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:

Tompkins. L. (2025). Beyond the facts: Evaluative language use in secondary CLIL History. *Journal of Research in Applied Linguistics*, *16*(2), 133-147. https://doi.org/10.22055/rals.2025.49979.3558

Shahid Chamran University of Ahvaz

Research Article

Beyond the Facts: Evaluative Language Use in Secondary CLIL History

Leah Tompkins

Corresponding author, Department of English Philology, Faculty of Philosophy and Letters, Universidad Autónoma de Madrid, Madrid, Spain; *leah.tompkins@uam.es*

Received: 15/06/2025 Accepted: 13/08/2025

Abstract

This study investigates how Spanish CLIL students in year 10 express evaluative meanings in written responses to a prompt on the French Revolution. Drawing on Cognitive Discourse Functions (CDFs) and the Appraisal framework, it analyzes 119 texts for students' use of evaluative language and compares performance across higher- and lower-SES schools. Findings show that while students broadly reproduce the value positions modeled in their textbooks, they often struggle to make evaluations explicit, rely on everyday lexis, or overstate disciplinary claims. Students at lower-SES schools more frequently used personal stance markers and made evaluations that were less contextually appropriate. These results suggest a need for explicit scaffolding in discipline-specific evaluation, especially in contexts where students may not recognize the subjective nature of historical knowledge.

Keywords: Appraisal Theory; CLIL; Cognitive Discourse Functions; Evaluative Language; History Education.

1. Introduction

Success in school History requires learning not only the facts, but how to think about the facts, as high-stakes assessment tasks often implicitly require students to construe perspectives that align with those of relevant historians (Martin et al., 2010) and to make discipline-specific use of language to do so (Coffin, 2006; Derewianka, 2007; McCabe & Whittaker, 2017; Whittaker & McCabe, 2023). However, not all students may be aware of these implicit requirements. While some students will intuitively grasp the value systems underlying accepted construals of the past, many, especially those from more marginalized social backgrounds, will struggle to identify these values and to express them through language (Martin et al., 2010). This challenge is heightened in Content and Language Integrated Learning (CLIL) contexts, where students study History in a foreign language. Varying levels of proficiency in the language of instruction may make recognizing and articulating value judgments more difficult. Indeed, findings from a broader PhD project encompassing the present study point to a disconnect between teachers' and students' ideas about achievement in CLIL History in Spain: while teachers emphasize interpretation and disciplinary thinking, students minimize their importance and focus on factual recall (Tompkins, 2024). This scenario raises the question of how successfully they will construe such values in assessment tasks that elicit them.

To explore students' development of academic language for History, this study investigates their evaluative language use in written texts broken into Cognitive Discourse Functions (Dalton-Puffer, 2013). Cognitive Discourse Functions (CDFs) refer to communicative patterns through which "subject-specific thought processes [are made] intersubjectively accessible" (Dalton-Puffer, 2013, p. 230). The model distinguishes seven key purposes for which students use language to demonstrate learning in content subjects – *categorize, define, describe, evaluate, explain, explore, report*[1]–, building on Bloom's (1956) taxonomy, Anderson and Krathwohl's (2001) revision, and Systemic Functional Linguistics (Halliday & Matthiessen, 2004). It has been widely used in CLIL research to explore students' academic language performance in subjects learnt through an additional language (e.g., Whittaker & McCabe, 2023; Evnitskaya & Llinares, 2022), and to identify the lexicogrammatical patterns characteristic of successful student productions (e.g., Nashaat-Sobhy & Llinares, 2023; Whittaker & McCabe, 2023).

To pinpoint the linguistic resources through which students construct their disciplinary evaluations, I turn to Appraisal Theory (Martin & White, 2005), a framework within Systemic Functional Linguistics (SFL) that has often been used for this purpose both generally and in CLIL contexts (e.g., Llinares & Nikula, 2024; Morton & Llinares, 2018). The Appraisal framework identifies three linguistic systems responsible for portraying writers' attitudes and assessments and attempting to align addressees with their perspectives. These systems are *attitude*, which comprises resources for conveying emotions (*affect*), for assessing human behavior (*judgment*), and for evaluating products and processes (*appreciation*); *graduation*, which fine-tunes the representation of these attitudes and assessments; and *engagement*, which indicates that the writer acknowledges other voices.

The present study analyzes 119 year 10 Spanish CLIL students' responses to a prompt on the French Revolution that was designed to elicit four CDFs: *explain, categorize, report* and *evaluate.* I compare students' performance on *evaluate*, which requires subjective assessments, with their performance on the more factual CDFs. I also examine students' perspectives on the revolution, how they are articulated, and how they align with those of the discipline. To do so, I analyze whole texts using the Appraisal framework, zoom in on the CDF *evaluate,* and highlight differences between students at higher and lower socioeconomic status (SES) schools where relevant. The remainder of the article is organized as follows: Section 2 reviews the relevant literature, Section 3 outlines the methodology, Section 4 presents key findings, and Section 5 offers concluding remarks.

2. Literature Review

From the perspective of SFL, factual texts, including those used in school contexts, are not objective, but portray an individual or collective interpretation of reality that may be made more or less explicit (Martin, 1989; Martin & White, 2005). This is especially the case in school History, where students must pick up on the 'right values' to align themselves with the dominant narrative of historical events that is presented in their school books. Brodovskaya et al. (2023) explore how such narratives reflect and shape historical memory, while Martin et al. (2010) and Oteíza (e.g., 2020) investigate how students are apprenticed into dominant value systems through classroom discourse, addressing the extent to which alternative perspectives are recognized.

As for students' subsequent representation of these values in assessment tasks, researchers have often turned to Appraisal Theory (Martin & White, 2005) to analyze how they are constructed. Coffin's (2006) seminal work on school History in Anglophone contexts demonstrated that the genres students must write vary across years of schooling and require different sets of Appraisal resources to achieve their rhetorical goals. In late primary, empathetic genres encourage direct *affect*; in early secondary, seemingly neutral genres suppress evaluation; in mid-secondary, interpretive genres elicit assessments of significance and social esteem; and by late secondary, arguing genres prompt assessments of morality alongside social esteem and significance. In other contexts, the nature of school History writing may vary, but researchers confirm a move away from more direct expressions of emotions (*affect*) in the earlier years of secondary education towards more frequent assessments of events (*appreciation*) and behavior (*judgment*) in the later years, as students gradually develop a voice closer to that of a historian (e.g., Derewianka, 2007; McCabe & Whittaker, 2017). In Spanish CLIL programs, this progression is characteristic of texts considered successful by History teachers (McCabe & Whittaker, 2017).

However, the expectation that student texts present perspectives in line with those of relevant historians is often tacit (Martin et al., 2010), and students are not typically taught how to use language to meet their rhetorical goals in content subjects (e.g., Andujar & Spratt, 2023; Martin, 2013; Morton, 2010; Whittaker & McCabe, 2023). As Martin (2013, pp. 33-34) points out, this scenario may complicate achievement for marginalized groups of students, as "only students from the right background are positioned to learn by osmosis what has to be learned but is never made explicit." The role of values in History seems especially opaque to students in the present study, as many consider History a matter of memorizing facts rather than of making interpretations or taking stances (Tompkins, 2024). This perception may be related to a lack of interpretive and argumentative essay writing in primary and secondary CLIL History in Spain, where the genres elicited by textbooks are not typically scaffolded by teachers, who instead assign classroom and homework tasks involving very short answers (Morton, 2010; see Dalton-Puffer, 2007 for similar findings in Austrian CLIL). The present study contributes to extant literature by exploring how successfully secondary students align their perspectives with the values of the discipline in a context where many students downplay the importance of making interpretations and taking stances.

3. Methodology

The data come from a larger project (PID2022-140718OB-I00) collecting cross-sectional and longitudinal data, primarily questionnaire responses and short texts, from 20 CLIL secondary schools in Spain over four years (2022-2026). This study analyzes a corpus of texts written by year 10 students in their history class, during the fall semester of 2022. The texts (n = 339) were produced at eight schools in response to a prompt on the French Revolution, which was prepared by the researchers and elicited four CDFs. Students were asked to write about 150 words on the revolution, including its causes *(explain)*, a comparison of the lifestyles of the monarchy and the peasants *(categorize)*, its main events *(report),* and an evaluation of its importance *(evaluate).* The full prompt is provided in Appendix 1. Of the 339 texts collected, those from three schools with higher average student SES and three with lower average student SES were selected for random sampling. Student SES had been previously determined by researchers in the project by applying Principal Component Analysis to participants' questionnaire responses that indicated parental occupation, parental education, number of books at home, and possession of material resources indicative of spending power and/or conducive to studying (Tompkins & O'Donnell, 2025). Average student SES was then calculated for each school, and the schools were divided into terciles, Tercile 1 being higher SES schools, Tercile 2 middle SES schools, and Tercile 1 lower SES schools, relative to other schools in the project. The sampling process produced a corpus of 119 texts (19-20 per school).

To identify the values into which students were being apprenticed, a complementary dataset of learning materials on the French Revolution was collected from three textbooks used at the schools in the study. These materials were coded for *attitude* using Martin and White's (2005) framework, with each instance recorded in a spreadsheet indicating its target, *attitude* type and subtype, polarity (positive, negative, or ambiguous), and whether it was *inscribed* into the text or *invoked* through ideational choices. Subtypes were coded following the distinctions established in the framework: un/happiness, dis/satisfaction, in/security, and dis/inclination under *affect*; propriety, veracity, normality, capacity, and tenacity under *judgment*; and reaction, composition, and social valuation under *appreciation*. This spreadsheet would inform subsequent analysis of the student texts.

Students' texts were analyzed in two phases using the UAM CorpusTool (O'Donnell, 2008). In the first phase, students' responses to each part of the prompt were coded for the CDF they responded to, as well as their appropriateness to that CDF in terms of language and content. Responses were marked as *appropriate* if they contained both the expected content and language, *partly appropriate* if they expressed the content without the necessary linguistic resources, *inappropriate* if they did not present the content as elicited, and *missing* if they hadn't responded to that part of the prompt. The criteria used to code each CDF are provided in Appendix 2.

In the second phase, an Appraisal analysis was carried out throughout the texts. All instances of *attitude* and *engagement* were coded, as these are considered "key to creating the genres and voices of the historian" (McCabe & Whittaker, 2017, p. 108). The third system, *graduation,* was left out of the analysis because it has been viewed as less relevant for investigating students' evaluative language use in previous CLIL research (e.g., McCabe & Whittaker, 2017; Whittaker & McCabe, 2023). Figure 1 below shows the coding scheme: instances of *attitude* were coded for the same variables as those in the textbooks, as well as whether they were found in these textbooks. Students' evaluations were considered 'from the textbook' when they assessed a target in the same way as the textbook did, with regard to *attitude* type, subtype, polarity, and explicitness, even when their wording varied. Evaluations 'not from the textbook' did not meet these criteria, meaning that they either represented similar values in a different way (e.g., by using *judgment* instead of *appreciation*), or represented different values coming from another source (e.g., personal assessments, their teachers' explanations, other teaching materials…). Within *engagement,* instances of heteroglossia, or phrases that recognize alternatives to the current proposition, were coded as either contracting the dialogic space or expanding it, and for their subtype therein.

Quantitative results on the length and appropriacy of students' responses to each CDF, as well as the proportion of *attitude* and *engagement* instances across each variable, were retrieved from UAM CorpusTool. The software also provided comparisons across higher- and lower-SES schools, t-tests and chi-squared tests to identify statistically detectable[2] differences, and links to the text and co-text of instances within each category. These results are presented in the following section.

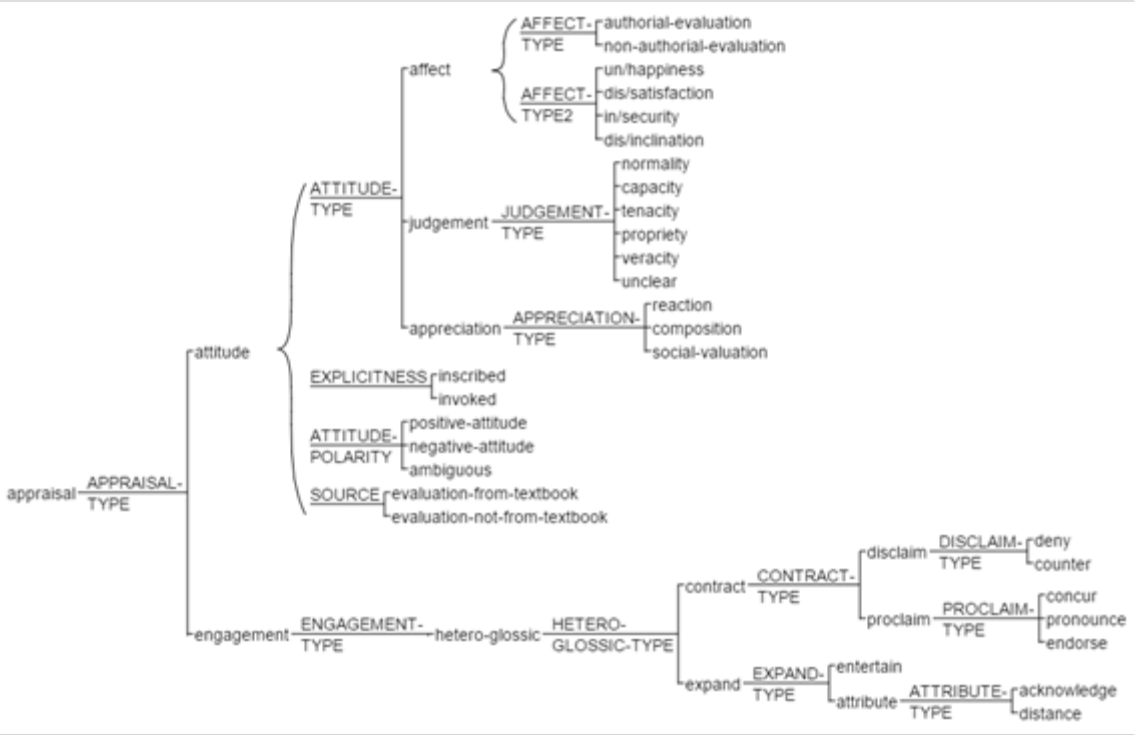


Figure 1. Scheme used to code Appraisal in UAM CorpusTool (O'Donnell, 2008)

4. Results and Discussion

4.1. Students' Performance Across the Four CDFs

Students wrote a total of 25,396 words in response to the French Revolution prompt, averaging 213 words per text. Students at higher SES schools wrote more on average (240 words) than those at lower SES schools (187 words), and this difference was statistically detectable (t-stat = 37432.39, p = 0.000). Both averages are higher than the 150 words required by the prompt. The total words written were divided unevenly across the four CDFs, as shown in Table 1 below. Students wrote the most about *report*, dedicating over half of all their words to reporting the events of the French Revolution, followed by *explain*, *categorize* and *evaluate*. Furthermore, 639 words within *categorize* also contributed to *explain*, as several students merged these two CDFs by listing the differences in the peasants' and monarchy's lifestyles as one of the causes of the revolution. Words in the texts that did not respond to any CDF, such as introductory information, are not counted here, leading to a slightly lower total word count within CDFs (24,278) than the overall total (25,396).

Table 1. *Word Count in Each CDF by School SES*

	Explain	Categorize	Report	Evaluate
SES1*	2885	1238	7968	1419
SES3	2621	1393	5570	1184
Total	5506	2631	13538	2603

*SES1 refers to SES Tercile 1, or higher SES schools. SES3 refers to SES Tercile 3, or lower SES schools.

Another way of measuring students' success in each CDF is by referring to response rates and the extent to which responses were appropriate in their use of language and content. Table 2 below shows that the CDFs with the highest response rates were *report* (97%) and *explain* (92%), followed by *evaluate* (63%) and *categorize* (47%). Table 3 shows that the responses themselves were also somewhat more appropriate in *report* (86%) and *explain* (86%) than *categorize* (82%) or *evaluate* (66%). Results for students' fluency, response rate, and appropriateness are thus broadly aligned across the four CDFs: students seem to perform better when reporting and connecting events or circumstances than when

comparing social groups or assessing significance. This was true of students at both higher and lower SES schools in the study, as there were no statistically detectable differences in the variables explored so far.

Table 2. *Presence of Responses to Each CDF*

Presence	Explain		Categorize		Report		Evaluate	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Present	109	92	56	47	115	97	75	63
Missing	10	8	63	53	4	3	44	37
Total	119	100	119	100	119	100	119	100

One possible explanation for these findings is that students may get more practice writing about events and their causes in their history classes, and thus feel more prepared to respond to these types of questions. As for *categorize* and *evaluate,* it seems they posed different challenges for students. As shown in Table 3, responses not marked as totally appropriate tended to be inappropriate within *categorize* but partly appropriate within *evaluate*. That is, students struggled more with the elicited content in *categorize*, as many failed to include and describe both groups, and more with the elicited language in *evaluate*, as many failed to use explicitly evaluative terms (see Appendix 2 for full criteria).

Table 3. *Appropriateness of Each CDF*

Appropriateness	Explain		Categorize		Report		Evaluate	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Appropriate	94	86	46	82	99	86	50	66
Partly appropriate	13	12	2	4	1	1	23	31
Inappropriate	2	2	8	14	15	13	2	3
Total	109	100	56	100	115	100	75	100

Within *evaluate,* appropriate responses to the prompt's request to "include an evaluation of the importance of the revolution" involved inscribed evaluations such as "important" or "significant," or even specific, positively charged ideational terms such as "legacy." However, many students (31%) chose to remain in more factual territory by referring to consequences of the revolution that are widely viewed as positive. For instance, one student writes that:

"In conclusion, the Third Estate did what they want, that was a Revolution to had a Republic. The consequences of the Revolution ended the absolutism, the stratief society and the feudalism in France." [School03-0019][3]

In this example, the student does not explicitly evaluate the revolution but rather implies that it was important by stating that it produced major changes in France's economy and society. Present-day readers who support democratic institutions may infer that the writer's reference to the end of absolutism and social stratification is a positive evaluation of the revolution, especially considering the text's earlier descriptions of pre-revolutionary France, where people questioned "unfair" and "inefficient" royal authority, and suffered high taxes and rising prices in a class-based society. However, relying on the reader's ability to infer one's perspective fails to address the prompt directly, which could have negative academic consequences. The fact that many students did not manage to make their perspectives explicit suggests that they may need additional scaffolding and support in this area, a point to which I return in 4.3.

4.2. Uncovering Students' Perspectives on the French Revolution: An Appraisal Analysis of Whole Texts

This section presents key findings from the Appraisal analysis of whole texts. Full results tables are provided in Appendix 3.

4.2.1. Use of Attitude

To uncover students' perspectives on the French Revolution and assess their alignment with those of the discipline, I examined their use of attitudinal resources. *Affect*, or expressions of feelings, accounted for 12% of all attitudinal meanings in the texts. Students used these resources to ascribe emotions to historical actors, especially as regarded their desires and reactions to the economic, social, and political situation in France before and during the revolution. The vast majority (87%) of *affect* was inscribed. Within inscribed *affect*, students mainly (50%) wrote about what different groups wanted (dis/inclination). For example, "the peasants wanted same rights for them as the nobles" and "the nobility didn't want to loss their privileges" [School02-0013]. They also included the dissatisfaction (27%) and unhappiness (16%) felt by the Third Estate, mentioning that they were "angry," "frustrated," and "tired of" the situation

in France, or that "they didn't like a King who didn't think about the peasants" [School02-0013]. Invoked *affect* was less common and typically expressed the Third Estate's insecurity (69%) or dissatisfaction (19%) via ideational verbs such as "suffer" or "complain."

Judgment accounted for 23% of all attitudinal meanings in the texts. Students used these resources to evaluate the behavior of human actors and groups in the revolution. About a quarter (27%) of all *judgment* was inscribed. Within inscribed *judgment*, students most often assessed normality (47%), propriety (24%), and veracity (14%). Assessments of normality typically referred to different revolutionary groups, i.e., "moderate" Girondins and "radical" Jacobins; assessments of propriety to the privileged estates, e.g., their "abuses" or disrespect;" and assessments of veracity to King Louis XVI's "treason." Invoked *judgment* was more common (73% of all instances) and echoed the inscribed assessments. For instance, some students framed the court as gluttonous and lacking empathy in statements such as "the king and queen were living his best life, without thinking in the people" [School11-0002] or "Marie Antoinette […] spend loads of money while the peasants were suffering of hunger" [School03-0016]. But even more frequent than judgments of the monarchy were judgments of the Jacobins, as students foregrounded their seizure of power, authoritarian rule, and response to opponents. For example, one student wrote about "Robespierre, which was the leader of a group of people who persuit absolutist people and kill them, they kill about 40k people in only one year" [School14-0020]. References to "dictators" or "dictatorships" were almost always linked to Robespierre and the Jacobins. By contrast, Napoleon Bonaparte, another authoritarian ruler, was described in terms of his capacity, as students foregrounded his efforts to improve society, for example by stating that he "solved the economic and social crises of France" [School13-0008].

Appreciation accounted for 65% of all attitudinal meanings in the texts. Students used these resources to evaluate products, processes, and historical figures in terms of their social value and emotional impact. Nearly half (45%) of all *appreciation* was inscribed. Within inscribed *appreciation*, students most often (84%) assessed social value. The most common tokens were "important," from the prompt, and "bad," referring to the peasants' harvests and living conditions. These living conditions were also described as "poor," sometimes in opposition to the "good" or "better" life of the privileged estates. Other common tokens of inscribed social valuation were "unfair" and "inefficient," referring to the pre-revolutionary royal authority, and "problem," referring to the causes of the revolution. Most remaining instances of inscribed *appreciation* fell into the reaction category (16%) and referred either to "the Terror," a disciplinary term referring to the Jacobins' rule, or to some aspect of life before the revolution, such as "exhausting taxes," "hard living conditions," and "stupids wars." Invoked *appreciation* (55% of all instances) echoed its inscribed counterparts, as students made ideational choices reflecting their negative assessments of pre-revolutionary France (e.g., "crisis," "inequality," "turmoil") and positive assessments of the revolution (e.g., it "gave liberties to all French people" [School13-0005]).

The preceding analysis of attitudinal resources in student texts reveals a perspective largely sympathetic to the revolutionary cause. Uses of *appreciation* depict pre-revolutionary France as characterized by stark inequalities, as the majority of the population lacked political power and struggled to meet its most basic needs. Uses of *affect* encourage the reader to empathize with this groups' negative feelings and desire for change, while uses of *judgment* portray the ruling class as unethical. However, the acceptability of the revolutionary response is filtered through the politics of its agents: the Girondins' approach is portrayed within the bounds of normality, while the Jacobins' is cast as beyond those bounds and its members are judged as a consequence. Finally, the significance of the revolution as a whole is assessed in terms of its contributions to present-day democratic ideals.

Thus, the perspective visible in the students' texts largely aligns with their history textbooks. Where differences emerged, they lay not in values but in how they were expressed: many students made stronger, more direct assessments of actors and events than their textbooks. In *affect*, students described the Third Estate as "angry," "frustrated," or "resentful," while the textbook nominalized negative feelings as "discontent," a more distant and indirect expression typical of historical discourse. In *judgment,* students mentioned "abuses," "discrimination," and "oppression" by the privileged estates, whereas the textbooks described an unequal society in which they held power. Students also directly and exclusively attributed mass killings to the Jacobins, while textbooks used passive voice and noted that political opponents were also killed under the Girondins. In *appreciation,* some students used "good" and "bad" to describe lifestyles and portrayed the revolution as remedying all pre-revolutionary ills, while textbooks took a more nuanced view. These more direct, everyday uses of *attitude* may reflect how teachers unpack content or students' own views. As shown

in Table 4, they were more common in lower SES schools, which suggests that all students, but particularly those with fewer economic resources, would benefit from more scaffolding in expressing values in discipline-appropriate ways.

Table 4. *Source of Attitude Across School SES*

Source	SES1		SES3		Comparison		
	N	*%*	*N*	*%*	*ChiSqu*	*P*	*Effect size*
Evaluation from textbook	351	67	210	50	29.57	0.000	0.357
Evaluation not from textbook	171	33	212	50	29.57	0.000	0.347
Total	522	100	422	100			

4.2.2. Use of Engagement

To assess the extent to which students' texts recognized different voices and perspectives, I examined their use of engagement resources. Dialogic expansion made up 53% (n = 191) of all instances of heteroglossia, indicating a balance with contraction. The most common type was *attribute* (75%), whereby many students attributed speech and thought to historical actors. For example, they wrote that "[the Third Estate] believed this was an unfair situation" [School03-0018]. These uses of *attribute* potentially helped align readers with certain actors by revealing their perspectives, complementing the use of *affect* discussed earlier. However, while they expanded the dialogic space of the past, they did not do so for the present. Attributions to disciplinary experts, such as "its is considered one of the most important European Revolutions" [School13-0031], were comparatively rare. As for *entertain* (25%), students often used it to signpost their subjective viewpoints in expressions such as "in my opinion" and to acknowledge the subjectivity in historians' identification of the causes and phases of the revolution, as in "It causes can be tracked at least 10 years before [...]" [School13-0041]. These uses suggest that some students are beginning to recognize that their reporting on the French Revolution contains elements of subjectivity.

Within dialogic contraction (47% of all instances, n = 166), the most common type was *disclaim* (95%), where students often used *counter* to present certain events as unexpected (e.g., "they tried to improve agriculture, but it failed" [School02-0011]), thus reflecting the ups and downs of the revolution, and to highlight the contrast between the lifestyles of the privileged and the non-privileged (e.g., "The life of the monarchy was full of luxuries, they lived very well and they ate the best food. However, the peasants lived in the street, they don't had eny food for eat […]" [School10-0075]). These uses of *counter* reinforced students' negative assessments of pre-revolutionary France, built up through attitudinal resources discussed earlier: in a situation where the few were living in luxury, they presented the suffering of the many as counter expected. Likewise, some students used *counter* to highlight the persistence of gender inequality as unexpected, as in: "But just half the population receive rights; as women's didn't start receiving them till 150 years later" [School10-0022].

Differences in engagement resources across school SES are presented in Table 5 below. The slightly higher proportion of *attribute* used at higher SES schools owes to the fact that these students were more inclined to use it as a means of opening up the dialogic space in the present day. While both groups of students attributed propositions to historical actors fairly often, most attributions to people reflecting on the past (e.g., "is considered") came from higher SES schools. Their lower SES counterparts more frequently marked their perspectives as personal, through expressions such as "in my opinion" or "I think" (within *entertain*). These patterns suggest that students at higher SES schools may have greater access to, or familiarity with, the disciplinary discourse practices of history, in which competing interpretations are routinely acknowledged and compared. By contrast, students at lower SES schools drew more heavily on personal stance markers, such as those commonly used in EFL writing, which may reflect differences in genre awareness or instructional emphasis. Nevertheless, the fact that acknowledgement of historians' interpretations was relatively rare overall suggests that disciplinary stance-taking could be supported through more explicit instruction in all school contexts.

Table 5. *Engagement Resources Across School SES*

Engagement type	SES1		SES3		Comparison		
	N	*%*	*N*	*%*	*ChiSqu*	*P*	*Effect size*
Entertain	18	10	30	18	5.31	0.021	0.245
Attribute	84	44	59	35	3.22	0.073	0.191
Disclaim	81	43	77	46	0.32	0.572	0.060
Proclaim	6	3	2	1	1.60	0.206	0.140
Total	189	100	168	100			

4.3. Assessments of Significance: Students' Approach to *Evaluate*

Bringing together the findings presented in 4.1 and 4.2, this section zooms in on students' responses to the *evaluate* portion of the prompt, which elicited "an evaluation of the importance of the revolution." Full results tables are provided in Appendix 4. Unlike the other CDFs, *evaluate* makes explicit the expectation that students include subjective assessments. However, as in other school subjects (see, e.g., Christie, 2016 on English), these assessments are not truly personal: students are expected to draw on disciplinary values and evidence (Whittaker & McCabe, 2023; Martin et al., 2010). Results show that students in the present study seem to have struggled with either one or the other. Of the 75 students who responded to *evaluate,* most (n = 66) included disciplinary evidence, but nearly half of these used the evidence to imply a positive evaluation of the revolution rather than making the evaluation explicit, as elicited by the prompt. Another nine students, seven of whom came from lower SES schools, made explicit evaluations but did not support them with evidence (e.g., "I think this was a very important event" [School13-0024]). These patterns suggest that students could benefit from additional support in constructing evaluations that link subjective opinions with factual evidence.

As for evaluative language use, nearly all instances of *attitude* (89%) were categorized as *appreciation-social valuation*, which is expected in assessments of the revolution's significance. Inscribed *appreciation* was most commonly expressed using the term "important" from the prompt, or similar evaluative adjectives such as "good," "key," "significant," "useful," or "influential." These assessments referred either to the revolution itself (e.g., "This Revolution was <u>important</u> because …" [School03-0012]), or to the changes it effectuated (e.g., "[It] led to <u>important</u> changes in France" [School13-0028]). Invoked *appreciation* appeared though ideational choices that implied a positive evaluation without using explicit attitudinal lexis, for instance, terms like "thanks to" and "legacy." Students also used more extended descriptions of the revolution's consequences by presenting it as an agent of societal change. The polarity of these evaluations depended on positions constructed earlier in the text; for example, if the revolution was described as ending something previously framed as harmful, this implied a positive assessment of its social value. These descriptions of the revolution's consequences can be broken into three main categories. First, some assessed its impact around the time of its completion, construing the revolution as ending the Old Regime, improving France and/or life for its citizens, spreading certain ideals, or inspiring subsequent revolutions. For example, one student wrote that "it ends up with the absolutism, the feudalism economy and the stratified society" [School03-0016]. Second, others assessed its impact on present-day societies, often referring to concepts such as democracy, rights, liberties, freedom, and equality. For example, "it stablished the ideas of freedom, equality etc… of our actual society" [School03-0002]. Finally, still others made broader generalizations, such as that the revolution "showed us the power of the people together" [School02-0024] or "bring equality between the women and the men" [School10-0047].

Students' use of *appreciation* in *evaluate* was broadly aligned with the perspectives in their textbooks, but about half (48%) were marked as not from the textbook because they used more everyday lexis like "good" or "great," or made generalizations that portrayed the revolution's effects as overwhelmingly positive. While the textbooks did discuss specific aspects of the revolution's lasting impact on contemporary society – such as granting citizenship and popular sovereignty, forming the basis for new political systems, and determining the roles of the state – they presented these outcomes in a more nuanced way that also highlighted persistent inequalities of gender and class and the post-revolutionary return to the Old Regime. By contrast, many student texts framed the revolution as a kind of 'cure-all' for the problems of pre-revolutionary France. Such generalizations appear to simplify disciplinary evaluations, perhaps because students are echoing how their teachers have unpacked the revolution's consequences in class, or because they are introducing their own views. These findings reflect varying degrees of alignment with disciplinary discourse: while

some students echoed textbook-based evaluations, others tended to overgeneralize them or drew on more everyday lexis to construct evaluative meanings.

As for students' use of *engagement* within evaluate, there were only 45 instances of heteroglossia, which came from 29 of the 119 texts, indicating that most students presented their evaluations as monoglossic bare assertions. Evaluations that did recognize alternative perspectives were roughly evenly divided between expanding the dialogic space (51%) and contracting it (49%). Expansions of the dialogic space were mostly instances of *entertain* (83%), which students used to frame their evaluations as personal (e.g., "in my opinion") and to speculate about hypothetical scenarios in which the revolution had never happened (e.g., "life as we know it would be totally different" [School10-0065]). Contractions of the dialogic space were mostly *disclaim* (96%), which students used either to delimit the positive effects of the revolution (e.g., "the Revolution didn't advanced at all in its subjects" [School10-0022]) or to amplify them by suggesting broader, international impacts (e.g., "[it] wrote a new beginning not only in France, but to all the countries" [School13-0044]). Boundaries on the revolution's positive effects appeared in only eight texts, but they suggest that some students are beginning to engage with nuance in evaluating a historical event that is widely viewed as positive and significant.

5. Conclusion

This paper has presented an in-depth analysis of students' expression of values in CLIL History texts on the French Revolution, to determine how they approach such a task in a context where many see rote memorization, not values, as crucial to success in the subject (Tompkins, 2024). In section 4.1, the CDF *evaluate* emerged as one of the most challenging for students, with only 50 of 119 responding appropriately. This finding raised the question of how well students were picking up on the discipline's values and/or means of expressing them. The Appraisal analysis in 4.2 showed that students do reproduce the perspectives presented in their textbooks, portraying the revolution as a necessary response to widespread inequalities, and one which moved Europe towards modern democratic ideals. Further, the distribution of their attitudinal resources – mostly *appreciation* (65%) followed by *judgment* (23%) and *affect* (12%) – is typical of historical discourse, which tends not to foreground personal feelings but instead to convey perspectives on the past through "institutionalized feelings," that is, by assessing phenomena in terms of their value and impact, and behavior terms of its adherence to social norms (Martin & White, 2005, p. 45; Whittaker & McCabe, 2023). By year 10, students seem to be learning to institutionalize their feelings in this way.

However, two key pedagogical implications arise. First, where students' evaluations were not drawn from the textbook, discrepancies largely reflected challenges with discipline-specific evaluation rather than differences in values. Many students, particularly those at lower SES schools, struggled to express their assessments in contextually appropriate ways. For instance, they tended to overuse *affect* and to judge historical actors by foregrounding their unethical actions rather than by using disciplinary terms (e.g., "radical," "treason") or *appreciation* (e.g., "excessive spending") to express comparable evaluations. Second, because most attitudinal assessments were either inscribed through evaluative terms of the discipline (e.g., "radical," "the Terror"), or invoked through ideational descriptions, it remains unclear the extent to which students recognize that the perspectives they are constructing are subjective. As we saw in 4.3, even where students were asked to give a subjective evaluation of the revolution, about half chose to remain in more factual territory by mentioning its consequences without clearly evaluative language. Further, the stance produced in response to *evaluate* was largely monoglossic and strongly positive, as many students presented the revolution's effects in unqualified terms, without adopting the more nuanced framing modeled in their textbooks. For example, claims that the revolution "brought equality between men and women" or "gave liberties to all French people" did not acknowledge the persistent inequalities discussed in the learning materials. These findings suggest that many students would benefit from more explicit scaffolding in constructing disciplinary evaluations, as they may not yet fully recognize the interpretive nature of historical knowledge and may struggle to assess social impact without overstating it.

One way to help all CLIL students recognize and express values in History is to guide them in analyzing the language of their learning materials. Collaborations between History and English teachers could use or adapt Rose and Martin's (2012) teaching and learning cycle to reveal evaluative language in assigned texts, co-construct evaluative texts with students, and thus better support their independent constructions. Such an approach would deepen students' understandings of non-fiction history texts, which are generally more difficult to read and comprehend than other text types (Yarmakeev et al., 2023), while also helping them to recognize the presence of subjective perspectives, to identify

how values are conveyed through language, and to articulate their own perspectives in ways that align with disciplinary conventions. Based on the findings presented here, interventions might focus on discipline-specific uses of *affect*, which students tend to overuse, and *judgment,* which they tend to express by foregrounding and exaggerating virtuous or blameworthy actions rather than through discipline-specific lexis. They might also guide students towards contextually appropriate alternatives for more everyday expressions like "good" and "bad." Within *evaluate*, students might be taught to make their evaluations explicit, to justify them with disciplinary evidence and to qualify them where appropriate. They would also stand to benefit from engaging with nuance and alternative views, perhaps by comparing competing interpretations of the same events. Such explicit teaching of the linguistic resources necessary to write successful texts has been long advocated by SF linguists (e.g., Coffin, 2006; Morton, 2010; Rose & Martin, 2012) and, in the context of CLIL in Spain, it could provide a bridge between content and language and move students beyond rote memorization and into analysis and evaluation, as is often advocated by proponents of CLIL (e.g., Asmolovskaya & Kondrateva, 2023; Maksimova et al., 2023).

Acknowledgments

I would like to thank my thesis supervisor, Ana Llinares, for her guidance during my analysis of the data.

Information on Informed Consent or any Data Privacy Statements

Following data protection laws in Spain (Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales), participants gave their informed consent to the processing of their pseudonymized data for research purposes.

Conflict of Interest

The author has no conflicts of interest to declare.

Ethics Board Approval Statements

The PhD project encompassing this study was approved by the Ethics Board of the Universidad Autónoma de Madrid, as was the state-funded research project of which it forms part.

Funding

This research has been funded by the author's institution, through a grant for doctoral studies.

References

Anderson, L.W., & Krathwohl, D.R. (Eds.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives.* Longman.

Andujar, A. & Spratt, M. (2023). Using AI to support CLIL teacher language. *Journal of Research in Applied Linguistics*, *14*(2), 7-19. https://doi.org/10.22055/RALS.2023.45267.3177

Asmolovskaya, M. V., & Kondrateva, I. G. (2023). Analyzing the importance of motivation in the context of English language learning. *Journal of Research in Applied Linguistics, 14*(3), 123–127. https://doi.org/10.22055/rals.2023.19467

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: The cognitive domain.* McKay.

Brodovskaya, L. N., Ibragimov, R. R., Buravleva, V. V., & Fazliev, A. M. (2023). The role of history textbooks in the formation of historical memory in educational literature and the modernization of the history education process in English. *Journal of Research in Applied Linguistics, 14*(3), 507–511. https://doi.org/10.22055/rals.2023.19597

Christie, F. (2016). Secondary school English studies: Cultivating a knower code. In K. Maton, S. Hood & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 158-175). Routledge.

Coffin, C. (2006). *Historical discourse: The language of time, cause, and evaluation.* Continuum.

Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms*. John Benjamins.

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

Derewianka, B. (2007). Using Appraisal Theory to track interpersonal development in adolescent academic writing. In A. McCabe, M. O'Donnell, & R. Whittaker (Eds.), *Advances in language and education* (pp. 142–165). Continuum.

Evnitskaya, N., & Llinares, A. (2022). (In)equity in CLIL programs? Classroom interaction and the development of higher order thinking skills across bilingual strands. *AILA Review, 35*(2), 227–249. https://doi.org/10.1075/aila.22026.evn

Gorard, S. (2013). *Research design: Creating robust approaches for the social sciences.* Sage Publications.

Halliday, M. A. K. & Matthiessen, C. (2004). *An introduction to functional grammar* (3rd ed.). Arnold.

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

Maksimova, J.O., Gilmutdinova, A.R., Yashina, M.E., & Shelestova, O.V. (2023). Integrated teaching of content and English language using CLIL method while learning. *Journal of Research in Applied Linguistics, 14*(3), 119–122. https://doi.org/10.22055/rals.2023.19466

Martin, J.R. (1989). *Factual writing: Exploring and challenging social reality.* Oxford University Press.

Martin, J.R. (2013). Embedded literacy: Knowledge as meaning. *Linguistics and Education, 24*(1), 23-37. https://doi.org/10.1016/j.linged.2012.11.006

Martin, J.R., Maton, K. & Matruglio, E. (2010). Historical cosmologies: Epistemology and axiology in Australian secondary school history discourse. *Revista Signos: Estudios de Lengua y Literatura, 43*(74), 433-463. https://doi.org/10.4067/S0718-09342010000500003

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

McCabe, A. & Whittaker, R. (2017). Genre and appraisal in CLIL history texts: Developing the voice of the historian. In A. Llinares & T. Morton (Eds.), *Language learning & language teaching* (Vol. 47, pp. 105-124). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.47.07mcc

Morton, T. (2010). Using a genre-based approach to integrating content and language in CLIL. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 81-104). John Benjamins. https://doi.org/10.1075/aals.7.05mor

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

Nashaat-Sobhy, N., & Llinares, A. (2023). CLIL students' academic language performance on a non-curricular topic: A comparison between high-exposure and low-exposure groups. *International Journal of Bilingual Education and Bilingualism, 26*(4), 567–583. https://doi.org/10.1007/s42321-023-00151-y

O'Donnell, M. (2008). Demonstration of the UAM CorpusTool for text and image annotation. In *Proceedings of the ACL-08: HLT Demo Session (Companion Volume)* (pp. 13–16). Association for Computational Linguistics.

Oteíza, T. (2020). Language resources to negotiate official and alternative memories of human rights violations in Chilean history: A study on classroom interactions. *Historical Encounters: A Journal of Historical Consciousness, Historical Cultures, and History Education, 7*(2), 26-49. https://www.hej-hermes.net/_files/ugd/f067ea_009616d2271f46cdb421d9aabba697d4.pdf

Rose, D. & Martin, J.R. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney school.* Equinox.

Tompkins, L. (2024, January 15-18). *Clarifying bases of achievement in content and language integrated learning* [Conference presentation]. LCT 5, Johannesburg, South Africa.

Tompkins, L. & O'Donnell, M. (2025). *Estimating socioeconomic status from student questionnaire data* [Unpublished manuscript]. Department of English Philology, Universidad Autónoma de Madrid.

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

Yarmakeev, I.E., Zamaletdinov, R.R., Pimenova, T.S., & Abdrafikova, A.R. (2023). Efficient methods of increasing the ability to read texts of different styles in English. *Journal of Research in Applied Linguistics, 14*(3), 204–208. https://doi.org/10.22055/rals.2023.19503

Appendices

Appendix 1. Prompt on the French Revolution

We have been studying the French Revolution. Write a short text about it. Make sure you include the **causes**, and also a **comparison** between the life of members of the monarchy (for example of Maríe Antoinette), and the peasants. Also include the main **events** which took place, and an **evaluation** of the importance of the revolution. Try to write about **150 words.**

Appendix 2. Criteria for Coding Students' Productions

Criteria for coding students' productions of *explain:*

Appropriateness	Indicators
Appropriate	The text explains the causes of the French Revolution using cause/effect language, e.g. "the causes," "because," "due to," "so," etc.
Partly appropriate	The text mentions the causes of the French Revolution without using cause/effect language, thus leaving the logical relationship implicit; OR the text mentions the causes using cause/effect language but without clause structure.
Inappropriate	The text portrays something that happened after the start of the revolution as its cause.
Missing	The text does not attempt to identify any causes of the French Revolution.

Criteria for coding students' productions of *categorize:*

Appropriateness	Indicators
Appropriate	The text compares the life of members of the monarchy and the peasants using comparative language (e.g., "on the one hand… on the other," "while," "differences/similarities," etc.)
Partly appropriate	The text describes the life of members of the monarchy and the life of the peasants but does not use comparative language.
Inappropriate	The text names groups but does not compare their lifestyles; OR The text describes the lifestyle of one group but not the other.
Missing	The text does not refer to the monarchy nor the peasants.

Criteria for coding students' productions of *report:*

Appropriateness	Indicators
Appropriate	The text reports the events of the French Revolution using clauses.
Partly appropriate	The text lists the events of the French Revolution without clause structure.
Inappropriate	The text names the stages of the French Revolution without reference to any events; OR the text names or reports events that were not part of the French Revolution.
Missing	The text does not refer to any events or stages.

Criteria for coding students' productions of *evaluate:*

Appropriateness	Indicators
Appropriate	The text evaluates the revolution's importance, either for the people of the time or for the future, using clearly evaluative language (e.g., "important," "relevant," "significant," "incredible,") or specific, positively charged ideational terms (e.g. "legacy").
Partly appropriate	The text refers to the revolution's consequences without clearly evaluative language.
Inappropriate	The text evaluates something besides the revolution's importance.
Missing	The text does not evaluate the revolution's importance nor refer to its consequences.

Appendix 3. Descriptive Statistics for Appraisal Analysis of Whole Texts

Type	Attitude	
	N	*%*
Affect	112	12
Judgment	215	23
Appreciation	617	65
Total	944	100

Explicitness	Affect	
	N	*%*
Inscribed	97	87
Invoked	15	13
Total	112	100

Subtype	Inscribed affect	
	N	*%*
Un/happiness	16	16
Dis/satisfaction	26	27
In/security	7	7
Dis/inclination	48	50
Total	97	100

Subtype	Invoked affect	
	N	*%*
Un/happiness	0	0
Dis/satisfaction	3	19
In/security	11	69
Dis/inclination	2	12
Total	16	100

Explicitness	Judgment	
	N	*%*
Inscribed	59	27
Invoked	157	73
Total	216	100

Subtype	Inscribed judgment	
	N	*%*
Normality	28	47
Capacity	7	12
Tenacity	0	0
Propriety	14	24
Veracity	8	14
Unclear	2	3
Total	59	100

Subtype	Invoked judgment	
	N	*%*
Normality	4	3
Capacity	47	30
Tenacity	43	27
Propriety	63	40
Veracity	0	0
Unclear	0	0
Total	157	100

Explicitness	Appreciation	
	N	*%*
Inscribed	278	45
Invoked	343	55
Total	621	100

Subtype	Inscribed appreciation	
	N	*%*
Reaction	44	16
Composition	1	0
Social valuation	233	84
Total	278	100

Subtype	Invoked appreciation	
	N	*%*
Reaction	57	17
Composition	0	0
Social valuation	286	83
Total	343	100

Type	Heteroglossia	
	N	*%*
Contract	166	47
Expand	191	53
Total	357	100

Subtype	Contract	
	N	*%*
Disclaim	158	95
Proclaim	8	5
Total	166	100

Subtype	Expand	
	N	*%*
Entertain	48	25
Attribute	143	75
Total	191	100

Appendix 4. Descriptive Statistics for Analysis of *evaluate*

Type	Evaluate-present	
	N	*%*
Disc-evidence-given	66	88
Disc-evidence-not-given	9	12
Total	75	100

Type	Disc-evidence-given	
	N	*%*
Evidence-justifies-explicit-evaluation	36	55
Evidence-is-implicit-evaluation	30	45
Total	66	100

Type	Attitude-in-evaluate	
	N	*%*
Affect	7	4
Judgment	11	7
Appreciation	150	89
Total	168	100

Source	Attitude-in-evaluate	
	N	*%*
From textbook	88	52
Not from textbook	80	48
Total	150	100

Explicitness	Appreciation-in-evaluate	
	N	*%*
Invoked	70	47
Inscribed	80	53
Total	150	100

Polarity	Appreciation-in-evaluate	
	N	*%*
Positive	143	95
Negative	7	5
Total	150	100

Type	Heteroglossia-in-evaluate	
	N	*%*
Contract	22	49
Expand	23	51
Total	45	100

Subtype	Contract-in-evaluate	
	N	*%*
Disclaim	21	96
Proclaim	1	4
Total	22	100

Subtype	Expand-in-evaluate	
	N	*%*
Entertain	19	83
Attribute	4	17
Total	23	100

© 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/).