The role of different tasks in CLIL students' use of evaluative language

Ana Llinares a, *, Christiane Dalton-Puffer b

a Universidad Autónoma de Madrid, Spain
b Universität Wien, Spain

ABSTRACT

Some consider CLIL to be a “particular pedagogic manifestation of the task-based approach” (Skehan 1998, 276). In this study we examine learners working on a range of naturalistic tasks following the rationales of CLIL social science subjects in three European contexts (Austria, Finland, Spain). The focus is on learners’ use of interpersonal resources, especially the language of evaluation. Using an integrative analytical framework that draws on systemic functional appraisal theory (Martin & White 2005), Goffman’s (1981) participation framework as well as educational-pragmatic notions of academic discourse functions (e.g. evaluating; Dalton-Puffer 2007, 2013), we analyze CLIL students’ evaluative language across five ecologically viable task-types (whole-class discussions, group-work discussions, individual interviews, oral presentations and role-plays). Findings show clear differences in the frequency and distribution of different appraisal types between different tasks, with role-play and whole-class discussion forming the opposite ends of a continuum. The analysis puts particular emphasis on the resources for engagement because of its significance for expressing learners’ epistemological stance on curricular content. In educational terms such resources are often framed in terms of higher order thinking skills or discourse functions like arguing, evaluating.

© 2015 Published by Elsevier Ltd.

1. Introduction

Task-based language teaching (TBLT) and content-based instruction (CBI) represent different approaches to second language teaching which have received wide attention from research but have seldom been explored in conjunction, even though both belong to the field of second language learning research. The two models share their focus on meaning-making activities geared to achieving a non-linguistic outcome in the interest of enhancing second language development. However, in some bilingual education (or CBI) programs like European CLIL because of the importance of content learning, an approach seems commendable that is less exclusively inspired by language acquisition models (as in TBLT) and takes a more general educational-linguistic perspective.

It is true that the rationale based on CLIL improving second language learning of school-level students has been very important as a trigger for the implementation of such educational programmes (European Commission, 1995, 2003).
However, these motivations have been met with understandable misgivings by content-subject specialists fearing for the status and weight of their intrinsic learning goals in the enterprise (cf. Badertscher & Bieri, 2009). In order to solve this potential conflict, subsequent conceptualizations of CLIL have claimed it represents an integrated approach that incorporates a focus on both content and language. In these conceptualizations, learner-centredness and learner activation is frequently invoked, as in TBLT. In CLIL this type of pedagogical practice is viewed as a pathway to achieve the desired integration, as well as heightened cognitive engagement from which both subject and language learning are expected to profit (e.g. Coyle, Hood, & Marsh, 2010; Marsh, 2013).

Despite the pervasive conceptualization of CLIL as a content and language integrated approach and the growing research on this construct, the realities of CLIL implementation clearly emphasize its content-side. Prototypical CLIL implementations (Dalton-Puffer, 2011; Dalton-Puffer, Llinares, Lorenzo, & Nikula et al., 2014) show that CLIL lessons are taught by content-specialists, who are often non-native speakers of the CLIL language. Sometimes these teachers also possess qualifications for foreign language teaching but this is a requirement only in a few national education systems. At school level, CLIL lessons are timetabled as content-lessons which are guided by content learning goals and pedagogical traditions inscribed in the content-subject curricula. Furthermore, assessment and certainly high stakes testing in CLIL classes is driven by the content-subject. Likewise, the pedagogical tasks set in CLIL lessons are typical of the respective content-subject (history, science etc.), thereby incidentally foregrounding students’ role as users rather than learners of the L2. Nevertheless, in a CLIL situation, students’ participation in content subject-specific tasks will necessarily require them to use the L2 both for epistemic and social purposes, involving the use of linguistic and pragmatic resources to express both factual and interpersonal meanings, all of which is ultimately expected to enhance language learning. The empirical focus of this study will be on interpersonal meanings and, more specifically, the type of evaluative language used to express those meanings across a variety of tasks.

Since the reality of CLIL thus combines the rationales and practices of two pedagogies and their principles and traditions (that of language pedagogy and that of subject pedagogy) it seems a pertinent question to pursue how far this intertwining of meanings and, more specifically, all of which is ultimately expected to enhance language learning. The empirical focus of this study will be on interpersonal meanings and, more specifically, the type of evaluative language used to express those meanings across a variety of tasks.

2. TBLT, CLIL and the notion of task

In this section, we elaborate on our rationale for exploring the interface between TBLT and CLIL. For this purpose, we need to discuss in detail the understanding(s) of task that underlie both educational approaches.

Let us start with the definition of task as developed in second language acquisition (SLA) research. Although the notion of task is not unequivocal in TBLT, ‘the following criteria outlined by Ellis (2003) are generally accepted as necessary (if not sufficient) for a classroom/learning activity to count as a task in the sense of TBLT:

- It has a clearly defined, non-linguistic outcome.
- It involves a primary focus on (pragmatic) meaning.
- It has some kind of ‘gap’ (information gap, reasoning gap, opinion gap, …).
- The participants choose the linguistic resources needed to complete it. (Ellis, 2003)

Interestingly, tasks that fulfill all the above criteria happen naturally in CLIL classrooms, in other words, there seems to be no need to design them specifically as in EFL classrooms. However, the rationale behind these four criteria may be different in EFL and CLIL contexts and the identification of their specific nature in CLIL is key for the understanding of the opportunities these tasks offer for language and content engagement and learning, and also to help design similar tasks which might also work in EFL contexts.

Quite naturally for CLIL teachers, the primary desired outcome of tasks is that learners reach the curricular goals of the content-subject and these are by definition not linguistic but content-subject goals. Even though language is clearly implicated in subject learning as initiatives like Language Across the Curriculum (the study of language in non-language subjects) have repeatedly underscored, it continues to be extremely difficult to change the understanding and self-image of subject-teachers in a fundamental way. Therefore, it is important to understand the nature of tasks in CLIL classrooms in order to identify to what extent they also contribute to language use and learning while students are engaged in the expression and understanding of academic content.

With regard to the second criterion, it is obvious that tasks in CLIL represent a focus on meanings which are connected with the concepts, notions, facts and skills of the content-subject. In other words, meaning orientation is there automatically via the content curriculum and concrete non-linguistic outcomes are required qua specific content learning goals (e.g. “present arguments for why this is a breach of the contract of sale”). Within TBLT in foreign language classrooms focus on

---

1 There is a sort-of-canon of TBLT tasks in TBLT research, much of which is done outside the classroom, while the notion of task in classroom settings has been less explored.
language is assumed to arise incidentally, from problems with executing the task. In the case of CLIL and immersion class-
rooms, there are now a number studies which show that explicit focus on the target language is actually a rare event. If focus
on language does happen in CLIL, it overwhelmingly orient towards lexical items that count among the subject-specific
vocabulary (e.g. Dalton-Puffer, 2007; Jakonen & Morton, 2014). The absence of language focus is so marked in some im-
mersion contexts that Lyster has proposed addressing this imbalance (Lyster, 2007) by introducing language-focused tasks in
immersion teaching, cf. his article in this special issue. Lyster argues, convincingly in our view, that within the totally meaning
focused context of French immersion in Canada, it is beneficial to introduce form-focused tasks in order to enhance learners’
acquisition of a second language. Llinares, Morton and Whittaker (2012) and Llinares and Morton (2012) refer to need for
language-focused tasks in bilingual education addressing not only formal but also functional features of the language. These
form-focused or language-focused tasks also imply a wider understanding of task than is customary in core TBLT research as
the criterion of exclusive meaning-centredness is questioned.

With regard to the third criterion, “some kind of gap”, we can observe that the tasks implemented by CLIL teachers for
working towards content learning goals of necessity include knowledge, opinion and reasoning gaps. But in this context it is
also necessary to distinguish the source of the gap: language or content; in other words, we need to know whether the gaps or
communication breakdowns identified in different oral tasks in CLIL are only related to the subject-curriculum or whether
there are also language gaps that affect the way these students learn, understand, remember, analyse, discuss issues of the
content-subject.

Finally, as far as the fourth criterion is concerned, students’ freedom of choice of linguistic resources, i.e. their acting as L2
users in solving the task, offers particular interest for research as differences in tasks are likely to be accompanied by dif-
fferences in linguistic resources to express what is in essence the same content.

Thus, all basic defining criteria of task in TBLT seem to be fulfilled in CLIL tasks naturally. However, they will require
adjusting the focus, as they present differences as compared to prototypical TBLT tasks in foreign language classrooms. As
CLIL, by definition, has a built-in language learning agenda, we think it is of great interest to investigate the extent to which
these different tasks — depending on the meanings, gaps, and participation roles involved—affect the linguistic resources
learners put to use and are able to learn in the process.

Since, as pointed out above, actual CLIL-classes are predominantly driven by the logic of the content-subjects, it is not only
logical but necessary to turn to those subjects for enriching/further concretising our notion of task before analysing language
use in naturalistic CLIL lessons. The question which needs to be solved, then, is what are the tasks that are part of the
pedagogical cultures in the particular subjects. For reasons of the economy of space required in a journal article we restrict
ourselves to the oral side of social science subjects (history, economics, geography) for the rest of this contribution.

In the following, we therefore briefly describe the types of tasks we identified in the database that was at our disposal. The
tasks under analysis in this study share the following features: they have a component of oral performance and they are
“naturalistic” in nature. By naturalistic we mean that these are tasks more or less commonly used in CLIL secondary class-
rooms at least in three European contexts (Austria, Finland and Spain). We are using the term task-type in a generic way. In
other words, the classroom activity we are discussing lacks specification regarding the concrete content-operations required,
in contrast with the characteristic descriptions of tasks in SLA research.

Altogether there are five types of tasks that can be identified in the CLIL classroom data accessible to us (for details see
section on data below). All activities shared an end-of-topic knowledge discussion agenda, but they did this in different
formats, involving the participants in different constellations and different types of subject positions. The task types were:
whole-class discussion, group-work discussion, student-presentation, and role-play. In addition, there was one task, an end-
of-topic interview, which was collaboratively planned by teachers and researchers. Although this was the only task that did
not happen naturally in the CLIL classrooms observed and was designed for other research purposes, it mapped the format of
individual oral assessment (as confirmed by the class teachers) and had the same purpose of checking end-of-topic knowledge
as whole-class discussions.

Teacher-fronted whole-class discussions involved the whole group of learners in a conversation led by the teacher. The
overall topic of the conversation was determined by the curriculum but it was for the teacher to decide how to lead the group
to cover the ground required. The roles in the interaction were a-symmetrical, granting automatic right to the floor to the
teacher as well as giving him or her control over the topics introduced and pursued. The teacher was expected to feed back to
the whole group whether a student contribution could be sanctioned as building legitimate subject content. These roles with
their different rights and obligations were deeply culturally ingrained and overlearned by both parts. Several studies on CLIL
teacher-fronted interactions have highlighted the teachers’ control of what is talked about and how and, for example, the fact
that students do not have opportunities to use certain speech acts like directives, as these are primarily considered to be
allocated to the teachers’ role (e.g. Dalton-Puffer & Nikula, 2006).

In group-work discussions students were gathered in smaller peer-groups in order to pool resources for solving some kind
of problem, for example, working through fresh material together (like comparing and analyzing images according to criteria
previously discussed with the teacher). Some studies have focused on the positive role of peer or group interaction and
collaboration among students for participation and language learning in both immersion (e.g. Ballinger, 2013) and CLIL
contexts (e.g. Llinares, Morton and Whittaker (2012) and Llinares and Morton (2012)). Ballinger’s (2013) study highlights the
importance of teaching students reciprocal learning strategies and shows the positive role of acknowledging partners’
contributions. Llinares, Morton and Whittaker (2012) and Llinares and Morton (2012) show the opportunities provided in
group work discussions for students’ participation not only as reproducers of the content learnt from the teacher or in

Please cite this article in press as: Llinares, A., & Dalton-Puffer, C., The role of different tasks in CLIL students’ use of evaluative
language, System (2015), http://dx.doi.org/10.1016/j.system.2015.05.001

A. Llinares, C. Dalton-Puffer / System xxx (2015) 1–11
textbooks (‘animators’ in Goffman’s (1981) participation framework) but also as ‘authors’, who choose and put together the words, and ‘principals’, by generating new content-related ideas (Bunch, 2009; Goffman, 1981).

Student presentations were student monologues relating knowledge gained outside of the whole-group plenary to the members of that plenary. The source which generated that knowledge could be in an in-class working group (i.e. the students reported on the results of a group-work session) or a homework assignment completed individually or jointly with classmates (cf. Ostojic, 2010; Zareva, 2013).

In role plays students took on the roles of fictitious or imagined, historical or present-day persons for whom some aspect of subject-content had “real-life” relevance. They acted out a dialogue, spontaneous or scripted by themselves, as it might occur between the carriers of those roles outside the classroom. In CLIL research, Gassner and Maillat (2006) and Maillat (2010) have studied the involvement of students in role-play situations as a result of students’ perception of their linguistic performance not being evaluated as in EFL classrooms.

Interviews were one-on-one situations where a researcher had an end-of-topic interview with a student asking the student to tell the interview partner about the content of the unit they had been studying during a sequence of CLIL lessons. In a study comparing CLIL secondary school students’ performance in teacher-fronted discussions and in individual interviews with a researcher, Llinares and Morton (2010) found that in the interviews students produced longer and more linguistically complex explanations, which could be due to differences in the role students had to take on in the task.

To sum up this section, our approach to task and TBLT differs from SLA-based TBLT studies, which study the linguistic behaviour of learners as they solve tasks such as picture stories or information gap activities either in the language classroom or in an experimental situation. Task in this article refers to the naturalistic learning situation (the classroom) but widens the angle to the CLIL classroom and thus to content-pedagogical tasks set by CLIL content teachers and to be worked upon by students in the L2.

3. Criticality and stance as learning goals in social science CLIL

In social science curricula, to a greater or lesser extent, students are expected to evaluate and adopt a stance on the content that they are learning as illustrated by the following short extracts from the Spanish CLIL curriculum (currículo integrado) for the social sciences (1a) and the Austrian upper secondary history and civics curriculum (1b):

1a. ... interpretar y relacionar información procedente de fuentes diversas. [...] aplicar las técnicas elementales de comentario de textos [interpret and relate information from diverse sources [...] Application of basic techniques of textual commentary] (BOCM, 2008)

1b. Entwickeln eines individuellen Handlungsrepertoires für die politische Auseinandersetzung und Meinungsbildung [Developing a personal activity repertoire enabling political engagement and formation of opinion] (BMUKK, 2005)

Breaking down such important but very general learning goals to a more fine-grained level, one can identify certain speech functions by means of which these goals are actually worked upon in the classroom situation. Recently, a coherent construct of such Cognitive Discourse Functions has been proposed (Dalton-Puffer, 2013) where the following speech functions are available for working towards the above-mentioned curricular goals of the content-subject:

<table>
<thead>
<tr>
<th>Type 4</th>
<th>“I tell you something that is potential”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 5</td>
<td>“I give you reasons for and tell you cause/s of X”</td>
</tr>
<tr>
<td>Type 6</td>
<td>“I tell you what is my position vis a vis X”</td>
</tr>
</tbody>
</table>

Thus it is through verbalization that participants can make accessible to others their cognitive positioning vis a vis some fact, event or state of affairs (e.g. the actions of a specific historical agent in a specific situation). More importantly, such verbalization can be employed to develop such positionings via exploring a number of possible alternatives. As such, cognitive-linguistic operations are central to working towards the educational goals of the (CLIL)-subject. Clearly then, it is of great interest to find out if, in fact, learners have opportunities to carry them out during classroom lessons, how second language users cope with their demands and whether some tasks encourage their use more than others.

Building on this more generally cognitive, educational and content-oriented discussion, the following section introduces the actual framework of analysis in this study. This framework is a fully articulated general linguistic model integrating functional categories with linguistic form.

4. Systemic functional linguistics (SFL) and appraisal theory in CLIL research

It has been widely acknowledged that the acts of evaluating people, things or events are not exceptional but must be seen as intrinsic to language (Halliday & Matthiessen, 2004; Sarangi, 2003). Within the systemic functional linguistic model (SFL), evaluating (by adopting a stance or expressing attitude) is part of the interpersonal metafunction of language (Halliday &
Matthiessen, 2004), which is part of any classroom culture, as it is necessary to establish social relations as well as to adopt a stance towards subject-related information. Several studies carried out within the systemic functional model have indicated the key role of students’ competence in conveying interpersonal meanings for their success at school (e.g. Christie, 2002; Coffin, 1997; Marshall, 2006; Martin, 2000; Polias, 2003; Schleppegrell, 2004).

One SFL model that has been used to identify speakers’ and writers’ linguistic resources to express evaluative meanings such as emotion or judgement is “Appraisal Theory”, defined by Martin and White (2005: 1) as

…concerned with the construction by texts of communities of shared feelings and values, and with the linguistic mechanisms for the sharing of emotions, tastes and normative assessments … with how writers/speakers construe for themselves particular authorial identities… (Martin & White, 2005:1)

There are three main types of appraisal (see Fig. 1 below):

- **attitude** (concerned with feelings and emotional reactions),
- **engagement** (concerned with the sources of attitudes),
- **graduation** (concerned with grading phenomena).

Appraisal theory, then, identifies the lexicogrammatical resources used to convey these types of interpersonal meanings. Following Martin and White (2005), graduation is concerned with either amplifying or weakening, i.e. force (e.g. a lot), or softening and sharpening, i.e. focus (e.g. sort of nice). Within attitude there are 3 categories: affect, related to emotions (e.g. I’m sorry); judgement, concerned with resources for assessing people’s behaviour (e.g. greatest king); and appreciation, for the evaluation of concrete and abstract things (e.g. good society).

Within the category of engagement, in this study we focus on heteroglossia or dialogic dimension where some propositions allow for alternative positions (expansion) while others challenge or restrict the scope of a proposition (dialogic contraction). Within the latter (i.e. dialogic contraction) the taxonomy of engagement includes the categories of disclaim and proclaim. In disclaim the voice rejects a contrary position through negation-deny (e.g. no, don’t) or concession-counter (e.g. but). Proclaim consists of presenting a proposition as well-founded and reliable, ruling out alternative possibilities. This can be done through concur, which expresses agreement through affirming or conceding (e.g. yes, yes... but); pronounce, which expresses authorial presence and emphasis (e.g. it is like that) and endorse, through which external voices are valued by the speaker as correct (e.g. the report proves...). Turning now to the expanding categories, entertain implies that other alternatives may be possible (Maybe) and attribute implies that the proposition may be presented as one out of other possible options (the Egyptians thought...). White (2003) distinguishes justify as a third category of engagement, which refers to non-factual propositions and expresses some kind of legitimation of the interpretation.

In school contexts, the category of attitude is particularly relevant for social science, particularly history, where characters and events are evaluated (e.g. Coffin, 2006). In the analysis of students’ participation in oral tasks, the category of engagement acquires special interest as it reflects participants’ positioning both towards content (How do we know this? Who said so? Can we be sure?) and towards their peers (Is this her/his opinion? What is it founded on? Do I agree? Do I have a counterargument?).

The engagement system used for the present study, set out in Fig. 2 below, will help us identify the types of interpersonal language triggered in different tasks, as well as the lexicogrammatical resources used by CLIL students in order to convey these functions.

5. The study

Based on the assumption that CLIL implementations are not driven by the rationale of foreign language teaching but by subject-specific curricular goals, and taking into account that evaluating and appraising academic content is part of the curriculum of any subject, the main aim of this study is to describe CLIL learners’ use of evaluative language across tasks arising in content-subject pedagogy. More specifically, the study aims to answer the following questions:

a) What types of evaluative resources are used by secondary school CLIL students across tasks?
b) To what extent are these resources task-related?
c) What can this tell us about the learners’ epistemological stance and content and language integrated learning goals?

5.1. Data

The data consist of classroom discourse video and/or audio-recorded in naturalistic tasks, in other words, tasks more or less frequently used by CLIL teachers in secondary social science classrooms (history, economics and geography). As the data

---

2 No example of endorse was found in our corpus. This example comes from Martin and White (2005).
were not collected specifically for this study, the teachers had not been asked to include any tasks that were not part of their usual lesson planning. The year groups range between 7 and 10 (students' age between 12 and 16). As Table 1 below shows, the data come from three European CLIL contexts: Austria, Spain and Finland. The total number of words analysed was 26,300.

The teacher-fronted or whole-class sessions, the interviews and group work sessions followed an end-of-topic discussion agenda in different formats and involving different participants (teacher-students in teacher-fronted sessions, student—in group work sessions and interviewer/researcher-student in interviews). A total number of 12 interviews with 4 different students were analysed, comprising a total of 8400 words. The students belonged to two different schools participating in the same bilingual program in Madrid (Spain). The interviews (3 per student) were carried out by a researcher following a prompt negotiated with the teacher and consisting in a one-to-one discussion at the end of a topic. The prompt included 4 or 5 questions designed by the researchers together with the teachers with the purpose of eliciting students’ knowledge of the topic, covering different history genres, such as recount, explanation or argumentation (Coffin, 2006). The group-work data included 3 sessions, one from each of the 3 European contexts (Austria, Finland and Spain), comprising a total of 5100 words. There were between 2 and 4 students in each group. The teacher-fronted class discussions consisted of 2 sessions, one from the Austrian context including 18 students and one from the Spanish context with 23 students (6200 words). The oral presentations involved defending a proposal. A total number of 8 presentations were analysed (4 from the Austrian context and 4 from the Finnish context), comprising a total number of 3600 words. The role plays expected the students to play the part of, for example, a German soldier in WWI, or a customer filling in a complaint in a business to business context, and participating in a dialogue adopting that role. A total number of 8 role plays with 2 students participating in each were analysed, two from the Austrian context and six from the Spanish context (3000 words).

Once the data had been transcribed, students’ evaluative resources were coded following Martin and White (2005) appraisal model, adapted for the analysis of CLIL classroom data (see Fig. 1). The UAM-Corpus Tool was used to code the data and to retrieve the quantitative results regarding the different types of appraisal used by the students. The coding procedure was carried out by both researchers, and 50% of data was checked by another expert. The few discrepancies in coding were discussed until a common solution had been reached.

Table 1
Data used for analysis.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Sessions</th>
<th>Subject</th>
<th>Context</th>
<th>Words approx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview</td>
<td>12 (10 min) interviews (3 × 4 students)</td>
<td>History</td>
<td>Spain</td>
<td>8400</td>
</tr>
<tr>
<td>Teacher-fronted</td>
<td>2 sessions (à 50 min)</td>
<td>History, Geography</td>
<td>1 Spain, 1 Austria</td>
<td>6200</td>
</tr>
<tr>
<td>Group work</td>
<td>3 sessions (à 50 min)</td>
<td>History</td>
<td>1 Spain, 1 Austria, 1 Austria</td>
<td>5100</td>
</tr>
<tr>
<td>Presentation</td>
<td>2 sessions (à 50 min), 4 presentations in each</td>
<td>Economics, History</td>
<td>4 Austria, 4 Finland</td>
<td>3600</td>
</tr>
<tr>
<td>Role play</td>
<td>8 role-play sessions (10 min)</td>
<td>6 History, 2 Economics</td>
<td>6 Spain, 2 Austria</td>
<td>3000</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>26,300</td>
</tr>
</tbody>
</table>

Please cite this article in press as: Llinares, A., & Dalton-Puffer, C., The role of different tasks in CLIL students’ use of evaluative language, System (2015), http://dx.doi.org/10.1016/j.system.2015.05.001
5.2. Analysis and findings

A global analysis of the data in terms of the three types of appraisal proposed by Martin and White’s (2005) model shows a marked difference in students’ use of evaluative language depending on the task they were engaged in. This is particularly clear when one considers the weighted totals (use of appraisal per 1000 words) that appear in the rightmost column of Table 2. Role Play shows more than 10 times as much appraisal as Whole—Class Discussion. Also, there is a clear grouping of high—mid—low appraisal tasks: Role Play and Interview are high in incidence, Group Work and Presentation occupy a middle position (ca. half of the top group) and teacher-led Whole—Class Discussion brings up the rear with very little observable appraisal going on. Although these findings are in line with previous studies which have shown the positive role of group work sessions for students’ use of the interpersonal function and evaluation compared to whole-class discussions (Llinares & Pastrana, 2013), this study shows that there are other tasks that prompt the students to use evaluative language even more than Group Work.

What is also visible from Table 2 is the fact that engagement is the most frequently used appraisal category in this set of classroom data overall and also in every task-type except Interview, where the three categories of appraisal are very evenly distributed. Engagement is also interesting because its linguistic and cognitive function is particularly important for the subject learning goals required by the task. Attitude depends on the type of content being talked about; and, thus for example, in history students are expected to appreciate and judge (attitude) historical events and actors. However, the use of engagement, in addition to the content, is particularly related to the characteristics of different tasks, as they may require the students to position themselves differently as participants. This, of course, may have interesting implications not only for how content is talked about but also for the linguistic resources needed to perform adequately in each task in the L2. After a brief review of the findings regarding graduation and attitude we will therefore focus on engagement.

Graduation, the least frequently employed appraisal type, involves speakers in intensifying or downtoning their utterances (force), as well as applying discourse focus (zooming in or out), through focus. Frequent examples of graduation include the use of “very” and “a lot/a lot of” as in the following extract from an interview on Feudal Europe:

The figures for appraisal-type attitude are shown in column 2 of Table 2. What immediately becomes apparent is that Role Play and Interview produce most attitude. As explained in Section 4 this appraisal type projects the speaker’s emotional subject position, their wants and obligations, as well as moral and aesthetic judgements. The data show that one-on-one interaction (Role Play and Interview, see extracts 2 ab & 3) seems to encourage making this kind of positioning explicit, while presentational monologue as well as the multi-voiced whole-class discussion seem to transpire almost devoid of this type of meaning as students present factual content with very little affect, judgement or appreciation. The following examples illustrate the attitude-rich language of Role Play and Interview:

![Table 2](http://dx.doi.org/10.1016/j.system.2015.05.001)

Please cite this article in press as: Llinares, A., & Dalton-Puffer, C., The role of different tasks in CLIL students’ use of evaluative language, System (2015), http://dx.doi.org/10.1016/j.system.2015.05.001
Let us now turn to engagement as the most frequent type of appraisal in the data analysed and one that has been less studied in school contexts. The following table (Table 3) sets out a selection of engagement resources from the Martin and White (2005) and White (2003) framework. Only the types that were used by the students are included in the table.

Viewing the table as a whole, it is noticeable that the distribution of engagement token frequency seems to be related to type frequency; that is to say token-rich Role Play exhibits all engagement types found in the data, while Whole-class shows only 5 out of 8 types.

As explained in section 4, engagement types can be classified into contract, expand and justify. The contract-subcategory disclaim is most frequently actualized in the one-to-one interactions, where participants both deny the truth value of propositions (see extract 4 from an interview) as well as counter them with alternatives (see extract 5 from a Role Play). This latter meaning is particularly frequent in the role-plays, all of which simulate situations where participants represent opposing interests and contrary positions (a customer complaining about a break of contract vs. a supplier maintaining they kept to it; a German vs. a French World War I soldier trying to maintain the rightfulness of their own views on the conflict).

The fact that poly-logic group-work and the mono-logic presentations show similar values with regard to the disclaim category may seem puzzling, but can be explained by a closer look at the content of these presentations. The Austrian students, for instance, were required to report on a problem-solving procedure concerning an unclear legal situation where different alternatives had to be considered before coming to a decision. In doing this they had to formulate various alternatives, thus representing different voices and ending with a justification of the final decision. This also explains the frequent use of entertain (15.58 per 1000 words, in Table 3 above).

Another interesting observation in the area of contract concerns the occurrence of Concur. Concur: concede (“I see what you mean; you are right, but”) seems to be almost beyond the reach of the students who participated in the interactions analysed, given that Concede occurs only in Role Play and there is only 1 instance in Group Work. The use of this type of appraisal is illustrated in extract 6 from a role play between Colonizers and American Indians:

One of the most frequent types of appraisal in group-work sessions was concur: affirm, as illustrated in extract 7 below, in which a group of students uses some pictures to discuss children’s conditions during the Industrial Revolution:

### Table 3

<table>
<thead>
<tr>
<th>Engagement type</th>
<th>Task</th>
<th>Role play</th>
<th>Interview</th>
<th>Group work</th>
<th>Presentation</th>
<th>Whole class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disclaim: Deny</td>
<td>15.73</td>
<td>12.34</td>
<td>8.03</td>
<td>7.19</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>Disclaim: Counter</td>
<td>20.78</td>
<td>9.28</td>
<td>6.45</td>
<td>5.19</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Proclaim: Pronounce</td>
<td>9.55</td>
<td>3.88</td>
<td>0.66</td>
<td>0.40</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Proclaim: Affirm</td>
<td>1.12</td>
<td>0.40</td>
<td>11.32</td>
<td>0.20</td>
<td>3.41</td>
<td></td>
</tr>
<tr>
<td>Proclaim: Concede</td>
<td>3.09</td>
<td>0</td>
<td>0.13</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Entertain</td>
<td>17.97</td>
<td>2.79</td>
<td>5.39</td>
<td>15.58</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Attribute</td>
<td>1.40</td>
<td>2.49</td>
<td>2.89</td>
<td>2.00</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Justify</td>
<td>9.27</td>
<td>6.37</td>
<td>0</td>
<td>3.20</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
The frequency of expressions of agreement in group discussions might seem superficially surprising since in group work situations the learners are supposed to negotiate a common position/solution from presumably different starting points. It becomes less so, if we consider the actual social dynamics of group–work sessions: the solution is to be found within a set time and often one or two group members assume or are given a position of authority, steering the talk while the rest of the group seem satisfied with taking a purely transactional view of the task and seem inclined to accept any solution which completes the task. This makes group–work less promising as a task where learners can ‘train’ their argumentative muscle than it might seem.

Regarding the expand category, attribute has a very low incidence in all tasks. Interestingly, entertain is the most frequent category in presentations and the second most frequent in Role Plays (see extract 9 below). As explained above, in the presentations, students had to present different voices in order to justify their decisions (see extract 8 below):

### Extract 8: Examples of entertain in a presentation on contracts of sale

<table>
<thead>
<tr>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>We can (entertain) insist on delivering what means that we give the seller a certain point of time where he can (entertain) deliver and if he doesn't do we can (entertain) still cancel the contract.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Extract 9 below illustrates the use of entertain in a role play.

### Extract 9: Examples of entertain in a Role Play: Conquistador-Indio

<table>
<thead>
<tr>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1: …we discover that the Sun is not a God, is a star, so…</td>
<td>S2: Maybe (entertain) it’s a star but we believe in that, maybe (entertain) that is what guide us</td>
</tr>
</tbody>
</table>

Finally, the act of justifying one’s position (justify) is not performed very often in the, otherwise appraisal rich, disputive role-plays. This, we suggest, is due to the higher cognitive complexity of proposing reasons for a position compared to denying the interlocutor’s proposition or countering it with alternative propositions. What must not be disregarded either is the actual formulation of the role-play task itself. The instructions of the role plays analysed did not specifically ask the students to justify their points as well as defend them. In this respect it would be interesting to obtain and analyse data from role-plays that follow the format of a debate.

### 6. Discussion and conclusion

This study started from the assumption that in order to reach the curricular goal of learning to evaluate or appraise subject content, CLIL students should actively carry out such evaluative acts, which means they need to use evaluative language in classroom tasks. In terms of conceptual framing, we have argued that the notion of task affords a valuable access point for gaining an integrated (content and language) research perspective on CLIL learning situations. In particular we have shown that the four basic criteria for tasks as defined by Ellis (2003) for TBLT also hold for tasks as implemented by subject educators in CLIL classrooms, but they are of a different nature, and this needs to be taken into consideration in CLIL research.

In the study, five types of activities or typical subject-classroom tasks were compared with regard to students’ use of evaluative language, using the framework of Appraisal for the purpose of analysis. The five task types were Whole—Class Discussion, Group—Work Discussion, Role—Play, Interview and Presentation. Our analysis has shown that clear differences exist among these five task types with regard to the affordances they offer for the learners’ use of appraisal to express evaluative meanings. Contrary to expectations, Group Work did not turn out to be the task-type which showed the highest incidence of evaluative language. In Group Work, students did use more appraisal than in Whole—Class interaction, in line with previous research (e.g. Llinares & Pastrana, 2013), but remained considerably below the frequencies reached in Role Play and Interview. We suggest that the more specific and ultimately more guided nature of Role Play and Interview contribute to this, as does the fact that both involve one-on-one speech situations, while Group-Work happens among more, usually four to six, participants. Interestingly, in the classroom contexts investigated, Group Work did not trigger strong speaker positioning, but showed a conciliatory tone, represented by a high frequency of the appraisal type affirm. We assume that the learners were orienting towards the group task as something that needed completion rather than as an opportunity to negotiate their subject positions in any explicit manner.

In considering the characteristics of the different types of tasks with regard to the evaluative resources used by the students it is particularly enlightening to draw on Goffman’s (1981) participation framework (see Section 2), as the various task types differ considerably in terms of which speaker roles they enable.

Table 4 shows how typical teacher-led Whole—Class discussion has students operate as ‘animators’, supplying the voice to contributions that are predetermined by the teacher’s display questions. This is the typical minimal student response in IRF (initiation-response-follow-up) sequences. Interview and Role Play, on the other hand, provide students with opportunities to play out the whole range of speaker roles. Role Plays give students the freedom to choose their own wording (‘author’) and in part also the content of their contributions (‘principal’), even though the overall topic is, of course, given by the...
specification of the task. Interview, though not a typical classroom task in that it involved a classroom-outsider, i.e. a researcher, is particularly interesting because it throws light on what happens when the clearly circumscribed context of the classroom lesson is opened up. Even though the learners were conversing with a knowledgeable adult (the researcher) who was asking them questions about the didactic unit that had just ended, we claim the students assumed a different subject position than they would have done with their teacher. The interviewer had not taken part in the didactic unit and was not an external examiner either, which apparently empowered the students to assume the position of a ‘knowledgeable other’, who decides on their own content and wording (‘principal’, ‘author’) (cf. Llinares & Morton, 2010).

The analysis we have conducted in this article uses functional categories of appraisal but does not detail the actual wordings. A natural next step to take will be to analyse in more detail the specific linguistic resources which are used by the students in expressing the functions they use. A further dimension which warrants scrutiny is determining which uses of appraisal are related to academic content work (the instructional register) and which for social and task organizational purposes (the regulative register, cf. Christie, 2002). From the point of view of student language use this is particularly relevant in the case of Group Work. A further avenue for future work, one that links our analysis to the core of TBLT-research vs. CLIL methodology in CLIL contexts, do indeed trigger different types of appraisal, thus providing different affordances for the use of evaluative language, especially engagement. This implies that different tasks have different strengths with regard to different aspects of the linguistic resources of evaluation. One consequence to be drawn from this is that CLIL content-teachers could be encouraged to use a wide range of tasks in their lessons, and especially to include Role-Play. Potential objections that teaching of evaluative language is not part of their brief as content-teachers can be dispelled by pointing out that engagement is highly relevant for content learning goals. In fact, evaluation is a notion which is frequently brought up in relation to ‘higher order thinking skills’ (Anderson, Krathwohl, & Bloom, 2001) and is a core component of cognitive discourse functions (Dalton-Puffer, 2013). Present-day curricula in many school-subjects put considerable emphasis on taking students towards being able to handle such higher order skills in order to ensure that (oral) argumentation skills emerge as an overarching educational goal. Such skills involve the use of engagement in fundamental ways.

Beyond the specific focus on evaluative language, the study has shown the interest of identifying the types of gaps and student participation roles generated in different tasks and how this affects the type of language needed for the expression of academic content, which varies not only across academic disciplines or subjects but also across tasks within the same academic domain, as this study has demonstrated.

Acknowledgements

Work on this article was supported by the INTERCLIL Project, financed by the Ministry of Economy and Competitiveness, Spain (FFI2010-20790) and the CONCLIL Project, financed by the Academy of Finland.

We are very thankful to Rachel Whittaker for her insightful comments on earlier versions of this article.

References


Please cite this article in press as: Llinares, A., & Dalton-Puffer, C. The role of different tasks in CLIL students’ use of evaluative language, System (2015). http://dx.doi.org/10.1016/j.system.2015.05.001

Table 4

Students’ possibilities for speaker roles in different task types.

<table>
<thead>
<tr>
<th>Roles</th>
<th>Whole-class</th>
<th>Presentation</th>
<th>Group work</th>
<th>Interview</th>
<th>Role play</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animator</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Author</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Principal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>


Please cite this article in press as: Llinares, A., & Dalton-Puffer, C., The role of different tasks in CLIL students’ use of evaluative language. System (2015), http://dx.doi.org/10.1016/j.system.2015.05.001