CLIL and Foreign Language Learning: A Longitudinal Study in the Basque Country

Yolanda Ruiz de Zarobe
University of the Basque Country (Spain)

Abstract

The aim of this paper is to analyse Content and Language Integrated Learning (CLIL) in the Basque Country in Spain, where many schools are progressively adopting a content-based methodology in the classroom to incorporate English as a third language in bilingual (Basque-Spanish) contexts.

This paper analyses the differences between CLIL and English-as-a-foreign language (EFL) instruction in the case of bilingual students learning English. More specifically, we will describe the differences between both approaches by means of a longitudinal project, where the main differentiating variable between the groups is the type of programme (CLIL vs. EFL). Our results aim to provide an understanding of why CLIL should be further implemented as a mediating approach in language education.

Key words: Content and Language Integrated Learning (CLIL), longitudinal study, bilingual community, speech production; English as a Foreign Language.
Introduction: Content and Language Integrated Learning

Content and Language Integrated Learning (CLIL), a generic umbrella term for bilingual, content-based education,1 has been spreading throughout Europe since the mid-nineties. This has been partly due to a commitment of the European Union to a multilingual Europe. In 1995 it was proposed that EU citizens should be proficient in three European languages, their first language (L1), a language of international communication, and a “personal adoptive language” (i.e. mother tongue + 2 objective)2 to ensure multilingualism as an essential characteristic feature of European identity. Since then, promoting multilingualism has become a cornerstone of the EU for tackling the challenge of creating a more integrated and inclusive society. If we add to this the migration flows from the Third World and the effects of globalisation, we easily understand how there must be a coexistence of different cultural identities and languages, not only where minorities live, but in many other countries and regions in Europe. Multilingualism, then, is viewed as a tool to foster integration and deal with the reality of multiethnic societies.

Due to this commitment, EU language policies have shown a growing need to adopt an educational model to account for the diversity of European programmes and to ensure that everyone can become proficient in several languages. As CLIL appears to comply with EU policies for multilingualism, it has rapidly been adopted as a dual-focused approach by various European professional networks. According to EU institutions, CLIL is regarded as a key instrument to create a multilingual population in Europe. This has been reflected in the adoption of CLIL by EU institutions and the support given to a number of CLIL projects, studies, organisations and experimental initiatives as an integral part of foreign language teaching. Among these are the 2006 Eurydice Report which describes 30 different European CLIL experiences, and the CLIL Consortium, a growing network of experts who work on specific problem-solving and development tasks across the world in the field of bilingual and content-based education.

CLIL, which is generally viewed as a descendent of French immersion education in Canada and of North American bilingual language teaching programmes, covers many different educational models which have been introduced with various degrees of success. Some European regions have had many years of experience with CLIL but others have had almost none. Despite this variation a great deal of commonalities exists between regions in terms of the rationales and goals for implementing CLIL. It is seen as a “flexible system which responds to a very wide range of situational and contextual demands” (Coyle, 2005, 23): it is an educational approach that integrates content and language, either by learning a content subject through the medium of a foreign language or learning a foreign language by studying a content-based subject (Lambert and Tucker, 1972; Genesee, 1987; Marsh, 2000); it creates conditions for naturalistic language learning and, therefore, develops communicative competence (Met, 1998; Marsh and Langé, 1999; Marsh and Marshland, 1999); it involves the learner in using the language of learning for learning and through learning (Coyle, 2000); it is efficient as it integrates language and content into the broad curriculum, which can be especially interesting in bilingual contexts, where two languages already need to be

---

1 Content and Language Integrated Learning (CLIL) and Content-based instruction (CBI) can be considered synonymous. The former is used more frequently in Europe while the latter has gained more popularity in the United States and Canada. Nevertheless, these labels coexist with many others. In www.content-english.org, there are more than 40 labels mentioned to describe the integration of content-learning with language learning, although sometimes reflecting subtly different approaches to the educational practice.

accommodated (Muñoz and Nussbaum, 1997; Muñoz, 2003); and it increases motivation and interest levels (Grabe and Stoller, 1997; Pavesi et al, 2001) by diversifying methods and forms of classroom teaching and learning.

On the other hand, CLIL as it stands still has to face obstacles to be implemented. First, this educational practice requires language teachers who are trained not only in the subject area, but also in how to exploit content-based materials for language. Furthermore, subject teachers may be unwilling to take on the responsibility of teaching through a foreign language, as competence in the target language is a necessity.

“effective multilingual education calls for qualified teachers. Teachers working in multilingual schools where language and content instruction are integrated face the dual challenge of teaching the academic curriculum and a second language at the same times… As a result, they require specialised training in language pedagogy, and especially second language pedagogy, along with the pedagogy required of all teachers who teach academic subjects. Moreover, immersion teachers must be qualified to teach the academic curriculum to second language learners in the second language – this means they must have qualifications to teach in that language. It is not sufficient simply to know a language in order to teach it or to use it or to use it effectively to teach academic subjects. This requires specialised training.” (Genesee, 1998, 256-257)

Unfortunately CLIL teacher-training programmes are scarce and so is the development of content materials and instructional resources.

Finally, there are still few research-based empirical studies that analyse the linguistic outcomes of CLIL in education. ³ Notably, there is a need to evaluate language competence in CLIL, and compare it with traditional instruction to support the benefits of this approach. “Academic research on CLIL remains embryonic and there is now a time-lag between practice and research-based validation” (Langé, 2007, 352). It is this issue that this paper seeks to address, as we will see in the second part of this contribution.

---

³ Recently, and based mainly on research about Canadian immersion programmes, there has been a concern for the role of focus on linguistic form in the development of learners’ language competence in immersion and related models like CLIL (see, for instance, Lyster, 2007, with his “counterbalanced approach” and Pérez-Vidal, 2007)
CLIL in the Basque Country

The foreign language teaching situation in Spain is currently under change. The debate on language learning is articulated under the title Competencias y saberes para la sociedad del siglo XXI (Competencies and knowledge for 21st century society)4 to ensure that every student has a knowledge of two foreign languages and to promote the use of languages in different areas of knowledge. On March 30th, 2005, the Minister of Education and Science presented the new Education Bill (Anteproyecto de la Ley Orgánica de Educación), which, among the most important changes, establishes the advancement of the starting age for foreign language learning. The teaching of a foreign language will begin in the second stage of Educación Infantil (pre-primary education), particularly during the final year, when children are 5 years old.

Nevertheless, research on the general trend towards beginning foreign language learning in primary education5 has not produced very good results (see García Mayo and García Lecumberri, 2003; Ruiz de Zarobe, 2005; Muñoz, 2006, among others). In these studies it is claimed that children’s language proficiency does not benefit much from an early start that does not involve an increase in the number of hours and meaningful exposure to the language. Research suggests that the intensity and timing of exposure may be more important than exposure itself. The proposal of adopting a content-based approach within foreign language learning can thus guarantee more and richer opportunities for using the language in meaningful ways.

The implementation of CLIL6 in Spain, particularly in the Basque Country, is relatively new, with initiatives taken by the Department of Education of the Basque autonomous government mainly in the last five years. Nevertheless, these programmes have benefited from the experience gathered in programmes for the normalisation of Basque as an official language.7 When the Basic Law on the Standardisation of Basque was passed in 1982, three linguistic models were established to ensure that every student had the possibility to learn in Spanish and/or Basque.

(i) Model A: all subjects, apart from the Basque language and literature and modern languages, are taught in Spanish.

(ii) Model B: both Spanish and Basque are used to teach all the subjects

(iii) Model D: all subjects, except Spanish language and literature and modern languages, are taught in Basque.8

---

5 After the Spanish Educational Reform was implemented in 1993, the study of foreign languages began at the age of eight (3rd grade), while traditionally English had begun to be taught in the sixth year of primary school, when pupils are aged eleven. Nevertheless, some bilingual schools began to introduce the study of English in kindergarten, when the pupils are aged four.
6 In Spanish, CLIL is usually translated as AICLE (Aprendizaje Integrado de Contenidos y Lengua Extranjera).
7 Since the 1980s these educational programmes have been set up in different bilingual communities such as the Basque and Catalan autonomous communities of Spain, Wales or France (Artigal, 1993; Sierra, 1994; Baker, 2001). In the case of Spain, the linguistic Standardisation Acts date from 1982 and 1983, although their application started at different times in the various Autonomous Communities concerned.
8 The application of these models is currently under review to reorganise the education system for the purpose of achieving a more qualitative system.
Apart from these models, the Department of Education of the Basque autonomous community has set up an experimental programme to implement CLIL models geared to promote the knowledge and use of a foreign language (English or French). As mentioned before, this approach provides the opportunity to learn a third language without requiring extra time in the curriculum.

With regards to CLIL subjects and the number of hours of participation per week, a number of centres in the Basque Country have begun to teach 1 or 2 non-language subjects in the foreign language. Any non-language subject can be taught in English, but social sciences, sciences, music and technology are the subjects that are usually offered in English. The programme is generally implemented at the levels of compulsory secondary education. Nevertheless, although the general trend has been to offer 1 or 2 non-language subjects in the foreign language, the number of individual institutions offering alternative CLIL curricula, with the integration of more subjects in the foreign language, has increased. This gives us the possibility to research how CLIL is being applied and, more importantly for us here, to analyse the competence attained through CLIL compared to more conventional linguistic programmes.

The study presented in the following section seeks to assess the linguistic competence attained by two groups of students following two different CLIL programmes, and another group enrolled in a traditional English as a Foreign language programme in order to examine the effects of CLIL on linguistic competence in the foreign language. Furthermore, we analyse the longitudinal progression of these three groups to offer a more prolonged perspective on CLIL.
The Study

This study, which is part of a larger longitudinal project aimed at providing support for CLIL, addresses the following question:

Can CLIL lead to foreign language learning faster and in a more articulated way than a more traditional educational model with regards to speech production?

Based on the findings from previous studies in this field of research, we entertain the following hypotheses:

1. Students enrolled on CLIL programmes will outperform students on the non-CLIL programme in all speech production categories.

2. There will be a positive relationship between the amount of content-based instruction and the speech production outcomes. In other words, the more content-based instruction there is, the better the results (more-content-is-better hypothesis).9

3. The participants on the CLIL programmes will obtain better results than the ones on the non-CLIL programme in the longitudinal evaluation of the results.

The question and hypotheses stated above led to the design of a study undertaken in the Basque autonomous community with the aim of exploring linguistic competence in different content-based programmes. In the next section we will explain the sample used for the research.

The Sample

The study was conducted in three school programmes in the Basque Country, selected to include different approaches to CLIL.

Table 1 provides information about the corpus analysed.

<table>
<thead>
<tr>
<th></th>
<th>SECONDARY 3</th>
<th>SECONDARY 4</th>
<th>PRE-UNIVERSITY (Baccalaureate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBER OF STUDENTS</td>
<td>Non-CLIL: 29</td>
<td>Non-CLIL: 18</td>
<td>Non-CLIL: 7</td>
</tr>
<tr>
<td></td>
<td>CLIL 1: 24</td>
<td>CLIL 1: 16</td>
<td>CLIL 1: X</td>
</tr>
<tr>
<td></td>
<td>CLIL2: 36</td>
<td>CLIL2: 17</td>
<td>CLIL 2: 14</td>
</tr>
<tr>
<td>AGE WHEN DATA</td>
<td>14-15</td>
<td>15-16</td>
<td>17-18</td>
</tr>
<tr>
<td>COLLECTION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOURS OF INSTRUCTION</td>
<td>Non-CLIL: 695</td>
<td>Non-CLIL: 792</td>
<td>Non-CLIL: 990</td>
</tr>
<tr>
<td></td>
<td>CLIL 1: 875</td>
<td>CLIL 1: 1120</td>
<td>CLIL 1: XXX</td>
</tr>
<tr>
<td></td>
<td>CLIL2: 910</td>
<td>CLIL2: 1155</td>
<td>CLIL 2: 1453</td>
</tr>
</tbody>
</table>

9 We have paraphrased the ‘more-English-is-better hypothesis’, used in Genesee and Jared (2008: 141).
All the participants started learning English at school when they were eight years old. In the three groups Basque was used as the main language for instruction. Nevertheless, Spanish is the majority language in the community and all the students in the study exhibited native competence in Spanish which did not overlap with the linguistic competence obtained by students whose instruction had been conducted entirely in Spanish. Thus, all of the participants were fully bilingual in Basque and Spanish and were learning English as their third language.

As Table 1 shows, the corpus was divided into three groups on the basis of the English programme. The first group, the non-CLIL group, had received 3 hours of English per week, following a conventional English as a Foreign Language (EFL) programme. These participants did not receive any extra-English classes outside school. The second group, CLIL1, had received instruction in EFL 3 hours per week. When they were 14 they entered a CLIL programme, in which one curricular subject (Social Science) was taught through English for 3 or 4 hours per week. The number of hours provided in Table 1 includes their EFL classes and the CLIL classes. Unfortunately we could not collect information for this group at pre-university level.

The third group, CLIL2, had received, apart from the EFL classes 3 to 4 hours a week, two curricular subjects through English (Social Sciences: 3/4 hours a week and Modern English Literature: 2 hours a week).

**Instruments and procedure**

In order to collect the data, participants were asked to complete a speech production test. The speech production task consisted of elicited narratives of the *Frog, where are you?* story by Mayer (1969), a story that has been used in a large amount of research and with a large variety of languages (Berman and Slobin, 1994; Stromqvist and Verhoever 2004). The oral narrative was elicited through a sequence of 24 pictures.

For the purpose of the analysis on speech production, five categories were used:

- **Pronunciation** (10 points): in this category phonetic accuracy and the communication effect of pronunciation was analysed.
- **Vocabulary** (10 points): this category deals with the selection or words and their usage.
- **Grammar** (10 points): accuracy and the use of different grammatical structures were considered in this scale.
- **Fluency** (10 points): in this category the communicative effect of the oral production and continuity were analysed.
- **Content** (10 points): the adequacy of the content was examined.

Furthermore, we also analysed the number of tokens, types and the type/token ratio in the three groups to look at the kind of productive vocabulary used in the speech production task.

The recordings were completed individually in the schools the participants attended, guided by a trilingual (Basque-Spanish-English) examiner/researcher. All recordings were audio-taped and later transcribed using the CHILDES programme (MacWhinney, 2000, 2008). Statistical analyses were then conducted using the SPSS (Statistical Package for Social Sciences) statistical package.

---

10 The “Frog Story” has been used extensively in different projects undertaken at the University of the Basque Country (Cenoz, 2001; Garcia Mayo and Garcia Lecumberri, 2003; Ruiz de Zarobe, 2003, 2005, among others).
Results and discussion

Results are summarised first by an overall evaluation of speech production in the three school programmes and then, by a description of the results according to the participants’ grade. Figure 1 shows the overall speech production of the whole sample depending on the educational approach.

Figure 1. Overall speech production depending on education approach: Whole sample

According to our first two hypotheses, the students enrolled in CLIL programmes would outstrip the non-CLIL group in every single speech production category analysed and there would be a positive relationship between the amount of content-based instruction and the speech production outcomes. After looking at the results in the overall speech production test apportioned in Graph 1, our first two hypotheses can be born out. The CLIL groups significantly outperform the non-CLIL group in every single one of the scales analysed. The nonparametric tests\(^\text{11}\) show that these differences are significant as regards pronunciation (Chi-square=13.752, \(p<0.001\)), vocabulary (Chi-square=31.359, \(p<0.000\)), grammar (Chi-square=22.031, \(p<0.000\)), fluency (Chi-square=21.122, \(p<0.000\)), and content (Chi-square=17.545, \(p<0.000\)). In all the scales, the results are significantly better in the case of CLIL students. Furthermore, there is a positive relationship between the content-based programme and the results. The students enrolled in the third programme, with more instruction through English, obtained significantly better results than the other groups in all the categories studied. Consequently, our two first hypotheses seem to be confirmed, reflecting more positive outcomes in the CLIL groups.

\(^{11}\) The Kolmogorov-Smirnov Test (K-S test) run for normality testing showed the sample did not have a normal distribution \((p<0.05)\). Thus, as the assumption of normality was not met, the Kruskal-Wallis non-parametric test was used to compare the three independent groups of sampled data.
Our third hypothesis claimed that the CLIL approach would seem to provide better longitudinal results than the non-CLIL traditional approach. Figure 2, 3 and 4 below show the results for each of the three educational programmes depending on the participants’ grade.

Figure 2. Speech production: Secondary 3

As expected, in the third year of secondary education the CLIL groups significantly outstrip the non-CLIL group in every single category: pronunciation (Chi-square=11.394, p<0.003), vocabulary (Chi-square=28.940, p<0.000), grammar (Chi-square=33.109, p<0.000), fluency (Chi-square=28.334, p<0.000), and content (Chi-square=23.802, p<0.000). Nevertheless, as Figure 2 shows, the resulting differences in the case of both CLIL groups are not so important.

Figure 3. Speech production: Secondary 4.

In the fourth year of secondary education the students enrolled in the more intensive CLIL programme (CLIL2) once again scored higher than the other groups, although these differences do not bear out in the other two groups under study. Even so, the differences are still significant in all the categories: pronunciation (Chi-square=7.683, p<0.021), grammar (Chi-square=5.066, p<0.005), fluency (Chi-square=9.649, p<0.08), and content (Chi-square=18.721, p<0.000), with the exception of vocabulary (Chi-square=5.066 p<0.79) where no statistical differences are found.
The participants enrolled in the pre-university grade only belonged to two groups: non-CLIL and CLIL2. The differences between both turned out to be significant in two of the variables: vocabulary (Chi-square=6.561, p<0.10) and grammar (Chi-square=3.899, p<0.048), with the CLIL group scoring higher than the non-CLIL counterpart. In the other three categories, statistically significant differences between both groups were not observed: pronunciation (Chi-square=1.247, p<0.264), fluency (Chi-square=3.603, p<0.058), and content (Chi-square=0.555, p<0.456).

In the case of the longitudinal evaluation of the results depending on the educational programme, the non-parametric Friedman test for correlated values was used. Results suggest that in the non-CLIL group there is a positive relationship between grade and linguistic outcomes with significantly better results on fluency (Chi-square=7.60, p<0.02) and content (Chi-square=7.60, p<0.02). Nevertheless, and contrary to expectations, the CLIL groups do not show a positive relationship between grade and proficiency. In the case of CLIL1,12 there are no significant differences between both groups, while in the case of CLIL2, significant differences are only observed in the case of vocabulary (Chi-square=6.33, p<0.04) in the longitudinal evaluation of the three grades. However, with respect to the linguistic outcomes in the three educational programmes, once again the CLIL groups perform more accurately than the non-CLIL participants, confirming the effectiveness of the CLIL approach on speech production outcomes.

Finally, with regards to the kind of productive vocabulary used in the three groups, the results indicate that, similar to other studies (Jiménez Catalán, R. et al, 2006),13 there are no significant differences in the number of tokens and types used among the three groups. Nevertheless, the type/token ratio is higher in the content-based groups, with significant differences observed (Chi-square=18.247, p<0.000). This suggests that the CLIL groups display more lexical richness and somehow a higher language level than the non-CLIL counterpart.

12 In this group, CLIL1, the Wilcoxon signed rank test for two related simples was used, as results were not obtained at pre-university level.
13 This study also showed how CLIL students presented significantly better results in receptive tasks as opposed to productive tasks, where no significant differences were found between both programmes (Ruiz de Zarobe, 2008). This mismatch between productive and receptive tasks is similar to that encountered in some Canadian immersion programmes (see, for instance, Genesee, 1987; Swain and Lapkin, 1986 for a review).
Conclusions
The aim of this study has been to compare the speech production outcomes in two groups following two different types of educational programmes: CLIL (CLIL1 and CLIL2) vs. non-CLIL. Our results reveal that the former score significantly higher than the latter in every single category of the speech production test. The CLIL participants also achieve a higher type/token ratio in their production, which suggests more lexical richness than their non-CLIL counterpart. Furthermore, there is a positive relationship between the amount of exposure through English and the linguistic outcomes (the more-content-is-better hypothesis). This implies that students with more exposure through English (CLIL2) achieve higher levels of proficiency on the speech production task than students with less exposure through English (CLIL1). Finally, with regards to the longitudinal evaluation of the results, although the CLIL groups score higher than the non-CLIL group in the different grades, there does not seem to be a significant increase in proficiency throughout the academic years.

These results further support other studies which show how CLIL students present significantly better results in receptive tasks such as reading comprehension tasks (Jiménez Catalán and Ruiz de Zarobe, 2007), global tests (e.g. cloze tests), or analyses regarding lexical richness and complexity (Jiménez Catalán, R. et al, 2006). Although more empirically driven research is necessary to confirm the benefits of the CLIL approach in different linguistic domains, this study serves as evidence that CLIL can be more effective than traditional foreign language teaching in promoting proficiency in the foreign language. This is particularly interesting in communities where two languages are already known and need to be accommodated in the curriculum, as is becoming increasingly the case in the European educational landscape.

Acknowledgements
This research was supported by the grant HUM2006-09775-C02-01/FILO from the Spanish Ministry of Science and Technology and the grant IT-202-07 awarded by the Department of Education, University and Research of the Basque Government. Any errors remain obviously my own responsibility.
Bibliographic References


