The Andalusian Bilingual Sections Scheme: Evaluation and Consultancy

Sonia Casal, Pat Moore
Universidad Pablo de Olavide (Spain)

Abstract
With around 8m inhabitants, and comprising eight provinces, the southern Spanish region of Andalusia is the largest in the state. In 2005 the Andalusian government launched the Plan de Fomento del Plurilingüismo (Plan to Promote Plurilingualism, henceforth the Plan), with a view to modernising linguistic policies in line with European initiatives. Among its major objectives, the Plan includes provision for bilingual sections in state schools. It also provides for monitoring schemes to evaluate the implementation of said sections. This article presents an outline of the rationale behind and organisation of the fact-finding component of such an evaluation project, detailing the main research questions and describing the tools which were designed to facilitate information gathering. It closes with the recognition of some of the shortcomings which emerged during implementation and suggestions regarding ways in which these obstacles could be overcome in future projects. As the study here outlined embodies one of the first large-scale multi-faceted CLIL evaluation and consultancy projects in Europe we believe it will be of interest and relevance to research groups involved in similar endeavours.

Keywords: bilingual sections, CLIL, evaluation project, questionnaires, skills-based approach

1. Introduction
This article outlines the elaboration and implementation of the fact-finding component of a large-scale, multi-faceted evaluation and consultancy project designed to provide an overview of the current state of play in the bilingual sections of Andalusia’s state school system. The first part outlines regional linguistic and educational policies – and thus sets the scene; the second part details research objectives, scope and informant profiles; the third part covers the battery of questionnaires employed and the fourth part presents the linguistic evaluation component. The article concludes by acknowledging shortcomings which were revealed during the implementation stage and discusses ways in which these could be remedied.
2. The Plan de Fomento del Plurilingüismo (Plan to Promote Plurilingualism)

European linguistic policies have become effective in Andalusia (southern Spain) through the Plan de Fomento del Plurilingüismo, a document designed by the regional government in order to promote the learning and use of other languages in line with the European 1+2 initiative. Research has underlined the need for this modernisation, given that more than half of the Spanish population (56%) does not speak any language other than their L1 (Eurobarometer, 2006).

The Plan, awarded the European Seal for Innovative Programmes in the Teaching and Learning of Foreign Languages in 2005, comprises five distinct programmes:

1) The bilingual schools programme regulates the allocation of L2 periods and school timetables, the curriculum, technological equipment and study abroad visits for both learners and teachers in bilingual schools.

2) The official language schools programme, covering a network of state-subsidised language schools, offers bilingual teachers the opportunity to update their language competence.

3) The plurilingualism and teachers programme awards grants for teachers in bilingual schools for language courses and school visits abroad and for work exchange schemes.

4) The plurilingualism and society programme intends to plan extra-curricular and complementary activities to foster bilingual section parents' participation and L2 competence.

5) The plurilingualism and cross-cultural programme is implementing a series of training actions related to L2 teaching and attention to cross-cultural diversity.

The implantation of bilingual sections in Andalusia began in 1998 with the setting up of eighteen French and eight German bilingual schools and, since the launching of the Plan, has been expanding steadily: from the twenty six schools running experimental schemes prior to 2004, there was an increase to 250 schools in 2006; in 2007 the figure rose to 403 and at the beginning of the 2008-9 academic year it stood at 518, involving more than 40,000 students. The sections employ an integrated curriculum approach (CLIL) in partnered primary and secondary schools across the region; the predominant L2 is English followed by French and German. The specific objectives of the bilingual schools programme are:

1: The learning of some content areas will be carried out in a language other than the L1. A range of the most widespread languages in the European Union will be encouraged.

2: The methodology implemented both at primary and secondary levels will be based on communication, interaction by means of language immersion and the balanced development of oral and written skills.

3: From a linguistic point of view, the goal is general skills development embracing the L1 as well as the L2, and at later stages an L3. This implies not only an increase in partial linguistic competences in different languages but also the development of a pan-linguistic consciousness.

4: Learners will be confronted with different codes which will lead them to reflect upon linguistic behaviour. This approach should foster a special development of learners’ metacognitive skills and a natural use of languages as distinct from an explicit knowledge of linguistic codes.

5: Students will manipulate language in relation to different areas and academic content, multiplying the contexts wherein they will be able to efficiently use languages linked to academic and professional fields.

6: Students will need to manipulate diverse linguistic codes in order to ‘do things’, developing cognitive flexibility towards analysis and observation of learning processes.

7: From a cultural viewpoint, students in bilingual sections will be in touch with other realities at an early age, being able to draw comparisons with their own surroundings and increasing their interest in different cultures with different traditions, customs, institutions and techniques.
3. Research Objectives, Scope and Informant Profiles
The European Commission's White Paper on Education and Training (1995) recommended that initiatives designed to promote plurilingualism be incorporated into national curricula and has prompted many countries to implement bilingual programmes. Research into bilingual learning has grown exponentially. Several strands may be identified: one looks at the effects on subject learning (e.g., Jäppinen, 2005; Seikkula-Leino, 2007; Van de Craen, Ceuleers and Mondt, 2007; Vollmer, 2008) and another looks at questions relating to language development (e.g., Admiraal, Westhoff and de Bot, 2006; Dalton-Puffer, 2007; Mewald, 2007; Merisuo-Storm, 2007). A third strand, meanwhile, concentrates on situating bilingual learning within the larger picture (e.g., Müller-Schneck, 2005; Serra, 2007; Zydatiβ, 2007).

Leung (2005) warns against the evaluation of bilingual programmes purely through quantitative analysis of learner linguistic gains and in accordance with this the research project here outlined seeks to merge quantitative and qualitative evidence. In doing so it crosses the boundaries between the second and third strands outlined above. The linguistic analysis (see section 5 below) provides empirically obtained statistics relating to language gains and the questionnaires (see section 4 below) contain both open and closed questions and are supplemented by the recorded interviews with co-ordinators.

3.1. Objectives
The study here reported focuses on the following:

1. Competence development: analysis of the competences acquired in L2 learning, including attitudes, procedures and concepts.

2. Teaching organization in bilingual sections: study of school organization as well as bilingual models adopted, characteristics of bilingual sections, intervening areas (including L1, L2 and non-linguistic areas) and type of coordination established.

3. Implementation of bilingual models in the classroom: examination of L2 use in the classroom, typology of activities in relation to skills, materials and assessment models encompassing teachers, teaching assistants and learners.

4. Levels of satisfaction: investigation of the perception of bilingual sections in the different participating sectors: parents, teachers and learners in relation to linguistic anticipation, L2 boost in the curriculum and the establishment of the bilingual schools network.
3.2. Scope and variables

The compilation of data and test administration for the present study began with piloting in May (2007) followed by visits to the different schools from November 2007 to April 2008. Two co-ordinators visited each school. They implemented language tests (reading, writing and listening) and questionnaires with learners and one of them (a native-speaker of the L2) conducted oral evaluation with a randomly selected sub-sample of learners (see below) while the other interviewed the bilingual co-ordinator. Both oral and co-ordinator interviews were digitally recorded.

The study was conducted across Andalusia (Spain). Table 1 shows the total number of participating schools (61 randomly selected English, French and German bilingual schools) and their distribution in the different provinces, languages and sections. At English schools mainstream control groups (MS) were assessed alongside experimental bilingual groups (BL) but the French and German schools were totally bilingual and so no control groups were involved.

Table 1. Participating schools (provinces and languages).

| Schools | TOTAL | Secondary | | | | | Primary | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | English | French | German | | | English | French | German | | | |
| | BL | MS | BL | BL | | BL | MS | BL | BL | | | |
| Almería | 6 | 2 | 2 | 1 | 0 | 2 | 1 | 1 | 0 | | | |
| Córdoba | 6 | 2 | 2 | 1 | 0 | 2 | 1 | 1 | 0 | | | |
| Granada | 6 | 2 | 2 | 1 | 0 | 2 | 2 | 1 | 0 | | | |
| Huelva | 6 | 2 | 2 | 1 | 0 | 2 | 2 | 1 | 0 | | | |
| Jaén | 6 | 2 | 2 | 1 | 0 | 2 | 1 | 1 | 0 | | | |
| Málaga | 12 | 2 | 2 | 2 | 1 | 3 | 0 | 2 | 2 | | | |
| Sevilla | 11 | 2 | 2 | 2 | 1 | 3 | 1 | 2 | 1 | | | |
| TOTAL | 61 | 16 | 16 | 10 | 3 | 17 | 8 | 11 | 4 | | | |

Figure 1 shows the first variable (provinces). Obviously population distribution is not even and so this was factored in and there were more schools involved in the most densely populated provinces and vice versa.

Figure 1.
Figures 2 and 3 present the second variable: primary or secondary education in English, French and German bilingual schools. From the overall total of 61 schools, 32 (52.5%) are primary and 29 (47.5%) are secondary schools. As can be observed below, English schools constitute the majority both in primary and secondary education.

Figures 2 and 3

Figures 4 and 5 display the third variable: urban / rural location. Towns with 20,000 or more inhabitants were classified urban, while those with less than 20,000 were considered rural. The majority of primary schools (79%) are situated in urban zones with 21% in rural areas. In secondary education, the same situation prevails: 66% of schools are urban and 34% rural.

Figures 4 and 5
3.3. Informant profiles

Subjects in the study comprise:

1) Teachers involved in bilingual schools (including co-ordinators and teaching assistants), with an average answer of 8 teachers per school (a minimum of 4 and a maximum of 21).

2) Students in their 4th year of primary and 2nd year of secondary school.

3) The families of these bilingual students.

Table 2 shows the overall number of participants:
Table 2 Subjects involved in the study.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TEACHERS</td>
<td>398</td>
</tr>
<tr>
<td>STUDENTS</td>
<td>1,768</td>
</tr>
<tr>
<td>FAMILIES</td>
<td>972</td>
</tr>
</tbody>
</table>

4. Questionnaires

Opinion questionnaires addressed to teachers, students and families were sent to each school prior to the visit. The bilingual co-ordinator at each school was also interviewed with the aid of a pre-established set of questions. The questionnaires were all completed anonymously.

4.1. Teachers’ questionnaire

This questionnaire was designed to address the variables outlined below. It comprised 8 parts and 46 questions, with a range of question types.

I. Professional details (gender, age, teaching experience, administrative situation, role in the bilingual programme)
II. Language level (official certificate of linguistic competence, certifying organism, language level)
III. L2 use (on the part of teachers, students, skills involved, frequency of use)
IV. Methods (types of activities including cloze tests, true-false and gap-filling exercises, matching activities, translation, dictation, etc. and topics dealt with such as academic content, culture, grammar, etc.)
V. Resources (CDs or tapes, songs, videos or DVDs, computer, games, magazines, etc.)
VI. Competence development (general competences, intercultural skills, motivation and attitudes, learning skills, linguistic, pragmatic and sociolinguistic competence).
VII. Curricular integration
VIII. Programme assessment

4.2. Students’ questionnaires

Different questionnaires were elaborated for primary and secondary students. The questionnaires, consisting of 20 questions for primary students and 32 for secondary students with a range of question types were designed to take around thirty minutes each and were divided into the following sections:

I. Personal details (gender, language in the section, age)
II. Use of L2 information (frequency divided into skills)
III. Method (types of activities in the different subjects)
IV. Information about resources
V. Competences
VI. Programme assessment
4.3. Family questionnaire
This questionnaire, also sent prior to the visit and consisting of 15 questions, comprises three sections:

I. Personal details (gender, age, language in the bilingual section)
II. Socio-educational surrounding (educational level, importance accorded to good marks, number of books in the house, place to study)
III. Knowledge and assessment of the bilingual programme

4.4. Co-ordinator Interview
The interview included questions referring to:

I. Details of the organisation of the bilingual section (languages, departments, teacher stability, teacher development, relationships with other institutions, bilingual model followed)
II. Subjects involved (including non-linguistic areas)
III. SWOT Analysis (strengths, weaknesses, opportunities and threats)

5. The Linguistic Evaluation Component
The linguistic evaluation adopted a skills-based approach. Adhering to national norms, it comprised Comprensión Escrita (Reading); Comprensión Oral (Listening); Producción Escrita (Writing) and Producción Oral (Speaking) each broken down into a set of tasks focusing on sub-skills, including specifically academic skills such as describing, understanding instructions, justifying and evaluating – in line with the intellectual level of informant groups.

5.1. Test Design
Rather than employ ready-made test batteries, the tests were elaborated specifically for the respondents. Test design was shaped by both the Common European Reference Framework (CEFR) and the Andalusian government documents Decretos 148 and 1513 which specified Curricular Objectives for each year of ESO (Obligatory Secondary Education) and Primary education. The two sets of documents proved to be highly compatible.

At the outset, the CEFR facilitated the conceptualisation of competence levels for the two groups: A1 (Breakthrough) at primary and A2 (Waystage) at secondary. Descriptors from the CEFR outlining communicative language activities, strategies and competences at the appropriate level served to inform the elaboration of the tasks in each test. The Reading and Listening featured a variety of exercise types including open and multiple choice cloze, matching, true/false, gap-filling and short open questions, all of which would be familiar to the learners. Wherever possible illustrations were included to facilitate comprehension and examples were given for each task. The Writing tests were organised so that they combined three mini-tasks each and the Speaking tests were broken down into a sequence of overlapping mini-topics with varied interaction patterns. Space precludes the inclusion of detailed information regarding each of the eight tests here but the A2 Writing will serve to exemplify.

Envisaging writing as an interactive activity the CEFR suggests that A2 learners should be able to write “short, simple notes” (2001, 84) and so this was the task chosen. In addition a couple of Speaking descriptors were extrapolated: “Can discuss what to do [and] where to go” and “Can describe places in simple terms” (2001,77 and 59) and Decreto 148 provided a functional angle as it states that learners should be able to make suggestions and even includes target structures: Let’s; Shall we…?; Why don’t we + inf. (2002, 115). The following task was elaborated:

Your teacher wants to organise a class trip but she can’t decide where to go. In the last class she asked for suggestions. Write her a note and give her your idea. Where do you want to go? What can you do there? Why is your idea good for the class?

Not only does this task provide an opportunity for learners to demonstrate linguistic competence, it also incorporates stylistics and the potential for content-related language.
5.2. Implementation of the Tests

Piloting was conducted at two urban Sevillan schools both of which had been among the first to implement English bilingual sections. Results were collated and a couple of small changes were made to the tests to bring them in line with learners’ demonstrated abilities. French and German versions of the test were then elaborated. They adhered to the basic format of the English tests with changes as appropriate – for example Task 3 of the Reading was based on authentic text in all three versions and cultural references were adapted.

On arrival in the classrooms co-ordinators explained the background to the research project and thanked the learners for their help. It was also made very clear that the tests were completely anonymous and would in no way affect grades. It was hoped that this would foster positive affective corollaries. At all times a bilingual environment was idealised and maintained; for example instructions and examples, both oral and printed, were provided in both Spanish and the L2. During all of the tests except Speaking, classroom teachers were present alongside the co-ordinators and learners were encouraged to ask for help if needed (although key lexis was not translated for them).

Secondary learners were given the Reading and Writing tests together – so that they could work at their own pace and early finishers were given a supplementary task to occupy them while their colleagues finished. (The resulting texts were not included in the evaluation project but provide the raw data for a parallel study.) Primary tests were sandwiched with more physical activities such as Brain Gym or games.

All informants were given the Reading, Writing and Listening tests and at each school six learners (three girls and three boys) were randomly chosen from each group (BL/MS) for the Speaking. The primary interviews were individual and the secondary interviews were held in pairs: ♀ + ♂; ♂ + ♂ and ♀ + ♂. The total number of tests administered at both levels and in all three languages was: Comprensión Escrita 1,768; Comprensión Oral 1,695; Producción Escrita 1,741 and Producción Oral 467.

5.3. Evaluation 1: The Linguistic Component

The Reading and Listening tests were purely objective and the Speaking and Writing more subjective. One individual was responsible for the marking of all of the subjective tests in each language. The piloting produced written and recorded materials which were used in a series of benchmarking exercises designed to guarantee inter-rater reliability. Once again the A2 Writing will serve to illustrate:

The A2 writing grade was calculated out of a possible total of twenty five. Fifteen marks were allocated to task fulfilment and ten to language. Task fulfilment covered quantity (5 marks); the extent to which the text resembled a note – with opening and closing salutations, for example, and from the perspective of layout, tone and register (5 marks) and the successful completion of the three mini-tasks – where, what and why (5 marks). The descriptors for Language come straight from the CEFR (with page numbers in brackets) and each is worth 2.5 marks: General Linguistic Range: “Can use basic sentence patterns and communicate with memorised phrases, groups of a few words and formulae about themselves and other people.” (p. 110); Coherence and Cohesion within Overall Written Production: “Can write a series of simple phrases and sentences linked with simple connectors like ‘and’, ‘but’ and ‘because’ (p.61 and p.125); Orthographic Control: “Can write with reasonable phonetic accuracy (but not necessarily fully standard spelling) short words that are in his/her oral vocabulary.” (p. 118) and Grammatical Accuracy: “Uses some simple structures correctly, but still systematically makes basic mistakes – for example tends to mix up tenses and forget to mark agreement; nevertheless it is usually clear what he/she is trying to say.” (p. 114).
5.4. Evaluation 2: The Results

In order to correlate overall results, equivalence needed to be established between numerical results and CEFR levels and the following table was drawn up as a guide. Each of the tests was accorded a mark out of 25.

Table 3. Equivalence between test marks and CEFR levels.

<table>
<thead>
<tr>
<th>Mark</th>
<th>Approximate CEFR equivalent</th>
<th>Global descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>primary</td>
<td>secondary</td>
</tr>
<tr>
<td>25</td>
<td>A2</td>
<td>B1</td>
</tr>
<tr>
<td>22-24</td>
<td>A1+</td>
<td>A2+</td>
</tr>
<tr>
<td>18-21</td>
<td>A1(+)</td>
<td>A2(+)</td>
</tr>
<tr>
<td>14-17</td>
<td>A1</td>
<td>A2</td>
</tr>
<tr>
<td>10-13</td>
<td>Pre-A1</td>
<td>A1+?</td>
</tr>
<tr>
<td>6-9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Several points need to be made here: in the first instance it should be emphasised that this was not a test to be passed or failed – it was purely evaluative. In order to fine-tune the evaluation, an intermediary level was introduced between standard CEFR levels – that of (+) (see CEFR, 2001, 32-3). It is difficult to confidently attribute levels to learners who scored below the key-stage levels (A1/A2) although it is probably safe to say that those who obtained 6-13 in the A1 test are pre A1. The distinction between 0-5 and 6-9 is designed to distinguish between ‘cannot’ and ‘will not’ (Van Lier, 1989).
6. Closing remarks

The implementation stage of the project naturally led us to revise certain aspects and in closing we would like to reflect upon some of these modifications in the hope that they will help other researchers engaged in similar projects:

- It was the first time that many of the students had participated in a similar project and thus the first time that they had had to deal with detailed academic type questionnaires. This was particularly problematic with the primary learners. Over the course of the school visits the co-ordinators began to conduct the primary questionnaire as a whole class activity reading out each question and clarifying comprehension before asking learners to complete that bit.
- Even though schools often wanted to put them together, we found that it was best to separate experimental and control groups when possible as they tended to work at different paces.
- Language assistants’ questionnaires had been written in Spanish but many of the assistants are overseas students (on year abroad schemes) and comments from co-ordinators suggested that it would have been better to produce the questionnaires in the appropriate L2s.
- The parents’ questionnaires focused on educational level as an indicator of social class but on reflection perhaps they should have included secondary indices for cross-referencing and thus greater reliability.
- Many of the school co-ordinators were uncomfortable with the idea of being recorded even though we assured them that it was primarily to avoid the interviewer having to take detailed notes during the interview.
- At the outset learner numbers were calculated and a corresponding number of family questionnaires were sent to each school. As the visits progressed, however, it became clear that this was not necessary and we settled on 15 family questionnaires per school. This meant that it was more likely that all questionnaires would be returned although it also meant that teachers would probably distribute questionnaires to those learners who presented a greater chance of return – and we cannot rule out the possibility that this might have influenced results.
- In the linguistic evaluation a mark of 25 was interpreted as signifying a performance at (or above) the next level of competence (see table 3, above). On reflection, however, this serves more for the subjective tests (Writing and Speaking) than for the objective tests (Reading and Listening) where it is possible that a learner who obtained very high marks but not 100% might still be at the next level. It is extremely difficult to accurately assess someone’s level of linguistic competence according to a 25 item test. For this reason we will be cautious in future publications in attributing CEFR levels on the basis of Reading/Listening performance.
References


