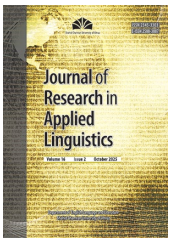
Journal of Research in Applied Linguistics
ISSN: 2345-3303 – E-ISSN: 2588-3887 – https://rals.scu.ac.ir
Published by Shahid Chamran University of Ahvaz

Please cite this paper as follows:

Llinares, A., Morton, T., & Nashaat-Sobhy, N. (2025). An analysis of presence in CLIL students' writing when defining and exploring in science. *Journal of Research in Applied Linguistics*, *16*(2), 115-132. https://doi.org/10.22055/rals.2025.50266.3612

Shahid Chamran University of Ahvaz



Research Article

An Analysis of Presence in CLIL Students' Writing When Defining and Exploring in Science

Ana Llinares[1], Tom Morton[2], & Nashwa Nashaat-Sobhy[3]

[1]Department of English Philology, Faculty of Philosophy and Arts, Autonomous University of Madrid, Madrid, Spain; *ana.llinares@uam.es*

[2]Department of English Philology, Faculty of Philosophy and Arts, Autonomous University of Madrid, Madrid, Spain; *tom.morton@uam.es*

[3]Department of Applied Linguistics, Polytechnic University of Valencia; Valencia, Spain; *nash@idm.upv.es*

Received: 25/07/2025 Accepted: 19/10/2025

Abstract

This paper investigates the written texts produced by students in Year 7 and Year 10 (beginning and end of compulsory secondary education in Spain) in a content and language integrated learning (CLIL) setting in Spain, with English as the language of instruction. The students responded to a prompt eliciting the same Cognitive Discourse Functions (CDF) on the same topic in a biology course. Using the SFL concept of presence, we examine its three aspects—iconicity (degree of abstraction or concreteness in encoding ideas), negotiability (arguability of propositions), and implicitness (how much meaning is recoverable from the text alone). Following previous studies that have applied SFL to the analysis of students' production of CDFs, the study compares the students' use of presence in the CDFs, DEFINE and EXPLORE, across levels of schooling, and relates this analysis to English and biology teachers' assessment of the texts using comparative judgment. Results show the onset of development in students' grammatical metaphors, with older learners displaying reduced iconicity, greater abstraction, higher epistemic negotiability, and more appropriate use of implicitness. DEFINE and EXPLORE differed systematically: Definitions were more factual and explicit, while Explorations exhibited higher negotiability and implicitness, though not always appropriately. Teachers' ratings, however, did not reflect these indicators of disciplinary literacy. We identify pedagogical implications regarding the role of concept dependency as a component of disciplinary literacy in bilingual education.

Keywords: Bilingual Education; Cognitive Discourse Functions (CDFs); Disciplinary Literacy; SFL.

1. Introduction

In many classrooms around the world, students are learning academic/school content through a second language (L2), an educational practice commonly referred to as Content and Language Integrated Learning (CLIL). These students must not only acquire disciplinary knowledge in subjects such as science, mathematics, or history, but also communicate that knowledge through the specialised discourses of each discipline. This requires the development of disciplinary literacy (DL), which refers to the specific ways of thinking, communicating, and meaning-making that are intrinsic to academic disciplines (Moje, 2015; Shanahan & Shanahan, 2012). DL is increasingly recognised as central to academic success, particularly when students are from linguistically and socially diverse backgrounds (Fang & Coatoam, 2013; Moje, 2007; Spires et al., 2018).

In CLIL classrooms, however, the development of DL is complicated by the fact that students are required to acquire discipline-specific literacy practices through an L2, making disciplinary literacy an inherently multilingual phenomenon in these contexts. While the first wave of CLIL research focused on L2 acquisition more generally, recent years have seen the emergence of DL as a specific research focus within CLIL scholarship. An early example was Llinares et al. (2012), which drew on Systemic Functional Linguistics to conceptualise how students in CLIL classrooms construct disciplinary meanings through oral and written language. More recently, Coyle and Meyer's (2021) pluriliteracies model

explicitly frames disciplinary literacy development as a multilingual, multimodal, and cognitively demanding process essential to deeper learning in CLIL contexts.

One response to the need for a firmer conceptual grounding for building DL in CLIL contexts is Dalton-Puffer's (2013) construct of Cognitive Discourse Function (CDF). Drawing on applied linguistics, education, and pragmatics, CDFs conceptualise how language expresses the cognitive processes involved in building disciplinary knowledge. The framework identifies seven core academic functions (categorize, define, describe, evaluate, explain, explore, report) that are found across disciplines. In this study, CDFs are used to examine how secondary CLIL biology students respond to prompts designed to elicit two of these functions: DEFINE and EXPLORE, providing insight into how learners use language to construct disciplinary meanings through an L2.

While CDFs offer a means of identifying how language expresses cognitive processes, they can be strengthened by engaging with broader principles of knowledge building within disciplines. From a knowledge-building perspective, as articulated in Legitimation Code Theory (LCT) (Maton, 2014), knowledge is real, structured, and consequential, shaping how educational practices are realised and how knowledge claims are recognised and legitimated. A key aspect of knowledge building involves managing the relationship between abstraction and concreteness by connecting general concepts to situated examples and controlling how meanings are embedded in communicative contexts. In LCT, this is theorised as shifts in semantic gravity, where weaker gravity allows for abstraction, while stronger gravity grounds meanings in concrete situations. This movement is particularly challenging in student writing, where learners must balance disciplinary precision with accessibility.

In Systemic Functional Linguistics (SFL), the concept of semantic gravity has prompted the development of the presence framework (Martin & Matruglio, 2020), which provides tools for analysing whether knowledge is made context-dependent, situated in the "here and now", or abstracted beyond immediate circumstances. The framework distinguishes three dimensions through which this relationship is managed: iconicity, the degree to which language reflects the material world; negotiability, the extent to which meanings invite interpersonal engagement; and implicitness, how much knowledge is backgrounded or taken for granted. Integrating presence with CDF analysis allows for a more nuanced investigation of how students construct disciplinary meanings, examining not only what cognitive functions are performed but how meanings are contextualised in ways that align with disciplinary knowledge-building practices.

Given the limited emphasis on DL in many CLIL classrooms, it is likely that teachers assess students' work without explicit awareness of the specific CDFs that students are expected to perform or of how language affects context-embeddedness. In the case of CDFs, teachers may not apply any explicit criteria to assess how well students produce definitions or evaluations, for example. Regarding context-embeddedness, students' ability to relate language to context by mirroring physical processes as they are observed, or to weaken semantic gravity by using grammatical metaphor (conveying meaning through an unexpected grammatical form) to express abstraction, may go unnoticed. One way to explore the tacit criteria teachers apply when assessing writing is through Comparative Judgement (CJ). CJ involves teachers making holistic, pairwise comparisons of texts to determine which is stronger, without relying on analytic rubrics. Research shows CJ produces reliable rankings of student work (Bouwer et al., 2024; Sickinger et al., 2025) while capturing teachers' intuitive, experience-based perceptions of quality. Combining CJ with teacher reflection can reveal the often-unspoken bases for recognising achievement in educational settings.

Building on these conceptual tools, this study investigates how secondary CLIL students use the linguistic resources of presence to vary context dependency in written responses to prompts which elicit the CDFs of DEFINE and EXPLORE. The study forms part of a larger project examining disciplinary literacy development within a national bilingual education programme in Spain. The dataset comprises 50 texts produced by Year 7 and Year 10 students responding to grade-appropriate prompts on the topic of cells, designed to elicit a range of CDFs, including DEFINE and EXPLORE. All texts were evaluated by English and biology teachers using CJ, generating ranked outcomes reflecting teachers' holistic perceptions of achievement.

The remainder of the paper is structured as follows: Section 2 reviews the theoretical background, outlining disciplinary literacy, Cognitive Discourse Functions, presence, and Comparative Judgement. Section 3 presents the methodology and data. Section 4 reports the results, followed by discussion in Section 5. Section 6 offers conclusions and implications for CLIL disciplinary literacy development.

2. Theoretical Background

2.1. Disciplinary Literacy and CLIL

Students in CLIL engage with subject matter through an additional language (L2), which is often English (Dalton-Puffer & Smit, 2013). Managing both the content and the language of instruction places dual demands on learners, who have to acquire and display disciplinary knowledge in English. This dual focus creates a distinctive learning environment that offers insight into how disciplinary literacy develops in multilingual contexts. The challenges that arise at the intersection of learning both epistemic content knowledge (in other words, the nature of knowledge in each discipline) and the language of that content in an L2 make it essential for teachers and students to negotiate how disciplinary knowledge is constructed, communicated, and evaluated.

In these contexts, linguistic mediation is particularly salient. Language not only supports access to content but also becomes an object of learning in its own right. Nikula et al. (2024) describe CLIL as an ideal testbed for studying and supporting disciplinary literacies because it makes visible the linguistic and epistemological demands that are often taken for granted in first-language instruction. They argue that CLIL draws attention to the kinds of meaning-making that different subjects require, including the use of academic language functions such as defining, explaining, evaluating, or hypothesizing. These functions demand more than surface-level linguistic accuracy. Where language is concerned, students are expected to manipulate the ideational (representing ideas), interpersonal (managing social relations), and textual (organizing coherent messages in texts) metafunctions, and the language resources to convey them, in ways that reflect disciplinary modes of reasoning, for instance, through abstraction, cohesive reference, or hedged claims.

However, the extent to which CLIL classrooms foster these disciplinary discourse functions remains uneven. Drawing on a systematic review of scaffolding practices in CLIL science education, Nashaat-Sobhy et al. (2024) report that transsemiotizing and translanguaging are frequently used by teachers to support students' access to, and expression of, epistemic knowledge, yet opportunities for students to engage in disciplinary language production depend largely on the cognitive and linguistic demands introduced through teacher questioning and probing. Similarly, the CLILNetLE WG2 report (Guzmán-Alcón & Lorenzo, 2025) indicates that while teachers across European contexts recognize the importance of disciplinary literacy, its development is often implicit. Instruction tends to prioritize subject content or general language skills, with limited focus on how language functions to construct disciplinary meanings. Moreover, coordination between subject and language teachers is limited, and training in integrating disciplinary discourse practices into classroom tasks remains scarce.

2.2. Disciplinary Literacy in Science

Literacy demands in content subjects, like science, become increasingly specialised in secondary education, with the result that general literacy strategies, such as activating prior knowledge or even summarising, are rarely sufficient to sustain learning complex knowledge (Elish-Piper et al., 2016). Advancing literacies in science requires students to express knowledge in ways that align with the epistemological norms and communicative practices of the discipline.

Scientific discourse conveys complex phenomena in technical and conceptually dense ways, often relying on nominalisation to condense dynamic processes into compact noun phrases (e.g., reproduction, fertilisation) (Lemke, 1990; Knain, 2015). Scientific texts also favour impersonal structures, the passive voice, and cohesive reference chains, all of which contribute to precision, objectivity, and coherence (Knain, 2015). These linguistic features in science writing are not just stylistic conventions but are central to the disciplinary practices of science and support students in constructing knowledge that is valued within the field. Students are expected to perform communicative tasks that range from defining and characterizing phenomena to exploring them and reasoning about possible hypotheses. These tasks demand a degree of abstraction and generalisation, which in turn rely on linguistic resources such as grammatical metaphor, modality, and cohesive reference. The ability to condense meaning into noun phrases, signal degrees of certainty, and maintain coherence across clauses becomes essential for constructing knowledge in ways that reflect the epistemological norms of the discipline.

However, research suggests that the development of these features is inconsistent across CLIL contexts. Based on teacher-generated texts to illustrate expected student output, and on curricular analyses across multiple European contexts, the CLILNetLE WG3 report (Ting et al., 2024) shows that nominalisation is expected in upper-secondary texts,

especially highly performing ones, but not at lower levels, aligning with earlier findings (e.g., Whittaker et al. 2011), and not in all contexts. In the Italian science samples, texts predominantly used active voice and lacked nominalisations, suggesting that students are not systematically taught to use abstract, discipline-specific grammatical forms. In Austria, English versions of teacher texts often displayed fewer details, less precision, and weaker cohesion compared to those written in German, though such differences were not consistent across prompts. The report also notes that some of the teacher model texts were not unified by a particular communicative intention, suggesting that texts written as isolated short answers appear to lack cohesion. The variation among teachers' texts in the CLILNetLE report is also seen in actual students' texts as shown in Llinares and Nikula (2024), who found differences between Spanish and Finnish groups in *appraisal* resources (Martin & White, 2005) and complex nominal groups when performing different CDFs. While the report helped clarify teachers' expectations for good texts, the patterns indicate that attention to the linguistic features necessary for expressing disciplinary knowledge can vary greatly, which may be due to the assumption that language development will occur incidentally (Guzmán-Alcón & Lorenzo, 2025).

2.3. Cognitive Discourse Functions (CDFs)

In addition to the language focus in disciplinary literacy, special attention is given to the cognitive habits involved in meaning-making and the types of texts that shape disciplinary thinking and practice (Manderino, 2012; Shanahan & Shanahan, 2008; Wilson, 2011). In this view, developing disciplinary literacy involves not only acquiring specialised language but also learning to think in ways aligned with disciplinary norms.

The framework of CDFs (Dalton-Puffer, 2013) offers a principled way to investigate how students engage in disciplinary thinking through language. They are defined as "verbalizations linked to cognitive processes routinely performed in the course of dealing with curricular content…" that are "analogues of thought" (Dalton-Puffer, 2018, p. 9), and that can be "identified through their lexico-grammatical realisation" (Dalton-Puffer, 2013, p. 217). In other words, they link communicative intention, cognitive process, and linguistic realisation. The framework consists of seven meaning-making operations observed to recur across school subjects: DEFINE, DESCRIBE, EXPLAIN, EVALUATE, CATEGORISE, REPORT, and EXPLORE. The way these functions are realised lexicogrammatically varies depending on the knowledge structures of the field (see Nashaat-Sobhy & Morton, forthcoming), whether by topic or subject. This variability makes CDFs particularly useful for analysing student discourse in CLIL contexts, especially where students are expected to produce relatively short written responses rather than staged genres.

A growing body of research has used CDFs to explore how students construct disciplinary meanings when writing. For example, EVALUATE has been examined in relation to how students build arguments or express value judgments in science and history (Whittaker & McCabe, 2023), while COMPARE has been analysed in tasks involving classification and contrast in biology and geography (Arias-Hermoso et al., 2025). DEFINE in Llinares and Nashaat-Sobhy (2021) has been analysed as a gateway to conceptual knowledge requiring students to categorise and specify key entities and processes in science. They also examined teacher evaluations of texts through comparative judgements, and reflections on desirable linguistic and conceptual precision in students' texts. Meanwhile, EXPLORE, associated with hypothetical reasoning and tentative explanation, has been examined in recent work by Llinares and Morton (2024) in both science and history. Their findings suggest that when students attempt to speculate or consider alternative possibilities, limitations in their linguistic resources can constrain their ability to engage in open-ended, exploratory reasoning, though more in English (L2) than in Spanish (L1).

This study builds on this research by focusing specifically on the latter two functions: DEFINE and EXPLORE in secondary school biology writing. These functions were specifically selected because they represent clear distinct ways of engaging with disciplinary knowledge; while DEFINE typically demands factual precision and categorical knowledge, EXPLORE requires students to formulate tentative claims and consider alternative possibilities. Although it is reasonable to hypothesize that these two CDFs would elicit different degrees of certainty and therefore different levels of context-dependency, it remains an empirical question about how students actually manage this in writing. In addition, we examine whether referents are recoverable within the boundaries of each CDF, and if they compact information through grammatical metaphor. By analysing these functions in students' productions, and how biology and English teachers evaluate them through comparative judgement, this study aims to shed light on how linguistic resources contribute to the construction and recognition of disciplinary knowledge in CLIL science classrooms.

2.3. Context in Disciplinary Writing: Semantic Gravity and Presence

CDFs offer valuable insights into how language realises different types of thinking across school subjects. However, CDFs are, by design, a general construct with broad applicability to disciplinary communication. Their roots lie in educational taxonomies such as Anderson and Krathwohl's (2001) revision of Bloom's framework, which map learning objectives as cognitive processes. While this is a productive starting point, it risks focusing primarily on what learners do with language, without addressing how disciplinary knowledge itself is structured and built. Maton (2014) argues that much education research suffers from "knowledge blindness", which he describes as a tendency to foreground processes and ways of knowing, rather than engaging with the organising principles of knowledge within disciplines. CDFs, if framed purely as cognitive operations, can risk falling into this blind spot. The aim here is not to critique the construct, but to extend its usefulness by adding a knowledge-sensitive lens, drawing on concepts from Legitimation Code Theory (LCT).

One powerful tool from LCT is semantic gravity, which captures the extent to which meanings depend on immediate context (Maton, 2020). When semantic gravity is strong, meanings are tightly tied to concrete, experiential, or situational references. When it is weak, meanings are abstracted, generalised, and less dependent on the surrounding context. However, to track how student writers manage context-dependence in their texts, we need tools that operate at the level of wording and discourse. This linguistic work can be complemented by bringing in presence, a framework from Systemic Functional Linguistics (Martin & Matruglio, 2020), which provides a way to track how meanings are projected into the material, social, and textual environment.

Presence captures how language positions meanings relative to the immediate context, focusing on three dimensions:

Iconicity refers to the degree of congruence between the wording and the material or experiential world. Highly iconic expressions closely mirror physical experience, while less iconic expressions rely on abstract, nominalised, or technical language. For example, in history the use of "*increase* taxes" is iconic, while "the *increase* of taxes" is less iconic.

Negotiability tracks the extent to which meanings are open to interaction, contestation, or dialogue. High negotiability involves hedging, questioning, or dialogic engagement, while low negotiability presents meanings as authoritative or fixed. For example, in science, "*I think* viruses are different from bacteria" has high negotiability, while "Viruses are different from bacteria" expresses no negotiability.

Implicitness concerns whether meanings are made explicit within the text or depend on contextual, shared knowledge. High implicitness requires readers or listeners to infer meaning, while low implicitness provides all necessary information within the text itself. For example, in geography, "*the Nile*" is explicit while "the river" is implicit.

Presence thus extends the notion of context beyond mode (spoken, written, visual) to include field (the nature of the activity or subject matter) and tenor (the social relationships and roles involved), offering a set of resources for analysing how language constructs meaning in context.

Previous studies have applied presence to spoken disciplinary discourse. Hood (2020), for example, used the framework to explore how university lecturers manage context-dependence in undergraduate biology lectures. Here, we extend the application to a secondary school setting, examining how presence operates in written student texts produced in response to tasks designed to elicit the CDFs DEFINE and EXPLORE. In student writing, we can expect different CDFs to display distinctive presence profiles. For instance, we might expect DEFINE to involve low iconicity, with meanings typically realised through abstract, nominalised language that departs from direct, experiential references. However, expansions or elaborations of definitions may increase iconicity by introducing concrete examples or describing features. Similarly, DEFINE tends to exhibit very low negotiability, as definitions are conventionally presented as fixed and authoritative. Implicitness is also expected to be low, as effective definitions should be self-contained, providing clear meanings without requiring readers to infer from surrounding context. In contrast, CDFs such as EXPLORE may show greater iconicity when examples are introduced, higher negotiability as possibilities are entertained, and higher implicitness as meanings unfold across extended stretches of text rather than being fully explicit upfront.

These expectations about presence profiles across CDFs lead directly into the next section, where we explore how teachers make evaluative judgements about student texts and whether constructs such as CDFs and Presence can illuminate those processes.

2.4. The Role of Presence in Teachers' Comparative Judgements of Student Text Quality

Measuring disciplinary literacy in CLIL is often approached using tools from education and applied linguistics, such as multidimensional analysis (Qiu et al., 2022) or automated web-based tools like Coh-Metrix 3 (Lorenzo et al., 2019). Qiu et al. (2022) use multidimensional analysis to explore register variation in student writing, focusing on how linguistic choices reflect shifts in formality, abstraction, and context-dependence, which are core phenomena in the development of disciplinary literacy. Importantly, Qiu et al. also demonstrate that such linguistic patterns are not merely theoretical constructs but are reflected in teachers' assessments, with greater abstraction and reduced context-dependence generally aligning with higher teacher evaluations of academic quality. Lorenzo et al. (2019) apply Coh-Metrix to analyse a range of textual features, including lexical density, syntactic complexity, and discourse coherence, providing insight into how students' linguistic choices contribute to fluency, precision, and cognitive demands in disciplinary writing.

While studies such as Qiu et al. (2022) provide important evidence of alignment between linguistic features and teacher evaluations, there is still more to understand about how these evaluations are formed, particularly in CLIL contexts where content and language demands interact. Whether teachers are consciously aware of it or not, their evaluations are based on underlying notions of achievement or quality. For example, in a CLIL context, Morton and Nashaat-Sobhy (2024) found that both content and language teachers articulated bases of achievement when judging students' written texts, often privileging content quality over quantity, and placing greater emphasis on formal aspects of language rather than its functional effectiveness for expressing academic content. Their study was situated within a broader project that used the method of comparative judgement (CJ) to explore how teachers evaluate student writing.

Comparative judgement is based on the principle that humans are generally better at making evaluative decisions when comparing examples rather than applying fixed, absolute criteria to individual cases. This insight dates back to Thurstone's work in the 1920s (Thurstone, 1927) and has since been developed as a reliable method for assessing open-ended student work, particularly writing tasks (Pollitt, 2012; Wheadon et al. 2020). Both Pollitt (2012) and Wheadon et al. (2020) argue that CJ offers a valid and reliable alternative to traditional marking for complex tasks such as writing, by harnessing human expertise through comparative rather than criteria-based judgements. Bouwer et al. (2024) provide further evidence that CJ can often outperform rubric-based approaches, particularly for open tasks. In a young learner EFL writing context, Sickinger et al. (2025) found that although some teachers preferred more familiar analytic rating methods, their judgements were more reliable when using CJ.

In the ongoing research project from which this study is derived, CJ has been used as a tool to elicit the bases of achievement that both content and language teachers draw upon when evaluating texts written by CLIL students. CJ has not been proposed as an assessment method to be implemented within the participating schools, although some teachers have expressed interest in its potential. An advantage of using CJ platforms such as No More Marking is that they allow researchers to separate and compare the evaluative behaviour of different groups of teachers, in this case, content specialists and language specialists. Metrics such as infit (a measure of consistency in a judge's rankings) can be used to assess the reliability of the judgements. Additionally, the ranked sets of student texts generated through CJ can serve as a stimulus for post-ranking discussions, in which teachers reflect upon and articulate the criteria or bases of achievement underlying their evaluations of text quality (see Llinares, Morton & Whittaker, 2024; Morton, 2022; Morton & Llinares, 2024; Morton & Nashaat-Sobhy, 2024).

In this study, we have drawn on CJ-generated rankings to investigate how the linguistic features associated with presence appear in texts that were ranked highly or lowly by both content and language teachers. This provides a means of exploring whether the presence dimensions of iconicity, negotiability, and implicitness, through which students manage context in disciplinary writing, are visible to teachers, and whether they influence perceptions of quality in CLIL student writing.

Bringing together the conceptual frameworks of cognitive discourse functions, semantic gravity as modulated through the resources of presence, and the method of comparative judgement, the study sets out to address the following research questions:

1. How do Year 7 and Year 10 students in CLIL biology use linguistic resources of presence to manage context-embeddedness when responding to prompts eliciting DEFINE and EXPLORE?

2. How do DEFINE and EXPLORE differ in terms of their presence profiles in student writing?

3. How are features of presence distributed across texts ranked high and low in CJ by biology and English teachers?

3. Methodology

3.1. Participants and Corpus

Our corpus comprised 50 written responses produced by secondary school students in CLIL science classes, conducted in English as an L2. The sample included 25 texts from Year 7 and 25 from Year 10. The data were selected from 10 different schools in Spain and corresponded to the texts that had been evaluated by biology and English teachers through CJ. All the schools participated in the same bilingual program (BEP-joint agreement between the British Council and the Ministry of Education in Spain).

Students completed a writing task during regular classroom hours in response to the prompts in Figure 1 without external intervention. The prompts elicited students' production of different CDFs: DEFINE, CATEGORISE, EXPLORE and REPORT (in the case of Year 7), and DEFINE, CATEGORISE, EXPLAIN and EXPLORE (in the case of Year 10). They were developed in collaboration with subject teachers, aligned with curriculum content and age-appropriate discourse expectations. In the present study, we focus on the students' responses to the part of the prompt eliciting DEFINE: What are cells? (Year 7), Define the two cell divisions: mitosis and miosis (Year 10), and EXPLORE: What would happen if someone discovered a living organism without cells? (Year 7), What do you think would happen if DNA replicated between the first and second meiotic divisions? (Year 10).

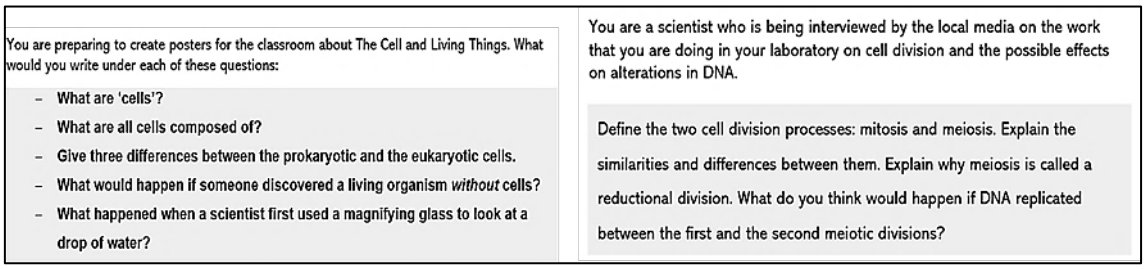


Figure 1. The prompts designed for Year 7 (Cells) and Year 10 (Cell division)

To evaluate the texts, teachers of biology and English from the BEP used comparative judgment (CJ) via the No More Marking platform (see 2.4.).

For qualitative and quantitative analysis, we focused on a subset of the corpus: the highest- and lowest-ranked texts per year group. This allowed us to explore how students with varying degrees of perceived achievement performed the targeted discourse functions and managed contextual dependency.

3.2. Coding & Data Analysis

All 50 texts were manually annotated using the UAM CorpusTool (O'Donnell, 2023). The analysis involved three successive layers:

CDF Coding: Clauses were first coded according to the CDF framework (Dalton-Puffer, 2013), with a focus on the functions DEFINE and EXPLORE. DEFINE was coded using the schema designed by Nashaat-Sobhy & Llinares (2023), where the CDF is composed of the core definition (defined term, class-word, and specifying features) and further expansions to elaborate, clarify, or to simply extend. In the same vein, a similar schema was created for EXPLORE, where the core is represented in a claim, and the elaboration (expansion) is the rationale for the claim when that was provided. Figure 2 shows the CDF coding system, where the response to the prompted CDF was coded as macro-CDF and expansions (for both DEFINE and EXPLORE) were coded as micro-CDF (as they add information to the core CDF).

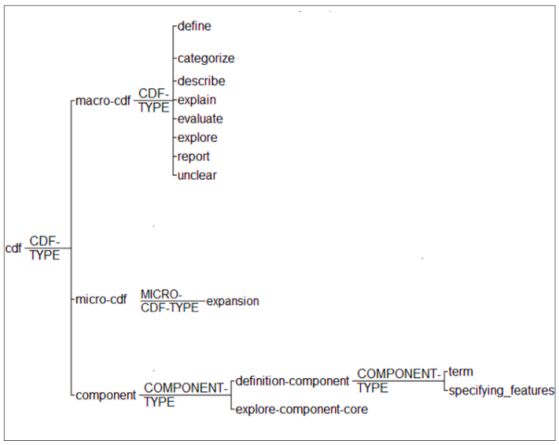


Figure 2. Framework for the coding of CDFs

Presence Analysis: Clauses performing these functions were then annotated for the type of presence, following Martin and Matruglio's (2020) categories—iconicity, negotiability, and implicitness—to assess the degree of contextual dependency (see Figure 3). For iconicity, we focused on the use (or not) of grammatical metaphor (iconicity_plus or minus). For negotiability, we focused on the presence (or not) of appropriate modality (negotiability_plus or minus). For implicitness, we focused on use of appropriate referential elements (implicitness_plus or minus). The categories were also coded for appropriateness/inappropriateness. For example, in the following student's response to the prompt for EXPLORE in Year 7 (*What would happen if someone discovered a living organism without cells*), which said: *It can't happen because all living organism contains cell,* we coded the use of *it* as inappropriate because identifying what *it* referred to required consulting the prompt.

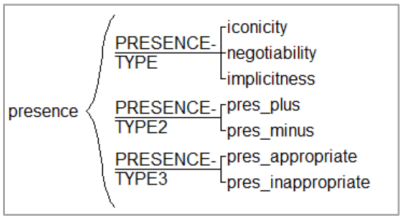


Figure 3. Framework for the coding of presence

CJ Ratings: Each text's CJ score was recorded to identify responses rated high or low, with texts scoring above 50/100 treated as high-rated and those below 50/100 as low-rated.

To ensure reliability in the CDF and presence coding, all data were coded by the three researchers to resolve discrepancies until mutual agreement was achieved.

For the quantitative analysis, we calculated the global percentage of each presence dimension. This allowed us to compare the relative distribution of presence features across school years and rating categories.

4. Results

The results report the use of presence (iconicity, negotiability and implicitness) across school years, across CDFs and in relation to the higher- and lower-rated texts by teachers. Note that, in general, there are more coded segments for year 10 than for year 7. This is due to the fact that year 10 students' texts were generally longer, reflecting the more complex nature of the content in the prompts.

4.1 Presence Across School Levels

For iconicity in DEFINE, year 10 students produced fewer segments with presence features than those in year 7 (Table 1), showing that their language increasingly used grammatical metaphor to express their ideas. As for iconicity in

EXPORE, there was a much greater balance between segments with and without presence features in year 10 (Table 2), suggesting that iconicity does not become a critical feature of this CDF as students gain disciplinary literacy in this subject.

Table 1. *Iconicity in DEFINE Across School Years*

	Year 7		Year 10	
	N	*%*	*N*	*%*
Segments with presence features	33	100.0%	71	68.9%
Segments without presence features	0	0.0%	32	31.1%
Appropriate use of presence features	33	100.0%	103	100.0%
Inappropriate use of presence features	0	0.0%	0	0.0%
TOTAL	33	100.0%	103	100.0%

Table 2. *Iconicity in EXPLORE Across School Years*

	Year 7		Year 10	
	N	*%*	*N*	*%*
Segments with presence features	0	0.0%	11	42.3%
Segments without presence features	4	100.0%	15	57.7%
Appropriate use of presence features	4	100.0%	26	100.0%
Inappropriate use of presence features	0	0.0%	0	0.0%
TOTAL	4	100.0%	26	100.0%

Examples 1 and 3 from Year 10 show the use of grammatical metaphor (in bold), while examples 2 and 4 from Year 7 show the use of verbs only. Interestingly, the expansions in DEFINE in Year 10 included more abstractions while the core definitions in year 7 were more iconic (see Example 2). All the uses were considered appropriate. In using grammatical metaphor (nominalisations), students were expanding their linguistic resources to meet the disciplinary literacy demands of the biology subject at this higher level of schooling.

Example 1 DEFINE:

The second one is meiosis in which two diploid cells are divided into four haploid cells, also in this process, there is sexual ***reproduction***, ***variability*** *of genetic information* (Year 10)

Example 2 DEFINE:

A cell ***is*** *a smallest living organism* (Year 7)

Example 3 EXPLORE:

we will have at the diploid cell that can't be use for sexual ***reproduction*** *because during* ***fertilization*** *we will have tetraploid and we need a diploid cell* (Year 10)

Example 4 EXPLORE:

It can't ***happen*** *because all living organism* ***contains*** *cell* (Year 7)

In the case of negotiability, as expected, Table 3 shows that instances of DEFINE were mostly factual, with over 90% of coded segments showing no negotiability in both years. In the case of EXPLORE, there was slightly more negotiability in Year 10, suggesting some tentativeness in the scenarios produced in the students' explorations, and this was seen as appropriate.

Table 3. *Negotiability in DEFINE Across School Years*

	Year 7		Year 10	
	N	*%*	*N*	*%*
Segments with presence features	2	6.1%	2	1.9%
Segments without presence features	31	93.9%	103	98.1%
Appropriate use of presence features	32	97.0%	103	98.1%
Inappropriate use of presence features	1	3.0%	2	1.9%
TOTAL	33	100.0%	105	100.0%

Table 4. *Negotiability in EXPLORE Across School Years*

	Year 7		Year 10	
	N	*%*	*N*	*%*
Segments with presence features	3	21.4%	7	26.9%
Segments without presence features	11	78.6%	19	73.1%
Appropriate use of presence features	13	92.9%	26	100.0%
Inappropriate use of presence features	1	7.1%	0	0.0%
TOTAL	14	100.0%	26	100.0%

There was also a difference in the function of negotiability across school years: year 10 examples are almost strictly epistemic outcomes (Example 5), whereas in year 7, the students used modality and modal expressions more often to express social outcomes (see Example 6). This also reflects the nature of the prompts, as the one in year 10 converged much more closely on disciplinary knowledge. Examples 5 and 6 also show how students at both levels were able to use hypothetical language (modal *would* and modal adjunct *probably*), thus demonstrating developing ability to use these linguistic resources. Part of disciplinary literacy is the ability to use linguistic resources to manage epistemic commitment to ideas related to the subject.

Example 5 EXPLORE:

If DNA replicated between the two processes of meiosis i ***would say*** *that the final formation* ***would not be*** *gametes or they* ***will be*** *gametes formed by the formation of n/2n* (Year 10)

Example 6 EXPLORE:

If someone discovered an organism without cells, the person who discovered it ***would probably*** *get a prize or a Guinness world record* (Year 7)

Finally, Tables 5 and 6 illustrate the level of implicitness in students' texts across school years.

Table 5. *Implicitness in DEFINE Across School Years*

	Year 7		Year 10	
	N	*%*	*N*	*%*
Segments with presence features	5	15.2%	30	29.4%
Segments without presence features	28	84.8%	72	70.6%
Appropriate use of presence features	31	93.9%	91	89.2%
Inappropriate use of presence features	2	6.1%	11	10.8%
TOTAL	33	100.0%	102	100.0%

Table 6. *Implicitness in EXPLORE Across School Years*

	Year 7		Year 10	
	N	*%*	*N*	*%*
Segments with presence features	4	100%	10	40.0%
Segments without presence features	0	0%	15	60.0%
Appropriate use of presence features	0	0%	19	76.0%
Inappropriate use of presence features	4	100%	6	24.0%
TOTAL	4	100.0%	25	100.0%

The percentages show that in year 10 in DEFINE there are higher levels of implicitness, probably due to the fact that students produced more expansions in their definitions and used anaphoric reference to pick up entities in the core. Example 7 shows appropriate use of anaphoric reference to refer to a process that has already been mentioned in the text and is easily retrievable for the reader. Example 8 refers to the prompt outside the text; thus, it is less obvious to the reader what "this" refers to. In dealing with implicitness, students need to control linguistic resources for endophoric and exophoric reference. A key part of disciplinary literacy is managing epistemic access – that is, what background knowledge a reader/interlocutor is expected to have.

Example 7 EXPLORE:

This process *is used for growth and for the replacement of cells that has been damaged* (Year 10)

Example 8 EXPLORE:

I think ***this*** *isn't happening never because all living organisms have cells* (Year 7)

4.2 Presence Across CDFs

Regarding the differences across CDFs, Tables 7, 8 and 9 show the differences and similarities in the levels of iconicity, negotiability and implicitness, respectively.

Table 7. *Iconicity Across CDFs*

	Define		Explore	
	N	%	N	%
Segments with presence features	104	76.5%	15	50%
Segments without presence features	32	23.5%	15	50%
TOTAL	136	100%	30	100%

Contrary to expectations, we found more iconic uses in DEFINE and more grammatical metaphors (less iconic) in EXPLORE, although that was particularly the case in year 7. We also find examples in year 10 such as the following (Example 9):

Example 9 DEFINE:

Mitosis ***creates*** *daughter cells with the same number of chromosomes as mother cells* (Year 10)

Regarding negotiability, as expected, it appeared more in EXPLORE (see Table 8 below) and its use was highly appropriate across both CDFs.

Table 8. *Negotiability across CDFs*

	Define		Explore	
	N	%	N	%
Segments with presence features	4	2.9%	7	23.3%
Segments without presence features	134	97.1%	23	76.7%
Appropriate use of presence features	135	97.8%	30	100%
Inappropriate use of presence features	3	2.2%	0	0%
TOTAL	138	100%	30	100%

Example 10 shows how a year 10 student uses linguistic resources of modality to refer to a hypothetical process. Disciplinary literacy at this level involves using linguistic resources to manage degrees of certainty, especially when students are invited to explore alternative scenarios.

Example 10 EXPLORE:

If this process ***could*** *happen, the zygote* ***would*** *die or be formed in a baby who* ***could*** *born with some kind of illness or syndrome* (Year 10)

Finally, Table 9 shows the levels of implicitness across CDFs. Again, as expected, instances of DEFINE were more explicit and self-contained, while EXPLORE included more instances of implicitness.

Table 9. *Implicitness across CDFs*

	Define		Explore	
	N	%	N	%
Segments with presence features	35	25.9%	14	48.3%
Segments without presence features	100	74.1%	15	51.7%
Appropriate use of presence features	122	90.4%	19	65.5%
Inappropriate use of presence features	13	9.6%	10	34.5%
TOTAL	135	100.0%	29	100.0%

These instances of EXPLORE were often inappropriate. As shown in the example below (Example 11), it is not clear what "it" is referring to.

Example 11 EXPLORE:

*I think that **it** would not earn gametes* (Year 10)

4.3 Presence and Teachers' Judgements

Tables 10 to 13 show the levels of presence (iconicity, negotiability and implicitness) in relation to the judgements by English teachers (E-high and E-low) and biology teachers (B-high and B-low) for those texts rated higher and lower, respectively.

Regarding iconicity (Table 10), the absence of this feature (which is seen in texts with more abstractions and nominalisations) did not seem to play a role in teachers' ratings. Although there were texts with grammatical metaphor among the texts that were rated higher by the biology teachers, this was not the case for the texts rated higher by the English teachers.

Table 10. *Iconicity and Comparative Judgment*

	B-high		B-low		E-high		E-low	
	N	%	N	%	N	%	N	%
Segments with presence features	82	68.3%	59	85.5%	101	75.4%	40	72.7%
Segments without presence features	38	31.7%	10	14.5%	33	24.6%	15	27.3%
Appropriate use of presence features	120	100.0%	69	100.0%	134	100.0%	55	100.0%
Inappropriate use of presence features	0	0.0%	0	0.0%	0	0.0%	0	0.0%
TOTAL	120	100.0%	69	100.0%	134	100.0%	55	100.0%

In a focus group session with English and biology teachers focusing on these texts, the use of nominalisations was clearly not a priority for them. For them, clarity and accuracy of the message was key whether students were using verbs or nouns, even highlighting the advantages of using verbs. Some of the biology teachers' comments were: "It is easier for the students to use verbs", "I like seeing science as a movie". Even English teachers prioritised communication over the use of disciplinary/academic language: "…as long as communication comes across…".

Regarding negotiability, as previously mentioned, definitions are factual and non-negotiable (high epistemic modality). However, the presence of inappropriate negotiability in definitions did not affect good ratings (see Table 11). Something similar happened in EXPLORE (Table 12). Although hypothetical language (e.g., *would*, *could*, *might*) is essential to distinguish real observed phenomena from imagined scenarios, and we would expect higher rated texts to have more instances of high negotiation, we found the opposite: texts rated relatively higher by biology and English teachers had in total more instances of low negotiation (high epistemic modality).

Table 11. *Negotiability in DEFINE and Comparative Judgment*

	B-high		B-low		E-high		E-low	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Segments with presence features	6	22.2%	4	30.8%	7	24.1%	3	27.3%
Segments without presence features	21	77.8%	9	69.2%	22	75.9%	8	72.7%
Appropriate use of presence features	27	100.0%	12	92.3%	29	100.0%	10	90.9%
Inappropriate use of presence features	0	0.0%	1	7.7%	0	0.0%	1	9.1%
TOTAL	27	100.0%	13	100.0%	29	100.0%	11	100.0%

Table 12. *Negotiability in EXPLORE and Comparative Judgment*

	B-high		B-low		E-high		E-low	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Segments with presence features	3	3.4%	1	2.1%	3	3.1%	1	2.5%
Segments without presence features	86	96.6%	47	97.9%	94	96.9%	39	97.5%
Appropriate use of presence features	88	98.9%	47	97.9%	96	99.0%	39	97.5%
Inappropriate use of presence features	1	1.1%	1	2.1%	1	1.0%	1	2.5%
TOTAL	89	100.0%	48	100.0%	97	100.0%	40	100.0%

Similarly, appropriate uses of implicitness did not seem to play a role in teachers' judgments (Table 13). The higher rated texts by biology and English teachers included more instances of inappropriate uses of pronouns than the lower rated ones, so again the appropriate use of referential elements does not seem to play any role. In fact, in the focus groups the participating teachers mentioned textual features, such as the importance of using connectors, but not the importance of explicitness or appropriate implicitness.

Table 13. *Implicitness and Comparative Judgment*

	B-high		B-low		E-high		E-low	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Segments with presence features	40	33.1%	16	23.2%	48	35.6%	8	14.5%
Segments without presence features	81	66.9%	53	76.8%	87	64.4%	47	85.5%
Appropriate use of presence features	99	81.8%	60	87.0%	108	80.0%	51	92.7%
Inappropriate use of presence features	22	18.2%	9	13.0%	27	20.0%	4	7.3%
TOTAL	121	100.0%	69	100.0%	135	100.0%	55	100.0%

5. Discussion

The present study focused on how CLIL students of science managed context-embeddedness in biology texts responding to prompts eliciting different CDFs. We compared students' production in the CDFs DEFINE and EXPLORE across school levels to characterize the different role of presence in these two CDFs, and explored the relevance of presence in the texts evaluated higher or lower by biology and English teachers.

The comparison of presence across school years yielded different results across the three dimensions: iconicity, negotiability and implicitness. There seems to be development in the use of grammatical metaphor (lower iconicity) from year 7 to year 10 in both CDFs. Interestingly, expansions included more abstractions than the core definition. As pointed out in Llinares & Nashaat-Sobhy (2021), expansions can either unpack or repack technical concepts depending on the type of expansion. Overall, the results suggest that students' expressions of the target phenomena gradually shift from being framed as concrete experiences to abstract conceptualisations, signaling that scientific meaning-making becomes increasingly disciplinary in the higher school years. Condensing processes into nominal groups (e.g., "*cell division*," "*genetic variability*") is not merely linguistic sophistication, but a way of construing entities as objects of knowledge. Hence, the increased use of grammatical metaphor in year 10 points to growing control of representing knowledge in science.

In the case of negotiability, definitions were mostly factual in both years and explorations showed higher negotiability, which increased slightly in year 10. Negotiability was more social in year 7 and more epistemic in year 10, either due to epistemic development in the field or perhaps due to the different prompts (a more common-sense answer would be expected in year 7 in contrast with a more epistemic answer in year 10). This shift from a rather social-oriented answer to an epistemic one also indicates increasing disciplinary maturity. Students in higher years move from providing personal stances to constructing claims about how biological systems operate. Year 10 students' execution of negotiability can also be seen as another indicator of their expanding capacity to engage with the epistemic and linguistic requirements of the prompts, moving between factuality and tentativeness, hypothesis-building, and reasoning with evidence.

In the case of implicitness, year 10 students' texts tend to be more implicit, as they wrote more, but the differences were slight. In EXPLORE, students in year 10 produced more instances of appropriate use of implicitness. Managing implicitness through references allows students to background and compress shared information, which is another indicator of maturing scientific discourse.

As for the characterization of CDFs, as expected, negotiability was higher in EXPLORE (where possibilities are expected to be entertained in contrast with the more factual approach to content in DEFINE). Implicitness was also more present in EXPLORE, as predicted, as meanings are developed in more stretched texts, compared to the more explicit textbook-like responses expected in definitions. However, implicitness was sometimes used inappropriately. Contrary to expectations, definitions depicted higher levels of iconicity than explorations. Grammatical metaphors were more present in the expansions of definitions, which have been shown to both enhance or reduce contextualisation depending on the type (Llinares & Nashaat-Sobhy, 2021). From an epistemological disciplinary literacy perspective, the CDFs did not vary solely due to their inherent linguistic expectations (e.g., DEFINE factual, EXPLORE hypothetical) but also in response

to the topic and prompt. The year 7 prompts ("*What are cells?*", "*What would happen if someone discovered a living organism without cells?*") elicited shorter definitions and common-sense reasoning, whereas the Year 10 prompts on mitosis and meiosis required longer definitions and clearly more epistemic reasoning grounded in disciplinary knowledge. Hence, variation across CDFs reflects not only functional but also topical and epistemological demands. What students were asked to DEFINE or EXPLORE, and their topical complexity in relation to students' ability to use the language, shaped how students employed presence in their answers.

Moving now to the discussion of the relation between the levels and appropriateness of presence and biology and English teachers' assessment, presence did not seem to play a big role in teachers' ratings. We found no relation between more frequent use of grammatical metaphor and teachers' higher ratings. In fact, in the focus groups that followed the analysis and teachers' assessments, both biology and English teachers prioritised communication and content delivery over students' use of a more academic/scientific language to convey their knowledge. The same applies to negotiability, where higher-rated texts included higher epistemic modality and less negotiability than lower-rated texts. Appropriate use of implicitness did not seem to play a role in teachers' higher ratings, either. This misalignment reveals that teachers' evaluative frameworks are not yet sensitive to linguistic features of context-embeddedness as indicators of disciplinary literacy.

6. Conclusion

The analysis of presence in CLIL students' definitions and explorations in science presented in this study has yielded contributions at the empirical, conceptual, methodological and pedagogical levels. Empirically, the analysis of CLIL students' responses to the same topic at years 7 and 10 (beginning and end of lower secondary in Spain) has shown development in the use of grammatical metaphor as well as more appropriate uses of pronouns to refer to implicit meanings. There also seems to be development in the higher use of negotiation in year 10 in EXPLORE. Although the study has controlled the topic in both years (cells), the number of texts was limited as they had to represent the sample that was judged by biology and English teachers. Future studies should focus on a larger number of texts.

Conceptually, the study has contributed to the characterization of the CDFs DEFINE and EXPLORE in terms of the expected levels of iconicity, negotiability and implicitness expected. Although the analysis of students' texts has shown our initial expectations of higher negotiability and implicitness in EXPLORE and lower in DEFINE, the characterization of these CDFs in relation to iconicity would require further research, taking into account educational levels and school disciplines. It seems that the presence or absence of expansions as well as the type may play a role in the levels of iconicity/grammatical metaphor produced by the students. Future studies should also attempt to characterize other CDFs taking into account students' productions, teachers' expectations and curriculum and textbook materials. In addition, CDFs could also be characterised in terms of mass (Martin, 2020).

In spite of the levels of development and appropriate use of CDFs by the students in the texts analysed, the teachers' judgments mostly did not align with the students' appropriate use of presence. The focus group seminar carried out by the teachers after their judgments shows that attention to disciplinary linguistic features as those discussed here were not a conscious priority, as opposed to communicating content knowledge. This lack of alignment underscores a crucial pedagogical implication: if disciplinary literacy is to be fostered systematically, teachers must learn to recognise and value the linguistic resources through which disciplinary knowledge is constructed. CJ-based reflection can be a powerful tool for making such criteria explicit, encouraging joint professional development among content and language teachers. We want to highlight the methodological contribution of relating content and language teachers' assessment through comparative judgment with our own analysis of presence and the role of the focus groups including content teachers, language teachers and researchers (see also Llinares & Morton, 2024; Morton & Nashaat-Sobhy, 2024). Pedagogically, more similar focus group seminars are needed to generate disciplinary literacy awareness leading to joint content and language teacher continuous professional development programs on the role of presence and other linguistic resources in CDFs in specific subjects and different educational levels. As Hu and Gao (2021, p. 57) indicate, content teachers need to "develop pedagogical skills to analyse and understand the features of the language used in their subjects". From this standpoint, integrating presence analysis within CJ discussions can extend such initiatives, demonstrating the value of attention to iconicity, negotiability and implicitness for disciplinary meaning making.

To conclude, bringing together CDFs, semantic gravity, and presence offers a more nuanced, layered account of how students engage with both the thinking and knowledge-building demands of disciplinary literacy. Beyond providing analytical insight, this combined approach has clear pedagogical relevance. It shows how linguistic development across school levels reflects the developmental trajectory of disciplinary literacy, while differences between CDFs trace the epistemological organisation of disciplinary knowledge. These tools may enable teachers to evaluate not only the content and linguistic accuracy of student writing, but also how effectively students manage context-dependence, abstraction, and meaning-making, which are all essential components of disciplinary literacy.

Acknowledgments

We would like to thank the participating schools and teachers in the study.

Author Contributions

The authors have contributed equally to the article.

Conflict of Interest

The authors have no conflicts of interest to declare.

Ethics Board Approval Statements

The study focuses on schools participating in the BEP project in Spain. The British Council obtained the approval from the participating schools (447465088 PFTC4) ensuring the anonymization and confidentiality of the data collected. Informed consents were obtained from the participants.

Funding

This article has been funded by the Ministerio de Ciencia, Innovación y Universidades (Spain). Programa Estatal de I+D+i (PID2022-140718OB-I00).

References

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives: Complete edition*. Longman.

Arias-Hermoso, R., Imaz Agirre, A., & Garro Larrañaga, A. (2025). Multilingual disciplinary literacies: Exploring developmental patterns of science writing across secondary education. *Journal of Multilingual and Multicultural Development*. https://doi.org/10.1080/01434632.2025.2481200

Bouwer, R., Lesterhuis, M., De Smedt, F., Van Keer, H., & De Maeyer, S. (2024). Comparative approaches to the assessment of writing: Reliability and validity of benchmark rating and comparative judgement. *Journal of Writing Research, 15*(3), 498-518. https://doi.org/10.17239/jowr-2024.15.03.03

Coyle, D., & Meyer, O. (2021). *Beyond CLIL: Pluriliteracies teaching for deeper learning.* Cambridge University Press.

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

Elish-Piper, L., L' Allier, S. K., Manderino, M., & Di, D. P. (2016). *Collaborative coaching for disciplinary literacy: Strategies to support teachers in grades 6-12.* Guilford Publications.

Fang, Z., & Coatoam, S. (2013). Disciplinary literacy: What you want to know about it. *Journal of Adolescent & Adult Literacy, 56*(8), 627-632. https://doi.org/10.1002/JAAL.190

Guzmán-Alcón, I., & Lorenzo, F. (2025). Aspects of disciplinary literacies in multilingual Europe: Rethinking policies for languages of schooling. *European Journal of Language Policy*, *17*(1), 45-64. https://doi.org/10.3828/ejlp.2025

Hood, S. (2020) Live lectures: The significance of presence in building disciplinary knowledge. In J. R. Martin, K. Maton, & Y. J. Doran, (Eds.), *Accessing academic discourse: Systemic functional linguistics and legitimation code theory* (pp. 211–235). Routledge.

Hu, J., & Gao, X. (2021). Understanding subject teachers' language-related pedagogical practices in content and language integrated learning classrooms. *Language Awareness, 30*(1), 42-61. https://doi.org/10.1080/09658416.2020.1768265

Knain, E. (2015). *Scientific literacy for participation: A systemic functional approach to analysis of school science discourses.* Springer. https://doi.org/10.1007/978-94-6209-896-1

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

Llinares, A., & Morton, T. (2024). Assessing CLIL students' expression of Explore across languages and school disciplines: An interdisciplinary approach. In Hüttner, J. & C. Dalton-Puffer (Eds.), *Building disciplinary literacies in content and language integrated learning* (pp. 101-118). Routledge. https://doi.org/10.4324/9781003403685-7

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

Llinares, A., Morton, T., & Whittaker, R. (2024). Fostering language awareness for integration through teacher-researcher collaboration in a Spanish bilingual education context. *Language Awareness, 33*(4), 751-771. https://doi.org/10.1080/09658416.2024.2385766

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. https://doi.org/10.1075/ltyl.20010.lli

Llinares, A., & Nashaat-Sobhy, N. (2023). CLIL students' academic language performance on a non-curricular topic: A comparison between high-exposure and low-exposure groups. *English Teaching & Learning, 47*(3), 337-358. https://doi.org/10.1007/s42321-023-00151-y

Llinares, A., & Nikula, T. (2024). CLIL students' production of cognitive discourse functions: Comparing Finnish and Spanish contexts. *Language and Education, 38*(3), 381-400. https://doi.org/10.1080/09500782.2023.2211049

Lorenzo, F., Granados, A., & Avila, 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*, 100767. https://doi.org/10.1016/j.jeap.2019.06.010

Martin, J. R. (2020) Revisiting field: Specialized knowledge in secondary school science and humanities discourse. In J.R. Martin, K. Maton, & Y.J. Doran (Eds.), *Accessing academic discourse* (pp. 114-147). Routledge.

Martin, J. R., & Matruglio, E. (2020). Revisiting mode: Context in/dependency in Ancient History classroom discourse. In J.R. Martin, K. Maton, & Y.J. Doran (Eds.), *Accessing academic discourse* (pp. 89-113). Routledge.

Martin, J. R., & White, P. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

Manderino, M. (2012). Disciplinary literacy in new literacies environments: Expanding the intersections of literate practice for adolescents. In P.J. Dunston, S. King, S. Fullerton, C.C. Bates, K. Headley, & P.M. Stecker (Eds.), *61st yearbook of the Literacy Research Association* (pp. 69– 83). Literacy Research Association.

Maton, K. (2014). *Knowledge and knowers: Towards a realist sociology of education*. London: Routledge.

Maton, K. (2020) Semantic waves: Context, complexity and academic discourse. In J. R. Martin, K. Maton, & Y. J. Doran (Eds.), *Accessing academic discourse* (pp. 59–85). Routledge.

Moje, E. B. (2007). Developing socially just subject-matter instruction: A review of the literature on disciplinary literacy teaching. *Review of Research in Education, 31*(1), 1-44. https://doi.org/10.3102/0091732X07300046001

Moje, E. B. (2015). Doing and teaching disciplinary literacy with adolescent learners: A social and cultural enterprise. *Harvard Educational Review, 85*(2), 254-278. https://doi.org/10.17763/0017-8055.85.2.254

Morton, T. (2022). Using cognitive discourse functions and comparative judgement to build teachers' knowledge of content and language integration for assessment in a bilingual education program. *Journal of Immersion and Content-Based Language Education, 10*(2), 302-322. https://doi.org/10.1075/jicb.21017.mor

Morton, T., & Llinares, A. (2024). Building teachers' knowledge of cognitive discourse functions to integrate content and language. In S. Ballinger, R. Fielding, & D. J. Tedick (Eds.), *Teacher development for content-based language education: International perspectives* (pp. 21-41). Multilingual Matters.

Morton, T., & Nashaat-Sobhy, N. (2024). Exploring bases of achievement in content and language integrated assessment in a bilingual education program. *TESOL Quarterly, 58*(1), 5-31. https://doi.org/10.1002/tesq.3207

Nashaat-Sobhy, N., & Llinares, A. (2023). CLIL students' definitions of historical terms. *International Journal of Bilingual Education and Bilingualism, 26*(3), 331-344. https://doi.org/10.1080/13670050.2020.1798868

Nashaat-Sobhy, N., & Morton, T. (forthcoming). Assessing knowledge complexity in science achievement in CLIL. In P. Gerns, L. Mortimore, & R. Breeze (Eds.), *Language and content in CLIL science and technology*. De Gruyter Mouton.

Nashaat-Sobhy, N., Mestre-Mestre, E. M., & Mac Donald, P. (2024). A review of a decade of scaffolding practices for learning in CLIL science classrooms. *Journal of Immersion and Content-Based Language Education*, *3*(2), 216-247 https://doi.org/10.1075/jicb.23010.nas.

Nikula, T., Nashaat-Sobhy, N., Minardi, S., Gülle, T., Yalçın, Ş., Kara Duman, S., Bozbıyık, M., Delibegović Džanić, N., Ellison, M., Gerns, P., Gómez, E., Hanušová, S., Kováčiková, E., Kääntä, L., Lin, A., Llinares, A., Lo, Y. Y., Lopriore, L., Meyer, O., . . . Žmavc, J. (2024, March). *Towards an initial operationalisation of disciplinary literacies*. Retrieved from on July 11, 2025 from https://hdl.handle.net/11353/10.2050621.

O'Donnell, M. (2023). *UAM CorpusTool* (Version 3.3) [Computer Software]. http://www.corpustool.com/index.html

Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice, 19*(3), 281-300. https://doi.org/10.1080/0969594X.2012.665354

Qiu, X., Lo, Y. Y., & San Teng, X. (2022). Development of L2 disciplinary literacy: A multidimensional analysis. *System*, *110,* 102907. https://doi.org/10.1016/j.system.2022.102907

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*, 40– 59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Shanahan, T., & Shanahan, C. (2012). What is disciplinary literacy and why does it matter? *Topics in Language Disorders, 32*(1), 7–18. https://doi.org/10.1097/TLD.0b013e318244557a

Sickinger, R., Brunfaut, T., & Pill, J. (2025). Comparative Judgement for evaluating young learners' EFL writing performances: Reliability and teacher perceptions of holistic and dimension-based judgements. *Language Testing, 42*(2), 137-166. https://doi.org/10.1177/02655322241288847

Spires, H. A., Kerkhoff, S. N., Graham, A. C., Thompson, I., & Lee, J. K. (2018). Operationalizing and validating disciplinary literacy in secondary education. *Reading and Writing 31*(6), 1401-1434. https://doi.org/10.1007/s11145-018-9839-4

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*(4), 273–286. https://doi.org/10.1037/h0070288

Ting, Y. L. T., Rieder-Marschallinger, S., & Dalton-Puffer, C. (2024). *The development of bi- and multilingual disciplinary literacies across educational levels: An exploratory survey* [CLILNetLE Working Group 3 Report]. https://www.clilnetle.eu/output

Wheadon, C., Barmby, P., Christodoulou, D., & Henderson, B. (2020). A comparative judgement approach to the large-scale assessment of primary writing in England. *Assessment in Education: Principles, Policy & Practice, 27*(1), 46-64. https://doi.org/10.1080/0969594X.2019.1700212

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. https://doi.org/10.1080/13670050.2020.1798869

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, *15*(3), 343-362. https://doi.org/10.1177/1362168811401154

Wilson, A. A. (2011). A social semiotics framework for conceptualizing content area literacies. *Journal of Adolescent and Adult Literacy, 54*(6), 435– 444. https://doi.org/10.1598/JAAL.54.6.5

odoulou, D., & Henderson, B. (2020). A comparative judgement approach to the large-scale assessment of primary writing in England. *Assessment in Education: Principles, Policy & Practice, 27*(1), 46-64. https://doi.org/10.1080/0969594X.2019.1700212

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. https://doi.org/10.1080/13670050.2020.1798869

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, *15*(3), 343-362. https://doi.org/10.1177/1362168811401154

Wilson, A. A. (2011). A social semiotics framework for conceptualizing content area literacies. *Journal of Adolescent and Adult Literacy, 54*(6), 435– 444. https://doi.org/10.1598/JAAL.54.6.5

© 2025 by the authors. Licensee Shahid Chamran University of Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution–NonCommercial 4.0 International (CC BY-NC 4.0 license). (http://creativecommons.org/licenses/by-nc/4.0/).