Máster Universitario en Estudios Ingleses Avanzados (Lingüística Cognitiva/Literatura) y Educación Bilingüe

TRABAJO FIN DE MÁSTER

Assessing academic language proficiency in CLIL primary education

Evaluación de la competencia lingüística académica en la educación primaria con enfoque AICLE

Blanca Córdoba Conde

Curso académico: 2022/2023

Abstract

In content and language integrated learning contexts, it is accepted that quality education is determined by an explicit focus on the language necessary for the meaning-making of disciplinary content. Therefore, there is a growing interest on providing insight into students' academic linguistic resources and how these influence their acquisition and verbalisation of conceptual knowledge. The examination of this process will allow for practitioners to be informed about language teaching and assessment needs in bilingual education. Thus, the main goal of this research was to assess learners' academic linguistic performance through their command of cognitive-discourse functions and features of genres. The specific objectives were to analyse the linguistic elements students presented when engaging in learning and to determine whether there were significant differences in the genres and skills observed. To this end, two specific tests were designed on oral and written production and reception, plus a focus-on-form activity, which dealt with four academic genres (description, explanation, narration and argumentation) in relation to topics from two primary education subjects. Results showed that narration and argumentation were the most challenging genres for students and that their writing skill was insufficiently developed. The main conclusions were that their academic language proficiency could be improved, and that the gap between oral and written competence should be addressed.

Keywords: academic language, language assessment, bilingual education, genre-based pedagogy

Resumen

En contextos de aprendizaje integrado de contenidos y lenguas extranjeras, se acepta que una enseñanza bilingüe de calidad depende de la atención explícita que se preste al lenguaje académico necesario para la construcción del significado de los distintos contenidos disciplinares. Por tanto, existe un creciente interés en conocer los recursos lingüísticos académicos de los/as estudiantes y cómo influyen en su adquisición y verbalización del conocimiento conceptual. El estudio de este proceso permitirá a los profesionales estar informados sobre las necesidades de enseñanza y evaluación de la lengua en la educación bilingüe. Así pues, el objetivo principal de esta investigación fue examinar el rendimiento lingüístico académico del alumnado mediante su dominio de

las funciones y características cognitivo-discursivas de los géneros textuales. Los objetivos específicos fueron examinar los elementos lingüísticos que los/as estudiantes presentaban para participar en el aprendizaje y determinar si existían diferencias significativas en los géneros y destrezas comunicativas. Para ello, se diseñaron dos pruebas de producción y recepción oral y escrita, además de una actividad centrada en la forma lingüística, en las que se abordaron cuatro géneros académicos (descripción, explicación, narración y argumentación) en relación con temas de dos asignaturas de educación primaria. En los resultados se observó que la narración y la argumentación fueron los géneros más desafiantes para el alumnado, y que su habilidad de escritura no estaba suficientemente desarrollada. Las principales conclusiones fueron que su competencia lingüística académica podría mejorarse y que habría que abordar la brecha existente entre las destrezas oral y escrita.

Palabras clave: lenguaje académico, evaluación del lenguaje, educación bilingüe, pedagogía de géneros textuales

Index

1. Introduction

Our globalised world demands increasingly multilingual and multicultural societies that will be able to face together the challenges that may arise in the future (Marsh, 2013; Pérez Cañado, 2021). Naturally, these requirements had a major impact at all levels of education (Marsh, 2006), giving rise to a series of European bilingual programs that are aimed at overcoming linguistic barriers so as to become a more cohesive, competitive and knowledge-based economy (Marsh, 2002, 2006; Mehisto et al., 2008). In the mid-1990s, Content and Language Integrated Learning (CLIL) was coined as an "umbrella term" (Mehisto et al., 2008, p. 12) to embrace many of these educational strategies that incorporate a "dual focused approach" (Coyle et al., 2010, p. 1) in which the goal is neither the language nor the content, but a combination of both. Due to the multiple benefits that CLIL reports alongside linguistic progression in terms of content learning, cognitive development, and at a social, cultural and pragmatic level (Pavón & Ellison, 2013; Pérez Cañado, 2013, 2021), together with the fact that it allows for flexibility in diverse contexts and circumstances (Marsh, 2002), CLIL has gradually embedded itself in mainstream education from pre-primary level to higher institutions, no longer being "the prerogative for the academic elite" (Pérez Cañado, 2020, p. 6). Hence, since its conceiving, CLIL has undoubtedly been "an escalating phenomenon" (Oxbrow, 2018, p. 137).

Its widespread popularity has run parallel to an increasing concern about the role of the language in the learning process, not just by researchers but also by institutions such as the Council of Europe (Council of Europe, 2014). As Vollmer (2006) noted, language education does not only happen in linguistic subjects, but pervades every academic discipline. Consequently, through the concept Languages Across the Curriculum (LAC), he raised awareness about the necessity of giving due attention to the different languages of schooling. Since then, numerous contributions have been made and today, language is considered a key element in providing equitable and quality teaching according to the Quality Education goal of the United Nations 2030 Agenda for Sustainable Development (UNESCO, 2016).

This claim is based on the intrinsic relationship that bonds together knowledge and language. In this sense, using the language for learning in academic contexts should not be regarded as a mere means of communication (Rolstad, 2005), but as a cornerstone for

meaning making (Vollmer, 2006), since it will determine students' comprehension, expression and acquisition of new information and skills in several subjects and for different communicative purposes (Bailey & Butler, 2003). These cognitive operations, therefore, require competence in what has been known as academic language (Scarcella, 2003), which represents a broad notion that encompasses the thinking processes, lexico-grammatical features and discourse aspects involved in the conveyance of knowledge in the different school subjects, concretely known as disciplinary languages (Beacco et al., 2015).

Thus, academic language proficiency is a fundamental requirement for learners to be able to develop their subject-specific literacy, which no longer refers to the linguistic accuracy involved in reading and writing, but rather to the ability "to become knowledgeable in a field of study, to get acquainted with its thinking and language conventions and to identify the contribution of the subject to society" (Beacco et al., 2015, p. 17). Furthermore, the concept of pluriliteracies has gained ground in bilingual educational contexts such as CLIL to emphasise the "literacy practices in sociocultural contexts, the hybridity of literacy practices afforded by new technologies and the increasing interrelationship of semiotic systems" (García et al., 2007, p. 215), since students are using more than one language to learn (Meyer et al., 2015).

Nonetheless, addressing the linguistic dimension of knowledge construction seems to pose a serious challenge for diverse educational institutions (Beacco et al., 2015), as there exists a vague understanding of the linguistic character of learning disciplinary content (Meyer et al., 2015). Moreover, due to the fact that academic language is found halfway between content and language, many CLIL content teachers consider themselves to be experts in the subject matter and they are not aware of the language used in their areas (Dalton-Puffer, 2016; Leontjev & deBoer, 2020); and on the other hand, subject-specific literacies in the foreign language classroom seem to be regarded as off topic (Meyer et al., 2015), as these lessons are usually more oriented towards a general language proficiency. Hence, academic language has traditionally been considered to evolve naturally through the study of school disciplines (Lorenzo, 2013) and it is not usually tackled at any level or context (Coyle, 2015).

There are several serious consequences of underestimating or neglecting students' academic language, such as an uncompleted literacy (Lorenzo & Rodríguez, 2014) or the belief that they present some kind of learning difficulty (Pavón, 2018).

Consequently, the interconnection between the achievement of educational objectives and academic language is undeniable: students' progression in an academic discipline will be hindered unless their subject-specific literacy is developed, which can only be achieved through a focus on their use of academic language (Lorenzo, 2013). That is why it cannot be rendered "an invisible component" (Llinares et al., 2012, p. 284) when assessing students' subject-specific knowledge, but it must be explicitly addressed and supported. However, CLIL practitioners seem to lack suitable resources to implement such approach (Meyer et al., 2015), which may jeopardise the quality of CLIL. After all, as Morton (2020b) puts it, if they are not clear about "the relative balance and roles of content, literacy and language objectives when planning and teaching" (p. 8), it is very unlikely that students' learning progression will be adequately assessed.

The idea that emerges in light of the above is that it is essential to determine what academic discourse features learners need in the different school disciplines in order to develop their full potential, guarantee their academic success and thus overcome the risks that may endanger the efficacy of CLIL (Beacco et al., 2015). Although there is some research on the importance of academic language in CLIL, most of it is mainly theoretical. In practice, some studies have analysed the discourse strategies used by teachers when co-constructing knowledge with students (e.g., Dalton-Puffer, 2007; Pérez Costa & Pavón, 2019), but there is a paucity of studies directly aimed at observing whether learners present adequate academic language resources for dealing with cognitively demanding content, as noted by Meyer et al. (2015), Pérez Cañado (2020) and San Isidro (2019).

This represents the research gap that this study will try to address by exploring the level of students' academic language proficiency in all communicative skills, in several textual genres, and in different subjects, using specifically designed tools. These tools constitute assessment instruments with pre-defined criteria created for CLIL instructors to observe their pupils' academic language and content knowledge, which have not been provided to date. Thus, it will contribute to the traditional inadequacies of language assessment for learning in CLIL. In particular, the present project is a case study on the academic language competence of primary education students from two schools in Cordoba, with the general objective of examining and describing participants' command of academic discourse features and functions. The results will show where students stand in terms of academic language proficiency and what gaps seem to require explicit

addressing in this context, which could be used to improve the teaching and learning process in CLIL by informing the different agents involved: educational authorities, school managers, teachers and students.

In what follows, firstly, a review of the theoretical background and research contributions in this field will be provided so as to ground this work on empirical principles. This review will shed light on how to address the academic dimension of language from a systemic-functional perspective, where lexico-grammatical features are considered according to the cognitive-discourse functions that the textual genres present in different subjects. As suggested before, this approach will develop students' subject-specific literacy, and thus their pluriliteracies. This idea will therefore be applied to the research method, where specific objectives will be defined and the context and participants, the instruments and data gathering process will be described. The instruments used in this study were two tests that deal with the academic language of four genres (description, explanation, narration and argumentation) according to the curricular content of two bilingual subjects (Natural and Social Sciences). Then, results will be presented according to the specific objectives, where the quality of students' answers is quantified to determine the genres and skills where they present the greatest problems, in this case, narration and argumentation, and writing competence. After that, the results will be discussed through the description of students' academic linguistic choices and in light of previous investigations. Finally, conclusions will be drawn, based on which pedagogical recommendations will be suggested in order to address the findings, and limitations and future research will be considered.

2. Theoretical background

This chapter will present the conceptual foundations on which the present research on academic language in CLIL contexts is based, namely the role of language in learning, what academic language entails and how it has been considered in bilingual education. The last section will be aimed at reviewing some contributions to the field of language-related assessment in CLIL and its current approaches.

2.1. Origins and underpinnings of academic language

One of the main concerns in CLIL is the overcoming of the traditional separation between content and language in order to achieve the desired integration (Coffin, 2017; Llinares & McCabe, 2020). Thus, it is essential to explore and understand the connection between language and content learning, and how communication mediates in this interaction. In this sense, there are two key tenets: Vygotsky's sociocultural theory of learning and Halliday's systemic functional linguistics, since they both consider language and cognitive development as mutually interdependent "social processes" (Llinares et al., 2012, p. 11).

Vygotsky (1986) argues that it is "the social means of thought, that is, language" (p. 94) that enables learning, and therefore intellectual growth, to occur. A child's conceptual knowledge is the result of his reflections on his sociocultural experiences and his interaction with others using linguistic tools, but when a child enters school, he will engage in more controlled experiences that offer new insights about the world. Vygotsky (1986), in consequence, distinguishes earlier concepts from these new ones that arise in instructional settings as a result of dealing with new ways of interpreting reality, referring to them as "spontaneous" and "scientific" concepts, respectively. Notably, both are developed through negotiation in social communication, which allows the child to progress cognitively and linguistically. This is called by Vygotsky (1986) the zone of proximal development (ZDP) in which a child can perform at better levels by means of cooperation with a more competent person, usually an adult. Naturally, his theory has been highly influential in applied linguistics and teaching practices since it presented the method of challenging students within their actual cognitive possibilities, which, in educational terms, is known as scaffolding (Llinares et al., 2012).

According to Schleppegrell (2004), the Vygotskyan approach to the socio-communicative dimension of education calls for the need of studying the uses and functions of the language that enable learning. In this respect, Halliday's (1978) linguistic approach to knowledge, his renowned systemic functional linguistics (SFL) theory, proposes analytical tools for such analysis. He considers language as a social semiotic system, that is, as a meaning-making set of choices that allows humans to construe experience and construct, develop and communicate knowledge (Halliday, 1978; Halliday & Matthiessen, 1999). Similar to Vygotsky's spontaneous and scientific

concepts, Halliday (1993) claims that the new forms of information emerging in school, "educational knowledge", are different from students' "common sense" knowledge since they require more abstract and technical ways of thinking, hence, of using language (p. 93).

Consequently, language is regarded as the heart of society and culture because it constitutes humanity's main resource to represent reality (Halliday & Matthiessen, 1999). This capacity of the language is the result of its 'metafunctional nature' (Llinares & McCabe, 2020, p. 2): it offers a wide range of linguistic options that are functional to express different kinds of meanings (Schleppegrell, 2012a), which Halliday (1978) classified into three broad abstract categories: ideational (content, representing the understanding of the world), interpersonal (activating and maintaining social relationships and taking a stance with regard to the proposition), and textual metafunction (arranging the other metafunctions into coherent and organised texts). It should be noted that these three metafunctions are always simultaneously operating in any manifestation of language (Llinares et al., 2012). In this sense, linguistic choices shape, and are shaped by, the social context of communication, which according to Halliday (1978) , is composed of the field (what is talked about), the tenor (social relationships) and the mode (expectations in the textual organization) (Schleppegrell, 2004, 2012a).

The binding interaction between the elements of the social context and the linguistic realizations of the metafunctions is what Halliday called register – "the patterns of instantiation of the overall system associated with a given type of context" (Halliday & Matthiessen, 2014, p. 29). These lexico-grammatical choices are governed by the rules of discourse-semantics (cohesion and coherence), which, in turn, build different types of texts (genres) depending on the structure and the social purposes that they are intended to perform (Coetzee-Lachmann, 2007; Schleppegrell, 2012b). The relation between language, register and genre is represented in the following table (Table 1):

<table>
<tr><td colspan="4">Genre</td></tr>
<tr><td rowspan="2">↑
realises
Register</td><td>Field of discourse</td><td>Tenor of discourse</td><td>Mode of discourse</td></tr>
<tr><td>↑
realises
Ideational meaning</td><td>↑
realises
Interpersonal meaning</td><td>↑
realises
Textual meaning</td></tr>
<tr><td rowspan="2">↑
realises
Language</td><td colspan="3">↑
realises
Discourse-semantics</td></tr>
<tr><td colspan="3">↑
realises
Lexico-grammar</td></tr>
</table>

Table 1. The relationship between language, register and genre (Coetzee-Lachmann, 2007, p. 59).

Therefore, SFL makes explicit how a text means what it does (Whittaker & Llinares, 2009), and thus, in educational contexts, it presents a pedagogic role for the intervention in the teaching-learning process, and an analytical role for written and oral text analysis (Morton, 2020a; Schleppegrell, 2004; Whittaker & Llinares, 2009), as it is going to be examined in next sections.

According to Llinares et al. (2012), these theories have had a great impact on education as they contribute to the understanding of the role of language in learning: students develop knowledge in school through new forms of social interaction (Vygotsky) and, for that, they need the necessary linguistic resources to construct and negotiate meanings in context (Halliday). Then, from this perspective, overall academic literacy is understood as the "learner's ability to use a specific register that is different from the registers embedded in the discourses of family life (…) and to engage with unfamiliar interpretations of experience" (Coetzee-Lachmann, 2007, p. 18).

On a similar note, albeit outside systemic-functional considerations of the language, it is worth mentioning Cummins' (1979) acclaimed binary construct of basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP), which is especially relevant for second language learners and is still considered a touchstone among field researchers. His distinction stemmed from the observation that immigrant children in the context of foreign language immersion could apparently understand and use English with ease, yet their academic performance lagged far behind their native peers (Cummins, 2013). As a result, these students were judged as cognitively disabled and inappropriately placed in special teaching programmes

(Cummins, 2013; Schleppegrell, 2004). Delving into this issue, Cummins noticed that age-appropriate levels of conversational fluency were attained several years earlier than academic aspects of the language. Thus, he suggested that BICS and CALP were "conceptually distinct components of the construct of language proficiency" (Cummins, 2013, p. 11), hence, they are "not reducible to each other" (Cummins, 2000, p. 84). Put differently, overall language proficiency should not be based on the linguistic performance of only one of its components.

This twofold nature of language competence was related to previously mentioned notional differentiations, such as Vygotsky's spontaneous and scientific concepts, in the sense that they recognise the formal and academic uses of the language that stand "in contrast to everyday informal speech" (Bailey & Butler, 2003, p. 9). However, in later studies, Cummins (2000) stressed the need to "go beyond a simple dichotomy" (p. 65) in order to understand the multidimensional character of academic language. Consequently, he brought to light the range of contextual support and cognitive demands of school tasks, arguing that academic linguistic difficulties will depend on the information provided by the context (if there exist different resources to negotiate meanings, such as gestures or images) and the degree of cognitive requirement (which depends on prior knowledge, motivation…). This analysis resulted in two intersecting continua with four quadrants that combine context-embedded/reduced and cognitive demanding/undemanding activities (Cummins, 2000, 2013). Understandably, students are expected to progressively use the language in content-reduced and cognitively demanding situations (i.e., academic language), which calls for "explicit instruction", even in the first language of speakers (Leontjev & deBoer, 2020, p. 14).

It should be highlighted several issues that were clarified by Cummins (2000) regarding his theory. Taking a Vygotskyan perspective to language learning, he claimed that both dimensions are acquired through social interaction in context: the fact that CALP tends to be more frequent in academic contexts does not mean that it is superior to BICS or that the latter is not an essential component of learning in educational situations. Moreover, although the natural progression is from BICS to CALP, Cummins (2000) did not state that it was the necessary path, as CALP may precede BICS or the two linguistic dimensions may advance in parallel via appropriate scaffolding (Meyer et al., 2015).

Thus, as suggested before, successful learners are those who understand and express meaning of cognitive-demanding disciplinary content for different academic purposes. In this regard, the most influential classification has been Bloom's Taxonomy of Educational Objectives, revised by Anderson et al. (2001), which stands as a categorization of cognitive operations, as they state what learners are supposed to do with their learning (to understand, to apply, to analyse…) (Dalton-Puffer, 2013). Originally understood as a hierarchical pyramid reflecting a linear process from lower-order (e.g., remembering) to higher-order (e.g., creating) thinking skills, today it is generally accepted that these cognitive processes are levelled and cyclical in the achievement of learning goals (Hemmi & Banegas, 2021; Morton, 2020b).

In conclusion, conceptual knowledge is inevitably presented and mediated by language (Lo & Fung, 2018; Schleppegrell, 2004). The theories discussed here represent the underpinnings of this intrinsic connection. As a result, they have been applied to research on the specific characteristics of the language used for learning, i.e., academic language. The following section shall focus on this topic.

2.2. Characterization of academic language

Quality education depends critically on the attention to the academic language needed to knowledge building, which can be foreseen and identified (Dalton-Puffer, 2013; Nikula et al., 2016) through the lexico-grammatical and discourse patterns that allow for subject-specific ways of meaning making (Achugar & Carpenter, 2014). Due to this relation between language and context, the systemic-functional notions of genre and register have been considerably helpful for researchers who have explored the features and functions of academic language (Coffin, 2017; Snow & Uccelli, 2009).

In fact, Scarcella (2003) defines academic English as a "register of English used in professional books" (p. 9), which is consistent across knowledge areas. Moreover, she claimed that this register is characterised by its multidimensional nature: it presents a linguistic dimension (phonological, lexical, grammatical, sociolinguistic and discourse components), a cognitive dimension (knowledge, high order thinking, strategic and metalinguistic awareness) and a sociocultural/psychological dimension. In brief, she identified key features of academic English that differ from everyday uses of language, such as lexical density and specialisation, highlighting the idea that expectations in the three dimensions need to be met in order to effectively communicate in academic

settings. However, Scarcella (2003) did not propose a conceptual, analytical and standard framework to enable the study of academic language, nor did she observe the linguistic differences according to the subject-matter or purpose of communication.

In this sense, a remarkable work was Schleppegrell's (2004) painstaking description of the language(s) of schooling. Based on SFL, her proposal started with the analysis of general grammatical features (register) in relation to the context elements (field, tenor and mode), thereby providing an overview of common linguistic traits that seem to be shared across the different languages of schooling. Broadly speaking, academic language users tend to "display knowledge authoritatively in highly structured texts" (Schleppegrell, 2004, p. 74) by means of abstract terms, declarative mood, modal verbs, clause-combining strategies such as condensation and embedding, theme position, nominalizations, etc. However, she also observed that each school subject has its own type of communication, typically seen in disciplinary text types, i.e., genres of schooling. Therefore, she considered different subject-specific genres separately: for instance, she identified the particular linguistic features of the register of four science genres (procedure, procedural recount, science report and science explanation) according to their communicative intention. Hence, should the purpose of a science procedure be "to provide instructions", the imperative mood will probably be found (Schleppegrell, 2004, p. 115). In a similar vein, Zwiers (2014) also provided an account of the academic language variations in different content areas depending on the social practice, for example, in science, he distinguished the language used in scientific inquiry, cause-effect, interpretation and comparison.

The bulk of studies in this area has tried to explore the linguistic schemata that co-occur with content units (Nikula et al., 2016), each reporting on different aspects that need to be considered so as to fully comprehend students' development of academic literacy. Besides the lexico-grammatical features of subject-specific genres, authors have also mentioned the social, cultural and learning backgrounds of students (Zwiers, 2014), knowledge of specialised subject matter (Krashen & Brown, 2007), reasoning/argumentative strategies (Snow & Uccelli, 2009), levels of granularity – at a discourse, sentence and phrase level (Gottlieb & Ernst-Slavit, 2014), thinking skills (Lemke, 1990), language skills (Uccelli & Galloway, 2016), language awareness (Achugar & Carpenter, 2014), etc. These contributions, among others, are useful to understand the different ways in which learners use language to engage in learning

activities so that teachers can provide them with academic linguistic guidance (Schleppegrell, 2012b).

Nonetheless, as noted by Nikula et al. (2016), most of them have been highly "concerned with aspects of linguistic form rather than function" (p. 7), and therefore, they seem more theories of language "than a framework for educational research" (Snow & Uccelli, 2009, p. 114), except for genre-based examinations of contextualized academic language, but even in this sense, there is no consensus in the literature about the types and number of genres in each subject. Thus, it can be argued that it may be confusing and challenging to apply all these varying theories to school practice and investigation.

As a result, Dalton-Puffer's (2013) cognitive-discourse functions (CDFs) proposal has become extremely useful as it integrates and condensates the above approaches to academic language into an operational toolkit for educational stakeholders. She observed that, in school contexts, students are expected to use the language in order to accomplish two basic goals: the learning of "new knowledge and skills" (Dalton-Puffer, 2007, p. 128) and being able to do "something with what they have learned" (Dalton-Puffer et al., 2018, p. 6). Therefore, a series of language functions are enacted and promoted in any learning situation as these are closely linked to the thinking skills that are constitutive of meaning making, which are known as academic language functions (Dalton-Puffer, 2007, 2013, 2016). In fact, as suggested before, most educational objectives in all subjects are expressed in terms of Bloom's taxonomy, i.e., as performative verbs that trigger these cognitive-linguistic functions (Dalton-Puffer, 2013; Morton, 2020b). Thus, using a framework of functions permits a more comprehensive and fine-grained approach to academic language use in different contexts.

In an attempt to reduce and organise the multiple academic linguistic functions that appear in literature, Dalton-Puffer (2013) created her CDFs construct, where she subsumed the many existing functions under seven categories according to their communicative intention: CLASSIFY, DEFINE, DESCRIBE, EVALUATE, EXPLAIN, EXPLORE and REPORT. Grounding her proposal on SFL and functional pragmatics, she claims that academic linguistic functions are verbal actions that represent thinking processes "in recurring and patterned ways" (Dalton-Puffer, 2013, p. 216), i.e., through identifiable "discursive, lexical and grammatical schemata" (Dalton-Puffer et al., 2018, p. 7). The

cognitive discourse functions arise in learning contexts from dealing with cognitive-demanding content in order to construct and exchange knowledge.

Therefore, for instance, if a student is classifying, they are linguistically performing the cognitive-communicative intention of "how we can cut up the world according to certain ideas" (Dalton-Puffer, 2013, p. 234), which entails a more abstract and complex type of knowledge because they are establishing connection or disjuncture between terms and facts (Dalton-Puffer, 2016). Hence, cognitive discourse functions offer an intersubjective, holistic, and integrated understanding of the academic language involved in learning subject-specific contents. The following table (Table 2) shows her construct:

Function Type	Underlying Basic Communicative Intention	Label
Type 1	I tell you how we can cut up the world according to certain ideas.	**CLASSIFY**
Performative verbs: Classify, compare, contrast, match, structure, categorise, subsume		
Type 2	I tell you about the extension of this object of specialist knowledge.	**DEFINE**
Performative verbs: Define, identify, characterise		
Type 3	I tell you details of what can be seen (also metaphorically)	**DESCRIBE**
Performative verbs: Describe, label, identify, name, specify		
Type 4	I tell you what my position is vis a vis X	**EVALUATE**
Performative verbs: Evaluate, judge, argue, justify, take a stance, critique, recommend, comment, reflect		
Type 5	I give you reasons for and tell you cause/s of X	**EXPLAIN**
Performative verbs: Explain, reason, express cause/effect, draw conclusions, deduce		
Type 6	I tell you something that is potential	**EXPLORE**
Performative verbs: Explore, hypothesise, speculate, predict, guess, estimate, simulate		
Type 7	I tell you about sth. external to our immediate context on which I have a legitimate knowledge claim.	**REPORT**
Performative verbs: Report, inform, recount, narrate, present, summarise, relate		

Table 2. Cognitive-discourse functions (Dalton-Puffer, 2013, p. 234; Dalton-Puffer et al., 2018, p. 9).

As Dalton-Puffer (2013) acknowledges, these categories present blurred limits, as "classifying is always part of DEFINE, but not all instances of CLASSIFY are" (p. 236). However, it is a matter of teachers being clear and explicit about the language that

pupils will be confronted with and will be producing in terms of key genres and CDFs. For instance, if teachers want students to inform about observable features, qualities or characteristics of something (content), they are asking them to produce a proper text in that discipline that belongs to the description genre (subject-specific literacy), therefore, learners will need linguistic resources to be able to DESCRIBE (cognition), which may include, according to Schleppegrell (2004), generic participants, timeless verbs in simple present tenses, complex sentences with relative clauses to specify properties, modifiers, such as adjectives (in the comparative and superlative forms), and sometimes, grammatical metaphors (language). This is how CDFs bridge the connections between the four dimensions of learning (content, cognition, literacy and language).

Hence, CDFs provide a principled framework for teaching and assessing academic language in all learning contexts, including CLIL. The following section will review how academic language has been considered in CLIL and how it is currently being addressed and promoted through genre-based and CDFs pedagogy.

2.3. The role of academic language in CLIL

As noted above, students in school will experience new ways of using language to engage in unfamiliar cultural and social practices, thus, they need to develop advanced literacy skills in subject-specific knowledge construction (Coffin, 2017). Notwithstanding, CLIL students are struggling with learning content in an additional language, which makes it even more necessary for content and language objectives to be clearly and coherently integrated (Leung & Morton, 2016). Thus, it has become generally accepted that careful observation of how participants express disciplinary knowledge through language is a matter of the utmost importance in bilingual education (Dalton-Puffer, 2007; Dalton-Puffer et al., 2018; Nikula et al., 2016), given the reason that being subject-specific literate in CLIL also entails "competences or skills that allow learners to translate content matter from and between all the different languages that are involved in literacy learning in a content subject" (Hallet, 2012, p. 196).

Therefore, if explicit instruction of academic language is essential for every learner, CLIL students may be said to be especially needy, mainly due to three reasons. Firstly, CLIL learners may not be proficient neither in their native academic language (especially in primary education), nor in English (general or academic) before entering

school (Otto, 2018). Secondly, the learning of a new register requires plenty and proper input (Schleppegrell, 2004), but in CLIL, teachers share students' status as non-native speakers of English, and thus, their linguistic resources are not as proliferous as in the first language, rendering them a less reliable source of academic language (Dalton-Puffer, 2016; Nikula et al., 2016). Lastly, CLIL lessons may be the only time of exposure to English for probably many students (Dalton-Puffer & Smit, 2013), thus, it is very difficult that "the language expectations of schooling fit seamlessly" (Schleppegrell, 2004, p. 153) into their primary social uses of language outside school, because in these cases, they are using their first language. These conditions suggest that the natural development of CLIL learners' academic linguistic resources can by no means be "assumed to happen by osmosis" (Morton, 2010, p. 82), as these will not be picked up as easily as social uses of the language (Scarcella, 2003; Zwiers, 2014).

However, true integration of language and content, where both are found in a joint and balanced way among the educational objectives, is quite rare across CLIL contexts (Cenoz et al., 2014; Llinares & McCabe, 2020). Especially in hard CLIL, the content-focused approach most extended in Europe (Ball et al., 2015), it is not usual that explicit attention to the development of academic literacy is provided (Coyle, 2015). Consequently, several theories on how to develop academic language-sensitive CLIL classrooms have emerged in recent decades which see the role of language as going beyond simply grammatical correction and general language proficiency (Coyle, 2015; Dalton-Puffer et al., 2010).

A keystone in this regard is Coyle et al.'s (2010) well-known four Cs framework as it binds together different elements involved in CLIL: thinking processes (Cognition) leading to the acquisition of subject matter (Content) are subjected to language learning and using (Communication) and to socio-cultural and contextual factors (Culture). Precisely, the language needed for learning in CLIL was conceptualised into a Language Triptych including the language of learning (to acquire subject-specific concepts and skills), for learning (to be able to interact and negotiate meaning with others), and through learning (to develop linguistic awareness and insight). Nevertheless, their proposal specified neither how these elements would be actually integrated in CLIL contexts (Meyer et al., 2015), nor an elaborated and well-defined set of linguistic goals in relation to content learning objectives (Dalton-Puffer et al., 2010). Thus, the operationalizing of the four Cs framework in concrete CLIL classroom

practices and even research remains a challenge (Lo & Jeong, 2018; Nikula et al., 2016).

According to Dalton-Puffer et al. (2010), the theoretical groundings for merging content and language goals are to be found in approaches that enunciate the social, linguistic, and contextual nature of learning, such as SFL, since it systematically relates meaning selections with their social functions (Halliday & Matthiessen, 1999, 2014). A case in point in this regard is the work by Llinares et al. (2012), since they proposed practical suggestions to integrate content and language, taking a systemic-functional approach to the language. They claim that subject-specific literacies are expressed through academic text types (genres) and lexico-grammatical features (register). Hence, content educators must be aware of their disciplinary genres and register so that they can teach the language to students using proper materials and classroom interaction. In this way, the attention of learners would be redirected towards language use, which in turn will result in a greater language development. According to them, this development presents three dimensions, which correspond to Halliday's metafunctions: lexico-grammatical forms to display subject-specific knowledge (ideational), appropriate expression of ideas (interpersonal), and coherent discourse arrangement (textual). Therefore, they shed light on how SFL could be used to observe the role of language in CLIL: "cognitive functions intrinsic to a subject become visible through a focus on genres and their stages" (Llinares et al., 2012, p. 147).

In this sense, learning a subject consists in being able to understand and produce typical academic discourses by means of which knowledge in that subject is transmitted (Morton, 2020b), thereby acquiring subject-specific literacy. Hence, a genre, understood as a "staged goal-oriented social process" (Martin, 2009, p. 13), could be used as a "guiding principle" (Meyer et al., 2015, p. 46) for teachers to map academic language progression onto content learning in classroom practices by making explicit the relation of "vocabulary, grammar, sentence structure and discourse organisation" (Lo & Jeong, 2018, p. 137) to the genre in question. Teachers providing scaffolded control of different subject-specific genres will lead, eventually, to the development of students' academic literacy (Morton, 2010), since they will be learning contextualised and purpose-oriented uses of the language (Halbach, 2018).

As a result, there have been several attempts to implement a genre-based pedagogy in CLIL, which consists of clarifying "how knowledge structures in the disciplines

(causality in history or multiplicity in maths) match textual structures" (Lorenzo, 2013, p. 378). To that end, according to Martin (2009), teachers and students must engage in a dialogue cycle of three phases: deconstruction of features of the genre, joint construction of the genre, and individual construction, when the learner writes the genre independently (p. 15). In fact, some basic genres, such as explanations or argumentations, tend to occur across subjects, although with slight variations. Thus, genre education will be completely useful for CLIL students as this knowledge can be transferred to other disciplines and between the different languages (Morton, 2020b), and when this transfer occurs, it means that deep and effective learning has taken place (Meyer & Coyle, 2017).

However, a genre-based pedagogy needs to take into account that the selection of genres cannot be random, but carefully chosen and adapted according to students' cognitive abilities and language competence (Lorenzo, 2013). In this sense, Dalton-Puffer (2013) realised that it was also needed a complementary "notion on an intermediate level of granularity" (p. 230) that was transdisciplinary enough to be applied to different subject-specific discourses at a sentence level. That is why she developed the previously discussed construct of cognitive-discourse functions, which serves to meaningfully bridge language- and content-learning goals. Thus, as Meyer & Coyle (2017) claim, CDFs "can be understood as micro genres which can be combined to build the larger genres representative of the various disciplines" (p. 211).

This represents, precisely, what is advocated by current educational approaches that aim to integrate content and language, such as the pluriliteracies model developed by Meyer et al. (2015). The term 'pluriliteracies' captures the multifaceted and complex language practices in CLIL – different and interrelated semiotic systems in more than one language for a wide range of purposes in different subjects (Pavón, 2018). In this sense, it should be noted that individual means of communication are viewed as dynamic, continuous and overlapping, and as Lin (2020) claims, the aim is not to replace the learner's means of meaning-making in their mother tongue with a foreign and academic linguistic register, but rather to "expand their holistic repertories without constructing a hierarchy of these different resources" (p. 7).

This idea casts doubt on how to map knowledge construction onto linguistic advancement so as to scaffold students' development of subject-specific literacies (Meyer & Coyle, 2017). That is why, based on Coyle et al.'s (2010) four Cs framework

and the Language Triptych, Meyer et al. (2015) visually represented in their pluriliteracies model what has been argued so far: disciplinary knowledge progression (facts, concepts, procedures and strategies) must inextricably co-occur with increasingly mastering of subject-specific literacies (purpose, mode, genre and style). Consequently, language has a two-fold role in learning: it makes students' understanding visible, and it represents the tool by means of which teachers can mediate and scaffold their knowledge construction (Meyer & Coyle, 2017). Thus, the only means of moving forward in this diagram is through promotion of the development of cognitive discourse functions, as they lie at the intersection between thinking and language. In other words, quality learning only occurs if students can "create links between the conceptual and the communicative continuum in increasingly more sophisticated ways" (Meyer & Coyle, 2017, p. 205), for which they need explicit instruction (Coyle, 2015).

Hence, according to Pérez Cañado (2020), language in CLIL should be redirected towards this pluriliteracies approach as it seems to be the pathway towards actual integration. Content teachers' implicit knowledge about the linguistic features of the textual genres of their subjects must be made explicit to learners in collaboration with language teachers, as they are co-responsible for the development of students' academic literacy (Lorenzo, 2013). In doing so, they will contribute to equity in education through helping students attain curricular goals, empowering them to develop their critical thinking, to communicate across cultures, and to bring their own voices to the subject matter (Achugar & Carpenter, 2014; Coyle, 2015; Moore & Schleppegrell, 2019), thus "forming successful citizens with a substantial contribution to make to society" (Pérez Cañado, 2021, p. 33) or as Beacco et al. (2015) put it, "the acquisition of competences in language is an essential foundation both for success in school and for participation in modern democratic and diverse knowledge societies" (p. 9). That is why academic literacy is also part and parcel of students' personal growth.

This explicit attention to language and content in integration needs to be understood in three interconnected dimensions: curriculum and pedagogy planning, participants and classroom practices (Nikula et al., 2016). From the perspective of classroom practices and with the learners in focus, the primary evidence of students' learning is language (Mohan et al., 2010; Otto, 2018), thus, it is essential to assess their subject-appropriate uses of language in order to determine "how they progress and the kind of mediation (scaffolding) that they need" (Nikula et al., 2016, p. 21). However, there seems to be a

slippery understanding of the essential role of academic language among CLIL teachers (Meyer et al., 2015), and therefore, most of them do not feel prepare to take on this responsibility (Dalton-Puffer et al., 2018; Otto & Estrada, 2019). The following section will focus on the problems that currently exist with regard to the assessment of academic language in CLIL.

2.4. Assessment of academic language in CLIL

Language-related assessment is one of the most contested and underexplored aspects of CLIL (Llinares et al., 2012; Lo & Fung, 2018; Morgan, 2006; Otto, 2018), where questions emerge regarding "what, who, when, with which methods and why" (Wewer, 2014, p. 102). That is why it has often been left "a blind spot" in many CLIL contexts (Massler et al., 2014, p. 138). Therefore, more research into assessment in CLIL is required so as to pave the way towards adequate, clear and systematic tools that guarantee the objectivity, reliability and validity of the assessment (Hönig, 2010).

Traditionally, assessment methods have been classified into summative assessment (also known as assessment of learning), which is aimed to observe what a student has learned after a period of time, and formative assessment (or assessment for learning), which is continuous and help identify how pupils are progressing in their learning (Bentley, 2010). However, any type of assessment should not be conducted with the purpose of punishing students' weaknesses and rewarding their strengths, but rather to enhance learning by means of scaffolding their abilities and skills (Wewer, 2014). As a result, lately, this binary view has been challenged through the proposal of the term learning-oriented assessment, which combine both formative and summative methods that "feed back into assessment planning and adjusting, which will promote further learning" (Xavier, 2020, p. 115).

Hence, learning-oriented assessment is essential in education, but due to the dual nature approach of CLIL, it seems to be particularly challenging in this context as both language and content goals need to be evaluated (Wewer, 2020). According to Bentley (2010), CLIL assessment practices may vary depending on their objectives: soft CLIL gives more emphasis to the language, whereas hard CLIL may focus either on both aspects or just on content. Consequently, two approaches to the assessment of content and language knowledge can be discerned: a discrete one, where these dimensions are

considered separately, and an integrated one, where these are observed simultaneously (Otto, 2018; Wewer, 2014). The following sections will elaborate on these two types of assessment procedures in relation to common classroom practices and research in CLIL.

2.4.1. Discrete assessment

Discrete assessment, which seems to be the most widespread approach to assessment in CLIL (Otto, 2018), involves considering content and language knowledge independently so that one does not affect the grading of the other. However, this strategy may not be as straightforward as it seems, since, as suggested above, it is doubtful that language and content can actually be separated.

On the one hand, regarding the assessment of content, students need to achieve the same objectives as their peers in monolinguals schools in order to advance in education (Wewer, 2020). Therefore, at least in Andalusia, official recommendations explicitly indicate that content knowledge is the main goal, and that language should be regarded as an additional value, without teachers from non-linguistic areas being able to downgrade marks due to linguistic aspects (Conserjería de Educación, 2013). Content teachers not only try to comply with this guidance (Barrios & Milla Lara, 2020), but also seem to welcome it, as some openly report uneasiness when correcting linguistic errors and do not consider it as part of their duties (Lorenzo, 2013; Otto & Estrada, 2019). Nonetheless, parents and students have the feeling that language proficiency does have an impact on their final grades (Barrios & Milla Lara, 2020). The reason for this mismatched perception seems to be that, although content teachers may believe that they are only assessing learners' disciplinary knowledge, the truth is that they are also evaluating language skills, albeit unconsciously (Lo & Fung, 2018). This fact was observed by Hönig (2010) in Austrian CLIL classrooms when she realised that eloquent students outperformed their peers with poor linguistic abilities, even though teachers claimed that they would not consider language use when grading. Hence, she concluded that pupils' inability to express their knowledge with adequate language negatively affected their grades on content mastery without teachers being aware of it (Wewer, 2014). Therefore, it seriously threatens the validity of trying to assess just content learning (Morton, 2020a; Otto, 2018) as it may not be accurately reflecting learners' real knowledge of the subject matter (Lo & Fung, 2018).

On the other hand, in relation to language assessment, two issues need to be highlighted. First, on those rare occasions when language is scaffolded by content teachers, it has been mostly related to lexico-grammatical questions and based on each instructor's intuition about students' needs, as there are no directions from the departments or official institutions (Otto, 2018). Second, when language is examined by language teachers, it has been much focused on the communicative competence of students according to the criteria of the Common European Framework of Reference (CEFR), following official recommendations (Conserjería de Educación, 2013). Notwithstanding the importance of an everyday language competence according to the standards of the CEFR, these descriptors do not seem to be completely appropriate for the subject-specific linguistic particularities presented in CLIL (Shaw, 2020; Wewer, 2020), as CLIL "includes a cognitive dimension not considered in the CEFR" (Shaw, 2020, p. 33), which is that young students are using the foreign language to learn disciplinary content, not only for general communicative purposes. Hence, as Morton (2020b) puts it, it may be that "measures of general English proficiency do not capture the specific academic language competences students need to be successful in bilingual programmes" (p. 13).

This is also acknowledged by Beacco (2010) when he claims that the CEFR descriptors do not either consider learners' age or the cognitive-linguistic functions of academic language as, for instance, it relates "more to reading as comprehension than as interpretation or critical response" (p. 17). Therefore, Evnitskaya & Dalton-Puffer (2020) question whether language testing methods based on communicative competence really "do justice to the specialist language" needed in CLIL (p. 1). In this regard, Shaw (2020) claims that "it is necessary to expand the familiar proficiency dimension by additional two dimensions relating to age and academic content area" (p. 34), in order to devise fine-tuned descriptors accordingly (Xavier, 2020).

2.4.2. Integrated assessment

Integrated assessment, therefore, seems to be a more suitable method for CLIL contexts. Integrated assessment of content and language consists of concurrently examining both by means of connecting content objectives of the different subjects with language goals, or in other words, determining the academic language needed to meet achievement standards (Otto, 2018; Wewer, 2014). This procedure requires identifying linguistic

forms in relation to their function, that is, a combination of both focus-on-meaning and focus-on-form (Massler et al., 2014; Nikula et al., 2016), as it has proven to be extensively beneficial in Canadian immersion programmes (Pérez Vidal, 2007). In doing so, the three dimensions on which evaluation should be based, according to Wewer (2014), will be considered: assessment of young language learners, which is related to students' cognitive and linguistic capacities, assessment of English for academic purposes and content-based assessment (p. 99).

In this sense, Halbach (2018) establishes three steps in preparing an integrated assessment: identifying desired outcomes, determining acceptable assessment evidence, and planning the learning experience (p. 210). Firstly, teachers from linguistic and non-linguistic areas should start from the requirements of content subjects (Pavón, 2014) in order to identify what language will be needed for students to understand and convey knowledge. Once the learning outcomes are clear, teachers should pre-define assessment criteria (Wewer, 2020), through the identification of relevant discourse features that shape a successful answer and that will be looked for in students' productions. Finally, they need to be explicit about these objectives and provide multiple opportunities to scaffold the learning progression towards the desired outcomes, since "assessment should mirror daily practice" (Otto & Estrada, 2019, p. 40). Moreover, Serragiotto (2007) also points at the need to be clear about the target of assessment, distinguishing between process assessment, which is aimed to observe students' engagement, learning strategies, reasoning, etc., and product assessment, which is focused on verifying students' skills and knowledge "by means of preset and standardized tests" (p. 271).

However, this integrated assessment seems to be a challenge for CLIL teachers due to four main aspects that need to be addressed:

(1) official curricula, guidelines and assessment criteria specifying the role of language in CLIL (Meyer et al., 2015; Otto, 2018; Otto & Estrada, 2019), which is currently missing in Andalusia (Barrios & Milla Lara, 2020);

(2) genre-based materials, tests and methods specifically designed for CLIL and not just translated and adapted from monolingual contexts (Meyer et al., 2015; Morgan, 2006; Otto & Estrada, 2019);

(3) teaching training (Morton, 2017; Otto & Estrada, 2019) and collaboration between content and language teachers (Pavón & Ellison, 2013; Szczesniak & Muñoz Luna, 2022);

(4) and further research into CLIL assessment (Lo & Fung, 2018; Otto & Estrada, 2019; Wewer, 2014).

This represents the theoretical background against which the present study is conducted. Summing up, due to the fact that language and learning are inseparable, educational success is determined by the students' ability to use the language for understanding and verbalising knowledge, which is known as academic language. Using a systemic-functional approach, it is possible to foresee and identify the linguistic features typically used in a disciplinary text, which will depend on the cognitive-discourse functions that present the genre. Hence, a genre-based teaching approach should be implemented, according to the pluriliteracies model, as it allows instructors to map content objectives to language goals by means of a focus on form and meaning, thus addressing them both in an integrated way when assessing students' learning progress. Lastly, the following section will provide a general review on research into students' language outcomes in CLIL in order to situate this project in the current academic landscape.

2.4.3. Research into students' language outcomes

As mentioned above, most research into CLIL students' educational advancement considers the language independently of the content. In this sense, the number of studies examining the impact of CLIL on students' general foreign language competence as compared to their non-CLIL peers is vast and proliferous. In Spain, it has been widely proven that CLIL favours foreign language development without detriment to their first language or content knowledge of students in primary and secondary education (Lasagabaster, 2008; Pérez Cañado, 2020; San Isidro, 2019; San Isidro & Lasagabaster, 2019). Specifically, the most benefitted linguistic aspect is considered to be spoken production and interaction, although research has shown a positive impact on all the other skills (Lasagabaster, 2008; Nieto Moreno de Diezmas, 2016).

However, as noted by several authors, such as Meyer et al. (2015) and San Isidro (2019), despite the large research into the beneficial effects of CLIL on second language development as compared to English as a Foreign Language (EFL) lessons, few studies

have addressed students' academic language competence and their subject-specific literacy development in CLIL according to age-appropriate standards. In this sense, Navés (2011) observed that "although CLIL learners' overall proficiency increased statistically significantly when compared to mainstream EFL learners, gains are insufficient from a language policy perspective" (p. 181). That is why Pérez Cañado (2020) claims that the conflict between CLIL and EFL instruction should no longer be considered an issue on the current CLIL research agenda, and that efforts should be directed towards "a more integrative stance which conflates academic literacy and disciplinary literacy" (p. 7).

In this regard, investigations into academic language in CLIL have been mainly focused on conversation analysis, examining classroom interactions between teachers and students when co-constructing knowledge. For instance, Dalton-Puffer (2007) observed in her research that, although explanations were somewhat more frequent, the "CLIL classrooms studied cannot be said to represent an environment that is conducive to the learning of academic language skills" (p. 170). Similarly, Pérez Vidal (2007) found that CLIL lessons were much concerned with focusing on meaning, but they did not show a focus on form, and Pérez & Pavón (2019) agree on the fact that classroom dialogues in CLIL present important deficits in the use of academic linguistic resources.

Among those studies regarding students' outcomes in terms of academic language, in Spain, those by Lorenzo & Rodríguez (2014), Falcón & Lorenzo (2015), Morton & Llinares (2018), Lorenzo et al. (2019) and Granados et al. (2022) can be highlighted, as they analysed from a functional perspective the linguistic resources that learners bring to the table when dealing with subject matter. Lorenzo & Rodríguez (2014) examined secondary education (from 9th to 12th grades) students' use of language when producing historical narratives, and they observed significant differences in length ratio, dependent clauses, verb diversity, nominalisations, sentence subordination and verb tenses, all of which in favour of older learners. Thus, they concluded that historical literacy is consolidated in higher grades. Falcón & Lorenzo (2015) agree in this respect, since they observed that the syntactic complexity of learners' linguistic expression over time remains stable in their mother tongue but develops considerably in their foreign language. Moreover, they also found that the genre in which students show the greatest syntactic complexity is explanation. Similarly, Morton & Llinares (2018), when examining the evolution of historical evaluations performed by secondary education

students over four years, also claimed that academic language develops over time, although they noticed that abstraction (nominalisations and grammatical metaphors) is challenging even for mature learners. The fact that abstraction is difficult for students and that academic language presents age constraints is further supported by Lorenzo et al. (2019) when assessing the evolution over three years of secondary pupils' historical narration, explanation and evaluation, and by Granados et al. (2022), who explored lexical richness progression again in historical writing by adolescent students for three years. Another worth mentioning investigation was the one conducted by Lo & Jeong (2018) into the impacts of a genre-based pedagogy on students' academic literacy in Hong-Kong, where they noted that, thanks to this approach, students perform better argumentative texts in terms of academic language.

Although this research sheds light on some aspects of academic language and its evolution, it focuses on the discipline of history and is concerned with secondary school students' writing competence. Recently, other studies have appeared aimed at observing students' strategies to express their academic knowledge using the CDF construct in primary education (Morton, 2020b) and in primary and secondary education (Nashaat-Sobhy & Llinares, 2020; Whittaker & McCabe, 2020), in subjects other than history, and in skills other than writing. Morton (2020b) studied the definitions on Natural Sciences in English and Spanish, and in spoken and written production, by students in the last year of primary education. His results showed more formal definitions when writing, but more expansions when speaking, and very similar definitions across languages and modes. Similarly, Nashaat-Sobhy & Llinares (2020) did not find significant differences in primary education students' definitions across languages, and they also confirmed that learners develop their academic language as they grow older, as did Whittaker & McCabe (2020) regarding evaluations across disciplines. These authors also noticed some differences in students' academic linguistic choices depending on the subject: for instance, secondary education students tend to be more emotionally involved in Art than in History, making personal and subjective evaluations that may not be appropriate in academic settings.

After having offered this review of research into academic language and students' language outcomes, several conclusions can be derived. Firstly, more investigation on learners' use of academic language is necessary in order to examine the development of their subject-specific literacy and be able to establish guidelines for improvement,

especially in primary education as this context has not been covered as much as that of secondary education. Secondly, many of the studies reviewed here are longitudinal, with the target of analysing of the evolution of academic language, but cross-sectional and descriptive projects are also needed in order to provide insight into students' level of academic language according to previously established criteria. Thirdly, as academic language permeates every subject-specific literacy, it should be subjected to scrutiny in different subjects, genres (and CDFs) and communicative skills so as to get a comprehensive view of what learners require. And finally, CLIL stakeholders seem to lack standardized and adequate tools that could be used for the assessment of students' academic language in relation to content. Hence, this project, described below, tries to address some of these gaps.

3. Research method

It should be noted that the present work is part of the Erasmus+ project "IBModel: Towards an integrated bilingual model", to which the results of this case study are intended to contribute. Therefore, the data have been collected thanks to the participants in this project, and the tests and assessment instruments have been designed and validated by experts in the field.

3.1. Objectives

This study has undertaken research into the academic language competence level of CLIL students in the last year of primary education. Its general purpose was to observe and describe participants' academic language proficiency. More specifically, it was aimed to assess their command of the cognitive-discourse features and functions characteristic of certain textual genres and how it influences their acquisition and verbalisation of disciplinary knowledge. To that end, the following specific objectives were addressed:

1. To estimate the appropriacy of learners' linguistic choices according to the genre in question and illustrate it with examples from the corpus.
2. To identify the textual genre(s) where students present the greatest and least learning needs.

3. To identify the communicative skill(s) in which they are most and least proficient when using academic language.

3.2. Context and participants

A total number of 84 CLIL students (46 boys and 38 girls) in the last year of primary education (6th grade, 11-12 years old) from two semi-private schools in the Andalusian province of Cordoba participated in this study, among which 30 were randomly selected to perform the speaking task. Both schools belong to the same institution, thus, the sample is balanced in terms of socio-economic backgrounds. Regarding bilingual education programmes in this context, according to regional legislation, in the last year of CLIL primary education it is compulsory to teach in the foreign language, at least, Natural Sciences and Social Sciences, although other subjects are usually included, such as Arts and Crafts and Physical Education.

3.3. Instruments and data gathering

The main instrument employed for eliciting pupils' academic language was a test, which consisted of two uniquely designed, piloted and validated original tests. These tests were created according to two topics from the official curriculum of the two subjects which must be taught in the foreign language in the last year of primary education: The Solar System in Natural Sciences and Climate Change in Social Sciences.

In order to design them, firstly, the language accreditation exams of various institutions which meet the descriptors of the CEFR, and which are aimed at students of this age group, were analysed. Afterwards, the most characteristic textual genres of these areas were determined, as well as the discursive functions and features of each one, so as to create engaging tasks in which students would have to use their academic language and which would be appropriate to their linguistic and cognitive abilities. As a result, The Solar System test (Natural Sciences) measured description and explanation, while the Climate Change one (Social Sciences) focused on narration and argumentation, so that students' linguistic expression was assessed in relation to a total of four textual genres. Regarding language skills, each test consisted of one activity for each communicative skill (reading, listening, writing and speaking) plus one activity focusing on the

linguistic form, i.e., explicitly aimed to observe expressive accuracy over communicative ability. Thus, each test had a total of five activities.

It should be noted that these tests were online, and students completed them independently in the classroom using individual tablets via the web application named notebooklearn.com (see Appendix A for a sample of the tests in document format). They had approximately 45 minutes to complete each test and they could not use any external references. Once the tasks were completed, the answers in the reading, listening, and focus on form activities were automatically corrected (as these were multiple choice or true/false type), while speaking and writing productions were assessed by means of a rubric (see Appendix B). This rubric presented assessment criteria at five levels (not satisfactory, weak, satisfactory, good and excellent) related to the use of target discourse features: syntactic accuracy, lexical richness, cohesion, and communicative achievement of the purpose of the genre. All productive responses were corrected by the same rater to ensure homogeneity of marks.

Finally, data were obtained for the difficulty and discrimination indices. The purpose of the difficulty index was to ensure that the test was aligned to the required level and allowed to obtain relevant and sufficient information, while the aim of the discrimination index was to confirm that the questions worked in distinguishing between more proficient candidates and those with more limited competence. In addition, the internal consistency index (Cronbach alpha) was also calculated to guarantee that the results were reliable, and that the percentage of error was within acceptable parameters.

4. Presentation of results

In order to respond to the specific objectives of this study, this section will present the quantitative results in terms of students' performance in each test regarding their command of academic discourse features and functions, then it will compare the overall results between the tests, and finally, it will contrast the specific results in each of the five activities (language skills). It should be mentioned that the speaking activity is not considered in the presentation of the results until the last section, as the sample was much smaller and therefore reliability and discrimination indices could not be

compared. Therefore, in the first two sections, test results are shown out of a total of 4, while in the last section the five activities are considered, each of which accounts for a total of 1.

4.1. Performance in each test

Regarding the test related to Natural Sciences (The Solar System), the difficulty index in every activity was above 0.5, which indicated that the level of the test was in line with the students' competence. Similarly, the discrimination index was above 0.6 in all cases, which was considered ideal for a proper item operation. The reliability data also showed that the activities were correctly measuring the information for which they were created, as Cronbach Alpha is 0.785.

Item	Skill	IF	r pbi
1	Reading	0,70	0,61
2	Listening	0,70	0,61
3	Writing	0,50	1,00
4	Focus on form	0,68	0,64

Table 3. Difficulty and discrimination indices for description and explanation.

In this test, students had to deal with the academic discourse features characteristic of the textual genres of description and explanation. The following table (Table 3) represents the overall results:

Total Score	
N Valid	84
Missing	0
Mean	2.3562
Median	2.3000
Mode	2.8000
Variance	1.0220
Standard Deviation	1.0109
Range	3.7750

Table 4. General results in description and explanation.

Considering that the results are displayed on a maximum score of 4, it can be noted that both the mean and the median were above 2 and very close to each other, so that their

performance seemed to have been sufficient, although not entirely satisfactory. However, it was determined that the mean could not be considered sufficiently representative of the overall level of students, since there was great disparity between them in terms of proficiency, as evidenced by the large magnitude of the standard deviation and the variance. This was further supported by the value of the range, which was also quite high. The following histogram (Figure 1) shows these results more visually and precisely:

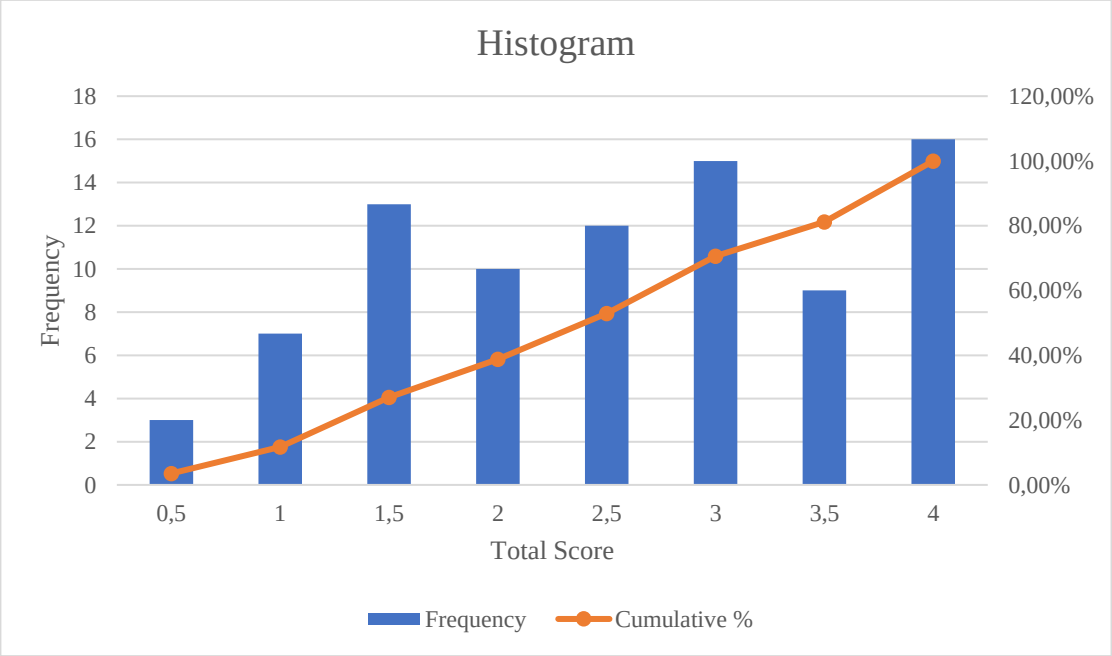


Figure 1. Histogram of description and explanation results.

The histogram shows on the horizontal axis (indicated by the Total Score legend) the minimum (0) and maximum (4) scores that could be obtained in this test. The left vertical axis (indicated by the legend Frequency) indicates the number of students who are found within each score range. Thus, blue bars refer to this axis, while the orange line (Cumulative percentage) is interpreted by means of the right vertical axis, where the percentage of students in that value with respect to the total is displayed.

As it can be observed, even though most cumulative frequency lied in the highest possible range, between 3.5 and 4 (16 students, almost 20 % of the population), it was closely followed by the interval between 2.5 and 3 (15 students), where the mode was found (2.8), and by the one between 1 and 1.5 (13 students). Hence, there was high heterogeneity in the level of students when describing and explaining, despite the relatively acceptable mean.

In relation to the test dealing with Social Sciences (Climate Change), the difficulty indices in the reading and listening activities were above 0.5, which meant that the level of these items was in line with the students' competence. Nonetheless, the values for the writing and focus on form activities were below 0.5, which indicated that these were slightly above their actual capabilities. Regarding the discrimination index, every item operated properly as the values were above 0.55 in all cases. The reliability data were also acceptable enough (0.65) to consider that the test was correctly measuring the information for which it was designed.

Item	Skill	IF	r pbi
1	Reading	0,70	0,61
2	Listening	0,70	0,61
3	Writing	0,50	1,00
4	Focus on form	0,68	0,64

Table 5. Difficulty and discrimination indices for narration and argumentation.

In this case, participants needed to deal with the academic linguistic resources necessary for engaging with the textual genres of narration and argumentation. The following table (Table 6) shows the general results of this test:

Total Score	
N Valid	84
Missing	0
Mean	1.7200
Median	1.5600
Mode	1.0500
Variance	0.6180
Standard Deviation	0.7861
Range	2.9100

Table 6. General results in narration and argumentation.

As in the previous test, the results are shown out of a total of 4. Thus, the performance of students in this case was not satisfactory enough, as the mean and median values were below 2, and the mode was also low. In fact, the standard deviation and the variance values were below 1, indicating that most of the data tended to be clustered around the mean. Consequently, the mean could be considered representative of their

actual performance, as the level of students appeared to be quite similar. This was also corroborated by the range value, which was quite low, suggesting that results were not dispersed. The following histogram (Figure 2) displays this idea:

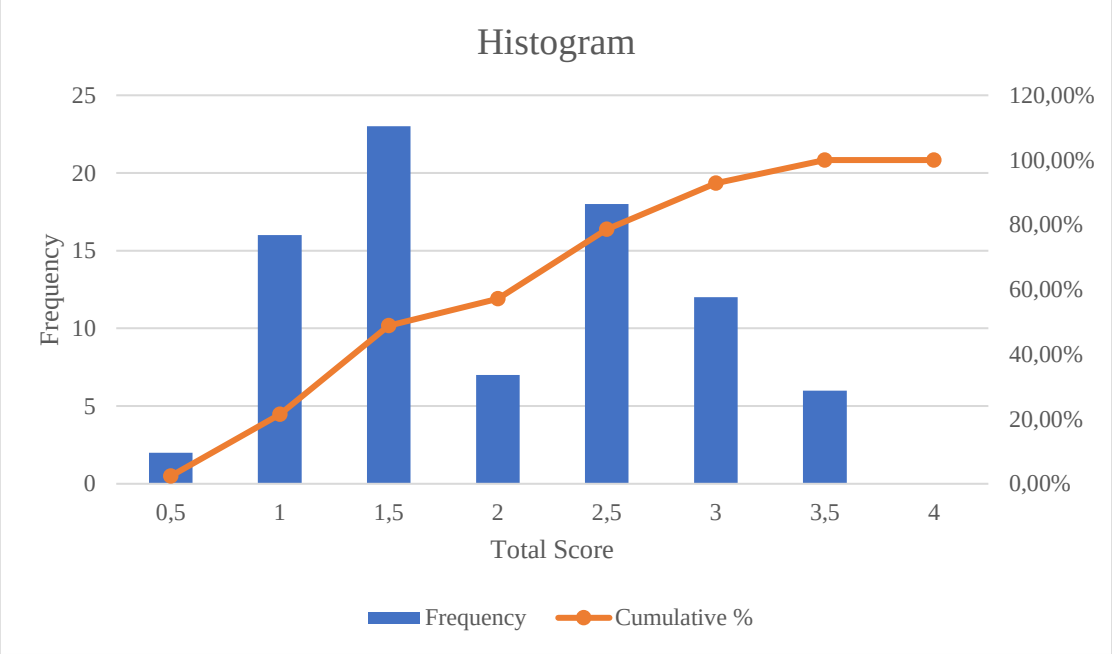


Figure 2. Histogram of narration and argumentation results.

It can be noted that the range between 1 and 1.5 was the most populated one by far (23 students, approximately 27 % of participants), and in fact, the mode was found here (1.05). It was followed by the intervals between 0.5 and 1, and the one between 2 and 2.5, both of which presented almost the same value (16 and 17 students, respectively). Noteworthy, only 6 students performed above 3 (7 %), and none of them scored above 3.5. Thus, apparently, pupils had great difficulty when facing narrative and argumentative discourses.

4.2. Comparison between both tests

Once the performance in each test has been shown separately, it is useful to present the following table (Table 7), which reflects the results of both tests together in order to compare them and to give a clearer picture of the overall performance of students:

	Natural Sciences (D+E)	**Social Sciences (N+A)**
N	84	84

Mean	2.3500	1.7200
Median	2.3000	1.5600
Mode	2.8000	1.0500
Variance	1.0220	0.6180
Std. Dev.	1.0100	0.7860
Range	3.7750	2.9100

Table 7. Comparison between the general results of both tests.

When comparing the general results of both tests, the mean, the median and the mode evidenced that students scored significantly higher when describing and explaining (D+E) than when narrating and arguing (N+A). Nevertheless, the rest of the statistical analysis (variance, standard deviation and range) showed that the students' level of competence in the case of description and explanation was much more heterogeneous and uneven than in the case of narration and argumentation. On the one hand, in the first case, although the mean seemed satisfactory, it could not be considered an accurate representation of the classroom level, since there were very competent students and others who found it more difficult to describe and explain. On the other hand, in the second case, the mean was low but descriptive, as apparently they shared a similar level when narrating and arguing.

4.3. Comparison between the language skills

The next table (Table 8) and its corresponding bar chart of means (Figure 3) show the specific results in each activity, i.e., in the different language skills assessed. It should be remembered that, in this case, results are presented out of a total of 1.

Skill	Test	N	Mean	Median	Mode	Variance	Std. Dev.	Range
Reading	D+E	84	0.6976	0.8000	1.0000	0.0842	0.2902	1.0000
	N+A	84	0.4837	0.5100	0.6800	0.0492	0.2218	1.0000
Listening	D+E	84	0.7035	0.8000	1.0000	0.0978	0.3127	1.0000
	N+A	84	0.4982	0.4300	1.0000	0.1297	0.3602	1.0000
Writing	D+E	84	0.4094	0.4000	0.4000	0.0973	0.3120	1.0000
	N+A	84	0.3286	0.3000	0.0000	0.0828	0.2877	1.0000
Focus on	D+E	84	0.5456	0.5000	0.5000	0.0731	0.2704	1.0000

form	N+A	84	0.4085	0.4000	0.6000	0.0542	0.2328	0.8000
Speaking	D+E	18	0.6055	0.5000	0.5000	0.0723	0.2689	0.8000
	N+A	12	0.7250	0.8000	0.9000	0.0365	0.1912	0.5000

Table 8. Comparison between the specific results in the different language skills.

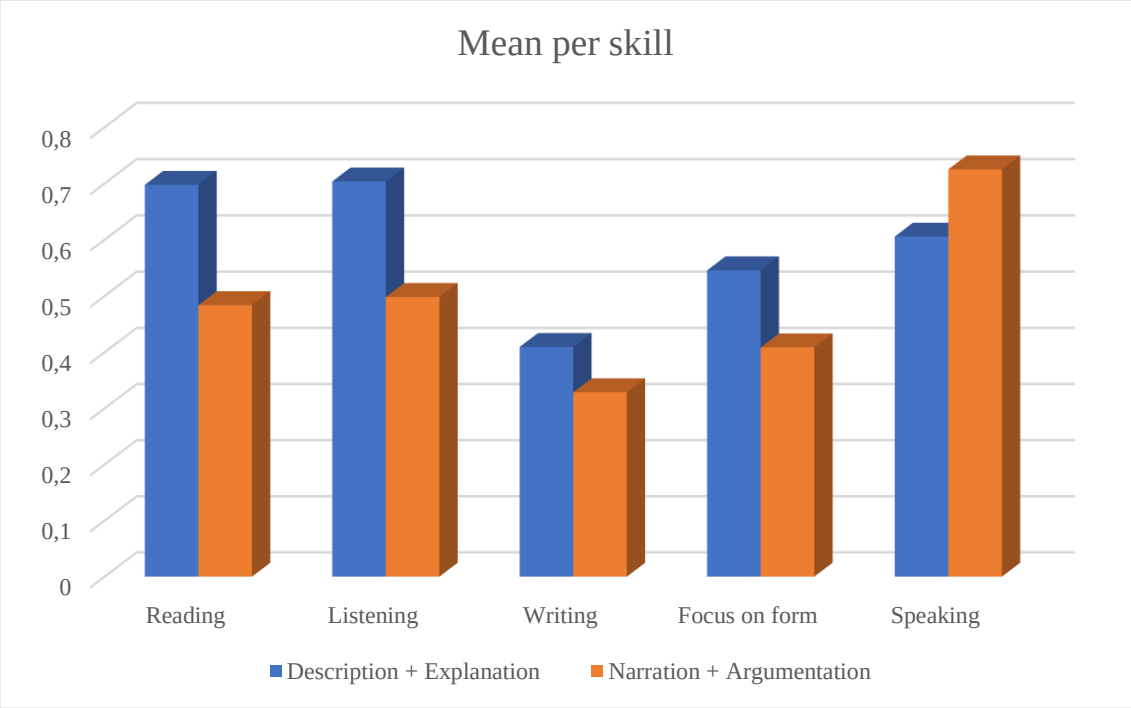


Figure 3. Means of language skills results.

Focusing on the results in the reading, listening, writing and focus on form activities, as they presented the same population, it is noteworthy that students performed better in the test about Natural Sciences (D+E) regarding all these four skills than in the one on Social Sciences (N+A). Interestingly, reading and listening (receptive skills) not only showed the highest scores, but these were also nearly identical (0.69 and 0.70 in Natural Sciences, and 0.48 and 0.49 in Social Sciences, respectively), suggesting that students performed similarly in both skills. Nonetheless, these relatively acceptable results may be considered with care due to the high magnitude of the variance, standard deviation and range, which evidenced that scores were spread out over the range (except for the reading skill in the Social Sciences test).

An important gap could be observed with respect to these receptive skills and those involving written production and attention to the linguistic form, especially regarding the former (the mean fell to 0.40 in Natural Sciences and 0.32 in Social Sciences). However, as in the previous case, the scores, in general, presented great dispersion.

Lastly, the results in the written production contrasted significantly with those in the oral production skill. In this sense, performance was remarkably satisfactory, and interestingly, it represented the only case where data were higher in Social Sciences than in Natural Sciences. In fact, the low magnitude of the variance indicated that those students who carried out the speaking task scored almost equally high. Yet, although the participants were randomly selected, it should be borne in mind that the population was much smaller and therefore the numbers were not entirely comparable.

These findings are visually represented in the histogram of each skill (see Appendix C), where results by frequency and cumulative percentage in each range of 0.1 are shown. Briefly commenting on them, in the reading task of Natural Sciences (Figure C1), the cumulative percentage slowly rose to the value 0.7, where the curve steepened, while its counterpart in the Social Sciences test (Figure C2) did not draw a curve as results tended to cluster around 0.5. In the listening skill, both histograms (Figures C3 and C4) showed that the results were either very low (between 0 and 0.3) or very high (between 0.7 and 1). A similar situation was found in the writing and focus-on-form activities (Figures C5, C6, C7 and C8), although in the Social Sciences test, the low values (between 0 and 0.3) were especially populated in the writing, so that the cumulative percentage was almost horizontal. Finally, the speaking skill contrasted significantly as its cumulative percentage rose sharply from 0.4 in Natural Sciences (Figure C9) and 0.6 in Social Sciences (Figure C10).

5. Discussion

This study has focused on CLIL primary education students' ability to comprehend and produce target discourse functions and features in different genres and subjects, i.e., their academic language competence. The results broadly showed that students' linguistic resources were sufficient for descriptions and explanations, although with high diversity, and limited for narrations and argumentations. Also, their written performance displayed the lowest numbers in both tests, whereas the highest scores were found in the reading, listening and speaking tasks.

In this section, the implications of these outcomes will be presented regarding each of the specific objectives. Firstly, answers in each test will be analysed and described

independently, using examples from the corpus to exemplify adequate and unsatisfactory responses to the prompt, and secondly, the comparative analysis of the quantitative results in terms of genres and language skills will be discussed considering prior investigation.

1. **To estimate the appropriacy of learners' linguistic choices according to the genre in question and illustrate it with examples from the corpus.**

In the Natural Sciences test, the first two tasks were aimed at assessing students' receptive competences (written and oral comprehension) by means of texts that describe and explain the Solar System, using a range of discourse features (e.g., relative clauses, comparative and superlative adjectives, place adverbs, rhetoric figures, sequence and logical connectors…), and in the questions, they were given the description of a celestial body for them to name it (e.g., Which is the name of the star that produces light and heat?). The results show that, in general, they successfully manage to understand the descriptions and identify the correct answer. This suggests that students in this context are familiarised to input in the form of description and/or explanation and do not seem to present problems in understanding oral or written texts with such features.

However, the situation changes dramatically when they face a descriptive and explicative written production. In this case, the prompt elicited short sentences describing some celestial bodies that appeared in the previous activities. The excerpts below show examples of two responses from student A and B (what they wrote is highlighted in bold):

- Student A. The Sun: It's **very hot**. It's **is colour orange**. It's **is big**. The Earth: It's the **is persons viving**. Jupiter: It's the **second planet**.
- Student B. The Sun: It's **a star**. It's **the centre of the solar system**. It's **bigger than any planet**. It gives **light and heat**. The Earth: It's the **third planet from to the sun**. It's called **the blue planet**. It's the only planet **that supports life.** It has **one natural satellite**. Jupiter: It's **the biggest planet in the solar system**. second **ice planet.** It's made of **gases**. It has **67 natural satellites**.

In student A's response, it can be noted that this student left unanswered many prompts and failed to produce a sufficient range of target features – he just used the present tense (simple and continuous), but not accurately. His vocabulary was rather poor ("viving" is

not an English word) and his syntax was frequently broken (he constantly repeats the verb 'to be'). As a result, his description was not comprehensible and elaborated. The answer of this learner was considered not satisfactory. On the other hand, student B not only provided a response to every prompt, but also, he accurately used an extensive set of appropriate discourse features: present tenses, relative clauses, complex sentences, comparative and superlative adjectives, etc. Moreover, he displayed a proper style and register, his vocabulary was rich, and he showed no mistake in the use of syntax. The result was a full and well-elaborated description; hence, his answer was judged to be excellent.

A common problem among students revealed in this activity is confusion between comparative and superlative adjectives, which in fact is again evidenced in the focus on form task, where they had to choose between one or the other depending on the context. As the results show, their performance in this respect is not entirely satisfactory.

Finally, regarding the speaking test, scores are again outstanding, reaching close to the level of the receptive skills activities. In this case, students had to describe the effects of a severe drought and explain possible solutions. Two excerpts from students' responses are provided below to reflect their linguistic choices (student C and D) and to justify the degree of appropriacy:

- Student C. Hello, guys. My name is X. I'm from Spain, in Cordoba. My school is (…). I talk solutions the saving the sequía. Primary, is the saving water, don't the unnecessary water (…). In conclusion, no the saving water unnecessary. Thanks for you see video.
- Student D. Severe drought has many effects. It is bad because if the severe drought is very big, we can die. The effects can be more climate change, there isn't enough water to drink or clean ourselves, so we will get very dirty, there won't be lakes, rivers (…). To help the severe drought, we cannot waste a lot of water when we are in the shower or when we clean the plates, don't use the car to avoid polluting the air (…).

As can be seen, these two examples are completely different. Learner C's response was interesting in terms of academic language, since he seemed to be more comfortable using the language to introduce himself, i.e., for more informal purposes. Thus, when he tried to address the answer to the question, he was not able to elaborate a coherent

explanation or description. Even though an attempt to order the discourse could be glimpsed through two connectors, they were not used correctly (primary instead of firstly), he lacked fluency and vocabulary, and his syntax was fragmented. Moreover, he used virtually no target discourse features. Therefore, his answer was deemed not satisfactory. Nonetheless, student D provided an excellent answer. He accurately produced an extensive range of target features: he expressed consequence (so), purpose (to avoid), cause-effect (because), there is/are, conditionals, etc. The result was an elaborated description of the effects of a severe drought and explanation of how it could be solved, using rich vocabulary and without significant syntactic mistakes.

Concerning the Social Sciences tests, students' receptive skills were assessed by means of questions 1 and 3. The first one consisted of a listening activity narrating what is happening as a result of the climate change and arguing why it is dangerous, in which students had to fill in the gaps in a text, while the third one is a reading about Greta Thunberg's life, based on which students had to decide if prompts were true or false. The students in this case seem to present comprehension problems since the overall results are low. This is probably because they are not used to deal with input presenting the syntactic-discursive characteristics of this type of text, such as qualifiers, expression of possibility, tense sequences, etc. Similarly, regarding the focus on form activity, students were supposed to match some terms with their definitions. In other words, they were asked to identify the *definiendum* based on the *definiens* followed by a relative clause that specifies its features, which is the form of a proper definition according to Dalton-Puffer (2016). However, their performance evidence some gaps in this regard.

It makes sense, therefore, that the answers in the written production question, in general, are not entirely satisfactory either. Here students had to argue the risks of climate change using conditionals (e.g., If earth gets warmer, plants will die). Once more, the passages below show two of students' answers (students E and F):

- Student E. If didn't person can life in cities. If the animals not can do respiration correct because will carbon dioxide. If the weather is no correct will because global warming. If the water no correct the moment can be the animals.
- Student F. If the planet gets warmer, cities won't have enought water to survive and cities will get warmer. If the planet gets too warm, the habitat of the animals

can be destroyed. If the climate change occurs, there will be stronger storms. If the planet gets warmer, the water will rise a lot because ice is melting.

Student E's answer was not satisfactory because he could not present and reason his ideas in a comprehensible way using target discourse features. He tried to use the first conditional, but it was not correct or understandable due to his scarce vocabulary and shattered syntax. However, student F performed excellently, despite the misspelled word: The vocabulary used in his arguments was rich (habitat, melting, survive, enough), the syntactic constructions showed no significant mistakes, he presented his views through analyses and evaluations, and he used a wide range of syntactic-discursive features, such as cause/effect expressions (first conditional, because), variety of simple tenses, intensifiers (too) and probability (can).

Finally, as far as the speaking activity is concerned, the performance is completely different, being the task with the best results in the whole test (although it should be remembered that only 12 students performed it). In this case, they were asked to give reasons for four climate change effects in their country. The following two answers exemplify students' linguistic choices:

- Student G. Hi, guys. I'm going to explain the climate change in my country. In Cordoba, the weather in summer is very hot and in winter is very cold. The climate change produces the fires, the drought and animals die. Bye.
- Student H. Hello. I'm going to tell you about the problems climate change is producing in my country. I'm going to explain you four because these are the most important ones in here. The first one is that some plants are blooming before, and other plants are drying. This is a very big problem because (…). The last and the most important problem is that we are having less and less water every year.

As can be seen, the response of student H is higher level than that of student G. Student H used a very rich vocabulary in his arguments (dry, blooming, health problems, carbon dioxide...), the syntactic constructions showed no significant mistakes, and he presented his views through analyses and evaluations (this is a very big problem because...). His ideas were clearly organised by means of discourse markers (the first one, the second one...), and he used quantifiers (very) and superlative adjectives to emphasise his opinion (this is the most important problem), along with cause/effect expressions (this is

because...). Thus, his answer was excellent. On the other hand, student G did not give such an elaborate argumentation and his vocabulary and syntactic resources were limited due to his short answer. However, his speech was organised and free of serious grammatical errors, so his response was considered weak rather than unsatisfactory.

2. **To identify the textual genre(s) where students present the greatest and least learning needs.**

Overall results in the Natural Sciences test (description and explanation) are significantly higher than in the Social Sciences test (narration and argumentation), which is in line with the study conducted in secondary education by Falcón & Lorenzo (2015) where they observed that the peak of students' syntactic complexity when expressing academic content related to Social Sciences in their foreign language is reached in the textual genre of explanation.

Specifically, the cognitive-discourse function of a descriptive text is to tell details of what can be seen (Dalton-Puffer, 2016). In this sense, most students show a good command of syntactic linguistic resources such as relative clauses and there is/are expressions, although they encounter greater problems in expressing verbs in their comparative and superlative forms. On the other hand, the textual genre of explanation presents the communicative intention of giving reasons or telling the causes for something (Dalton-Puffer, 2016). Students seem to be able to express cause through the conjunction "because" but find it more difficult to express consequence (so), to give recommendations using modal verbs in addition to "can" (such as should, could...), and to organise discourse appropriately. All these limitations hindrance some students' academic literacy development, in this case, carrying out adequate descriptions or explanations. However, it should be noted that this situation does not apply to all students, as those with a wide range of discourse features produced excellent descriptions and explanations. Thus, as was observed in the results section with the standard deviation, there is great diversity in terms of students' abilities in this respect. Furthermore, at a lexical level, in general they show rich vocabulary as they use different adjectives to modify nouns, which conforms to the results from study conducted by Lorenzo & Rodríguez (2014). Hence, as these authors suggest, it seems that lexical diversity develops before grammatical accuracy and "paves the way for advanced syntactic structures" (p. 70).

As for narration and argumentation, the results are more consistent but much lower. This may be because, according to Coffin (2006), argumentative genres involve a higher level of cognitive development. Her work, focusing on the discipline of History, shows that genres range from chronological narrations, through explanations and finally argumentations. This idea is also shared by Lorenzo (2013), who claims that the linguistic demands of this genre hinder the capacity to produce a text of this kind at intermediate levels, and by Morton & Llinares (2018), who also state, regarding evaluations, that there seems to be a component of cognitive involvement that has a strong impact on learners' linguistic expression. Thus, it is not surprising that primary education learners perform better overall in the less-cognitively demanding description/explanation test. Another fact that may explain these results is that the genres of description and explanation are more frequent in daily classroom practice than that of argumentation, hence, students seem to be more familiarised with these discourse functions (Dalton-Puffer, 2007).

In particular, pupils when arguing present difficulties especially in using simple conditionals to express consequence (hence the low results in the written production test) and in making suggestions using modal verbs. On these points, therefore, both tests coincide and reveal academic language needs of students across genres and subjects. Another worth mentioning fact is related to the communicative intention of an argumentative text, which belongs to the cognitive-discourse function of EVALUATE: I tell you what my position is vis a vis X (Dalton-Puffer, 2016). In science, as in other academic disciplines, the authors' position is presented as a fact rather than an opinion (Zwiers, 2014). Nonetheless, learners' arguments in the test are very much related to their appreciations (good, bad, important…) rather than objective reasons more suitable for academic contexts, which was also observed in the research by Whittaker & McCabe (2020). The authors suggested that this could be due to the fact that the academic task was presented as an everyday activity in the students' lives, which may have influenced their register. However, in the present test, the instructions are clearly objective and academic, so this seems to indicate that learners, especially in the primary school, always tend to make affective evaluations and argue using the first person regardless of the register presented in the prompt.

Perhaps the main reason is that they do not yet possess the necessary linguistic tools to express objectivity, such as the passive voice, nominalisations, grammatical metaphors,

etc. (Zwiers, 2014). This idea is also noted by Lorenzo & Rodríguez (2014) and Lorenzo et al. (2019), who observed that the linguistic resources for dealing with generalisation, evaluation and argument develop over time. In fact, Morton & Llinares (2018) found that linguistic abstraction is challenging even for more mature writers. Another test result that is in line with Lorenzo & Rodríguez's (2014) conclusions is that, in the textual genre of narration, students show hardly any tense sequence. Nevertheless, participants ordered the speech more logically and accurately in this test than in the previous one, probably because the question was phrased in a more explicit way (they were asked to argue for four climate change effects).

Finally, although it was not one of the objectives of the present project to examine students' responses from the point of view of content knowledge, it should be noted that those children who were able to use more cognitive-discourse features appropriately, delved into the subject-matter clearly and elaborately, and the other way around, as language deficits hampered demonstration of knowledge. Therefore, the impact of academic language on content performance is evident, as noted also by Lo & Jeong (2018), Whittaker & McCabe (2020) and Lorenzo et al. (2019). This clearly reflects that explicit attention to students' academic language is the way towards an integrated assessment of language and content.

In short, the results were as expected in terms of the most challenging discourse genre (argumentation) and the least demanding ones (description/explanation), but not in terms of the overall level, which can be considered low. It should be borne in mind that the participants in this study are primary school students, and as Pérez Cañado (2018) claims, "time is a crucial factor to ascertain the effects of CLIL on foreign language attainment" (p. 61), an idea on which Nieto Moreno de Diezmas (2016) and other papers related to academic language seem to agree, such as Navés (2011), Lorenzo & Rodríguez (2014), Lorenzo et al. (2019), Nashaat-Sobhy & Llinares (2020), Granados et al. (2022), etc., since all of them resolve that academic language levels are higher the older students are. While it is true that the age factor may be important, the tests in this study were adapted to the cognitive and linguistic levels of students. Thus, time should not be relied upon, and primary education learners should be taught in an explicit and scaffolded way what is expected of them in terms of academic language from the very beginning. This idea, together with pedagogical suggestions, will be further developed in the conclusion section.

3. **To identify the communicative skill(s) in which they are most and least proficient when using academic language.**

With respect to specific communicative skills, the results in this study conform to most research into the impact of CLIL on students' general language proficiency in that the skill that is most favoured is oral production, i.e., speaking (e.g., Dalton-Puffer, 2011; Lasagabaster, 2008; Nieto Moreno de Diezmas, 2016), although it should be remembered that the sample was much smaller. Spoken answers were longer and more elaborated, which had a great impact on their grading. This is in line with Morton's (2020b) observation that primary education students use more expansions when speaking than when writing. Thus, it ascertains the fact that CLIL boosts learners' oral fluency, both in its general and academic dimension. The oral production skill is followed by the oral comprehension one (listening), where participants also obtained high scores. In this regard, findings in the academic literature are mixed: the present results are consistent with those of Lasagabaster (2008), who reported positive data on students' listening skill, whereas Navés (2011) did not draw such conclusions.

The favourable outcomes in the oral activities (speaking and listening) may be due to strategies employed for the co-construction of knowledge by teachers and students, as the latter seem to be used to dealing with academic content in English orally, probably thanks to the interaction that occurs during the lesson. Studies focusing on conversation analysis, such as the ones conducted by Dalton-Puffer (2007) and by Escobar Urmeneta & Evnitskaya (2013), reveal that the cognitive involvement demanded by oral discourses in CLIL positively feeds back into second language acquisition. Although these works refer to general English proficiency, the outcomes in this project suggest that this effect is also present in terms of academic language, which may explain why the skills most mastered by students were oral comprehension and production.

After speaking and listening, the reading skill follows closely behind. Focusing on receptive abilities (reading and listening), the results are very similar to those obtained by Dalton-Puffer (2007), who stated that CLIL students benefit especially regarding their comprehension abilities, and by San Isidro & Lasagabaster (2019), who also reported high outcomes in the reading task. As far as academic language is concerned, this means that learners are used to academic input and do not seem to present serious problems in understanding this register in the oral or written form.

However, contrary to the favourable results usually yielded from most research into learners' writing outcomes (e.g., Dalton-Puffer, 2011; Lasagabaster, 2008), the achievement of participants with respect to academic writing was not satisfactory enough, as also observed by Vollmer (2008), nor was their performance in the focus-on-form activity. According to Pérez Vidal (2007), a focus on form involves attention "towards language form in order to develop linguistic awareness which may result in uptake and subsequently intake" (p. 44). In this context, it seems that participants' attention is drawn to content instead of language, which is a communicative resource rather than an end in itself, as noted by Lorenzo & Rodríguez (2014). This makes sense with current approaches to the teaching of foreign language from a communicative point of view through teacher-student interactions or between learners through pair/group work, where the language is not restricted to certain structures, but ad-lib use is enhanced to achieve the communicative purpose at the expense of the linguistic accuracy. Nevertheless, according to Whittaker & McCabe (2020), a more linguistically oriented lesson, where the CDF of the writing task and the linguistic resources for performing it successfully are explicitly linked and explained to students, is likely to have a strong positive impact on pupils' writing skills.

Briefly, the results of this test focusing on students' academic language are in line with most studies dealing with their general language proficiency in speaking, listening and reading. Thus, it seems that CLIL enhances both the academic and general oral skills of primary school learners, and that it also has a positive effect on their reading ability. Nonetheless, their linguistic awareness and writing competence do not coincide with research in this regard, as these seem to be lagging behind. Therefore, there is a gap between general and academic writing skills that needs to be overcome, probably by means of a focus-on-form instruction. Further pedagogical implications of these observations will be presented in the next section.

6. Conclusions and pedagogical suggestions

The present project has focused on the academic linguistic dimension of CLIL. Firstly, a review of the underpinnings of academic language has been provided so as to justify why the systemic-functional notion of genre, together with the cognitive-discourse

functions framework developed by Dalton-Puffer (2013), can offer a structured model for teachers to integrate content and language in bilingual education. Its transdisciplinary character allows CLIL practitioners to match the textual genre or the cognitive-discourse functions with the necessary linguistic resources according to each field of knowledge, thus developing students' subject-specific literacy and overall academic language proficiency. This process is crucial for performing the thinking operations involved in meaning-making and knowledge construction in academic settings. Therefore, a focus on academic language through a genre-based pedagogy paves the way for an effective content and language teaching and learning in CLIL.

One of the potentialities of such approach is related to the assessment of students' academic language outcomes in connection with disciplinary content, which represents the specific topic under scrutiny in this study. Through the creation of original tests, specifically designed for pupils in the last year of primary education, a case study has been carried out with the general goal of assessing participants' academic linguistic proficiency. The specific objectives were as follows: to estimate the appropriacy of learners' linguistic choices according to the genre in question and illustrate it with examples from the corpus, to determine the textual genre(s) where students present the greatest and least learning needs, and to identify the linguistic skill(s) in which they are most and least proficient when using academic language.

The overall conclusion of the present case study is threefold: Firstly, that analysing the academic language competence of students allows for the identification of learning needs in terms of both language and content, i.e., to carry out an objective and systematic assessment in different subjects, and thus to address them in a more targeted way. Secondly, that academic language proficiency of the learners in this study is not entirely satisfactory and seems to be still developing, as the results were lower than expected. Therefore, they need to improve their command of the cognitive-discourse functions involved in the different genres, mainly regarding narration and argumentation, but also with respect to certain features and functions of the four genres observed. And thirdly, there is great asymmetry between students' written and spoken dimension of academic English in favour of the latter. Thus, as mentioned before, despite the importance of encouraging oral interaction in the classroom, language should also be considered as an end in itself. The requirement of explicitly addressing

the linguistic form in writing cannot be underestimated, so that learners are provided with the discourse resources that are critical for academic success.

In particular, according to the findings with regard to the first specific objective, it seems that the most satisfactory answers are long and present a wide range of target discourse features, whereas the poorest answers are short and linguistically limited. Moreover, when primary school students are dealing with description and explanation genres, their linguistic choices are much related with present tenses, positive adjectives and relative clauses, and they show an acceptable lexical richness. Nonetheless, they seem to lack a solid linguistic basis for comparing (comparative and superlative adjectives), expressing consequence (so, therefore), expressing recommendations (modal verbs) and for ordering discourse coherently. These results conform to the research conducted by Lorenzo & Rodríguez (2014) in that lexical diversity occurs before syntactic complexity. Considering the corpus from the test on narration and argumentation genres, some of the same shortcomings can be also identified (consequence and recommendations), thus revealing a linguistic need across genres and disciplines. Other limitations are found regarding tense sequences or evaluations, as students tend to be too appreciative and subjective, as Whittaker & McCabe (2020) also witnessed in their study. Nonetheless, the ordering of speech is less problematic in these genres, probably due to the languaging of the prompt. This analysis is in line with most investigation into learners' academic linguistic resources, especially in two respects: that young students struggle to deal with academic language, which seems to develop over time (Granados et al., 2022; Lorenzo et al., 2019; Nashaat-Sobhy & Llinares, 2020), and that expressing abstraction, generalisation and evaluation, regardless of the age, genre or subject-matter, seems to be challenging (Lorenzo & Rodríguez, 2014; Morton & Llinares, 2018). Thus, this sheds light on learners' language needs, both at a general and subject-specific level, as well as implications for improvement.

Regarding the second specific objective, it was observed that the performance in the genres of description and explanation is higher overall than in argumentation and narration. On the one hand, a better performance when describing and explaining may be due to the fact that they are more commonly used in daily classroom practice (Dalton-Puffer, 2007). On the other hand, argumentation involves a greater level of cognitive abstraction, which hinders students' capacity of arguing appropriately (Lorenzo, 2013; Morton & Llinares, 2018). In fact, argumentative genres more

frequently demand linguistic resources that, as suggested before, young learners do not seem to possess, such as modal verbs or nominalisations. Thus, it is not surprising that learners tend to perform best at description and explanation, which is consistent with the results observed by Falcón & Lorenzo (2015), while argumentation is the most challenging genre for them. Nevertheless, it is also noteworthy that description and explanation performance shows greater disparity among pupils' level than argumentation and narration. This suggests that more explicit scaffolding that caters for the linguistic diversity of learners is needed, in order to ensure that no one is left behind, together with greater support regarding those genres that seem to be more demanding through input, output and feedback.

The data relating to the third specific objective led to the conclusion that students' performance in oral and receptive skills conforms to most studies on general language outcomes from CLIL learners in the sense that this approach seems to enhance speaking and listening abilities, especially the first one, both in their general (Lasagabaster, 2008; Nieto Moreno de Diezmas, 2016) and academic dimensions (Morton, 2020b). The higher results in oral activities are reasonable as students frequently use the language to construct knowledge with both teachers and peers (Dalton-Puffer, 2007). The reading activity also reports considerably satisfactory results, as in the study carried out by San Isidro & Lasagabaster (2019). Nonetheless, what contradicts the positive observations often found in the scientific literature, such as the research conducted by Dalton-Puffer (2011) or Lasagabaster (2008), are the results in terms of writing production and focus on form, where the gap between the academic and general language of students seems more evident. Consequently, it is argued that these two skills are closely related and that more attention to the linguistic form (features, functions and structure of genres) in classroom will have a positive impact on the improvement of writing skills, as also advocated by Pérez Vidal (2007) and Whittaker & McCabe (2020).

Based on these findings, a number of pedagogical suggestions are proposed below for CLIL stakeholders with regard to three dimensions: educational policies, school management, and classroom methodology and materials. Concerning the first dimension, curriculum design should be grounded on genre-based models, linking the CDFs of the genres to the language needed to accomplish them (Morton, 2010; Whittaker & McCabe, 2020). In this sense, Lo & Jeong (2018), based on Lin (2016), propose a "horizontal curriculum mapping", where the same genre is taught at the same

grade level but in different subjects, each providing its own specific literacy (e.g., a description can be part of a report on a scientific experiment, a characterisation of a painting in Art, etc.), and a "vertical curriculum mapping", where genres in the same subject are introduced gradually, even in different years (e.g., explanation is introduced first and revisited the following year, but argumentation is introduced later) (p. 44). As observed in this study, argumentation is more challenging for students than other genres such as description or explanation, so perhaps a staggered and coordinated introduction of the genres can be of great help for them.

This genre-based curriculum, which envisages the coordination of the different subjects together with the connection between the different school years, should come from the educational institutions in charge of its design. However, it can hit a brick wall if teachers are not properly prepared to take on a genre-based pedagogy that addresses students' academic literacy. That is why teacher training should be both linguistic and methodological and involve all content and language departments (Alcaraz Mármol, 2018; Pavón & Ellison, 2013). For instance, linguistically, conversation analysis studies reveal little or no meta-talk on academic linguistic features and functions, despite its importance in meaning making (Dalton-Puffer, 2016; Whittaker & McCabe, 2020). Therefore, it seems that teachers need conversational strategies that allow them to provide quality, abundant and explicit input, since as Escobar Urmeneta & Evnitskaya (2013) point out, these strategies "determine the quality of each conversation and its outcomes in terms of affordances for the integrated learning of content and language" (p. 111). Methodologically, it would be useful to develop what Berger (2012) calls "assessment literacy" (p. 57) but in the specific context of CLIL, i.e., how they can assess and scaffold students' academic language and subject-specific literacy in order to provide them with effective, well-grounded and integrated feedback that promotes their pluriliteracies.

An underlying idea in what has been suggested is that teacher collaboration is essential for CLIL to be successful, so both educational policies and schools' management should consider the time and resources that teachers need to be coordinated (Szczesniak & Muñoz Luna, 2022). Thus, in relation to the second dimension, school management should plan teachers' cooperation at three levels, according to Pavón et al. (2015): linguistic assistance for content subjects, coordination between languages and coordination between content disciplines. Firstly, language teachers are responsible for

identifying the linguistic resources needed to acquire and express knowledge in different disciplinary subjects, so they can support the teaching of content instructors and the learning of students (Pavón & Ellison, 2013). A case in point would be a collaboratively creation of tasks, tests and assessment rubrics, such as those proposed in this study, which take into account how students integrate language and content more or less successfully according to the textual genres or the CDFs involved. Secondly, coordination between teachers from different language departments through a language integrated curriculum is also a crucial aspect. As noted by Pavón et al. (2015), the best approach is to work on similar linguistic resources and communicative strategies in the different linguistic subjects to support students' acquisition of academic literacy. For instance, as it was observed in this study that learners' weakest skill is the academic writing, it would be beneficial to implement a plan to promote reading and writing in the different languages aimed at working on textual genres, placing special emphasis on those cognitive-linguistic resources that are more demanding. In this way, learners will not only develop the academic dimension in each of the languages they use, but the information can be transferred between them, with the improvement of one being equally beneficial to the others. And finally, the coordination between content teachers is not only related to making connections between their subject-matter, but also to their consideration of language and methodology: using similar strategies and techniques to scaffold students' learning of content knowledge (Pavón et al., 2015).

Lastly, regarding the third dimension, the previously mentioned genre-based and coordinated curriculum will favour the creation of CLIL materials that support this approach (Ball, 2018). Pimentel & Pavón (2020) claim that these materials should give visibility to the language and place it "at the service of the comprehension and oral and written verbalization of the academic contents of the subjects" (p. 326). These authors also suggest several recommendations for the creation of CLIL materials, such as presentation of contents in relation to prior knowledge, authentic and motivating tasks, presentation of the necessary language to perform those tasks, working with text genres, scaffolding and assessment of language and content, etc. In turn, this will facilitate the work of instructors in designing their teaching, which is in fact one of the concerns raised in the study by Szczesniak & Muñoz Luna (2022), as genre-based materials will provide them with the resources to conduct a more language-centred content lesson. It will also benefit students' processing of language and content as they will be clearer

about what it is expected of them. As mentioned before, in hard CLIL, the role of language is to support content learning (Ball, 2018). Hence, materials design should begin with the identification of objectives regarding subject-matter knowledge and the tasks through which students will achieve them, and based on these, determine the linguistic resources that students will be using in terms of genres and CDFs so as to scaffold their subject-specific literacy (Ball, 2018; Nikula et al., 2016).

As a result, classroom methodology should be both focused on form and on meaning, and in the oral and written dimension of the languages. Content teachers must be aware of the academic features and functions according to the communicative intention of the discourse so as to be able to incorporate them in their speeches and support students' acquisition. An increase in language awareness will lead to sufficient and adequate input (Escobar Urmeneta & Evnitskaya, 2013), the development of students' consideration of the language as a means for learning (Pavón et al., 2015), and an appropriate assessment of the content and linguistic outcomes in an integrated way.

In this sense, for instance, Whittaker & McCabe (2020) argue that learners should be informed of the linguistic choices and type of content expected by the academic community (e.g., impersonal structures, specialised vocabulary, generalisations…). Morton (2020b) also suggests showing an anonymous example from a peer so that the class can jointly identify strengths and areas for improvement. Similarly, Breeze & Dafouz (2017) propose to give learners explicit instructions and models regarding the linguistic realisation of each CDF and how to link them together. Their study is related to English as a medium of instruction in higher education, but it is perfectly suitable for primary education: as observed in the results, a more precise and explicit statement helps pupils to order the discourse and be clearer about what they are supposed to do, even more so in primary school where students' cognitive development is still in process. Another practice that seems to be relevant according to the conclusions in this study is peer-collaboration or peer-assessment, as there were both outstanding and very poor answers within the same classroom. Hence, fast-finishers can help slow-learners to understand what they lack in linguistic terms, thus developing their own language awareness. Lastly, it may also be helpful to use student-friendly assessment rubrics that they can refer to before performing a task.

In short, this study hopes to have contributed to the research gap identified by Meyer et al. (2015), Pérez Cañado (2020) and San Isidro (2019), among others, regarding academic language and subject-specific literacies in CLIL. It has provided CLIL practitioners with guidelines and tools on how to support and assess students' academic language, in this case from primary education but equally relevant for secondary education, in order to carry out systematic, fair and integrated assessment by means of textual genres and CDFs in relation to subject-specific contents. As mentioned before, time cannot be relied upon as the remedy for the language needs of primary school students, but rather their resources for meaning making and knowledge verbalisation must be appropriately scaffolded from the outset according to their age, the subject-specific content and the genre, for which assessment of their learning is fundamental. This will improve the teaching and learning process in bilingual education.

Finally, some limitations must be acknowledged. Firstly, the results of this study should not be considered generalisations of students in CLIL due to the reduced number of the sample (84 participants, of which 30 performed the oral task) and the fact that they belonged to a very similar context, thus the situation may be different elsewhere. Secondly, this project has just analysed students' use of their foreign language regarding four genres in two similar disciplines (Natural and Social Sciences), but it would be useful to examine other academic areas and genres. Hence, future research may reduplicate this study so as to confirm the statistical relevance of the results and be able to draw far-reaching conclusions in terms of academic language practices and gaps in CLIL. It may also be interesting to consider a larger sample immersed in diverse contexts (e.g., in secondary education), and to examine other subject-specific literacies to get a more comprehensive view of students' academic language and pluriliteracies development. Further studies can also observe learners' use of the first and foreign language and compare the results in order to understand more precisely the transfer process that exists between the academic register of these two languages.

7. References

Achugar, M., & Carpenter, B. D. (2014). Tracking movement toward academic language in multilingual classrooms. *Journal of English for Academic Purposes, 14*, 60–71. https://doi.org/10.1016/j.jeap.2013.12.002

Alcaraz Mármol, G. (2018). Trained and non-trained language teachers on CLIL methodology: Teachers' facts and opinions about the CLIL approach in the Primary Education context in Spain. *Latin American Journal of Content & Language Integrated Learning, 1*, 39–64. https://doi.org/10.5294/laclil.2018.11.1.3

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. Allyn Bacon Longman.

Bailey, A., & Butler, F. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document*. University of California, Los Angeles: CSE Report.

Ball, P. (2018). Innovations and challenges in CLIL materials design. *Theory Into Practice, 57*, 222–231. https://doi.org/10.1080/00405841.2018.1484036

Ball, P., Kelly, K., & Clegg, J. (2015). *Putting CLIL into practice*. Oxford University Press.

Barrios, E., & Milla Lara, M. D. (2020). CLIL methodology, materials and resources, and assessment in a monolingual context: An analysis of stakeholders' perceptions in Andalusia. *Language Learning Journal*, *48*(1), 60–80. https://doi.org/10.1080/09571736.2018.1544269

Beacco, J. C. (2010). Items for a description of linguistic competence in the language of schooling necessary for learning/teaching history (end of obligatory education). An approach with reference points. In *Language and school subjects: Linguistic dimensions of knowledge building in school curricula*. Council of Europe.

Beacco, J. C., Fleming, M., Goullier, F., Thürmann, E., & Vollmer, H. (2015). *The language dimension in all subjects: A handbook for curriculum development and teacher training*. Council of Europe.

Bentley, K. (2010). *The TKT course CLIL module*. Cambridge University Press.

Berger, A. (2012). Creating language-assessment literacy: A model for teacher education. In J. Hünter, B. Mehlmauer-Larcher, S. Reichl, & B. Schiftner (Eds.), *Theory and practice in EFL teacher education: Bridging the gap* (pp. 57–82). Multilingual Matters.

Breeze, R., & Dafouz, E. (2017). Constructing complex Cognitive Discourse Functions in higher education: An exploratory study of exam answers in Spanish- and English-medium instruction settings. *System, 70*, 81–91. https://doi.org/10.1016/j.system.2017.09.024

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*(3), 243–262. https://doi.org/10.1093/applin/amt011

Coetzee-Lachmann, D. (2007). *Assessment of subject-specific task performance of bilingual geography learners: Analysing aspects of subject-specific written discourse*. Universität Osnabrück.

Coffin, C. (2006). Learning the language of school history: The role of linguistics in mapping the writing demands of the secondary school curriculum. *Journal of Curriculum Studies*, *34*(4), 413–429. https://doi.org/10.1080/00220270500508810

Coffin, C. (2017). Introduction to part II. Systemic functional linguistics: A theory for integrating content-language learning (CLL). In A. Llinares & T. Morton (Eds.), *Applied Linguistics Perspectives on CLIL* (pp. 91–103). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.47.06cof

Conserjería de Educación. (2013). *Guía informativa para centros de enseñanza bilingüe*. Junta de Andalucía. https://www.juntadeandalucia.es/export/drupaljda/Guia_informativa_centros_ense%C3%B1anza_bilingue_.pdf

Council of Europe. (2014). *Recommendation CM/Rec(2014)5 of the Committee of Ministers to member States on the importance of competences in the language(s) of schooling for equity and quality in education and for educational success*.

Coyle, D. (2015). Strengthening integrated learning: Towards a new era for pluriliteracies and intercultural learning. *Latin American Journal of Content and Language Integrated Learning*, *8*(2), 84–103. https://doi.org/10.5294/laclil.2015.8.2.2/84-103

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

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19*, 121–129.

Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.

Cummins, J. (2013). BICS and CALP: Empirical support, theoretical status, and policy implications of a controversial distinction. In M. R. Hawkins (Ed.), *Framing languages and literacies: Socially situated views and perspectives* (pp. 10–23). Taylor & Francis Group.

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) classrooms* (Vol. 20). John Benjamins Publishing Company.

Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of Applied Linguistics*, *31*, 182–204. https://doi.org/10.1017/S0267190511000092

Dalton-Puffer, C. (2013). A construct of cognitive discourse functions for conceptualising content-language integration in CLIL and multilingual education. *European Journal of Applied Linguistics*, *1*(2), 216–253. https://doi.org/10.1515/eujal-2013-0011

Dalton-Puffer, C. (2016). Cognitive discourse functions: Specifying an integrative interdisciplinary construct. In T. Nikula, E. Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising integration in CLIL and multilingual education* (pp. 29–54). Multilingual Matters.

Dalton-Puffer, C., Bauer-Marschallinger, S., Brückl-Mackey, K., Hofmann, V., Hopf, J., Kröss, L., & Lechner, L. (2018). Cognitive discourse functions in Austrian

CLIL lessons: Towards an empirical validation of the CDF construct. *European Journal of Applied Linguistics*, *6*(1), 5–29. https://doi.org/10.1515/eujal-2017-0028

Dalton-Puffer, C., Nikula, T., & Smit, U. (2010). Language use and language learning in CLIL: Current findings and contentious issues. In C. Dalton-Puffer, T. Nikula, & Smit U (Eds.), *Language use and language learning in CLIL classrooms* (pp. 279–293). John Benjamins Publishing Company.

Dalton-Puffer, C., & Smit, U. (2013). Content and language integrated learning: A research agenda. *Language Teaching*, *46*(4), 545–559. https://doi.org/10.1017/S0261444813000256

Escobar Urmeneta, C., & Evnitskaya, N. (2013). Affording students opportunities for the integrated learning of content and language: A contrastive study on classroom interactional strategies deployed by two CLIL teachers. *Reviving Catalan at School: Challenges and Instructional Approaches*, 158–182. https://doi.org/10.21832/9781783090266-010

Evnitskaya, N., & Dalton-Puffer, C. (2020). Cognitive discourse functions in CLIL classrooms: Eliciting and analysing students' oral categorizations in science and history. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2020.1804824

Falcón, E., & Lorenzo, F. (2015). El desarrollo de la sintaxis compleja en L1 y L2 en entornos educativos bilingües (CLIL). *E-Aesla*, *1*, 1–9.

García, O., Bartlett, L., & Kleifgen, J. A. (2007). From biliteracy to pluriliteracies. In P. Auer & L. Wei (Eds.), *Handbook of Applied Linguistics, vol. 5: Multilingualism* (pp. 207–228). Mouton De Gruyter.

Gottlieb, M., & Ernst-Slavit, G. (2014). *Academic language in diverse classrooms: English Language Arts, grades 6–8*. Sage Publishing.

Granados, A., López-Jiménez, M. D., & Lorenzo, F. (2022). A longitudinal study of L2 historical writing: Lexical richness and writing proficiency in content and language integrated learning. *Ibérica*, *43*, 129–154. https://doi.org/10.17398/2340-2784.43.129

Halbach, A. (2018). A literacy approach to language teaching: A proposal for FL teaching in CLIL contexts. *Pulso, 41*, 205–223. http://hdl.handle.net/10017/35322

Hallet, W. (2012). Semiotic translation and literacy learning in CLIL. In D. Marsh & O. Meyer (Eds.), *Quality interfaces: Examining evidence & exploring solutions in CLIL* (pp. 175–190). Eichstaett Academic Press.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education*, *5*(2), 93–116. https://doi.org/10.1016/0898-5898(93)90026-7

Halliday, M. A. K., & Matthiessen, M. I. M. (1999). *Construing experience through meaning: A language-based approach to cognition*. Continuum.

Halliday, M. A. K., & Matthiessen, M. I. M. (2014). *An introduction to functional grammar* (Fourth Edition). Routledge.

Hönig, I. (2010). *Assessment in CLIL: Theoretical and empirical research*. VDM Verlag Dr. Müller.

Krashen, S., & Brown, C. L. (2007). What is academic language proficiency? *Language & Communication Review*, *6*(1), 1–4. www.joanwink.com/newsletter/2007/Krashen-Brown-ALP.pdf

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, *1*, 31–42. https://doi.org/10.2174/1874913500801010030

Lemke, J. (1990). *Talking science: Language, learning, and values*. Ablex.

Leontjev, D., & deBoer, M. (2020). Conceptualising assessment and learning in the CLIL context. An introduction. In M. deBoer & D. Leontjev (Eds.), *Assessment and learning in Content and Language Integrated Learning (CLIL) classrooms: Approaches and conceptualisations* (pp. 1–28). Springer.

Leung, C., & Morton, T. (2016). Conclusion: Language competence, learning and pedagogy in CLIL – deepening and broadening integration. In T. Nikula, E.

Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising integration in CLIL and multilingual education* (pp. 235–248). Multilingual Matters.

Lin, A. M. Y. (2016). *Language across the curriculum: Theory and practice*. Springer.

Lin, A. M. Y. (2020). Academic languages and literacies in content-based education in English-as-an-additional-language contexts. *Oxford Research Encyclopedia of Education*, 1–15. https://doi.org/10.1093/acrefore/9780190264093.013.224

Llinares, A., & McCabe, A. (2020). Systemic functional linguistics: The perfect match for content and language integrated learning. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2019.1635985

Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of languages in CLIL*. Cambridge University Press.

Lo, Y. Y., & Fung, D. (2018). Assessments in CLIL: The interplay between cognitive and linguistic demands and their progression in secondary education. *International Journal of Bilingual Education and Bilingualism*, *23*(10), 1192–1210. https://doi.org/10.1080/13670050.2018.1436519

Lo, Y. Y., & Jeong, H. (2018). Impact of genre-based pedagogy on students' academic literacy development in Content and Language Integrated Learning (CLIL). *Linguistics and Education*, *47*, 36–46. https://doi.org/10.1016/j.linged.2018.08.001

Lorenzo, F. (2013). Genre-based curricula: Multilingual academic literacy in content and language integrated learning. *International Journal of Bilingual Education and Bilingualism, 16*(3), 375–388. https://doi.org/10.1080/13670050.2013.777391

Lorenzo, F., Granados, A., & Ávila, I. (2019). The development of cognitive academic language proficiency in multilingual education: Evidence of a longitudinal study on the language of history. *Journal of English for Academic Purposes*, *41*. https://doi.org/10.1016/j.jeap.2019.06.010

Lorenzo, F., & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System, 47*, 64–72. https://doi.org/10.1016/j.system.2014.09.016

Marsh, D. (2002). *CLIL/EMILE. The European dimension. Actions trends and foresight potential*. University of Jyväskylä.

Marsh, D. (2006). *English as medium of instruction in the new global linguistic order: Global characteristics, local consequences*. UNICOM, Continuing Education Centre, University of Jyväskylä.

Marsh, D. (2013). *The CLIL trajectory: Educational innovation for the 21st century iGeneration*. Servicio de Publicaciones de la Universidad de Córdoba.

Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, *20*(1), 10–21. https://doi.org/10.1016/j.linged.2009.01.003

Massler, U., Stotz, D., & Queisser, C. (2014). Assessment instruments for primary CLIL: The conceptualisation and evaluation of test tasks. *Language Learning Journal*, *42*(2), 137–150. https://doi.org/10.1080/09571736.2014.891371

Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL: Content and language integrated learning in bilingual and multilingual education*. Macmillan Education.

Meyer, O., & Coyle, D. (2017). Pluriliteracies teaching for learning: Conceptualizing progression for deeper learning in literacies development. *European Journal of Applied Linguistics*, *5*(2), 199–222. https://doi.org/10.1515/eujal-2017-0006

Meyer, O., Coyle, D., Halbach, A., Schuck, K., & Ting, T. (2015). A pluriliteracies approach to content and language integrated learning – mapping learner progressions in knowledge construction and meaning-making. *Language, Culture and Curriculum, 28*(1), 41–57. https://doi.org/10.1080/07908318.2014.1000924

Mohan, B., Leung, C., & Slater, T. (2010). Assessing language and content: A functional perspective. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 217–240). Multilingual Matters.

Moore, J., & Schleppegrell, M. (2019). A focus on disciplinary language: Bringing critical perspectives to reading and writing in science. *Theory into Practice*, *59*(1), 99–108. https://doi.org/10.1080/00405841.2019.1685337

Morgan, C. (2006). Appropriate language assessment in content and language integrated learning. *Language Learning Journal*, *33*(1), 59–67. https://doi.org/10.1080/09571730685200121

Morton, T. (2010). Using a genre-based approach to integrating content and language in CLIL: The example of secondary history. In C. Dalton-Puffer (Ed.), *Language use and language learning in CLIL classrooms* (pp. 81–104). John Benjamins Publishing Company.

Morton, T. (2017). Reconceptualizing and describing teachers' knowledge of language for content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, *21*(3), 275–286. https://doi.org/10.1080/13670050.2017.1383352

Morton, T. (2020a). Afterword: SFL, theoretical pluralism and content and language integration at the levels of curriculum, pedagogy and assessment. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2020.1806781

Morton, T. (2020b). Cognitive discourse functions: A bridge between content, literacy and language for teaching and assessment in CLIL. *CLIL. Journal of Innovation and Research in Plurilingual and Pluricultural Education*, *3*(1), 7–17. https://doi.org/10.5565/rev/clil.33

Morton, T., & Llinares, A. (2018). Students' use of evaluative language in L2 English to talk and write about history in a bilingual education programme. *International Journal of Bilingual Education and Bilingualism*, *21*(4), 496–508. https://doi.org/10.1080/13670050.2016.1192101

Nashaat-Sobhy, N., & Llinares, A. (2020). CLIL students' definitions of historical terms. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2020.1798868

Navés, T. (2011). How promising are the results of integrating content and language for EFL writing and overall EFL proficiency? In Y. Ruiz de Zarobe, J. M. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning: Contributions to multilingualism in European contexts* (pp. 103–128). Peter Lang.

Nieto Moreno de Diezmas, E. (2016). The impact of CLIL on the acquisition of L2 competences and skills in primary education. *International Journal of English Studies*, *16*(2), 81–101.

Nikula, T., Dalton-Puffer, C., Llinares, A., & Lorenzo, F. (2016). More than content and language: The complexity of integration in CLIL and bilingual education. In T. Nikula, C. Dalton-Puffer, A. Llinares, & F. Lorenzo (Eds.), *Conceptualising integration in CLIL and multilingual education* (pp. 1–25). Multilingual Matters.

Otto, A. (2018). Assessing language in content and language integrated learning: A review of the literature towards a functional model. *Latin American Journal of Content & Language Integrated Learning*, *11*(2), 308–325. https://doi.org/10.5294/laclil.2018.11.2.6

Otto, A., & Estrada, J. L. (2019). Towards an understanding of CLIL assessment practices in a European context: Main assessment tools and the role of language in content subjects. *CLIL. Journal of Innovation and Research in Plurilingual and Pluricultural Education*, *2*(1), 31. https://doi.org/10.5565/rev/clil.11

Oxbrow, G. L. (2018). Students' perspectives on CLIL programme development: A quantitative analysis. *Porta Linguarum*, *29*, 137–158. https://doi.org/10.30827/digibug.54026

Pavón, V. (2014). Enhancing the quality of CLIL: Making the best of the collaboration between language teachers and content teachers. *Encuentro*, *23*, 115–127.

Pavón, V. (2018). Innovations and challenges in CLIL research: Exploring the development of subject-specific literacies. *Theory into Practice*, *57*(3), 204–211. https://doi.org/10.1080/00405841.2018.1484035

Pavón, V., Ávila López, J., Gallego Segador, A., & Espejo Mohedano, R. (2015). Strategic and organisational considerations in planning content and language integrated learning: A study on the coordination between content and language teachers. *International Journal of Bilingual Education and Bilingualism*, *18*(4), 409–425. https://doi.org/10.1080/13670050.2014.909774

Pavón, V., & Ellison, M. (2013). Examining teacher roles and competences in Content and Language Integrated Learning (CLIL). *Linguarum Arena*, *4*, 65–78. http://193.137.34.195/index.php/LinguarumArena/article/viewFile/3967/3715

Pérez Cañado, M. L. (2013). Introduction. *Revista De Lenguas Para Fines Específicos*, *19*, 12–27. https://ojsspdc.ulpgc.es/ojs/index.php/LFE/article/view/9

Pérez Cañado, M. L. (2018). CLIL and educational level: A longitudinal study on the impact of CLIL on language outcomes. *Porta Linguarum*, *29*, 51–70. https://doi.org/10.30827/Digibug.54022

Pérez Cañado, M. L. (2020). What's hot and what's not on the current CLIL research agenda: Weeding out the non-issues from the real issues. A response to Bruton (2019). *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2020-0033

Pérez Cañado, M. L. (2021). CLIL and ELF: Friends or foes? In M. L. Pérez Cañado (Ed.), *Content and language integrated learning in monolingual settings* (Vol. 38, pp. 31–51). Springer.

Pérez, S., & Pavón, V. (2019). Un estudio comparativo de las estrategias discursivas inglés-español utilizadas en la impartición de áreas no lingüísticas en un contexto AICLE. *CLIL. Journal of Innovation and Research in Plurilingual and Pluricultural Education*, *2*(1), 43. https://doi.org/10.5565/rev/clil.10

Pérez Vidal, C. (2007). The need for focus on form (FoF) in content and language integrated approaches: An exploratory study. *Revista Española de Lingüística Aplicada*, *1*, 39–54.

Pimentel, C., & Pavón, V. (2020). The pedagogical dimension and the use of materials in English-taught programs in higher education. In M. Sánchez Pérez (Ed.), *Teacher training for English-medium instruction in higher education* (pp. 320–339). IGI Global.

Rolstad, K. (2005). Rethinking academic language in second language instruction. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 1993–1999). MA: Cascadilla Press.

San Isidro, X. (2019). The multi-faceted effects of CLIL: A literature review. *Neux Aedean Journal*, 34–49.

San Isidro, X., & Lasagabaster, D. (2019). The impact of CLIL on pluriliteracy development and content learning in a rural multilingual setting: A longitudinal study. *Language Teaching Research*, *23*(5), 584–602. https://doi.org/10.1177/1362168817754103

Scarcella, R. (2003). *Academic English: A conceptual framework*. The University of California Linguistic Minority Research Institute. Technical Report 2003-1.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Lawrence Erlbaum Associates, Inc.

Schleppegrell, M. J. (2012a). Systemic functional linguistics. In J. P. Gee & M. Handford (Eds.), *The Routledge handbook of discourse analysis* (pp. 21–34). Taylor & Francis Group.

Schleppegrell, M. J. (2012b). Academic language in teaching and learning. *Elementary School Journal*, *112*(3), 409–418. https://doi.org/10.1086/663297

Serragiotto, G. E. (2007). Assessment and evaluation in CLIL. In D. Marsh & D. Wolff (Eds.), *Diverse contexts-converging goals: CLIL in Europe* (pp. 271–286). Peter Lang.

Shaw, S. D. (2020). Achieving in content through language: Towards a CEFR descriptor scale for academic language proficiency. In M. deBoer & D. Leontjev (Eds.), *Assessment and learning in Content and Language Integrated Learning (CLIL) classrooms: Approaches and conceptualisations* (pp. 29–56). Springer.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge University Press.

Szczesniak, A., & Muñoz Luna, R. (2022). Teachers' perceptions of Content and Language Integrated Learning (CLIL) in primary schools in Andalusia. *Porta Linguarum*, *37*, 237–257. https://doi.org/10.30827/portalin.vi37.18414

Uccelli, P., & Galloway, E. P. (2016). Academic language across content areas: Lessons from an innovative assessment and from students' reflections about language. *Journal of Adolescent and Adult Literacy*, *60*(4), 395–404. https://doi.org/10.1002/jaal.553

UNESCO. (2016). *Incheon declaration and framework for action for the implementation of sustainable development goal 4*. https://uis.unesco.org/en/files/education-2030-incheon-framework-action-implementation-sdg4-2016-en-pdf-1

Vollmer, H. J. (2006). *Language across the curriculum*. Council of Europe.

Vollmer, H. J. (2008). Constructing tasks for content and language integrated learning and assessment. In O. Eckerth & S. Siekmann (Eds.), *Task-based language learning and teaching – Theoretical, methodological, and pedagogical perspectives* (pp. 227–290). Peter Lang.

Vygotsky, L. (1986). *Thought and language* (A. Kozulin, Ed.). The MIT Press.

Wewer, T. (2014). *Assessment of young learners' English proficiency in bilingual content instruction CLIL* [Doctoral dissertation, University of Turku]. https://www.utupub.fi/handle/10024/96838

Wewer, T. (2020). Assessment of young learners' English proficiency in bilingual content instruction (CLIL) in Finland: Practices, challenges, and points for development. In M. deBoer & D. Leontjev (Eds.), *Assessment and learning in Content and Language Integrated Learning (CLIL) classrooms: Approaches and conceptualisations* (pp. 137–166). Springer.

Whittaker, R., & Llinares, A. (2009). CLIL in social science classrooms: Analysis of spoken and written productions. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 215–234). Multilingual Matters.

Whittaker, R., & McCabe, A. (2020). Expressing evaluation across disciplines in primary and secondary CLIL writing: A longitudinal study. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2020.1798869

Xavier, A. (2020). Assessment for learning in bilingual education/CLIL: A learning-oriented approach to assessing English language skills and curriculum content in Portuguese primary schools. In M. deBoer & D. Leontjev (Eds.), *Assessment and learning in Content and Language Integrated Learning (CLIL) classrooms: Approaches and conceptualisations* (pp. 109–136). Springer.

Zwiers, J. (2014). *Building academic language: Meeting common core standards across disciplines, grades 5-12* (2nd ed.). John Wiley & Sons, Inc.

Appendix A. Tests

Topic: The solar system

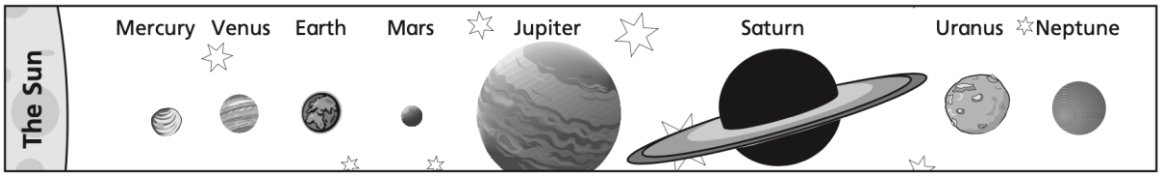
Level: 6th Year, primary education

Subject: Natural sciences

Text type: Description and explanation

TEST 1

1. Read the text about our Solar system:



Taken from Onestopenglish.com

Our Solar System is located in the Milky Way Galaxy. The Solar System is made up of the **Sun** and the celestial bodies that revolve around it. There are different types of celestial bodies: planets, satellites, comets and asteroids. Our solar system formed about 4.6 billion years ago.

The Sun is a star at the centre of the Solar System. It is much bigger than any planet. The Sun is very important because its gravity holds the Solar System together, keeping everything in its orbit. Also, it is very important for life on Earth because it gives out light and heat. There are eight planets which orbit the Sun.

Nearest the Sun are four small, rocky planets - Mercury, Venus, Earth and Mars.

Mercury is the closest planet to the Sun. Mercury is very small and it has no atmosphere and no water.

Venus is the second planet from the Sun and is very bright. It is very hot because its atmosphere traps most the heat from the Sun.

The Earth is the third planet from the Sun. It is special because it is an ocean planet; for this reason, it is called the `blue planet´. It is the only planet that supports life and it is the planet we live on.

Mars is the fourth planet from the Sun. It's called the `Red Planet´ because it as an orange-red planet. It is much smaller than The Earth, it's very cold and it has two very small moons.

Beyond the orbit of Mars are the four gas giants - Jupiter, Saturn, Uranus and Neptune. These planets are much bigger than the Earth, and they are made of gases, mainly hydrogen and helium.

Jupiter is the largest of the Solar System. It has 67 moons. Saturn is the second largest planet in the Solar System. It has got a lot of moons as well and 7 rings made of rock and ice.

Uranus and Neptune are the most distant from the Sun and they are very cold. Uranus is the seventh planet from the Sun, it has 27 moons; Neptune is the eighth planet from the Sun and it has 13 moons.

Adapted from: https://www.geogebra.org/m/UrjHxrjU#chapter/202488

Answer the questions:

1. What is the name of the star that produces light and heat?:
2. Which are the rocky planets?
3. Which planet doesn't have an atmosphere?
4. Which planet is very hot?
5. Which planet is the only one that supports life?
6. Which is the `blue planet´?
7. Which planet is called the `Red planet´?
8. Which is the largest planet?
9. Which planet has rings made of rock and ice?
10. Which planets are the most distant from the Sun?

Answers: 1The Sun, 2Mercury, Venus, Earth and Mars, 3Mercury, 4Venus, 5Earth, 6Earth, 7Mars, 8Jupiter, 9Saturn, 10Uranus and Neptune

2. You are going to watch a video about the Solar system. Before you watch, match the names of the celestial bodies (planets, stars, etc.) below to their characteristics in the table. Now, watch the video and check your answers.

(Link: https://youtu.be/zzbCEF37MfU)

Celestial bodies: The Sun, Mercury, Venus, the Earth, Mars, Jupiter, Saturn, Uranus and Neptune, the Asteroid Belt, Pluto.

Celestial bodies	Characteristics
1.	It's the second closest planet to the Sun.
2.	It's a star and a big ball of fire.
3.	It's a read-looking planet. It has two natural satellites.
4.	It's the only planet that has life. The Moon is its natural satellite.
5.	It's the nearest planet to the sun, it's also the smallest planet of the solar system.
6.	It's a crowd of asteroids.
7.	These two planets are called `ice-giants´.

8.	It's the biggest planet of all and it's made of gases.
9.	It has 62 natural satellites, it's also made of gases.
10.	It's one of the dwarf planets.

Answers: 1Venus, 2The Sun, 3Mars, 4the Earth, 5Mercury, 6 the Asteroid Belt, 7Uranus and Neptune, 8Jupiter, 9Saturn, 10Pluto

3. Writing. Use the information from activities 1 and 2 to complete the unfinished sentences about celestial bodies with one up to five words:

The Sun	**Mercury**	**Venus**
1. It's a…	1. It's the…	1. It's the…
2. It's…	2. It's…	2. It's…
3. It's…	3. It has no…	3. It's very…
4. It gives…	4. It has no…	4. It's very…
the Earth	**Mars**	**Jupiter**
1. It's the…	1. It's the…	1. It's the…
2. It's called…	2. It's called...	2. It's made of…
3. It's the only planet…	3. It has…	3. It's has…
4. It has…	4. It's…	
Saturn	**Uranus**	**Neptune**
1.It's the second…	1. It's the…	1. It's the…
2. It has…	2. It has…	2. It has…
3. It´s made of…	3. It's very…	3. It's very…

Answers: The sun: 1. It's a… star; 2. It's… a big ball of fire; 3. It's… much bigger than any planet/ is very important for life on Earth; 4. It gives… out light and heat. **Mercury**: 1. It's the… closest planet to the Sun; 2. It's… very small; 3. It has no… atmosphere; 4. It has no… water. **Venus:** 1. It's the… second planet from the Sun; 2.

It's… a rocky planet 3. It's very… bright; 4. It's very… hot. **The Earth**: 1. It's the… the third planet from the Sun; 2. It's called… the `blue planet´; 3. It's the only planet… that has life/ that supports life; 4. It has… one moon. **Mars** 1. It's the... the fourth planet from the Sun**;** 2. It's called... the `Red Planet´; 3. It has… two very small moons; 4. It's…much smaller than The Earth/ it's very cold. **Jupiter:** 1. It's the… largest of the Solar System**;** 2. It's made of… gases; 3. It's has… 67 moons. **Saturn:** 1. It's the second… largest planet in the Solar System; 2. It has… 62 satellites/7 rings made of rock and ice; 3. It´s made of…gases/ mainly hydrogen and helium. **Uranus** 1. It's the… seventh planet from the sun; 2. It has… 27 moons**.** 3. It's very… cold/distant from the sun. **Neptune:** 1. It's the… eight planets from the sun 2. It has… 13 moons 3. It's very… cold/very distant from the Sun.

4. Focus on form. Complete the sentences with the correct form of the adjectives in brackets:

1. Uranus is_____(big) than the Earth.
2. Mars is_____(small) than the Earth.
3. Jupiter is the_____(large) of all the planets.
4. Which planet is the_____(hot)?
5. Which planet is the_____(cold)?
6. Venus is_____(close) to the Sun than the Earth.
7. Mercury is the_____(small) planet.
8. The gas giants are _____(cold) than the Earth.

Example: Mars is <u>bigger than</u> Mercury.

5. Speaking. Record your voice (approx. 1 m.). Describe the effects of a severe drought and explain possible solutions.

Topic: Climate change

Level: 6th Year, primary education

Subject: Social sciences

Text type: Argumentation and narration

TEST 2

1. **You are going to watch the first part of a video about climate change; its causes and some of its consequences. Complete the sentences with the correct words from the box:**

heat, gas, snow, heating, rising, flying, fossil fuels, carbon dioxide, using, warm, driving, warmer, melting, blooming.



Climate change basics

To understand climate change, you first need to know about the greenhouse effect. The Earth gets_____from the sun. In the atmosphere gases like_____trap this heat and keep it from escaping back to outer space. Trapping some heat in the atmosphere is a good thing because it keeps the planet_____ enough for us to live. But there's a problem, people all over the world are adding extra carbon dioxide to the atmosphere. That's because today we burn_____ like coal, oil and_____to do many of our everyday activities, like _____our cars,_____our computers and _____ our homes. All this extra carbon dioxide is trapping more heat in the atmosphere, making the Earth warmer and causing other climate changes too. The signs of climate change are all around us: temperatures are getting_____, giant ice sheets are_____ and oceans are_____. In many places, flowers are_____ earlier, _____ is melting sooner and birds aren't_____ as far as south for the winter.

Answers (in order): heat; carbon dioxide, warm, fossil fuels, gas, driving, using, heating; warmer, melting, rising, blooming, snow, flying.

(Link up to 1:23) https://www.youtube.com/watch?v=ScX29WBJI3w&t=164s)

2. **Writing: Think about the negative effects that climate change can have on cities/people, plants, animals, the weather or water. Write 5 sentences about each one of these effects, use the bullet points:**

A negative effect on cities/people: •	A negative effect on animals: •	A negative effect on the weather: •
A negative effect on water: •	Can you think of another negative effect? •	

Examples:

If climate change occurs, some cities **will** get warmer.

If earth gets warmer, plants **will** die.

If we don't fight climate change, **then** we **won't** have any water.

3. **Reading: You are going to read a text about climate activist "Greta Thunberg". Answer the questions below.**

Greta Tintin Eleonora Ernman Thunberg (born on 3 January 2003) is a Swedish activist. She is famous for her work against climate change. Greta was only eight when she first learned about the climate crisis.

On August 20th 2018, she sat alone outside the Swedish parliament in protest against climate change. Her protest soon spread on social media to other school children of her age. It also received a lot of media coverage.

After Greta became popular on social media many other students around the world began to strike along with her. On 15 March 2019, 2200 strikes were organized in 125 countries.

In May 2019 Greta Thunberg published a book – *No one is too small to make a difference* to inspire other young people to stand up for what they believe in.

Greta spoke at the Unite Nations Climate Action Summit on 23rd September 2019 in New York City. She travelled by yacht instead of flying so she would not contaminate air with carbon dioxide to get to the summit. At the summit she gave a very powerful speech:

"This is all wrong. I shouldn't be standing here. I should be back in school on the other side of the ocean," she said with tears in her eyes. "Yet you all come to me for hope? How dare you! You have stolen my dreams and my childhood with your empty words."

Answer the questions below, say if they are TRUE (T), FALSE (F) or DOESN'T SAY (DS):

1. Greta is a climate change activist from Norway:	**T/F/DS**
2. She was very young when she first knew about climate change:	**T/F/DS**
3. She became famous in 2018 for protesting outside parliament with other school children:	**T/F/DS**
4. She soon received attention from the media:	**T/F/DS**
5. In 2019, many protests were organized around the world along with her:	**T/F/DS**
6. Greta wrote a book entitled "No one is too small to fight against climate change":	**T/F/DS**
7. Greta travelled by boat instead of taking a plane in 2019 because she didn't like flying:	**T/F/DS**
8. At the Unite Nations she spoke about climate change to the leaders of the world:	**T/F/DS**

Answers: 1F, 2T, 3F, 4T, 5T, 6F, 7F, 8DS

4. Focus on form. Match the following terms on the left (a-e) to their definitions on the right (1-5):

a. Carbon dioxide	1. The warming of the earth that is caused by gases in the air that trap energy from the sun.
b. Climate change	2. A fuel such as coal, oil, or natural gas.
c. Fossil fuel	3. A raising of global temperatures as a result of high levels of certain gases e.g. carbon dioxide
d. Global warming	4. A gas that is formed by burning fuels.
e. Greenhouse effect	5. A change in the planet's weather and temperatures.

Answers: a4, b5, c2, d3 ,e1

5. **Speaking. Record your voice (approx. 1 m.). Give reasons for four effects of climate change in your country.**

Appendix B. Rubric

RUBRIC **Primary Education**						
Target discourse features		Not satisfactory (1-2)	Weak (3-4)	Satisfactory (5-6)	Good (7-8)	Excellent (9-10)
Writing & Speaking	***Description*** **Writing:** • Present tenses (facts) • Simple Relative clauses • Compound/complex sentences • Appropriate style/register **Speaking:** • Comparative/superlative • Connectors/discourse markers (e.g. on the one hand… whereas… while) • There is/are	Ss fail to produce a sufficient range of target features. Ss fail to produce target features accurately. Ss cannot produce comprehensible descriptions. Vocabulary is poor and the syntax is frequently broken and inaccurate.	Ss produce a limited range of target features. Ss rarely produce target features accurately. Ss cannot produce elaborated descriptions. Vocabulary is reduced and the Ss make common	Ss produce a sufficient range of target features. Ss sometimes produce target features accurately. Ss can elaborate basic descriptions. Vocabulary is not sophisticated	Ss produce a wide range of target features. Ss frequently produce target features accurately. Ss are able to produce elaborated descriptions. Vocabulary is varied and appropriate, and	Ss produce an extensive range of target features. Ss consistently produce target features accurately. The vocabulary of the descriptions produced by the Ss is rich, the syntactic constructions

			mistakes with the use of syntax.	but correct. The level in the use of syntactic constructions is average.	they can also produce accurate syntactic constructions with few mistakes.	show no mistakes, and the Ss make use of metaphors and similes.
	Explanation **Writing**; • Compound/complex sentences • Appropriate use of linkers (functions e.g. cause and effect) – be specific. • Purpose (e.g. to + infinitive) • Recommendations (should, could) • Consequences (so)	Ss cannot produce comprehensible explanations. Vocabulary is poor and the syntax is frequently broken and inaccurate.	Ss cannot produce elaborated explanations. Vocabulary is reduced and the Ss make common mistakes with the use of syntax.	Ss can elaborate basic explanations. Vocabulary is not sophisticated and use of syntactic constructions is average.	Ss can produce elaborated explanations. Vocabulary is varied and appropriate, and they can also produce accurate syntactic constructions with few mistakes.	The vocabulary is rich, the syntactic constructions show no mistakes. The Ss can adequately describe the stages in a process and how things are made.

	Argumentation **Writing 1** • Cause/effect (because) • Qualifiers/intensifiers (very) • Variety of simple tenses • Possibility/probability **Writing 2** • Purpose (e.g. to + infinitive) • Recommendations (should, could) **Speaking** • Comparing/contrasting (but, however) • Open conditionals (if the government does X, then Y)	Ss cannot present their ideas in a comprehensible way. Vocabulary is poor and the syntax is frequently broken and inaccurate.	Ss cannot produce elaborated argumentations. Vocabulary is reduced and the Ss make common mistakes with the use of syntax.	Ss can present their arguments comprehensively. Vocabulary is not sophisticated and the use of syntactic constructions is average.	SS can present their arguments fluently and accurately. The ideas are clearly organised, the vocabulary is correct and there are not significant mistakes in the use of syntax.	The vocabulary used in their arguments is rich, the syntactic constructions show no mistakes, and the Ss present their views through analyses and evaluations.

	• Cause/effect/reasons/consequences (because)					
	Narration **Writing** • Ordinal discourse markers (first, second, next, after that, at the end) • Tense sequences (The boats were arriving while I was….) **Speaking** • Ordinal discourse markers (first, second, next, after that, at the end) • Tense sequences. • Vocabulary for describing feelings/emotions.	Ss cannot produce comprehensible descriptions. Vocabulary is poor and the syntax is frequently broken and inaccurate.	Ss cannot produce elaborated descriptions. Vocabulary is reduced and the Ss make common mistakes with the use of syntax.	Ss can elaborate basic descriptions. Vocabulary is not sophisticated and the use of syntactic constructions is average.	Ss are able to produce elaborated descriptions. Vocabulary is varied and appropriate, and they can also produce accurate syntactic constructions with few mistakes.	The vocabulary of the descriptions produced by the Ss is rich, the syntactic constructions show no mistakes, and the Ss make use of metaphors and similes.

Appendix C. Skills histograms

Natural Sciences (description and explanation)

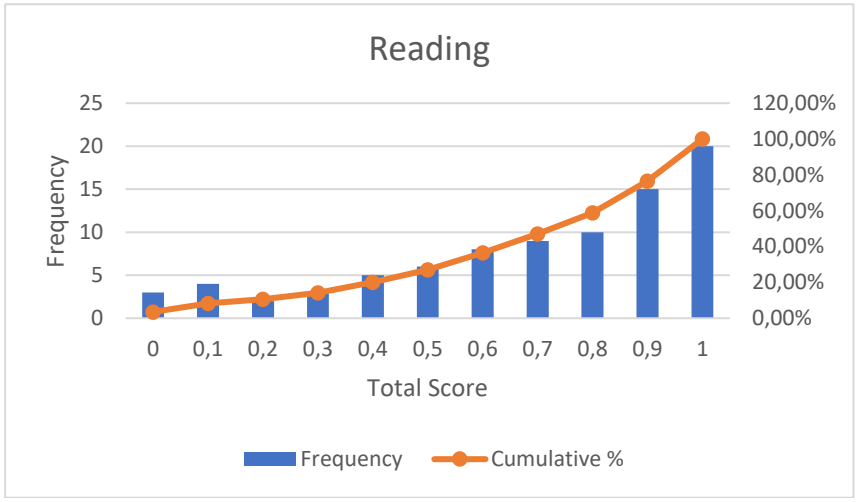


Figure C1. Reading histogram in Natural Sciences.

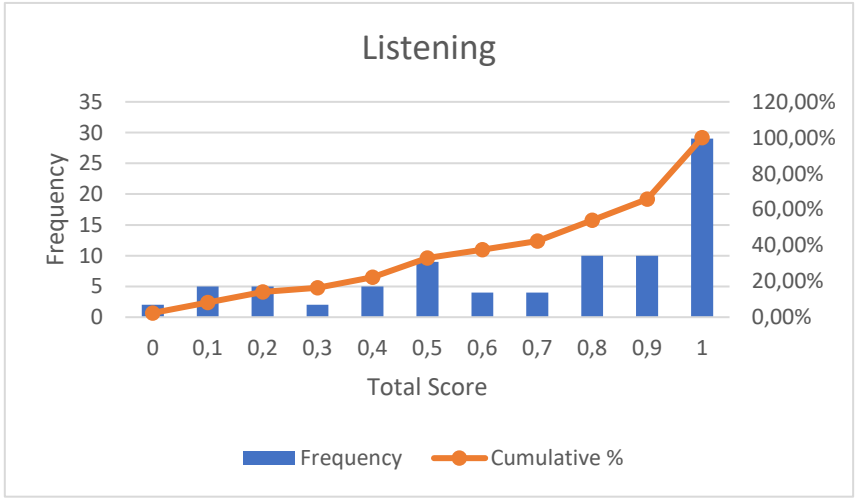


Figure C3. Listening histogram in Natural Sciences.

<u>Social Sciences</u> (narration and argumentation)

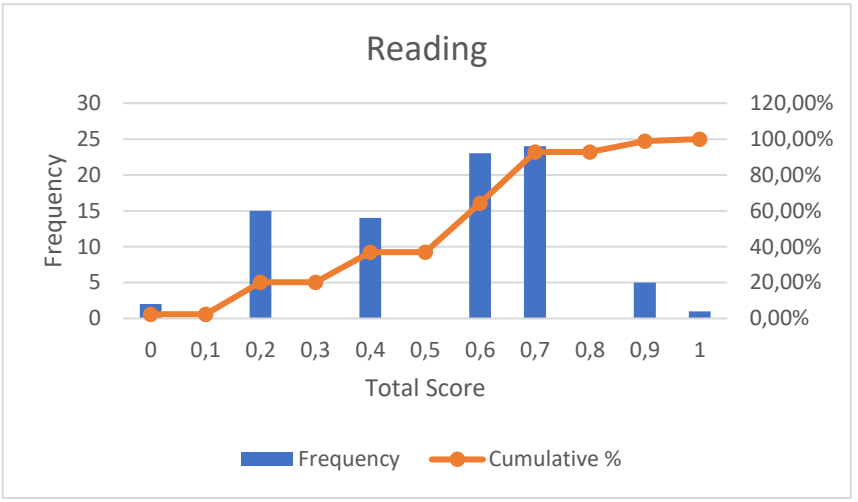


Figure C2. Reading histogram in Social Sciences.

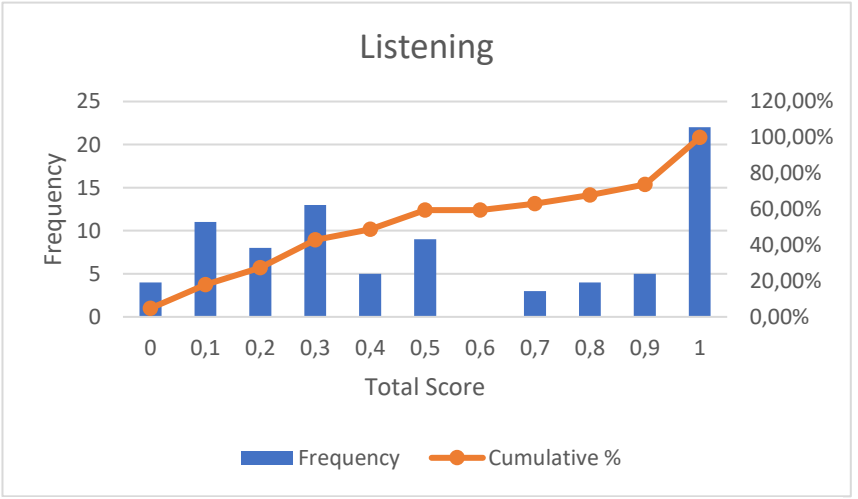


Figure C4. Listening histogram in Social Sciences.

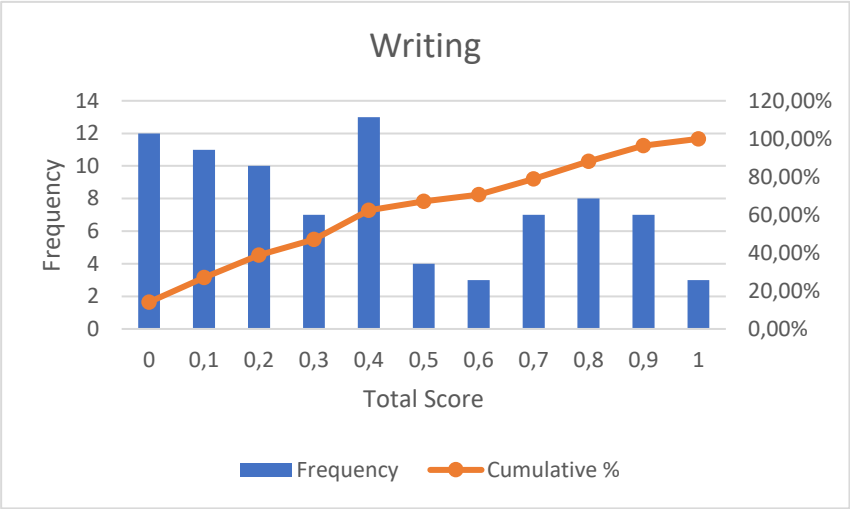


Figure C5. Writing histogram in Natural Sciences.

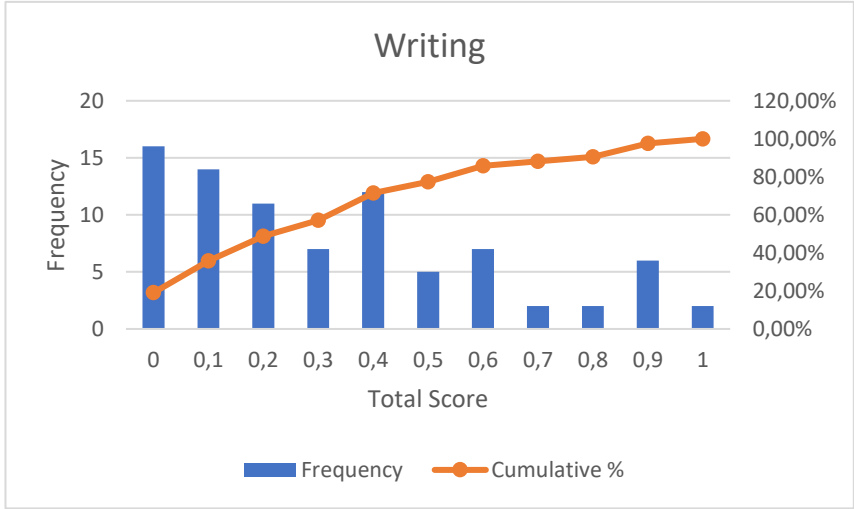


Figure C6. Writing histogram in Social Sciences.

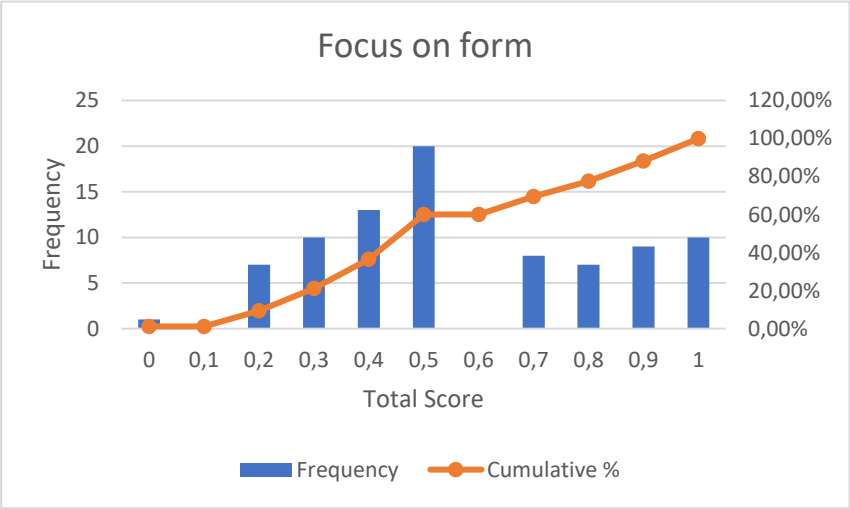


Figure C7. Focus-on-form histogram in Natural Sciences.

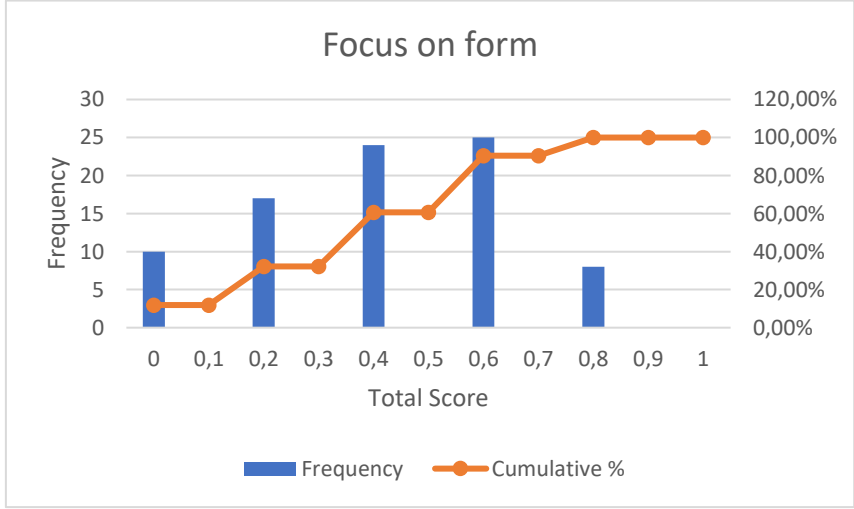


Figure C8. Focus-on-form histogram in Social Sciences.

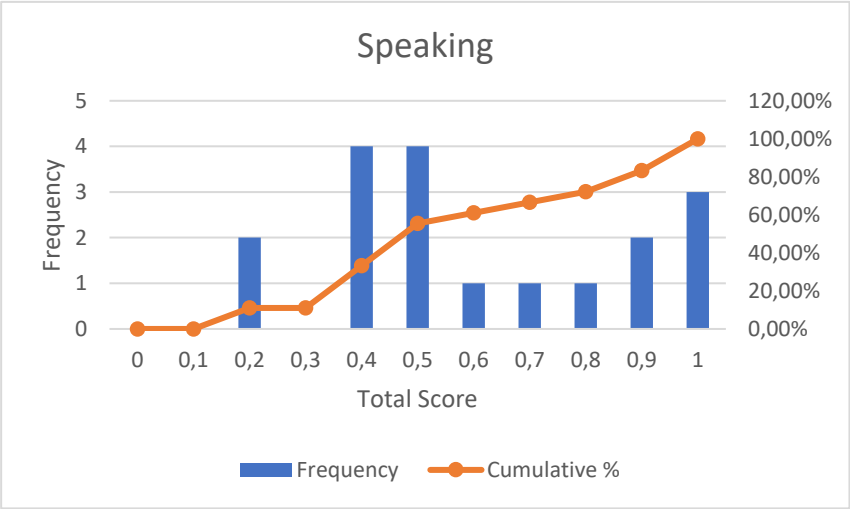


Figure C9. Speaking histogram in Natural Sciences.

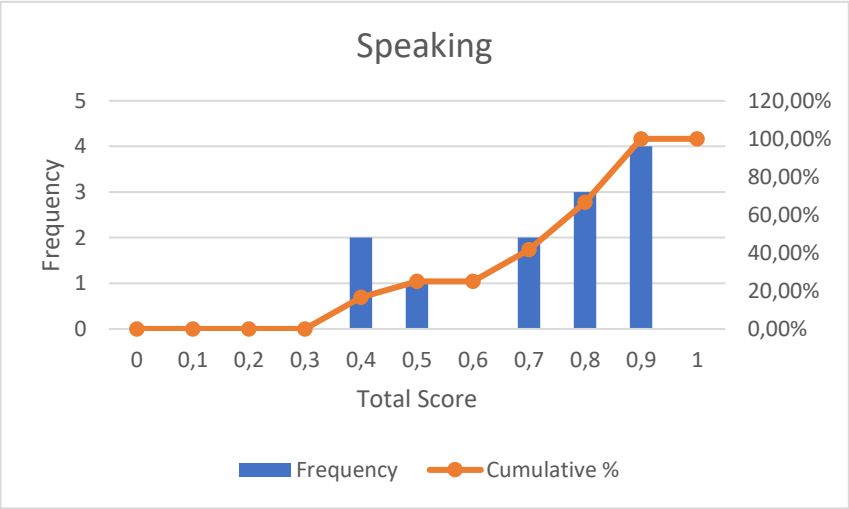


Figure C10. Speaking histogram in Social Sciences.