



# CLIL Students' Academic Language Performance on a Non-curricular Topic: a Comparison Between High-Exposure and Low-Exposure Groups

CLIL學生非課內相關主題之學術語言表現:高接觸組與低接觸組之比較

Ana Llinares<sup>1</sup> · Nashwa Nashaat-Sobhy<sup>2</sup>

Received: 15 December 2022 / Revised: 3 May 2023 / Accepted: 9 May 2023 /

Published online: 10 October 2023

© The Author(s) 2023

## Abstract

In Content and Language Integrated Learning (CLIL) contexts, when students are prompted to express content in a second language, they are expected to draw on their knowledge of the topic at hand and use linguistic resources they have learnt implicitly or explicitly in the content and/or language class. In some CLIL programs, while some students study academic subjects like history in the L2, other students study fewer and more practical subjects such as art. This is the case with high-exposure (HE) and low-exposure (LE) groups, respectively, in the Madrid bilingual secondary school programs. Thus, HE students are expected to perform better in the expression of academic content in the L2, but it is not clear if this is also the case when prompted to write about a topic that is not part of their curriculum. In this study, we compare the texts written by groups of CLIL students with different degrees of exposure to English (HE and LE) in response to a prompt on the Women's Movement (Feminism) eliciting Cognitive Discourse Functions, such as defining or evaluating, among others. The aim of the study was to compare how HE and LE students make specialist-knowledge claims and use their voices in this topic, which they have not been instructed in, across different CDFs. Results show differences both across groups (HE and LE) and across CDFs, showing an advantage for knowledge transfer in the HE group.

## 摘要

在學科內容與語言整合學習 (CLIL) 的情境下，當學生使用第二語言來表達內容時，他們被期待利用現有的學科知識，及在學科課堂和/或語言課堂上隱式或顯式學到的語言資源。在某些CLIL課程中，有些學生使用第二語言來學習學術性科目，例如歷史課程，其他學生則學習較少且更實用的科目，例如藝

術課程。在馬德里的雙語中學學程的高接觸組(HE)和低接觸組(LE)就分屬這樣的情況。因此，高接觸組的學生被期待在使用第二語言表達學術內容時有更好的表現，但當他們使用文字表達一個非課程內的主題時，能否有同樣表現就不得而知。本研究比較了接觸英語不同程度的CLIL學生(HE and LE)所寫的文章，文章回應了關於婦女運動(女權主義)的寫作提示，這些提示引導出認知性言談功能，例如定義、評價等。本研究旨在比較HE和LE組學生如何針對這個他們沒有接受過指導的主題，在不同的認知性言談功能中提出專業知識的主張和觀點。研究結果顯示了不同組別(HE and LE)和不同認知性言談功能之間的差異，顯示出HE組在知識移轉方面的優勢。

**Keywords** Content and Language Integrated Learning · Legitimation Code Theory · Specialisation codes · Appraisal theory · Academic lexis · Cognitive Discourse Functions

**關鍵詞** 內容與學科整合學習(CLIL) · 合理化符碼理論 · 專業符碼 · 評估理論 · 學術詞彙 · 認知性言談功能

## Introduction

Since CLIL programs started to grow in Europe in the 1990s, considerable research has focused on students' language performance and development. In line with the original motivation in the implementation of CLIL programs, most research at the beginning investigated the effect of CLIL programs on students' general language proficiency in the target language, often yielding positive results (for an overview, see, for example, Goris et al. 2019). The realisation that CLIL is not merely a bi/plurilingual context for language learning but an approach with pedagogic and academic demands for quality education (Meyer & Coyle, 2017) shifted the focus in CLIL research from CLIL as a 'context' for language learning to CLIL as a content and language integrated learning and teaching 'approach' (e.g. Llinares, 2015; Dalton-Puffer et al., 2022). This brought about models and frameworks for researching and teaching content and language in integration and, at the same time, a growing interest in pluriliteracies, and the specificities of different disciplinary cultures and different CLIL programs in different parts of the world.

One model that has proved useful for the operationalization of content and language integration is Dalton-Puffer's (2013) framework of Cognitive Discourse Functions (CDFs hereafter), a construct of 7 prototypical functions (such as *defining*, *describing*, and *exploring*). This model resulted in subsequent research studies of CDFs in Science (e.g. Llinares & Nashaat-Sobhy, 2021; Whittaker & McCabe, 2023), History (e.g. Dalton-Puffer & Bauer-Marschallinger, 2019; Lorenzo, 2017; Whittaker & McCabe, 2023), Business (Breeze & Dafouz, 2017), or in Pre-service Teacher Education (Nashaat-Sobhy, 2018), to mention a few. These studies have helped conceptualise how content and language are learnt through integration (e.g. Llinares, 2015) and explore ways of implementing CLIL as an approach (Dalton-Puffer et al. 2022). However, the identification of the disciplinary and contextual

specificities involved in the enactment of these CDFs is the key to guarantee an understanding of the generalizability of the findings to other contexts by the research community (Cenoz et al. 2014) and hence the need for research studies across different domains of academic study and to explore further the linguistic resources that characterise each CDF (Dalton-Puffer et al. 2018).

Another useful model to investigate CLIL as ‘approach’ is Pluriliteracies Teaching for Learning (Meyer & Coyle, 2017), which stresses the interrelation of language, content, cognition, and academic literacies. This model draws attention to literacy practices to support learners’ development within and across disciplinary subjects. It proceeds from the premise that academic progress requires general academic and subject-specific strategies (e.g. reasoning), in addition to knowledge of facts and concepts, hence the importance of organising, explaining, and arguing for creating knowledge. This model was the result of the Graz Group’s call to promote pedagogic processes that encouraged intra- and interdisciplinary knowledge transfer for knowledge building and deeper learning, i.e. making meaning of new knowledge items in light of previously acquired knowledge and skills. This proposal is echoed across education systems through the OECD agenda for Future Education and Skills 2030, which calls for the interrelation of epistemic and procedural knowledge and skills across subjects of study, highlighting the importance of thematic learning and the transfer of key concepts and procedures (OECD, 2019a). This agenda also aims to promote the development and expression of values and attitudes towards cultural and societal issues in light of the current global challenges (OECD, 2019b), which are found reflected in varying degrees in the learning outcomes of different school subjects.

When CLIL students are prompted to express disciplinary content in the target language, they are expected to draw on their knowledge of the topic (whether they have learnt it through the target language or through any other language) and to use the L2 linguistic resources they have explicitly or implicitly learnt (in the content or the language class) to express academic content. However, it remains to be seen if students’ academic language performance is transferred when they are asked to write or speak in the L2 about a topic that is not part of the curriculum of specific disciplines.

In the CLIL program in the Community of Madrid (Spain), the context of the present study, the learning outcomes of English as a subject of study (EFL) in grade 10 (ages 15 to 16) state that students will use language accurately (language forms), but also that they will give opinions on the latter and on current affairs (voice), and that they will use not only common expressions, set phrases, and vocabulary related to topics of general interest but also to the contents of different parts of the curriculum (academic lexis). It would be interesting to explore if these language learning outcomes are different across groups of CLIL students who do not only study different subjects in the target language but who also have different degrees of exposure to the L2.

Against this backdrop, the aim of the present study is to compare two groups of CLIL students on their written productions which were triggered by a prompt based on the CDF model. Both groups are compared on their use of voice, academic lexis, grammatical forms, and meaning making. One of the groups had more hours of

exposure to English in more varied content subjects (science, history, geography, art, PE and technology), which we shall be referring to throughout as the high exposure group (HE), whereas the other group, the low exposure group (LE), had less hours and less varied exposure to English (in art, PE and/or technology, but not in science or social science). For the analysis of students' language production for the expression of meanings, we use the specialisation dimension in Legitimation Code Theory (Maton, 2014), where in all fields of practice (EFL and CLIL included) specialists (teachers/researchers) validate certain forms of knowledge over others when considering the bases of achievement, either of knowledge as knowing (what is learned as a generic product) or knowledge as knowers (individual's knowledge brought on by personal attributes, dispositions, and experiences). Following from this view, knowledge is regarded in this study to be the product of knowing (epistemic relations or ER) that is brought about by knowers (social relations or SR). In our study, we focus on language forms and academic lexis as representations of 'knowing' and voice as the representation of 'the knowers'. We draw on Martin and White's (2005) appraisal theory to analyse voice in the students' texts and on Coxhead's (2000) Academic Word List to measure academicness. Finally, we observe the distribution of appraisal and academicness in the different Cognitive Discourse Functions.

## The CLIL Advantage

When CLIL students are asked to express ideas in the target language, they are expected to draw not only on content matter and language resources from different subjects to which they have been exposed, and from which they have learned ideas and meaning-making resources, but also on knowledge of the world that they gain socially and from the media (Shor, 1980, p. 140). The ability to transfer knowledge is an underlying outcome of education in general and of schooling in particular, and the effects of subject learning are expected to extend to other subjects and world applications (DePalma & Ringer, 2011; James, 2006). De Palma and Ringer (2011) argue that knowledge transfer is adaptive, i.e. students not only reuse prior knowledge but also reshape it, and thus define knowledge transfer as a process that is conscious and intuitive, whereby learners reshape learned knowledge and apply it in new unfamiliar situations.

As a subject of language study, English is concerned with socialising students in language as code for the purpose of communication and action (Gee & Handford, 2012: 1) so that the students may create new meanings, take positions, and construct identities (Jordão & Fogaça, 2012). These actions, by definition, are conducive to critical literacy, the pathways to which depend on text and media and engaging students in writing activities to explore the students' thoughts and words: how they represent their knowledge of the world and how they use language as a system (e.g. syntax, diction/lexis) to project their voices (expressing opinions, evaluations, and critiques).

Comparisons of CLIL and non-CLIL students' texts show that studying in CLIL programs generally favours writing competence. Ruiz de Zarobe (2010) showed that CLIL learners scored significantly higher marks in vocabulary,

language use, and mechanics; Järvinen (2010) found that they used a wider range of vocabulary and grammatical structures, which were more accurately executed than in their non-CLIL counterparts' performances. Similar observations were made in Jexenflicker and Dalton-Puffer (2010, p. 180) as well as in Pérez-Vidal and Roquet (2015), who respectively commented on CLIL students' error free performances and their wider range of morphosyntactic resources.

However, in a review of CLIL and non-CLIL comparative studies, Muñoz (2015) concluded that the hours of additional immersion that some groups had were fundamental in the studies where significant differences between groups were found. Muñoz also adds that other variables, such as having an initial language advantage and extramural exposure to the L2, may affect the CLIL variable, but these are very difficult to control. In a more recent study by Artieda, Roquet, and Nicolás-Conesa (2020), the researchers gathered data from two groups of secondary school students where one group received CLIL classes in Science for 3 years (from grade 5 onwards) in addition to English as a subject of study from pre-school, and the other group studied English as a subject only, also from preschool. Language practice in Science was mostly meaning-oriented and dependent on reading L2 texts to compose texts about science topics. The data were gathered by means of several instruments, including multiple choice and grammaticality judgement tests and a picture story for which the students needed to write an imagined dialogue between three characters. The difference in the number of hours of exposure to English in the CLIL group was reflected in the group's heightened control of grammar (although no significant differences were reported), vocabulary, task fulfilment, and organisation, when compared to the English-only group. That is, students in a reduced/partial CLIL immersion in one subject only still showed greater linguistic gains in L2 skills.

Former research within the UAM CLIL group (<http://www.uam-clil.org>) has contributed to students' academic language performance, motivation towards CLIL, and teaching practices this time comparing groups in the same program with higher (HE) and lower exposure (LE) to CLIL. The difference between these two strands is not only the amount of exposure to CLIL but also the variety and types of disciplines that students learn in the L2. While LE students usually learn one or two disciplines in English (apart from the English subject), often with a practical orientation (PE, technology or art), the HE groups are exposed to the same subjects as the LE ones plus the science and social science disciplines (biology, history, geography, etc.). Studies on these types of programs have shown differences across these strands regarding motivation and anxiety towards CLIL (e.g. Fernández-Agüero & Hidalgo-McCabe, 2022; Somers & Llinares, 2021) and potential inequity issues in teachers' interactional practices, where students in HE groups are engaged in more higher-order thinking skills than their LE counterparts (see Llinares & Evnitskaya, 2021). The students in these groups, then, seem to experience different knowledge practices. Against this backdrop, HE students are expected to perform better than their LE counterparts in the expression of academic content in the L2, yet whether this would also be the case when they write about a topic that has not been pedagogized remains to be empirically tested. It would, then, be interesting to explore whether students' language performance on

a general topic may be influenced by the type of CLIL program (or strand in this case) that they participate in.

## Cognitive Discourse Functions

Dalton-Puffer's (2013) model of Cognitive Discourse Functions was designed to relate the knowledge expected to be acquired and the language required for CLIL learners to access that knowledge. As mentioned above, the framework consists of 7 different functions which derive from the analysis of CLIL classroom lessons: define, describe, classify (changed to categorise in Evnitskaya & Dalton-Puffer, 2023), report, explain, evaluate, and explore. While the construct was designed to 'enable content-subject educators and language-subject educators to find a common language in talking about and researching student-learning in CLIL' (Dalton-Puffer, 2016: 51), no study (to our knowledge) has applied the CDF model to students' performance on a non-curricular topic.

## Appraisal

The appraisal system is a Systemic Functional Linguistics model for analysing the language of evaluation (Martin & Rose, 2003; Martin & White, 2005). Its application in education emerged in response to observed genre-specific literacy demands in different school subjects. Appraisal theory allows for studying learners voice through the dimension of *Attitude*, which is differentiated into statements of *affect*, *judgement*, and *appreciation*. These three types of attitude express writers' positions and views, understood to be determined in relation to social norms with which the learners (un-)identify themselves. Statements including *affect* express how to feel and involve expressions of satisfaction, confidence, inclinations, and their opposites (e.g. 'Girls *want to* be the same as boys, and have the same salaries as boys.').<sup>1</sup> Statements of *judgement* evaluate human behaviour against recognised social norms, showing what is considered moral, ethical, and legal and what is admired, criticised, praised, or condemned in a society (e.g. '*We're considered sluts...* for wearing shorts or short skirts.'). Finally, statements of *appreciation* evaluate tangible outcomes or objects, also in relation to systems of social value, for quality, complexity, and functionality (e.g. 'it would be *an excellent change*').

A second dimension of appraisal that is used in this study is *engagement*, which foregrounds the source of the attitude and through which writers either allow or restrict positions that may be different from theirs (e.g. '*I don't agree* with this idea'). The third and final appraisal dimension used in this study is *graduation*, by which writers intensify or reduce the force of an evaluation (e.g. '*...even though all genders are equal*'), and by which they may also soften or sharpen the focus of the evaluation (e.g. 'The economy would stay *pretty much* the same').

<sup>1</sup> The examples are from the students' texts.

## Academic Lexis

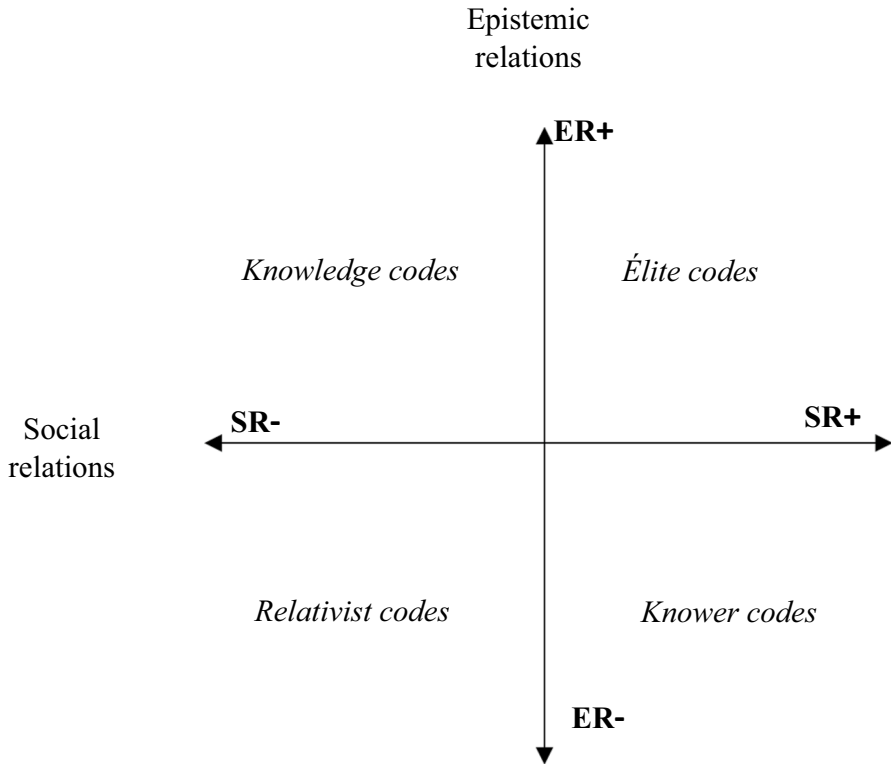
CLIL students are expected to attain a number of academic language competences through their study of content subjects in English. These include the accurate use of grammatical structures that are typical of academic writing, such as nominalisations and register (e.g. Whittaker, Llinares & McCabe, 2011) and general academic lexis (Olsson, 2015). Academic lexis supports attention to register and contributes to the development of cognitive academic language proficiency (CALP), defined as ‘the extent to which an individual has access to and command of the oral and written academic registers of schooling’ (Cummins, 2000, p. 67). It also influences how CLIL students are judged and evaluated, as observed in Olsson (2015).

Coxhead (2000) distinguishes academic words as those found in academic texts, irrespective of the subject area, and where the following three principles apply: (1) words that appear outside the first 2000 most frequent words of English; (2) words (or members of the words family) that appear 10 times or more in each of the fields —arts, commerce, law, and science—drawn on for compiling the Academic Corpora, each of which subsumes seven subject areas (e.g. the arts subcorpus is composed of topics from education, history, linguistics, philosophy, politics, psychology, and sociology); and (3) words that occur a minimum of 100 times in the Academic Corpus. Coxhead’s (2000) academic word list (AWL) contains 570 word families that make up about 10.0% of the tokens found in academic texts, as opposed to only 1.4% of those found in a collection of fiction of the same size. Academic words, in comparison to technical subject-specific words, appear in lower frequency, the reason why students, in general, tend to struggle more with their use (Worthington & Nation, 1996).

Some studies have addressed the role of technical and academic language in specific subjects in CLIL, either focusing on the different use of technicality by the same teachers teaching in CLIL and non-CLIL classrooms (Bieri, 2019) or on the role of certain tasks in generating subject-specific and academic language (Nikula, 2015). Another study that compared the use of academic lexis using AWL in CLIL and non-CLIL students’ written essays (Olsson, 2015) has shown no significant differences between groups in the Swedish context. However, the students were given prompts from two subject areas (Social Sciences and Natural Sciences) and of different genres (argumentative and expository essays), which the authors believe could have influenced the outcome.

## LCT: Knowledge-Knower Structures

Legitimation Code Theory (LCT) is a conceptual toolkit and an analytical methodology from the field of Sociology of Education that enables learning productions and artefacts (e.g. students’ texts) to be analysed in light of different practices and knowledge structures and whose framework provides researchers with additional explanatory power to make sense of the data. According to Maton



**Fig. 1** Specialisation codes (Maton, 2014, 30)

(2014), knowledge is often equivocally considered either epistemological—absolute, decontextualized, and value free—or ontological—socially constructed, shaped by history and culture, and reflecting social interests. LCT thus allows for studying the practices of individuals in any field from both views, thus avoiding the reduction of knowledge to a specific form and ignoring others that also construe bases of achievement.

One of the powerful tools in LCT is *specialisation codes*, which rests on the premise that practices and beliefs are about something and are by someone. Its two organising principles are epistemic relations (ER), which are about a target object of study or practice, and social relations (SR), which are about those who enact the target practice. The specialisation dimension of LCT is particularly useful for exploring students' advancedness through, for example, (1) the use of specialised knowledge of an object of knowledge or practices gained from formal study (i.e. ER) and (2) personal dimensions as experiences and opinions that influence how individuals socially relate to an object or a practice (i.e. SR). Both types of relations co-occur and can vary in strength. When a practice emphasises epistemic relations, these are labelled ER+, and are inversely labelled ER- when epistemic relations are downplayed. Similarly, when a practice emphasises social



relations, these are labelled SR+, and are inversely labelled SR– when these are downplayed. Practices vary across both continua (ER+ to ER–/SR+ to SR–), and their intersection yields four specialisation codes (Fig. 1).

- (i) Élite codes (ER+, SR+) emphasise both specialised knowledge of an object of study and the presence of the knowers' personal dimensions in the practice. Élite here refers to having the right type of knowledge and showing the role of the knower (or the right type of knower).
- (ii) Knowledge codes (ER+, SR–) emphasise specialised knowledge of objects of study and downplay the personal dimension of the actors involved.
- (iii) Knower codes (ER–, SR+) downplay specialised knowledge and shift the emphasis to the personal dimensions of the actors involved (e.g. experiences, opinions, and personal engagement with a topic).
- (iv) Relativist codes (ER–, SR–) downplay both specialised knowledge and personal dimensions. Neither the practices nor the actors are important, and anything goes, which is rare.

Each of the specialisation codes is a cline of degrees of strength. A stronger emphasis on epistemic relations can then be represented as ER++ and social relations as SR++.

As CLIL students at secondary school level, and above, are expected not only to demonstrate adequate language use for making meaning but to also project their voice, the specialisation dimension can shed light on how the students manipulate the combination of these structures and where instructional interventions are required.

## The Purpose of the Study

The purpose of the study is twofold. First, we investigate and compare the language resources used by HE and LE students to project their voice and to reflect their knowledge of the target language. Secondly, we investigate students' use of voice and academicness across the different CDFs.

The three research questions related to these two objectives are the following:

RQ1. How do HE and LE students use their voice?

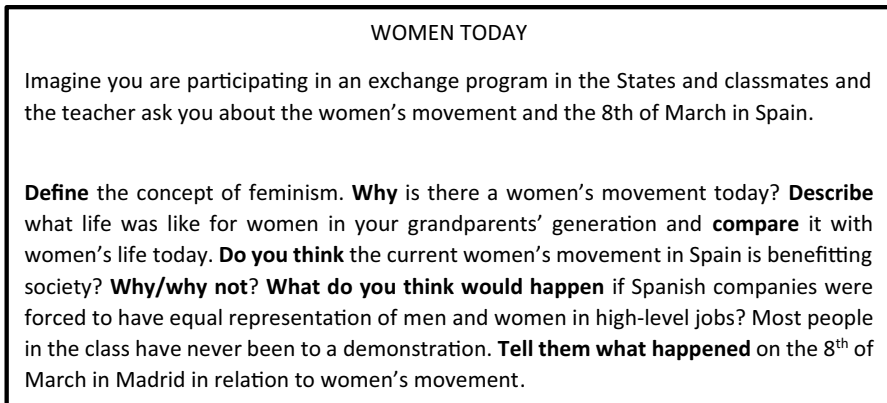
RQ2. How do HE and LE students represent their knowledge?

RQ3. How do HE and LE students represent their knowledge and use their voice across different cognitive discourse functions?

## Methods and Data Analysis

### The Context

The context of the study is the bilingual program in the Madrid region, which started in 2004 and expanded quickly, with more than 50% of the state primary and



**Fig. 2** Prompt eliciting students' production of CDFs on the topic 'Women Today'

secondary schools participating in the program in 2020–2021 (399 primary schools and 187 secondary schools). A specificity of this program compared to other similar ones is that, while at the primary school level all the students participate in CLIL lessons to the same degree (between 30% and 50% of the syllabus is taught in the target language, mostly English), at secondary school level (grades 7 to 10), the students are divided into two strands: *sección*, with higher exposure (HE) to CLIL, and *programa*, with lower exposure (LE). This division is done on the basis of students' level of general English demonstrated in an external test that all the students take at the end of Primary Education. In the LE group, students study at least one subject in the target language of a more practical nature (e.g. PE, art and/or technology). In the HE group, students take these same subjects as their LE peers but, in addition, they also learn the natural science courses (biology and/or chemistry and physics) and social science courses (geography and history) in the target language. In other words, while all the students in a bilingual secondary school participate in CLIL, for those students that are expected to have a higher level of general English, the number of hours of exposure is higher, and they study more, and more varied, academic subjects in the target language.

### The Data

The data consist of the essays written by 2 groups of students in grade 10 (aged 15–16) from the same school. One was a HE-group (with 22 students) and the other one was a LE-group (with another 22 students). The students were asked to write about a topic that was not part of their curriculum, *Women Today*. Following Maton (2014), then, the topic cannot be treated as 'pedagogised knowledge', but 'common-sense' 'world knowledge'.

The prompt was designed to elicit the students' response to the 7 CDFs (Dalton-Puffer, 2013) in relation to the topic, as illustrated in Fig. 2 below. This was done to explore to what extent students in both groups were able to define, explain, etc. as they are expected to do when they work in their disciplines taught in English.

## The Analysis

We used O'Donnell's (2008) UAM Corpus tool to create a schema with several sub-layers for coding the instances of appraisal (attitude, graduation, engagement) and their correctness in form and meaning.  $T$ -statistics were calculated for significance in this layer at  $P < .05$ . The corpus tool was also used to calculate the 'academicness' of texts using Averil Coxhead lists of academic words<sup>2</sup>, measured in terms of the percentage of the lexical words in the text that are in the Academic Word List (AWL); e.g. if 50% of the lexical words used by the students were anywhere in the AWL, it would score 50%. For the identification of the specific words used by the students, Antconc (Anthony, 2010) was used to generate a keyword list to identify characteristic words from the AWL list that would appear in one group and not the other, also allowing us to determine the range of students that produced these characteristic words.

Finally, these linguistic measures were interpreted in view of the specialisation dimension (Maton, 2014) for the analysis of students' writing performance. Table 1 illustrates the 'translation device' (as Maton, 2014 puts it) that relates students' language performance and their potential achievement as knowledge and knowers.

Our coding of form and meaning targeted the parts of the texts that combined both knowledge and knower codes (that is, language performance that included appraisal). Naturally, the task required that all the students used their knowledge of the L2 (accurate and clear language as well as instances of academic language), and that they projected their voice in some parts of the prompts (the parts eliciting the CDFs *evaluate* and *explore*). Our interest here is to see if there are differences across the two groups under comparison in their ability to project knowledge and project themselves as knowers in the L2.

## Results

In response to RQ 1, the first comparison focused on HE and LE students' use of voice through the use of Appraisal, which relates to how students present themselves as knowers. A similar distribution of appraisal types was observed across both groups (Table 2). The students' texts presented a similarly high frequency of *engagement*, representing 50 or more instances per 1000 words in both groups, which was substantially more frequent than the other two appraisal types. With regard to *attitude*, it was statistically more frequent in the LE group due to the higher use of *affect* among the students in that group.

In response to RQ 2, the second comparison focused on both groups' display of language knowledge in the segments where the students presented themselves as knowers (through the use of Appraisal), as illustrated in Table 3. Students' accurate use of appraisal resources was measured, both in terms of form and meaning. The three examples below, respectively, illustrate instances of inaccurate form, unclear

<sup>2</sup> <https://www.wgtn.ac.nz/lals/resources/academicwordlist/information/howto>

**Table 1** Translation device between students' performance (appraisal, form, and meaning) and the specialisation dimension

Descriptions of performances	Examples	The specialisation dimension
Texts with accurate forms, clear meaning, including words from the AWL and that use voice (evaluations) where needed show that their writers have the right type of language knowledge and project themselves as knowers.	Women have been <i>discriminated (Judgement)</i> and seen as sexual objects for decades.	ER++/SR++
Texts where the forms are inaccurate and the meaning is not coherent, but where students attempt to project their voice show that the writers have acquired unequal forms of knowledge and knowing, which is likely to affect the strength of their overall performance.	She <i>think (Engagement)</i> that the womans are superior and this womans <i>do a bat think (Judgement)</i> the feminist.	ER+/SR+

meaning, and finally inaccurate form and unclear meaning in the students' expression of social relations (SR+) through appraisal.

Example 1: '...and now woman will be doing whatever *she want*'

Example 2: 'I think that in other countries like Irak or Arabia Saudita *it is better to benefit society*'

Example 3: 'The politisian people said that *women want*'

The results presented in Table 3 show that the students in the HE-group used more accurate and clearer performance of all the appraisal types. In spite of the

**Table 2** Appraisal in students' texts in HE and LE exposure groups [In CorpusTool, +++ indicates significance at 98% level, ++ at 95% level, + at 90%.]

Feature	High exposure students		Low exposure students		T-Stat	Signif
	N	Per 1000 words	N	Per 1000 words		
Appraisal	N = 604		N = 349			
Attitude	124	21.27	97	28.54	2.57	+++
Engagement	318	54.55	170	50.01	1.17	
Graduation	162	27.79	82	24.12	1.13	
Attitude	N = 604		N = 349			
Affect	24	4.12	28	8.24	2.66	+++
Judgement	42	7.21	22	6.47	0.39	
Appreciation	58	9.95	47	13.83	1.84	+

**Table 3** Forms and meanings in students' use of appraisal in HE and LE exposure groups

	<i>N</i>	HE group per 1000 tokens	<i>N</i>	LE group per 1000 tokens	<i>T</i> -Stat	Signif
Quality affect	<i>N</i> = 604		<i>N</i> = 349			
Accurate/clear-affect	20	3.43	17	5.00	1.20	
Inaccurate/unclear-affect	4	0.69	11	3.24	2.99	+++
Inaccurate/unclear-affect	<i>N</i> = 604		<i>N</i> = 349			
Form-affect	4	0.69	10	2.94	2.73	+++
Meaning-affect	0	0	0	0.00	0.00	
Form and meaning-affect	0	0	1	0.29	0.00	
Quality judgement	<i>N</i> = 604		<i>N</i> = 349			
Accurate/clear-judgement	37	6.35	11	3.34	2.02	++
Inaccurate/unclear-judgement	5	0.86	11	3.24	2.70	+++
Inaccurate/unclear judgement	<i>N</i> = 604		<i>N</i> = 349			
Form-JUDGEMENT	5	0.86	6	1.77	1.24	
Meaning-JUDGEMENT	0	0	1	0.29	0.00	
Form and meaning-judgement	0	0	4	1.18	0.00	
Quality appreciation	<i>N</i> = 604		<i>N</i> = 349			
Accurate/clear-appreciation	51	8.75	28	8.24	0.23	
Inaccurate/unclear-appreciation	7	1.20	19	5.59	3.94	+++
Inaccurate/unclear appreciation	<i>N</i> = 604		<i>N</i> = 349			
Form-appreciation	4	0.69	11	3.24	2.99	+++
Meaning-appreciation	1	0.17	1	0.29	0.39	
Form and meaning-appreciation	2	0.34	7	2.06	2.58	+++
Quality-engagement	<i>N</i> = 604		<i>N</i> = 349			
Accurate/clear-engagement	240	41.17	90	26.48	4.40	+++
Inaccurate/unclear-engagement	78	13.38	80	23.54	4.03	+++
Inaccurate/unclear-engagement	<i>N</i> = 604		<i>N</i> = 349			
Form-engagement	53	9.09	54	15.89	3.17	+++
Meaning-engagement	11	1.89	4	1.18	0.81	
Form and meaning-engagement	14	2.40	22	6.47	3.12	+++
Quality-graduation	<i>N</i> = 604		<i>N</i> = 349			
Accurate/clear-graduation	142	24.36	67	19.71	1.55	
Inaccurate/unclear-graduation	20	3.43	15	4.41	0.78	
Inaccurate/unclear-graduation	<i>N</i> = 604		<i>N</i> = 349			
Form-graduation	17	2.92	14	4.12	1.00	
Meaning-graduation	2	0.34	0	0.00	0.00	
Form and meaning-graduation	1	0.17	1	0.29	0.39	

small numbers, the difference seems to be very significant in almost all appraisal types, where LE students show a more inaccurate/unclear use of appraisal resources compared to their peers in the HE strand. This relates mostly to formal accuracy (in *affect*, *appreciation*, and *engagement*) and also to the combination of

**Table 4** Measurement of differences between the HE and the LE groups in the use of academic lexicon using UAM Corpus tool

Academicness	HE group	LE group	ChiSqu	P val	Signif.
Academic word use	5.82%	3.75%	17.22	0.0000	+++

accurate form and clear meaning. In both aspects, HE students perform significantly better in two types of appraisal (*appreciation* and *engagement*).

With regard to academic lexis, the comparison of academicness between groups yielded significantly better results in the HE-group as well  $\{X^2 = 4.14\}$ , as shown in Table 4.

We used the keyword list from Antconc and the AWL from Coxhead to generate a list of academic lexis that was characteristic of the HE-group. Table 5 shows those words which appeared a minimum of two times in the target group (HE) and includes the range of students who generated these words in both groups.

**Table 5** Keyword list of academic lexis and the range of students responsible for their generation

Academic lexis	Frequency in HE	Frequency in LE	Range in HE	Range in LE
gender* (-s)	14	4	7	3
achiev*(-edl-mentl-ing)	11	0	7	0
benefi*(-cialltsltd)	6	0	6	0
ideolog* (-yl-ical)	6	0	5	0
economic* (-al)	5	0	5	0
objective	5	0	4	0
role	4	0	4	0
affect* (-ed)	3	0	3	0
create* (-d)	3	0	2	0
demonstrate	3	1	3	0
furthermore	3	0	3	0
radical* (-ized)	3	0	2	0
sex* (-esl-ism-ual)	5	5	4	3
specifi* (-cl-cally)	3	1	3	1
tasks	3	0	3	0
adapt	2	0	2	0
aspects	2	0	2	0
decades	2	1	2	1
domestic	2	0	2	0
goal* (-s)	2	0	2	0
participate	2	0	2	0
principle	2	0	2	0
sources	2	0	1	0
violation	2	0	1	0

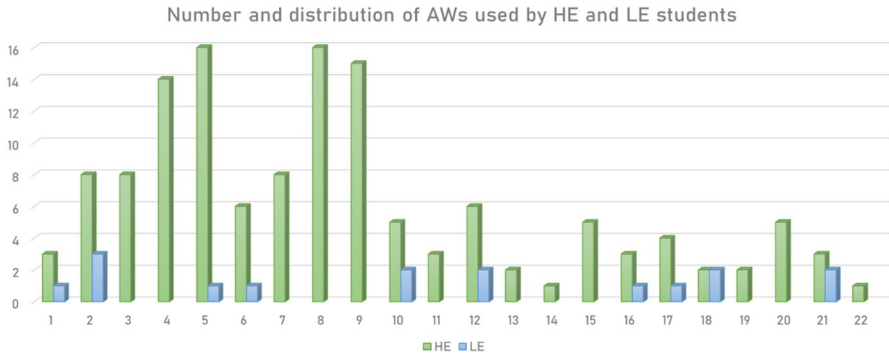


Fig. 3 Distribution of academic words by individual students in the HE and LE groups

Again, the HE group is at an advantage. In all cases, the words were either used more frequently than in the LE group or with the same frequency (only in one case). Also, the maximum range of students generating academic lexis in the HE was 7 (in two occasions), in comparison to a maximum of 3 students in the LE group. In addition, the other academic words that were used once ( $n = 38$ ) by the HE group were not used at all by the LE group. Figure 3 shows individual students’ use of academic lexis in each strand.

The results show that 12 students in the HE group (more than 50%) use more than 5 academic words. In contrast, none of the LE students use as many academic words, and more than 50% use no academic words at all.

Regarding the students’ performance in terms of their knowledge of academic words and as knowers through their expression of voice (RQ 3) in different CDFs, Fig. 4 shows that academic words appear mostly not only in *Define* (24%) but also in response to all the other CDFs.

*Report* is the CDF which includes the lowest instances, probably as this part of the prompt asked the students to refer to the events that took place on the 8th of

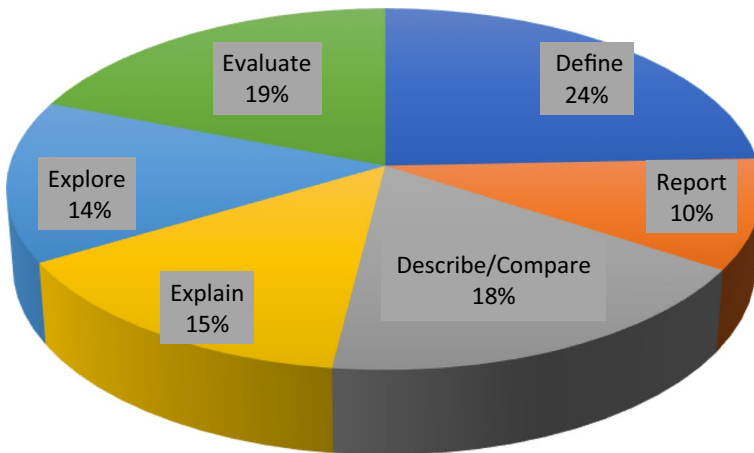
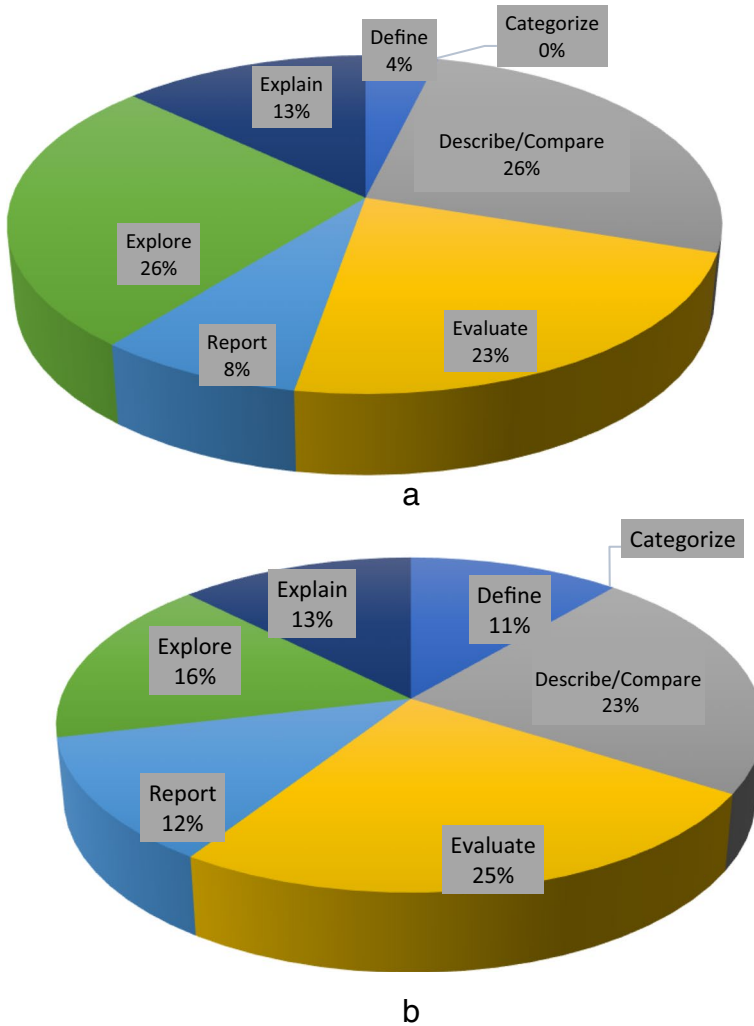


Fig. 4 Academic words across CDFs



**Fig. 5** **a** Appraisal in CDFs in the HE group. **b** Appraisal in CDFs in the LE group

March and, thus, did not elicit so many academic words. The use of academic words across CDFs shows that students are capable of using academic language even when they are asked to write about a general topic.

Regarding the second part of RQ 3, related to how students use their voice across CDFs, Fig. 5a, b shows the results in the two different strands.

The results show that, as expected, appraisal appears mostly when students *evaluate*, in the part responding to ‘**Do you think** the current women’s movement in Spain is benefitting society? **Why/why not?**’, also in *describe/compare* as the prompt required *attitude* (mainly *appreciation*): ‘**Describe** what life was like for women in your grandparents’ generation and **compare** it with women’s life today’. The cases of *explore* and *define* are interesting as differences are found across groups, where



**Table 6** Use of appraisal in *explore* in the HE and the LE groups

Feature	HE		LE		T-Stat	Signif
	N	Per 1000 tokens	N	Per 1000 tokens		
Appraisal type	N = 159		N = 52			
Attitude	37	6.91	11	5.06	0.31	
Engagement	89	16.61	26	11.97	0.75	
Graduation	33	6.16	15	6.92	1.21	

appraisal is more frequent in *define* in the LE group, whereas it is more frequent in *explore* in the HE group. We, then, investigated further the types of appraisal used across groups in response to each of these two CDFs.

The part of the prompt eliciting *explore* said: ‘**What do you think would happen** if Spanish companies were forced to have equal representation of men and women in high-level jobs?’ Table 6 shows the instances and types of appraisal used in response to *explore* across HE and LE groups.

The results show that the students in the HE group used more instances of *engagement* and *attitude* per 1000 words in comparison with their LE counterparts. *Engagement* is expected in order to hypothesise about a potential situation (‘*What do you think would happen...?*’), and *attitude* is also expected when referring to ‘*equal representation of men and women*’. Figure 6 illustrates the use of appraisal by a HE student in response to *explore*.

The same analysis was done in the case of the CDF *define*, where the students were expected to define the concept of Feminism (see Table 7).

The results show that students in the LE group used more of the three types of appraisal (*attitude*, *engagement*, and *graduation*) than in the HE group. The use of appraisal to define in the LE group is illustrated in Extract 2 (Fig. 7).

*I think, companies must search for the most **qualified people** without taking account of their gender, if a big company is looking for a person that knows how to promote their product and reach the most people they can, they won't look specifically for a man or a woman, they would look for the person with the **better CV in marketing**, So i think that companies were forced to hire the same amount of men than women they will lost the opportunity of hiring the best one*

**Fig. 6** Appraisal in CDFs in the LE group (‘engagement’ is in italics, ‘attitude’ is in bold, and ‘graduation’ is underlined)

**Table 7** Use of appraisal in *define* in the HE and the LE groups

Feature	HE		LE		T-Stat	Signif
	N	Per 1000 tokens	N	Per 1000 tokens		
Appraisal type	N = 23		N = 52			
Attitude	8	2.24	11	6.18	0.33	
Engagement	9	2.25	22	12.36	1.66	
Graduation	6	1.68	3	1.69	1.87	+

## Discussion

As observed in the results, regarding the students' response to the prompt as knowers (SR), the distribution of appraisal types showed similarities and differences. *Engagement* was the most frequent appraisal type in both HE and LE groups, probably linked to the part of the prompt in which students had to give an opinion (*evaluate*) and hypothesise (*explore*). However, the LE-group used *affect* detectably more, which is an emotion-based type of appraisal. According to Cavasso and Taboada (2021), objective opinions do not rely on emotions (*affect*) but on *judgements* and *appreciation*. In the same vein, in the case of CLIL students' academic writing in science, history, and art, Whittaker and McCabe (2023) consider evaluation-statements based on *affect* to be personal reactions, which are not expected in their expression of knowledge. This is in line with Martin's (1989) identification of the overuse of *affect* as a feature of immature writing. As knowers then, though both groups exhibit an equal capacity for the production of evaluation statements, the reliance of the LE-group on *affect* depicts them as being less mature as writers.

The results related to the students' (language) knowledge for the expression of meanings, measured through their use of academic words and the accuracy of forms and clarity of meanings conveyed through appraisal, revealed clear advantages for students in the HE-group. The students in this group used language more accurately when conveying their evaluations, and they were also able to integrate a more varied and broader set of academic lexis in their writing. This result is in line with previous research overviews by Dalton-Puffer (2011) and Nikula, Dalton-Puffer, and Llinares (2013), which show that CLIL often has a positive effect on students' general language use and vocabulary. This is also witnessed in our study, where the

*In my opinion, feminism is a movement to defend women's right's and to express the message that women are equal to men. I consider myself as a feminist person, because no one should be discriminated because of their ideas, ethnic group, sex, sexual orientation.*

**Unfortunately, there are people who don't think that way and they are stuck in the middle ages** (LE group)

**Fig. 7** Appraisal in CDFs in the HE-group ('engagement' is in italics and 'attitude' in bold)

LE exposure group assumes the position of the ‘non-CLIL’ groups of these former studies. The HE students’ attention to language form and structure rendered their texts as more linguistically accurate and meaningful. Given their much wider use of academic lexis, we take it that HE students were able to retain and recycle academic lexicon to a much higher extent than the students in the LE group.

Finally, moving to the use of appraisal for the expression of different CDFs, as previously noted in the results, some CDFs triggered the generation of academic lexis more (*define*) and others triggered the generation of appraisal forms more (*explore*). The latter was visible in the HE group’s performances. It was unorthodox to find statements of *engagement* in many of the LE groups students’ realisations of *define*, as it is a factual language function that does not require leaving the arena open for positions and views, unlike *explore*, which specifically requires this move. The inadequate use of appraisal statements in certain CDFs (e.g. offering opinions when defining) leads to the inadequate use of voice in the performance of academic language functions.

To sum up, the students in the HE group were not only able to choose the right appraisal resources expected for different CDFs. They also used more accurate language forms and were able to convey their meanings more clearly. In terms of knowledge building (ER, SR), the LE group appears weaker in terms of epistemic knowledge (ER+/SR+) than the HE group (ER++/SR++). Engaging with opinions and views is apparently not too challenging for the students, even for those who struggle to formulate messages that can be easily understood (mostly in the LE group). However, the choice of appraisal types was more consistent with the conventions of experienced writers in the HE group.

## Conclusion

Teachers of English as a Foreign Language teach the particularities of the language (structure, vocabulary, essay writing...) using contexts that are relevant to the students’ worldly interests (e.g. social media) or to the students’ other areas of studies (e.g. the historical importance of women’s roles). The students then bring to the activity their trajectory of acquired knowledge, guided by the language principles from the language classes, knowledge of related topics from other subjects, out-of-school experiences, and their personal dispositions and qualities as authors.

In line with the studies that have reported better general language results for CLIL students in comparison with non-CLIL students, this study shows that this seems to be the case when different types of CLIL groups are compared in the same school. In line with Muñoz (2015), advantages may not apply, then, equally to all types of CLIL. However, if we focus on students’ language resources for knowledge building (both epistemologically and ontologically), we argue that the difference may not just lie in the amount of CLIL but in the type of subjects (more or less academic) that students learn in the L2. If, in addition to being more accurate in the language forms to express appraisal meanings, HE students also choose the correct appraisal resources as well as academic words for the expression of different CDFs, it could be argued that these students may be transferring the academic language resources

that they use in history, science, etc. to their writing performance on a more general topic, which they have not studied as part of the curriculum. Particularly where CLIL is taught in more practical/less academic subjects, our findings point to the importance of promoting pluriliteracies (general, academic, and subject-specific skills and strategies) (Meyer & Coyle, 2017) and language resources in the L2 to aid students recontextualise knowledge and transfer it across subjects.

Finally, identifying and describing both ER and SR relations in students' practices could help teachers avoid reducing knowledge to ER without considering SR, for example, establishing formal knowledge as the basis of achievement and disregarding how students engage with a topic. On the other hand, the emphasis on the knower (SR) dimension where unneeded may get in the way of communicating content in a clear academic stance. It is, then, important that teachers familiarise students with the discourse functions that require 'voice' and those that do not.

**Availability of Data and Materials** Not applicable.

**Author Contributions** The authors have contributed equally to the analysis of the data and the writing of the manuscript, so the names are listed in alphabetical order (Llinares and Nashaat-Sobhy).

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by the Ministerio de Ciencia e Innovación (Spain) (Award Number: RTI2018-094961-B-I00).

## Declarations

**Ethical Approval and Consent to Participate** This study is part of a research project granted by Ministerio de Ciencia e Innovación (Spain) Award Number: RTI2018-094961-B-I00, which involved an ethical approval of the project. By participating in the study, students had consented to using their anonymised texts for research purposes.

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Anthony, L. (2010). *AntConc (Version 3.2.1)*. [Computer Software].
- Artieda, G., Roquet, H., & Nicolás-Conesa, F. (2020). The impact of age and exposure on EFL achievement in two learning contexts: formal instruction and formal instruction + content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 23(4), 449–472.
- Bieri, A. S. (2019). Technicality and translanguaging in CLIL biology lessons in Switzerland. *Bulletin vals-asla*, 110, 109–124.

- 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.
- Cavasso, L., & Taboada, M. (2021). A corpus analysis of online news comments using the Appraisal framework. *Journal of Corpora and Discourse Studies*, 4, 1–38.
- Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: taking stock and looking forward. *Applied Linguistics*, 35, 243–262.
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213–238.
- Cummins, J. (2000). *BICS and CALP* (pp. 76–79). Encyclopedia of language teaching and learning.
- Dalton-Puffer, C. (2011). Content-and-language integrated learning: from practice to principles? *Annual Review of applied linguistics*, 31, 182–204.
- 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.
- 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. (2019). Cognitive Discourse Functions meet historical competences: towards an integrated pedagogy in CLIL history education. *Journal of Immersion and Content-Based Language Education*, 7(1), 30–60.
- 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.
- Dalton-Puffer, C., Hüttner, J., & Llinares, A. (2022). CLIL in the 21st century: retrospective and prospective challenges and opportunities. *Journal of Immersion and Content-Based Language Education*, 10(2), 182–206.
- DePalma, M. J., & Ringer, J. M. (2011). Toward a theory of adaptive transfer: expanding disciplinary discussions of “transfer” in second-language writing and composition studies. *Journal of Second Language Writing*, 20(2), 134–147.
- Evnitskaya, N., & Dalton-Puffer, C. (2023). Cognitive discourse functions in CLIL classrooms: eliciting and analysing students’ oral categorizations in science and history. *International Journal of Bilingual Education and Bilingualism*, 26(3), 311–330.
- Fernández-Agüero, M., & Hidalgo-McCabe, E. (2022). CLIL Students’ affectivity in the transition between education levels: the effect of streaming at the beginning of secondary education. *Journal of Language, Identity & Education*, 21(6), 363–377.
- Gee, J. P., & Handford, M. (Eds.). (2012). *The Routledge handbook of discourse analysis*. Routledge.
- Goris, J., Denessen, E., & Verhoeven, L. (2019). Effects of content and language integrated learning in Europe: a systematic review of longitudinal experimental studies. *European Educational Research Journal*, 18(6), 675–698.
- James, M. A. (2006). Teaching for transfer in ELT. *ELT Journal*, 60(2), 151–159.
- Järvinen, H. (2010). Language as a meaning making resource in learning and teaching content. Language use and language learning in CLIL classrooms. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 145–168). John Benjamins.
- Jexenflucker, S., & Dalton-Puffer, C. (2010). The CLIL differential: comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 169–190). John Benjamins.
- Jordão, C. M., & Fogaça, F. C. (2012). Critical literacy in the English language classroom. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 28, 69–84.
- Llinares, A. (2015). Integration in CLIL: a proposal to inform research and successful pedagogy. *Language, Culture and Curriculum*, 28(1), 58–73.
- Llinares, A., & Evnitskaya, N. (2021). Classroom interaction in CLIL programs: offering opportunities or fostering inequalities? *TESOL Quarterly*, 55(2), 366–397.
- Llinares, A., & Nashaat-Sobhy, N. (2021). What is an ecosystem? Defining science in primary school CLIL contexts. *Language Teaching for Young Learners*, 3(2), 337–362.
- Lorenzo, F. (2017). Historical literacy in bilingual settings: cognitive academic language in CLIL history narratives. *Linguistics and Education*, 37, 32–41.
- Martin, J. R. (1989). *Factual writing: exploring and challenging social reality*. OUP.
- Martin, J. R., & Rose, D. (2003). *Working with discourse: meaning beyond the clause*. Continuum.

- Martin, J. R., & White, P. R. (2005). *The language of evaluation: appraisal in English*. Palgrave MacMillan.
- Maton, K. (2014). *Knowledge and knowers: towards a realist sociology of education*. Routledge.
- Meyer, O., & Coyle, D. (2017). Pluriliteracies teaching for learning: conceptualizing progression for deeper learning in literacies development. *European Journal of Applied Linguistics*, 5(2), 1–24.
- Muñoz, C. (2015). Time and timing in CLIL: a comparative approach to language gains. In I. M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based learning in multilingual educational environments* (pp. 31–51). Springer.
- Nashaat-Sobhy, N. (2018). Operationalizing ‘defining’ from a cognitive discourse perspective for learners’ use. In S. M. Anwaruddin (Ed.), *Knowledge mobilization in TESOL: teachers’ experiences of connecting research and practice*. Leiden.
- Nikula, T. (2015). Hands-on tasks in CLIL science classrooms as sites for subject-specific language use and learning. *System*, 54, 14–27.
- Nikula, T., Dalton-Puffer, C., & Llinares, A. (2013). CLIL classroom discourse: research from Europe. *Journal of Immersion and Content-Based Language Education*, 1(1), 70–100.
- OECD. (2019a). *In brief. Knowledge for 2030*. [https://www.oecd.org/education/2030-project/teaching-and-learning/learning/knowledge/Knowledge\\_for\\_2030\\_concept\\_note.pdf](https://www.oecd.org/education/2030-project/teaching-and-learning/learning/knowledge/Knowledge_for_2030_concept_note.pdf)
- OECD. (2019b). *In brief. Attitudes and Values for 2030* [https://www.oecd.org/education/2030-project/teaching-and-learning/learning/aar-cycle/in\\_brief\\_AAR\\_Cycle.pdf](https://www.oecd.org/education/2030-project/teaching-and-learning/learning/aar-cycle/in_brief_AAR_Cycle.pdf).
- O’Donnell, M. (2008). Demonstration of the UAM CorpusTool for text and image annotation. In J. Lin (Ed.), *Proceedings of the ACL-08: HLT Demo Session Companion Volume* (pp. 13–16). Association for Computational Linguistics.
- Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna språk*, 109(2), 51–74.
- Pérez-Vidal, C., & Roquet, H. (2015). The linguistic impact of a CLIL Science programme: an analysis measuring relative gains. *System*, 54, 80–90.
- Ruiz de Zarobe, Y. (2010). Written production and CLIL: an empirical study. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 191–209). John Benjamins.
- Shor, I. (1980). *Critical teaching and everyday life*. Black Rose Books Limited.
- Somers, T., & Llinares, A. (2021). Students’ motivation for content and language integrated learning and the role of programme intensity. *International Journal of Bilingual Education and Bilingualism*, 24(6), 839–854.
- Whittaker, R., & McCabe, A. (2023). Expressing evaluation across disciplines in primary and secondary CLIL writing: a longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 26(3), 345–362.
- Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, 15(3), 343–362.
- Worthington, D., & Nation, P. (1996). Using texts to sequence the introduction of new vocabulary in an EAP course. *RELC journal*, 27(2), 1–11.

## Authors and Affiliations

Ana Llinares<sup>1</sup>  · Nashwa Nashaat-Sobhy<sup>2</sup> 

✉ Ana Llinares  
ana.llinares@uam.es

Nashwa Nashaat-Sobhy  
nash@idm.upv.es

<sup>1</sup> Universidad Autonoma de Madrid-Campus de Cantoblanco, 28049 Madrid, Spain

<sup>2</sup> Universitat Politècnica de Valencia, c/Paranimf, 1, 46730 Gandia, Spain