabíais del tema?	Interviewer	Ok, so... could you show all you knew about the unit?
(Todos)	Sí	(All)	Yes
Entrevistador	¿Hay algo que no hayáis podido demostrar en las actividades?	Interviewer	Is there anything that you could not show in the activities?
Participantes 1, 3 y 4	No	Participants 1, 3 and 4	No
Participante 2	Bueno, en una actividad a mí no me dio tiempo de pensar demasiado…	Participant 2	Well, in an activity I didn´t have time to think too much...
Entrevistador	¿En cuál?	Interviewer	In which one?
Participante 2	En la de las tarjetas (refiriéndose a Plickers)	Participant 2	In the one with the cards (referring to Plickers)

Entrevistador	Vale… por otro lado, ¿Ha sido difícil tener que escuchar y hablar en inglés durante estas actividades de evaluación?	Interviewer	Ok… on another note, was it difficult to listen and speak in English throughout these activities of evaluation?
Participante 1	No	Participant 1	No
Participante 2	Un poco sí	Participant 2	Yes, a bit
Participante 3	A veces	Participant 3	Some times
Participante 4	Para mí no	Participant 4	For me it wasn´t
Entrevistador	¿Qué os ha costado más; escuchar o hablar en inglés?	Interviewer	What was harder for you; listening or speaking in English?
Participante 1	Hablar cuesta más	Participant 1	Speaking is harder
Participante 2	Hablar	Participant 2	Speaking
Participante 3	Sí, hablar	Participant 3	Yes, speaking
Participante 4	Creo que hablar, sí	Participant 4	I think that speaking, yes

Entrevistador	Vale, por último… en las asignaturas bilingües, ¿os gustaría que os evaluaran más veces a través de los ordenadores?	Interviewer	Ok, and finally… in your bilingual subjects, would you like to be assessed more often through the computers?
(Todos)	¡Sí!	(All)	Yes!
Entrevistador	¿Por qué?	Interviewer	Why?
Participante 2	Porque es más divertido	Participant 2	Because it´s more enjoyable
Participante 1	Y nos gusta más, es como si no estuviésemos haciendo un examen	Participant 1	And we like it more, it´s as if we weren´t taking an exam
Participante 4	Es verdad	Participant 4	That´s right
Participante 3	Es verdad, es más fácil que hacer un examen	Participant 3	That´s right, it´s easier than taking an exam
Entrevistador	Vale, ¡muchas gracias por vuestra participación!	Interviewer	Fine, thanks a lot for you participation!
(Todos)	¡De nada!	(All)	Welcome!

Table 19. Transcription of a semi-structured interview

Source: Own elaboration