Expressing evaluation across disciplines in primary and secondary CLIL writing: a longitudinal study

Rachel Whittaker & Anne McCabe

To cite this article: Rachel Whittaker & Anne McCabe (2020): Expressing evaluation across disciplines in primary and secondary CLIL writing: a longitudinal study, International Journal of Bilingual Education and Bilingualism, DOI: 10.1080/13670050.2020.1798869

To link to this article: https://doi.org/10.1080/13670050.2020.1798869

Published online: 10 Aug 2020.

Submit your article to this journal

Article views: 10

View related articles

View Crossmark data
Expressing evaluation across disciplines in primary and secondary CLIL writing: a longitudinal study

Rachel Whittaker\textsuperscript{a} and Anne McCabe \textsuperscript{b}\textsuperscript{b}

\textsuperscript{a}Department of English, Universidad Autónoma de Madrid, Madrid, Spain; \textsuperscript{b}Departments of English and Communication, Saint Louis University, Madrid Campus, Madrid, Spain

ABSTRACT
The construct of cognitive discourse functions (CDFs) has been proposed as a bridge between linguists and educationalists, linking ‘subject specific cognitive learning goals with the linguistic representations they receive in classroom interaction’ (Dalton-Puffer 2013. “A Construct of Cognitive Discourse Functions for Conceptualising Content-Language Integration in Q4 CLIL Multilingual Education.” EuJAL 1 (2): 216–253, 220). We focus on the CDF evaluate, using the Appraisal model to analyze evaluative language in a longitudinal corpus of student texts written in L2 English across disciplines (natural science, history, art), collected from the same students at the end of primary school (aged 11+) and at the beginning and end of secondary year 2 (aged 13–14). We trace students’ control of meaning-making resources for the CDF evaluate across disciplines and over time through their ability to ‘couple’ interpersonal, or evaluative, meanings with their ideational, or field-specific knowledge. The findings show some development towards appropriate field + evaluation couplings, and suggest ways teachers can focus students’ attention on the language of evaluation across disciplines, aiding development of cognitive discourse competence. Our study further supports the contributions of Systemic Functional Linguistics to educational contexts, as the Appraisal framework discriminates types of evaluation for creating disciplinary knowledge.

ARTICLE HISTORY
Received 5 February 2019
Accepted 13 July 2020

KEYWORDS
Content and language integrated learning (CLIL); appraisal; everyday/ disciplinary knowledge; writing

Introduction
The teaching and learning of school subjects through a foreign language has become a regular feature of primary and secondary education in the majority of countries in the European Union (European Commission 2017), following a policy aimed at creating multilingual citizens (Official Journal of the European Communities 1995). In Spain, since the late nineties, the regional educational authorities have gradually implemented content and language integrated (CLIL) programmes in a number of European languages, at primary and secondary level, and in one or more content subjects. In the Madrid region alone, where the data for the study reported on here was collected, almost half the primary and secondary state schools had joined the Bilingual Project at the time of writing, which means that 170,000 students from this region learn two subjects through a foreign language at primary level and from two to four at secondary.

This exponential growth in numbers of schools introducing the teaching of subjects through a foreign language reflects a search for a solution to the traditionally poor skills in foreign languages reported by Spanish citizens (European Commission 2006) together with recognition of the requirements of today’s labour market with its European, if not worldwide, scope. Learning content subjects...
through a foreign language increases intensity of exposure to that language considerably (Muñoz 2008) at the same time as it gives an opportunity for ‘natural’, implicit, language learning, since the language is used for real communication, rather than to practice forms or display lexical knowledge (e.g. Ellis 2005). At the same time, the language of school subjects is not everyday language about the here and now, but specialized, abstract language building new knowledge for students in a variety of complex text types. For teachers who offer to join the bilingual projects, or are required to do so, and so have to rethink and re-present their subjects in a foreign language, the challenge is immense. Some go into it enthusiastically, some, understandably, with resistance.

If these attractive, wide-spread and costly projects are to succeed – costly not only financially but in, teacher enthusiasm, effort and time –, the two agents involved, teachers and students, need support. While the general foreign language competence of teachers is tested before allowing them to participate, and opportunities to improve skills are offered by the authorities with in-service language training, teachers recognize that something more is required. Teachers call for help with pedagogy to face the challenge, since language difficulties tend to be seen in terms of terminological equivalences, which, for them as content experts, do not present a problem. The view from the perspective of linguistics, of course, is different. Language as a resource for making meaning is realized through several levels – lexical, syntactic, and discoursal, for example. At the level of lexis much more is involved than an isolated orthographic or phonological form. Studies into the language of school disciplines have shown that many of the features of academic language and of the texts in which knowledge is transmitted and transformed in classrooms are taken for granted by teachers, although they represent almost insuperable obstacles for many students, as international assessments of literacy (PIRLS, PISA, PISA-D) or school-based studies such as that of Chall, Jacobs, and Baldwin (1990) show.

Much research into the language of school texts has been inspired in Halliday’s educational linguistics and its applications (Halliday 1993, 1987). For Halliday, language forms are inseparable from the meanings they make, so language development is inextricably linked to learning, both about the world which children experience directly, and to what they learn about the world as recon textualized through language in the classroom. In short ‘the ontogenesis of language is at the same time the ontogenesis of learning’ (Halliday 1993, 93). Thus, for Halliday, ‘educational failure is primarily linguistic failure’ (Doughty and Thornton 1973, 3, emphasis original), which applies both to L1 education and to learning in a foreign language. This view of teaching/learning is a far cry from supplying a glossary of equivalences for specialized terms developing content topics; necessary as those are, learning terms needs to go beyond equivalences into learning concepts themselves. Content teachers’ rejection of responsibility to focus on language is related to perceiving it as working on correct isolated forms, rather than conceiving language as the resource for learning their subjects, as the tool to access and make discipline-specific meanings.

An attempt to overcome this blind-spot was triggered by the need to establish a space in which dialogue between content-teacher educators and linguists could take place fruitfully (Dalton-Puffer 2013). An instrument permitting mediation between the terminologies found in curriculum documents, well-known to those in education faculties, and the language used in textbooks and classroom discourse could be a basic tool to communicate with the world of content teachers. A bridge to link these territories, with its base in the education side, has been proposed by Dalton-Puffer (2013) in the form of a manageable number of ‘Cognitive Discourse Functions’ or CDFs, involved in learning, like explain, report, explore. Their appearance in different school subjects, and their linguistic realizations in different modes are being studied by functional linguists (Dalton-Puffer et al. 2018; Naashat-Sobhy & Linares, Dalton-Puffer & Evnitskaya, this issue) with the aim of contributing to the linguistic education of bilingual teachers.

This article focuses on the CDF evaluate as it appears in a longitudinal corpus of texts written in English in different subjects, spanning the transition from primary (grade 6) to secondary (end of grade 8) education in state bilingual schools in the Madrid area. The analysis uncovers the resources pupils use to express different types of evaluation as they move into the disciplinary areas of secondary school and weighs their appropriacy for different fields studied. Results can be applied to teacher
education, used in discussion sessions and workshops with content teachers, and incorporated into teaching materials.

Theoretical background

Cognitive discourse functions

Dalton-Puffer’s goal in working with the notion of cognitive discourse functions was to make ‘subject-specific thought processes intersubjectively accessible and thus available for learning.’ (2013, 230). These discourse functions ‘mark cognitive operations and their verbal performance at the same time; they are at the interface between cognition and verbalization’ (Vollmer 2006, 21). Her model distils seven cognitive discourse functions from a number of sources which provide taxonomies of school knowledge processes, such as Bloom’s (1956) taxonomy of cognitive functions (see Anderson and Krathwohl 2001; Krathwohl 2002), work from the Council of Europe on Language across the Curriculum (Beacco 2010, 2017; Vollmer 2006), and models such as Mohan’s (1986) Knowledge Framework. Dalton-Puffer’s proposal is presented in Table 1:

The labels are not meant to be taken as the CDF construct itself; rather, given that a term in some language could not be avoided, an English version of the most proto-typical member of each CDF type was chosen, and exemplified by a range of approximately synonymous terms found in the documents mined in her study. In the same way, in any language, each of the categories/types would be identified by a more or less wide range of terms. For example, for the CDF evaluate, central to our study, further terms include: judge, argue, justify, take a stance, critique, recommend, comment, reflect, and appreciate. Evaluate is related to higher order thinking skills, according to Bloom’s taxonomy, and involves learners in ‘[m]aking judgments based on criteria and standards’ (Krathwohl 2002, 215), which are discipline-specific and, therefore, must be learnt. For example, when historians evaluate, they use explanations which tend to be impersonal and abstract, while ‘students often remain focused on human “wants and desires”’ (Coffin 2006, 117). On the other hand, science writing tends to remain dispassionate (Christie and Deriwianka 2008). Thus, over the years of schooling, learners need to develop their abilities to provide discipline-specific forms of evaluation and also present support justifying their stance based on knowledge of the field. The next subsections describe the linguistic model used for the analysis of the types of evaluation and disciplinary content in this study.

Evaluation in school writing in content areas: Appraisal

Studies of language in school disciplines can draw on a wealth of research into the language produced and consumed across subjects through the years of schooling carried out by the Sydney school using the systemic-functional linguistic (SFL) model of language and school genres (see Rose and Martin 2012 for an overview). The model provides a framework to analyse ‘how a culture maps ideational, interpersonal and textual meaning onto one another in phases as a text unfolds’ (Martin 2007, 55). Analysis of the development of students’ expression of the CDF evaluate in content subjects requires a focus on both ideational and interpersonal meanings, since the ideational

Table 1. CDFs: Types, Communicative Intention, Label.

<table>
<thead>
<tr>
<th>Type</th>
<th>Communicative Intention</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I tell you how we can cut up the world according to certain ideas</td>
<td>CLASSIFY</td>
</tr>
<tr>
<td>2</td>
<td>I tell you about the extension of this object of specialist knowledge</td>
<td>DEFINE</td>
</tr>
<tr>
<td>3</td>
<td>I tell you details of what can be seen (also metaphorically)</td>
<td>DESCRIBE</td>
</tr>
<tr>
<td>4</td>
<td>I tell you what my position is vis a vis X</td>
<td>EVALUATE</td>
</tr>
<tr>
<td>5</td>
<td>I give you reasons for and tell you cause/s of X</td>
<td>EXPLAIN</td>
</tr>
<tr>
<td>6</td>
<td>I tell you something that is potential</td>
<td>EXPLORE</td>
</tr>
<tr>
<td>7</td>
<td>I tell you about something external to our immediate context on which I have a legitimate knowledge claim</td>
<td>REPORT</td>
</tr>
</tbody>
</table>

Dalton-Puffer 2013, 234.
develops the content, or field, while the interpersonal realizes the different types of evaluation. We first present the model for the linguistic analysis of evaluation, with examples from our data, and later relate this function to the expression of content.

The Appraisal framework (Martin 2000; Martin and White 2005) is a development of the interpersonal metafunction in SFL. The linguistic systems involved in evaluation have been teased out and systematised at the level of discourse semantics to create the Appraisal system, which models the way ‘evaluation is established, amplified, targeted and sourced’ (Martin and White 2005, 9). Three quite different types of meanings are involved: those related to the evaluation itself and its types, the sub-system of attitude; those related to its source and reception, through, for example, acknowledging other viewpoints (‘according to X’) or strengthening arguments (‘it is clear that’), which is the sub-system of engagement; and those which intensify or tone down an evaluation through intensifiers (‘very’) and mitigators (‘slightly’), the sub-system of graduation. This framework has been used to trace adolescent school writing by Derewianka (2007), who chronicles changes in the linguistic resources over the four years of secondary school, in the field of history. She notes a move away from direct expression of emotion and of explicit judgements of moral values towards a greater tendency to express evaluations in terms of events, a move which results in a more dispassionate, more factual text. McCabe and Whittaker (2017) also report on changes over the years of secondary schooling in a CLIL context, showing the learners’ ability to adapt to appropriate evaluative meanings for the history topic under study.

Especially relevant to this study is the system of attitude, which consists of three further subsystems: affect, the linguistic resources used to express emotional reactions and states, judgement, those that express evaluations of people’s abilities, behaviours and actions, and appreciation, those that assess things, events and phenomena, including aesthetic perspectives (Martin and White 2005). Examples of each of these three subsystems are provided from the data:

1) affect:
   a. I don’t like the people how through food or toys to the ocean. <ECO-1Any-GC-77>
   b. The drawing I liked the most was his selfportrait on a crystal ball. <ART-2nd-B-86>

2) judgement:
   a. Julius Caesar was more important than any other gladiator. <HIST-2nd-B-35>
   b. The author it is very genius and has a very extensive imagination. <ART-2nd-B-88>

3) appreciation:
   a. Their clothes were poor clothes, and horrible clothes. <HIST-1Any-GB-46>
   b. In the A there is a very good landscape, but the B has a very good linear perspective <ART-2nd-C-53>
   c. a very important activity for Rome also was mining and production of food or materials for daily life. <HIST-2nd-B-86>

This system, distinguishing different classes of attitude, reveals the type of appeal the student makes to evaluate in different school subjects. As to the other main systems, reference will be made to engagement below, while the system of graduation, not relevant to this paper, will not be described.

Ideational targets of evaluation in school writing

For an expression of evaluation to be successful in the context of a particular subject, the target (i.e. the content) evaluated should be selected among those recognized by experts, or the ideal knowers (Maton 2007), in the field. Students have to learn, then, both which values are appreciated by a discipline and what type of content the discipline studies. This combination can also be analyzed using the systemic-functional model. In the same way as the model identifies those parts of the linguistic system involved in expressing different types of evaluation (Appraisal, from the interpersonal
metafunction), it recognizes others which organize our experience of the world. These ideational meanings (from the experiential metafunction), may construe experience either from everyday understandings or from a scientific perspective. In education, technical understandings together with disciplinary evaluations have been analyzed using Martin’s concept of coupling, defined as ‘the way in which meanings combine – across strata, metafunctions, ranks, and simultaneous systems… and across modalities’ (Martin 2010, 19); this concept has been applied by Humphrey and Hao (2019), for example, to demonstrate field-specific couplings of ideational meanings with Appraisal in the field of school science.

The type of knowledge (ideational meanings) appealed to in a text can be seen in the register, which ‘varies along a dimension of technicality’ (Eggins 1994, 71). It may use more common-sense knowledge, learnt through personal experiences, as in example 1a above, where the student writer draws ideationally on the notion of people throwing trash into the ocean, part of his lived experience; similarly, in 3a, the primary student draws on everyday knowledge of clothes during an imaginary journey to the past. On the other hand, the type of knowledge may be more specialized, dependent on school learning (Bernstein 1999; Martin 2007), as in example 3b, where painting (A) is classified as a ‘landscape’, demonstrating knowledge of taxonomies in art, and the technicality with which it has been created is encoded through the nominal group ‘linear perspective’. This nominal group, organized by ‘Classifier^Thing’ (or a pre-modifier classifying a head noun) (Martin 1992, 312) is typical of scientific writing, in fields such as biology, or, in this case, the technical construction of art, as classified nouns build the taxonomies of specialized fields. History also relies on taxonomies, by dividing time into periods or societies into groups, for example. However, it has little technicality in terms of a build-up of hierarchical knowledge structures (Maton 2007; Wignall 2007), relying instead on abstractions such as eras, movements and ‘-isms’ (such as Liberalism, Marxism, etc.).

For fields such as ecology and art appreciation, then, learning relies on the build-up of taxonomic knowledge of concrete entities, and, in terms of the interpersonal meanings, values often emerge from the technical explanations themselves:

4) In painting B the colours of the leaves of trees are warm colours, from brown to yellow. So they combine harmoniously. <ART-2nd-D-21>

Example 4 presents a classification of colours (‘warm’), leading to a disciplinary evaluation, the creation of harmony. In 3b the writer appreciates the ‘good linear perspective’, which suggests a specialized understanding of the technicality of painting. In fields such as history, however, ‘[l]earning the right values is trickier’ (Martin, Maton, and Matruglio 2010, 441). Evaluative meanings may be attached to people, time, events and activities in terms of significance, as in examples 2a and 3c. These meanings do not emerge from knowing the historical characters, time periods or abstractions in and of themselves, but rather from the cultural and ideological values of historians (Coffin 2006), of the knowers in the field. At the same time, 2a and 3c, with their focus on the significance of historical characters and activities, contrast with 3c, in which vague, subjective epithets ‘poor’ and ‘horrible’, of everyday experience appear. Thus, students need to learn how ideational and interpersonal meanings are expressed and coupled appropriately in the different school subjects.

In addition to discipline-specific couplings of ideational entities and interpersonal evaluations, another type of coupling involves justification of evaluations, to which we now turn.

**Justifying evaluations**

In the previous section, we have seen that couplings are created through evaluation of field-specific entities (ideational) using resources from the system of Appraisal (interpersonal). Couplings may also extend across two systems within Appraisal, those of attitude and engagement. The system of engagement classifies the dialogic resources used by writers to open up or close down the dialogic space they offer their readers, by, for example, justifying a position, to, in a way, prevent readers
from continuing the dialogue with a query or challenge. Justification is included in the system of ENGAGEMENT by White (2003) as a means of contracting the possibilities for (dis)agreement with a proposition. For example, logical ‘because’-type clauses can provide a justification for an evaluation, as illustrated in examples 5–7 from the data (the attitudinal meaning is underlined and the justification italicized):

5) **AFFECT + justification**
   a. I like that animal because *is a strange animal and I like it*. <ECO-1ary-GA-24>
   b. I don’t like very much because the painter is using dark colours. <ART-2nd-B-65>

6) **JUDGEMENT + justification**
   a. Cristobal Columbus was a very important person because he discovered America. <HIST-1ary-GB-46>
   b. in the Ancient Rome I think they had more imagination because they don’t have any tecnology. <HIST-2nd-B-84>

7) **APPRECIATION + justification**
   a. Its better to be a patrician because they are powerfull. <HIST-2nd-E-43>
   b. The colours are better because *is a family colours and the colours look similar*. <ART-2nd-B-19>

The theoretical framework, then, permits a view of development of a key function in learning, that of evaluating. It combines analyses which distinguish different classes of evaluation (ATTITUDE) and their targets, which may be more or less discipline-specific, and so more or less appropriate for the context, while noting the presence or absence of suitable justification for that evaluation. Such a combination reveals different aspects of developing subject knowledge.

**Objectives and research questions**

The functional linguistic perspective just described was applied to the analysis of a corpus of texts to discover how learning is verbalized, and how this verbalization of subject knowledge develops from late primary to mid-secondary school in a group of pupils in bilingual classes in different subjects. Focusing on the CDF evaluate, we analyzed the types of interpersonal meanings in the texts and their coupling with ideational meanings in order to discover how students are developing the ability to achieve ‘the infusion of field with value’ (Martin 2007, 55), as they appeal to subject-specific criteria. Given the exploratory nature of the study, our general research question was:

What resources for evaluating subject knowledge do students studying in English L2 show at different moments in their schooling?

This was broken down into four specific research questions:

1) Do choices from the Appraisal system differ depending on the subject students are writing about?
2) Are justifications included to support evaluations? Do these show changes over time?
3) What type of knowledge – everyday or specialized – do students draw on in their evaluations and justifications?
4) Are the students beginning to develop the resources which produce appropriate field + evaluation couplings?
Participants, data collection, method

The data was collected in a project studying the transition from primary to secondary school for pupils learning several subjects through English (see uam-clil.org for information on the project). Base-line data (written and spoken, in Spanish and English, on history and ecology) was collected from all the sixth-grade students (three groups) in a state primary school a few days before the end of the school year (Time 1). The project followed the same pupils into secondary school, with the second data collection taking place just over a year later, when the pupils were starting grade eight (Time 2) in the subject of history, again in Spanish and English, and the third at the end of that year (Time 3) in art, only in English. No intervention took place.

For each collection, to avoid the problems of longitudinal research which often finds few examples of features under study when unprompted (Ortega and Byrnes 2008), tasks were prepared eliciting the seven CDFs in relation to some aspect of a topic from the syllabus studied before the students did each of the tasks. The prompts were based on content in the textbook and designed in collaboration with the teachers. In order to motivate the students, and encourage a meaningful text in response, the task was framed as a blog written by the students for their followers. The CDF evaluate in English was triggered by the following instructions:

- End Grade 6 primary science (ecology): Do you think it is important to protect the environment? Why? Why not? Explain why it was important.
- End Grade 6 primary history: Name one very important person, event or invention of the period you have chosen. Explain why it was important.
- Start Grade 8 secondary history: Was it better to be a patrician or a plebian? Why?
- End Grade 8 secondary art: Choose one of the paintings and comment on how effective the use of colour is and why.

Students were given about twenty minutes for the complete task (i.e. covering the seven CDFs; for an example of a full prompt, see the Appendix).

Texts were transcribed and uploaded into the UAM CorpusTool (O’Donnell 2008). The tool includes a layer with Martin and White’s (2005) scheme for the analysis of Appraisal, relevant to the analysis of the CDF evaluate, as explained. We coded for the least delicate level in the system of ATTITUDE and added the features ‘+ justification’ for each (Figure 1), which allowed us to register inclusion or absence of this move, as well as recover it and analyse it. All four prompts required explanations justifying an evaluation.

While the CDF evaluate was expressly triggered in one section of each prompt, ATTITUDE often appeared elsewhere in the students’ responses. All examples of evaluation, whether prompted or not, were coded. Appraisal analysis requires examination of context, so doubts were discussed by the two researchers and resolved together, sometimes in collaboration with the rest of the research

---

**Figure 1.** Appraisal Framework, System of ATTITUDE.
team in general meetings. The qualitative analysis of the degree of disciplinarity in student responses was carried out using concordances generated by CorpusTool from the coding.

For this study, a sub-corpus of texts was selected from the written bilingual English/Spanish UAM-CLIL corpus. It included only texts in English by those students who had been present at all the data collections (Table 2).

Differences in number of texts are normalized by CorpusTool for the treatment of data, and results are presented as counts of features selected per 1000 tokens. CorpusTool also calculates the t value and significance of differences between selections of features.

**Results and discussion**

Results are presented following the four specific research questions. Questions one and two, related to types of Appraisal choices and to the inclusion of justifications based on subject and age, are answered quantitatively using statistical data from CorpusTool. Questions three and four, related to the types of knowledge evaluated and to development of appropriate couplings, are answered qualitatively, examining text segments the tool provides for evidence of subject-specific lexis (knowledge) and evaluations which take the perspective of the discipline. In all cases, we look for evidence of development, based on the requirements of the subjects.

**Attitude by subject and age**

As regards differences by subject, Table 3 shows students’ selections from the different types of evaluation in the system of attitude, and presence or absence of justification at time 1.

Notably more bare appreciation was found in history, while the combination of appreciation and justification was detectably higher in the ecology texts. The table also shows that judgement and affect were used, although with minimal presence, the former in history, and the latter in ecology.

Table 4 compares the students writing in primary and secondary school on the same subject, history.

Table 4 shows a statistically detectable difference in the use of bare appreciation by primary pupils. As to judgement, while notably more instances appeared in history texts written by the students when in secondary school, the primary texts displayed a statistically detectable higher frequency of justified judgements. Affect with justification was somewhat more in evidence in the texts written at secondary level.

Table 5 gives the distribution of types of attitude by field and the appearance of justifications in the texts written when students were in grade 8, at time 2, the beginning of the school year (history) and at time 3, the end of that year (art).

Notable differences were registered between subjects for each of the types of attitude, with more appreciation and affect appearing in art, and more judgement in history. Table 5 also shows there were no notable differences between the two data sets in use of justifications for evaluations.

Thus, in response to our first research question, the effect of field on choices from the Appraisal system, appreciation is the main choice in the corpus studied, making up around 1.2 – 1.3 items per 100 words of text when all instances (both with and without justification) are totalled. This choice from the system of attitude shows the apprentice writers evaluating things or concepts using criteria.

<table>
<thead>
<tr>
<th>Table 2. Data analyzed.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longitudinal written sub-corpus</strong></td>
</tr>
<tr>
<td><strong>Primary</strong></td>
</tr>
<tr>
<td><strong>Time 1 History</strong></td>
</tr>
<tr>
<td>Number of texts</td>
</tr>
<tr>
<td>Total words</td>
</tr>
<tr>
<td>Average text length</td>
</tr>
</tbody>
</table>
appropriately, based on logic, not on feelings (Christie and Deriwianka 2008). Subject differences were also found. History produced JUDGEMENT, a very infrequent choice in art, and not found at all in ecology, an effect of both field (history is often developed around key players) and prompt (evaluations of historical actors or groups were triggered). In the secondary data collections, the choice of AFFECT with no justification in art marks a notable difference between the two fields, history and art.

This type of evaluation simply based on a personal reaction is considered an immature response in student writing (Martin 1989, 53).

With respect to our second research question, the inclusion of justification, we can discern no pattern related to development through this analysis. In fact, the highest ratio of responses to the prompt asking for justification for an evaluation were found in primary ecology (Table 3) and could be an effect of the students’ sensitivity to the topic of protecting the environment, since they had just returned from a week’s trip to an area of high ecological value.

Of course, numbers are only a part of the story. We now take a closer look at the data for couplings of evaluation + field and also the type of knowledge appealed to in justifications, in order to answer the third and fourth research questions. For this qualitative analysis, concordances were generated and examined, as explained above.
Couplings of Appraisal and disciplinary lexis by school subject and age

In the primary data, in terms of the types of entities appraised (ideational) and evaluative lexis used (interpersonal), in science, the students evaluated both superordinate entities related to the topic (‘animals’, ecosystem’, ‘environment’) and instances of those entities (‘fish’, ‘wolves’, ‘the shark eel’, ‘the food chain’, ‘seas and oceans’), with general and subjective evaluations revolving around the meaning ‘good/better’. More clearly showing knowledge of the scientific field, evaluations related to the effects on the ecosystem in terms of, for example, ‘danger/safety’, were also found, although not as frequently. In history, several appreciated items appeared in students’ observations during a time travel to the past, with everyday entities (‘clothes’, ‘food’, and ‘houses’) and subjective evaluative lexis related to ‘beauty/ugliness’, and ‘strangeness’. Less frequently, historical events or concepts were appreciated as ‘important’ (from the prompt), such as ‘the discovery of America’, ‘the French revolution’ or ‘liberalism’.

In the secondary texts, as in the primary, there are evaluations of everyday types of entities when looking at the past, such as ‘food’, ‘houses’, ‘clothes’. At the same time, appreciation of more general entities, or abstractions, such as ‘Roman society’, ‘conditions’, ‘craftwork’, ‘technology’, and ‘culture’ appeared. Finally, in the art data, everyday entities in the paintings, such as ‘snow’ and ‘grass’ were occasionally appreciated; however, more frequently, writers selected items belonging to the taxonomies of the subject, such as ‘painting’, ‘abstract painting’, ‘landscape’ ‘sculpture’, or entities of art and its study, such as ‘colour’ and ‘tone’. In answer to research question three, then, this overview of lexical selection in the longitudinal corpus can be seen to indicate slowly emerging disciplinarity.

Developing appropriate evaluation in different disciplines: evidence from justifications

In this section, we present a selection of justifications accompanying the evaluations to offer contextualized examples of the range of interpersonal and ideational couplings, as indicators of different degrees of development. Here we include only those evaluations which responded to the part of the prompt designed to elicit evaluation. Furthermore, only the choices of APPRECIATION and AFFECT have been included as they played a role in all three subjects, with JUDGEMENT practically limited to history. As to the identification of justifications, while the vast majority were encoded in ‘because’ clauses, there were other resources used, such as the conjunction ‘so’ and simple juxtaposition, as in examples 8 and 9:

6) But the plebeians fought with the patricians for 100 years for their political rights. So the two are good but I prefer to be a patrician. <HIST-2nd-B-50>

7) It was unfair, gladiators often were slaves who had no choice and had to fight for their lifes. <HIST-2nd-A-82>

We present and discuss representative examples by field, which allows us to evaluate the extent to which the language choices develop disciplinary knowledge with the evaluative perspective of a member of the community, that is, of biologists, historians or art critics. This, we contend, is evidence of development in writing in content areas.

Primary ecology

Table 3 in the previous sub-section showed a significantly higher number of justified appreciations in primary ecology, representative examples of which are shown in Figure 2:

The examples show a range of meanings, from the more scientific justification included in example 10, with its use of scientific terms ‘CO2’, and ‘oxygen’, to the more everyday knowledge in example 14. Except for example 10, the justifications themselves include APPRECIATION (through epithets such as ‘good’, ‘important’, ‘bad, dirty and poor’) and even AFFECT in example 14.
‘happy’), an unlikely inclusion in a scientific text. Example 11 also appeals to the field of ecology, but with unstable control of scientific expression (use of clauses such as ‘plants continuous growing’ rather than nominalized ‘continued plant growth’, for example). Example 12 uses typical vague language of everyday spoken conversation (‘things like that’), and appeals to the reader’s ability to make inferences, rather than showing the student’s field-specific knowledge. Example 13 provides no evidence, but rather builds up everyday appreciations of personal surroundings.

**Primary and secondary history**

In primary history, writers were prompted to choose a very important person, event or invention and explain their importance. As mentioned, few students chose to judge a historical figure. Figure 3 provides a sampling of important events and their justifications:

While all the students use a very limited selection of evaluative lexis (overwhelmingly ‘important’, from the prompt), most are able to give very basic historical information to justify their appreciations, with example 15 arguably providing the strongest historical evidence through the more complex nominal groups which incorporate classifiers, such as ‘scientific’, ‘technical’ and ‘commercial’, through which the student classifies the ‘advantages’ and ‘routes’. Example 18 shows an attempt to build a justification by hypothesizing, using the grammar of the CDF explore.

In secondary history, while the prompt asked students whether it was better to be a patrician or a plebeian, a number of students interpreted this personally, putting themselves in the role of a member of Roman society, and so encoded their responses as a personal preference, choosing from the system of AFFECT; examples appear in Figure 4:

While the justifications tended to include historical knowledge, the coupling with AFFECT brings the writer into the text, reducing its perceived objectivity. Example 19, like example 15, incorporates a classifier ‘economic’ for the abstract ‘situation’ of the patrician in Roman society, thus building in more technicality while justifying the preference. Example 22, however, while drawing on what

<table>
<thead>
<tr>
<th>Expression of APPRECIATION</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>10) It is very important to protect the environment</td>
<td>because if trees died we won’t breathe because they turn CO2 into oxygen. &lt;GA-22&gt;</td>
</tr>
<tr>
<td>11) It is very important to protect our environment</td>
<td>because: - It is not good to contaminate. - helps perseve animal’s species. - plants continue growing. - we can drink clean water. &lt;GC-87&gt;</td>
</tr>
<tr>
<td>12) I think keeping clean, safe...</td>
<td>the environment it's very important you know the global warming and things like that &lt;GC-70&gt;</td>
</tr>
<tr>
<td>13) I think is important to protect our environment</td>
<td>because we live, and if we don’t protect it, we would live in a bad, dirty and poor place &lt;GC-73&gt;</td>
</tr>
<tr>
<td>14) I think is important</td>
<td>because everybody will be more happy and the plant will be more clear. &lt;GC-85&gt;</td>
</tr>
</tbody>
</table>

**Figure 2.** APPRECIATION + Justification in Primary Ecology.

<table>
<thead>
<tr>
<th>Expression of APPRECIATION</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>15) I think the Discoverd of America is very important</td>
<td>because after it we made many scientific and technical advantages and we discovered new commercial routes. &lt;GB-40&gt;</td>
</tr>
<tr>
<td>16) I think the French Revolution its important</td>
<td>because it started our age. &lt;GB-34&gt;</td>
</tr>
<tr>
<td>17) I think the most important is the industrial Revolution</td>
<td>because we use a lot the industry. &lt;GB-39&gt;</td>
</tr>
<tr>
<td>18) In my opinion the discovered of America is very important</td>
<td>because all will chang if Colombus didn't found America &lt;GB-49&gt;</td>
</tr>
</tbody>
</table>

**Figure 3.** APPRECIATION + Justification in Primary History.
the student knew of the life of patricians, presents the voice of a present-day teenager (enjoying ‘free
time’) rather than that of a historian. Historians do use AFFECT, but coupled typically with countries as
historical actors having feelings (Coffin 2006), a choice also found in older CLIL student texts with
examples like ‘the anxiety of Germany’ (McCabe and Whittaker 2017).

More responses picked up the extra-posed structure of the prompt, so that students expressed the
attitudinal meaning through the system of APPRECIATION. A sampling appears in Figure 5:

Encoding attitudinal meanings through APPRECIATION is a way of institutionalizing feelings (Martin
and White 2005), and these examples show how this choice, in which the writer appears as
interpreter of history, is more successful in disciplinary terms than those in Figure 4. However, the
ideational meanings coupled with the APPRECIATION show the students to have a very unstable identity
as historians. The writer of example 23 is the only one able to avoid bringing the second person into
the presentation of evidence in the justification move. This student expresses the content in an objective,
arguably more disciplinary way, listing the advantages the patricians had during the time, while
other justifications range from offering some historical knowledge of the lives of the two collectives
to a simple appreciation, as in 26. The developed justification in example 25, while showing historical
knowledge through a personal reflection, presents it using the grammar of spoken language, chain-
ing together a series of subordinate and coordinate clauses (Halliday 1989).

Secondary art
The art prompt for CDF evaluate directed students to use knowledge of the discipline to select the
more effective of two paintings and give reasons for its success. As reflected in the Appraisal
results, here the main choice from the system of Attitude is that of APPRECIATION. Several writers,
however, took a personal stance, choosing to use AFFECT. Figure 6 presents sample responses
showing AFFECT + Justification for art.

Figure 6 offers a range of couplings of subjective stance and of technical or everyday knowledge.
In the examples, students positioned themselves using like and prefer. Their justifications, in some

<table>
<thead>
<tr>
<th>Expression of AFFECT</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>19) I obviously prefer to be a patrician</td>
<td>because they are in a better economic situation.&lt;C-2&gt;</td>
</tr>
<tr>
<td>20) I would like better to be a Patrician,</td>
<td>because I would have money, land and power &lt;A-73&gt;</td>
</tr>
<tr>
<td>21) I prefer to be plebeian to</td>
<td>don’t die in the revolution. &lt;C-23&gt;</td>
</tr>
<tr>
<td>22) I prefer to be a patrician because</td>
<td>I would have a lot of free time &lt;B-35&gt;</td>
</tr>
</tbody>
</table>

Figure 4. AFFECT + Justification in Secondary History.

<table>
<thead>
<tr>
<th>Expression of APPRECIATION</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>23) I think it would be better to be a Patricians</td>
<td>because the had all the rights, controlled the government and they didn’t have to paid taxes &lt;A-46&gt;</td>
</tr>
<tr>
<td>24) It was better to be a patrician</td>
<td>because you could have more lands, and slaves working for you. You would also live in a better place. &lt;C-6&gt;</td>
</tr>
<tr>
<td>25) I think, that was better to be a plebeian.</td>
<td>Because, if you were a plebeian, you don’t were so ambitious (ambicioso), because you were the poorest people of the empire, you could only get more things (same rights as patricians), so you could be completely happy, because you have your family, a work that could maintain it or maybe not, but you would have a family. And if were a patrician, you could only get less, so you would be care about what can you lose in stead of what do you have. &lt;B-76&gt;</td>
</tr>
<tr>
<td>26) I think it was better to be a patrician</td>
<td>because you live better. &lt;C-31&gt;</td>
</tr>
</tbody>
</table>

Figure 5. APPRECIATION + Justification in Secondary History.
cases, appealed to subject knowledge through technical meanings, as in example 27, through the nominal group ‘sensation of size, distance and depth’, creating a taxonomy of ‘dimensions’ (Hao 2015; Humphrey and Hao 2019), as art would be perceived by experts. Others, however, offered personal reactions as their reasons, (examples 30 and 31), an appeal which gives an impression of immaturity in the writer. This type of personal response is a genre found in primary school (Martin 1989). Interestingly, in example 27, the informal register of graduated or intensified JUDGEMENT (‘really good’) commenting on the artist’s technical capacity to achieve the effects in the painting diminishes its effectiveness as a scientific explanation.

These choices of AFFECT with justification to encode the CDF evaluate contrast with the samples of APPRECIATION, which appear in Figure 7:

Examples 32–38 show some knowledge of the technicality of art, through ‘dimensions of entities’ (Humphrey and Hao 2019, 224), evaluations and justifications involving the entities of art: colour, size, temperature, weight and depth. Example 36 taxonomizes colour (‘lighters colours’), evaluating its use in painting A as important, and then justifies that importance as having an impact in terms of ‘armony’, ‘peace’ and so on, moving away from an evaluation rooted in the technicality of art and into its effect on the viewer. Examples 37 and 38 take a representational approach to the painting, appreciating similarity to reality, with bare mention of disciplinary entities (only ‘colour’). In this data set (end of grade 8), notwithstanding the type of knowledge brought to their justifications, the students have reduced their use of features of spoken language found in the earlier texts.

### Table 1: Expression of AFFECT and Justification

<table>
<thead>
<tr>
<th>Expression of AFFECT</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>27) I prefer painting A</td>
<td>because I think he chose really good the colours to create sensation of size, distance and depth. &lt;C-31&gt;</td>
</tr>
<tr>
<td>28) I prefer the painting a)</td>
<td>because its colours are agreeable to eye because it is a harmonious painting &lt;B-75&gt;</td>
</tr>
<tr>
<td>29) I prefer the second one</td>
<td>because the colours and the perspective. &lt;B-35&gt;</td>
</tr>
<tr>
<td>30) I like mostly the A one</td>
<td>because I like the colors in it &lt;D-24&gt;</td>
</tr>
<tr>
<td>31) I like this one</td>
<td>because I love when winter ends and everything is green &lt;C-33&gt;</td>
</tr>
</tbody>
</table>

**Figure 6.** AFFECT + Justification in Secondary Art.

<table>
<thead>
<tr>
<th>Expression of APPRECIATION</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>32) In picture A the use of color has been very effective,</td>
<td>because the objects painted in darker colors, appears closer and smaller and lighter color appear further away and bigger. &lt;B-76&gt;</td>
</tr>
<tr>
<td>33) The colour is very usefull in this picture,</td>
<td>because appears to be autumn thanks to the red leaves, and it shows the distance of each object, the weight they have, the size they have, etc. &lt;B-55&gt;</td>
</tr>
<tr>
<td>34) There is a good use of the colours in picture A,</td>
<td>because they all give the sensation of the same temperature. &lt;C-6&gt;</td>
</tr>
<tr>
<td>35) In painting B the colour is very important</td>
<td>because give depth to the picture and said in what season is the painting inspired. &lt;C-23&gt;</td>
</tr>
<tr>
<td>36) In painting A it use lighters colours like yellowish green, white, green and light blue. the use of this colours are important</td>
<td>because transmit armony, peace, beauty, goodness etc. &lt;E-51&gt;</td>
</tr>
<tr>
<td>37) In painting A the colours are placed very well</td>
<td>because it seem like a photo. &lt;C-13&gt;</td>
</tr>
<tr>
<td>38) Painting B: it is very effective the use of colours</td>
<td>because they describe perfectly the dark trees and the path. &lt;B-7&gt;</td>
</tr>
</tbody>
</table>

**Figure 7.** APPRECIATION + Justification in Secondary Art.
Here we find few first and second person pronouns and a trend towards incorporating the abstractions of the subject.

In sum, the students’ texts demonstrate a range of abilities in using interpersonal and ideational coupling in writing. In response to research questions 3 and 4, first of all, it is clear that not all of the students show ability to use the more field-specific knowledge to respond to the prompts, since many texts draw on everyday knowledge. At the same time, examples appear which provide justifications rooted in disciplinary knowledge, which is especially evident in the final collection, in the texts on art. Development in ideational meaning is also shown in the amount and quality of language produced; there is a clear difference between the baseline data collection from the end of primary school and students’ production in secondary school, where justifications are longer and include more reasons based on subject knowledge. They also show a wider range of lexico-grammatical structures, and include more abstract nouns related to the two fields in the secondary data. At the same time, students’ control of the interpersonal metafunction when writing on topics learnt at school is, perhaps, less developed, as shown by the choices of AFFECT, rather than APPRECIATION as a strategy for evaluation in an institutional context, and the inclusion of ATTITUDE in justifications. We may say, then, that the students are slowly beginning to develop the resources which produce appropriate field + evaluation couplings.

Conclusions and pedagogical implications

In this final section, we sum up what the Appraisal analysis of the CDF evaluate has shown in three data collections spanning the transition from primary to secondary studies for pupils using a foreign language in three disciplinary areas, and relate the findings to pedagogical implications.

First, as regards choices from the system of ATTITUDE, in the whole corpus, the use of APPRECIATION is similar in frequency at all data collection times and in all subjects. APPRECIATION represents use of resources of evaluation to ‘assess the relative weight and social value of the things being interpreted’ (Christie and Deriwianka 2008, 123), and thus is an appropriate choice. JUDGEMENT appears in response to sections of the prompts on history, also appropriately. As to AFFECT, which appears in different subjects, a number of factors seem to be involved in this choice, which can be viewed as less appropriate for an academic context. One is related to the organization and design of the data collection, which led student writers to encode personal implication in the tasks, partly in response to their wording. Since the data collection was related to school subjects, but external to their teaching and assessment, the prompts were designed so that students could relate to the situations and be motivated to produce some language, in the form of a text and including the CDFs under study. With this in mind, the roles of writer and readers were presented as informal, rather than academic, with writers as bloggers and readers as followers, in the framing of all the tasks, which could have led to the personal reactions when evaluating. This has implications both for researchers and for the teaching community. Presenting an academic task as part of students’ everyday activities is a type of recontextualization which has an effect on register.

At the same time, we interpret the choice of AFFECT as evidence of developing but immature writers. In the case of history at Time 2, when asked to evaluate the classes in Roman society, some students made a personal interpretation, showing inability to use the historian’s perspective (Coffin 2006). This may also reflect a history teaching strategy, that of encouraging learners to empathize with the people of the times studied. In art, despite the prompt explicitly requiring a technical response in the comparison of two paintings, here, too, AFFECT was strongly present. Of course, the function of art is to engage the viewer, so in this subject, a personal response is more difficult to overcome.

As to the production of justifications of evaluations, which we expected to develop as the writers matured, it seemed that contextual factors tangentially related to the subject due to students’ experiences, and unrelated to age, intervened to produce the highest ratio at Time 1, in one of the subjects, ecology. However, the linguistic analysis of the coupling of discipline-appropriate criteria as a base for
evaluation, on the one hand, with concepts representing specialized knowledge, on the other, was more revealing of development, and offers an entry point into the application of this study.

In terms of application to the classroom, the range of academically more or less appropriate couplings provides material for awareness-raising class discussions around language and content, since studies on CDFs in class interaction find very little meta-talk around the functions, despite their role in constructing content (Dalton-Puffer et al. 2018). Teachers could use the meanings discriminated in the Appraisal framework and the range of more to less field-specific lexis analyzed to make explicit to students what is required in a task involving evaluation, that is, which choices from the Appraisal system would best express the required academic perspective for the target of evaluation, and what type of content would justify their view for the academic community. A series of text fragments like those analyzed here could be used to show how academic evaluations tend to be impersonal – that is, with third person focus – and typically choose Appreciation rather than Affect, and are supported by justifications which incorporate the specialized lexis of the subject, abstractions and generalizations, and excluding evaluative language.

Working with teachers on texts from their disciplines using categories from a functional linguistic model such as that applied here can bring the key meaning-making resources of the language of school subjects to consciousness – from implicit to explicit knowledge – in a short time, as shown by teacher education projects in different parts of the world (see, among others, Gebhard 2019; Rose 2011; Schleppegrell 2011; Schleppegrell and de Oliveira 2006; Schleppegrell, Achugar, and Oteiza 2004; Whittaker and Acevedo 2016). An important and complex CDF like evaluate is required in many of the questions students have to answer, and central to writing argument and discussion texts.

Curriculum designers, then, need to incorporate a developmental path linking CDFs and the language needed to realize them. This approach would trigger a focus on language, one which is relevant to the teaching of specific subjects. Content is inextricable from the language in which it is expressed, so that ‘teachers need to understand the kinds of language choices that are relevant to accomplishing the writing tasks they are asking their students to perform’ (de Oliveira and Schleppegrell 2015, 108). The transition to secondary studies – in Butt’s (2004; 218) apt and suggestive metaphor ‘the strait to specialization and beyond’ – can be supported, speeded up and, above all, made more meaningful for students when teachers give them explicit help so that they take control of the language making meaning in their texts.

Notes
1. Voices of CLIL teachers in seminars with the authors; see Cammarata and Tedick 2012 for voices from a different context.
2. To distinguish between technical and non-technical uses of terms, Appraisal systems are in small caps.
3. Student examples are presented verbatim. Tags read as in example 1: HIST-2nd-B-35: subject, primary/secondary, student group, student number.
4. The sub-corpus includes fewer texts on history at Time 1 (26 texts). As the three groups of students in the primary data collection alternated subject (history/biology) and language (English/Spanish), the combination of history in English was only written by one group, with the other two writing on ecology in English. For the secondary data, collected as students were starting grade 8 (history: 47 texts) and at the end of that year (art: 49 texts), all students did the tasks in both languages.

Acknowledgements
We would like to thank the students who wrote the texts, their teachers, and especially the heads of the schools; without them this study would not have been possible. Grateful thanks to Mick O’Donnell for UAM Corpus tool and for his invaluable help. This paper is part of a project that was funded by the Ministerio de Economía y Competitividad: Integrating and assessing content and language in the transition from primary to secondary bilingual education TRANS-CLIL (MINECO, FFI2014-55590-R) (Past projects financed by MINECO and CAM/UAM.). Project webpage: http://www.uam-clil.org
Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Ministerio de Economía y Competitividad: [grant number MINECO, FFI2014-55590-R].

Notes on contributors

**Rachel Whittaker** is honorary lecturer in the English Department of the Universidad Autónoma, Madrid, and is active in teacher education for CLIL and EFL, her areas of research. Co-founder of the UAM-CLIL research group (uam-clil.org), she is co-author of *The Roles of Language in CLIL* (2012, CUP), a book on language use and development in CLIL classes. She recently co-edited a monograph of *Lenguaje y Textos* (2017) on a project she coordinated, which brought the Australian literacy pedagogy, Reading to Learn, to Spain. She has published a number of articles and chapters on CLIL and EFL student writing and on Reading to Learn and its results.

**Anne McCabe** holds a joint appointment in the Departments of English and Communication at Saint Louis University’s Madrid Campus, where she teaches courses in rhetoric, academic writing, linguistics, public speaking, and ESL. She has published numerous articles and book chapters related to language teaching/learning using Systemic Functional Linguistics (SFL). She has also edited book collections, *Language and Literacy: Functional Approaches and Advances in Language and Education* (Bloomsbury, with Rachel Whittaker and Mick O’Donnell), and *An Introduction to Linguistics and Language Studies* (Equinox); she is currently completing a monograph titled *A Functional Perspective on Developing Language* (Routledge).

**ORCID**

Anne McCabe [http://orcid.org/0000-0001-6395-1442](http://orcid.org/0000-0001-6395-1442)

References


Appendix: Sample prompt

Secondary History. Imagine that you arrive in Ancient Rome in the time of the Republic and you write your impressions in a blog. Don’t forget to refer to all the following ideas in the blog: -What was life like in Ancient Rome (people, houses, activities, food …)? Compare it with life now. -You find that the organization of society was different too, so you want to tell your readers about that. Define patricians and plebeians for them. Was it better to be a patrician or a plebeian, in your opinion? Why? -Now imagine that you can stop at either (A) the Coliseum to watch a gladiator fight or (B) the Curia on the day of Caesar’s death. Choose one of these situations and write what happened. If you chose (A), don’t forget to explain to your readers why the gladiators had to fight. If you chose (B), explain why Julius Caesar died. -And finally, give your readers your reflections on this question: What would Europe, and, in particular, Spain, be like, if the Roman Empire hadn’t been so powerful and successful?