Integrating a corpus of classroom discourse in language teacher education: the case of discourse markers

CAROLINA P. AMADOR MORENO
Universidad de Extremadura, Departamento de Filologías Inglesa y Alemana
Facultad de Filosofía y Letras, Avda. Universidad s/n 10071 Cáceres, Spain
(email: camador@unex.es)

ANGELA CHAMBERS
Centre for Applied Language Studies, Department of Languages and Cultural Studies
University of Limerick, Limerick, Ireland
(email: angela.chambers@ul.ie)

STÉPHANIE O’RIORDAN
IRCHSS Government of Ireland Scholar, Centre for Applied Language Studies
Department of Languages and Cultural Studies, University of Limerick, Ireland
(email: stephanie.oriordan@ul.ie)

Abstract

While language teacher education programmes and language syllabi in secondary education encourage the use of the target language in the classroom, resources to support teachers in this endeavour, such as books with useful phrases, do not state that the examples they provide are corpus-based, i.e. drawn from actual language use rather than invented phrases. This paper investigates whether consultation of a corpus of classroom discourse can be of benefit in language teacher education. The paper describes a project involving the creation of corpora of classroom discourse in French and Spanish, and the use of these corpora with student teachers. After setting the research in the context of corpora and classroom interaction, it examines issues such as the content of the corpora, the type of consultation (direct or mediated by the teacher), and the student teachers’ evaluation of the activity. Special attention is paid to one particular aspect of classroom interaction, discourse markers.

1 Introduction

Within the steadily increasing number of publications devoted to the application of corpora in language learning, language teacher education (LTE) has received little attention to date, with a few exceptions (Coniam, 1997; Farr, 2004; Hunston, 1995;
O’Keeffe & Farr, 2003; Renouf, 1997). When one considers the emphasis on teaching through the target language in national documentation and academic publications concerning language teaching (Macaro, 1997; Guest & Pachler, 2001), there is clearly scope for corpora of classroom interaction to be made available to student teachers as part of their preparation for this activity. The corpora would serve not as prescriptive models to be followed, but as stimuli for the student teachers’ reflection on their future practice, using small corpora as Tribble (1997:109) recommends, namely as “a collection of expert performances (Bazerman, 1994:131) in genres which have relevance to the needs and interests of the learners”. There is of course a substantial body of research in classroom interaction, which Sinclair and Coulthard (1975) trace back to the late 1940s in their seminal publication on the subject. The application of corpora in LTE, however, remains largely unexplored, despite recommendations by researchers such as Conrad (2002) and Sinclair (2004) that familiarising student teachers with corpus consultation and analysis is a prerequisite for establishing the use of corpora in primary and secondary education.

This study investigates this area by asking the question, “Can consultation of a corpus of classroom discourse be of benefit in language teacher education?” After initially setting the research in the context of corpora and classroom interaction, we describe a project involving the consultation of corpora of classroom discourse in French and Spanish by student teachers in the University of Limerick, Ireland. Issues examined include the content of the corpora, the type of consultation (direct or mediated by the teacher), and the student teachers’ evaluation of the activity. While a number of aspects of classroom interaction were examined in the course, it is beyond the scope of this study to provide a detailed analysis of them all, and we shall therefore focus on one, namely discourse markers. This will enable us to determine to what extent the student teachers benefited from the activity, while also raising questions relating to the integration of this resource in the language-learning environment.

2 The classroom as a discourse community: the role of the corpus

Researchers in discourse analysis (Tannen, 1981; Swales, 1990; Schiffrin et al. 2001) emphasise the need to master the different types of behaviours and language use which are acceptable within a specific discourse community. The student teacher’s aim is thus, through observation of classes and consultation of corpus data, to acquire the linguistic features and discursive practices which characterise the discourse community of language teachers and learners in the target language. This is, however, no easy task, for a number of reasons. Firstly, as Macaro (1997:55-57) observes, the classroom is not only a distinct community, but a strange one, with the language classroom winning “the prize for being the strangest”. He lists a number of features which highlight the teacher’s dominant role:

[The language classroom] is a place:

• where a learner is asked to operate in a state of almost total dependence on the teacher; […]

• where the topic of discourse, the linguistic interaction, the pace of delivery, the intensity of language and action, the management of the physical environment, the establishment of social norms and of relationships are all dominated by one
member (the teacher) speaking a language both foreign to him/her and to the pupils.

In addition, the student teacher’s task in developing the competence to teach through the target language is complex for another reason. The roles of teacher and learner are constantly changing, with the teacher’s status moving from that of transmitter of knowledge to passive learners, to that of tutor guiding the pupils to their learning, and coach facilitating the pupils in their autonomous learning process. The language practices which characterise each of these approaches will of necessity vary considerably, with the result that the list of useful expressions provided in the teachers’ resource book (such as Macdonald, 1993), the observed class, or the transcribed corpus data can serve only as an example, as the basis for the student teacher’s reflection on the approach which s/he wishes to develop, and the type of language use which is most appropriate for the implementation of that approach.

Outside LTE there is a growing body of research and empirical studies into the use of corpora in the language classroom, starting in the early 1990s, and these are relevant to LTE in a number of ways. In their quantitative studies on the benefits of corpus consultation in the contexts of the acquisition of lexis and grammar respectively, Cobb (1997:303) and Gaskell & Cobb (2004:304) note that in the language learner’s daily life there is not enough time to encounter the multiple occurrences of certain terms, structures and phrases to internalise them. For Cobb (1997:303) the concordance printout, which gives all the occurrences of a given word or sequence of words in a text, makes possible what he terms “multicontextual learning”. For Widdowson (2000:7; 2004:71), on the other hand, the concordance provides decontextualised language, more aptly named co-text, as the context is not immediately available. Pedagogic mediation by the teacher is necessary in his view for corpus data to be made meaningful to learners (Widdowson, 2003). Taking this concept of mediation as her starting point, Braun (2005) uses video recordings of native speakers of English with her learners, supplementing the study of individual texts with concordances of the other occurrences of relevant items in her corpus. As we shall see, the methodology of Braun’s experiment, although not set in the context of LTE, seems most appropriate for the case of student teachers.

3 Methodology and corpus data

Teaching through the target language was integrated into the programme of study in Limerick (a one-year postgraduate qualification for teachers of two languages) in a number of ways. In a module devoted to language pedagogy it was the focus of one lecture and one seminar. It was also addressed in two modules, one in each semester, aiming to maintain and develop the student teachers’ language skills and knowledge of

---

1 The authors are aware that ‘learner’ is more commonly used, but ‘pupil’ is used throughout this article to avoid confusion as both the student teachers and pupils are learners.

2 In the Academic Year 2004–2005 there were seven student teachers in the French class and three in the Spanish class, 8 students in all as two studied French and Spanish.
the target cultures. It was within these modules that the corpus-based study of classroom interaction took place, although a very limited amount of time was available, approximately three hours in each language over 4–5 weeks.Three aspects of classroom interaction were selected. Firstly, discourse markers were chosen, partly because of their importance in relation to the teacher’s role as the person controlling and organising the discourse, and partly because they formed a relatively easy starting point for the student teachers. The other two aspects of the discourse studied were corrective feedback and learner discourse. In addition to an essay which the student teachers were required to write on teaching through the target language, their reactions to the corpus-based study were also examined through a questionnaire completed after the classes studying the corpus data, and a semi-structured group interview lasting approximately one hour.

The corpus data used here form part of corpora which are being created, consisting of audio recordings of French and Spanish classes involving pupils over 16 years old. The corpora will include recordings of classes in the following categories:

- Nonnative speaker teachers of French/Spanish teaching the language to nonnative speaker pupils in Ireland
- Native speakers of French/Spanish teaching the language to nonnative speaker pupils in Ireland
- Native speakers of French/Spanish teaching the language to native speaker pupils in a country where the language is spoken.

The first category is undoubtedly the most authentic, corresponding closely to the situation of the vast majority of the student teachers, who are nonnative speakers of the languages they are teaching. The second category provides them with data on the language use of native speaker teachers. The third category is included primarily as it provides the student teachers with examples of the language use of native speaker pupils, which may be of use to them when offering guidance to their pupils on using the target language in the classroom. Although only two classes had been recorded in each language in the 2004-2005 Academic Year, corresponding to the second and third categories above, this still formed a useful resource for teaching and research, as the corpora are not considered to be representative in any way of classroom interaction in general. Rather they are intended simply as examples of such interaction, which can serve as stimuli for the student teachers in their reflection and preparation for teaching through the target language. Aston (2001:75) notes the importance of this distinction when small corpora are used for pedagogical purposes.

In the publications on the use of corpora as a resource in language learning cited in this article, the researchers are reporting on their own experience of using corpora with language learners in higher education, and it is clear that a certain amount of time was available for training in corpus consultation and analysis. What characterises many LTE programmes, however, is their intensive nature, combining as they do several disciplines in education, alongside language pedagogy, language studies to maintain and develop language skills, and a substantial period of teaching practice. It is therefore not surprising that it proved impossible to allocate time to train the student teachers to consult and analyse the corpus data using concordancing software. Thus, while they were given a brief introduction to corpus linguistics, corpora and concordancing
software, their activity consisted, as we shall see, in analysing extracts, consulting concordances prepared by the teacher and, for the purposes of their assessment, requesting concordances of the terms they had chosen to study. While this is a very limited form of corpus consultation, it may well represent the only form possible in many LTE programmes for primary and secondary school teachers, given the competing demands of other disciplines. Indeed, the only foreseeable solution would be for corpus consultation and analysis to be included in undergraduate degree programmes in modern languages. Despite its limitations, this type of corpus consultation, as McEnery and Wilson (1997:6) point out, nonetheless demands active participation from the learner, in this case analysing and interpreting the data and reflecting on their relevance. Two types of corpus analysis are involved, firstly corpus-based analysis relying on concordancing software to provide quantitative data instantly on the occurrences of individual items. This is then complemented by a discourse-based approach, what Henry and Roseberry (2001:100) describe as “whole-corpus reading and study”, in which longer extracts of the corpus are studied, similar to the methodology employed by Braun (2005).

4 Analysing the data: the example of discourse markers

As the definition and analysis of the discourse markers (DMs) in the recorded classes was necessary before the student teachers’ study of them could be carried out, this section will provide a brief definition of DMs and an examination of the occurrences of them in the four classes.

Despite the large body of research dealing with the group of expressions commonly referred to as discourse markers, difficulties arise in the discussion of these linguistic expressions when it comes not only to issues of classification and function⁴, but also terminology. Brinton (1996:29) lists more than twenty terms including, for example, the terms “sentence connectives” (Halliday & Hasan, 1976), “semantic conjuncts” (Quirk et al., 1985), and “pragmatic markers” (Fraser, 1990). Discourse marker is employed here in the sense used by Schiffrin, who describes these expressions as “sequentially dependent elements that bracket units of talk” (1987:31). By discourse markers, then, we mean expressions such as well and but in English, alors⁴ and bon in French, and bueno, venga and claro in Spanish, which are used in conversation to indicate how a message should be interpreted. These linguistic expressions are syntactically heterogeneous (Schiffrin, 2001:57), that is, “comprised of members of word classes as varied as conjunctions (e.g. and, but), interjections (oh), adverbs (now, then), and lexicalised phrases (y’know, I mean)”. DMs⁵ have been studied in a variety of languages.

---

³ For an account of studies in various fields of linguistics see Schiffrin (2001: 54–55).
⁴ Literal English translations of the DMs studied in the article are provided in Tables 1 and 2.
⁵ Space/time does not allow us to concentrate on the large group of interjections such as ah, oh, aha, in French and Spanish, which Yngve (1970) considers back channels, Schegloff (1981) turn-continuers and Fraser (1999) pause markers. We have also excluded response tokens (McCarthy & Carter, 2000; McCarthy, 2002), as well as items such as d’accord?, OK? ¿verdad, ¡vale?, ¡oui?, ¡non?, ¿sí?, ¿no? etc. These expressions, followed by a question mark, are used to check understanding and therefore have an important classroom interaction, but are not strictly DMs (see Ortega, 1998).
including French (Cadiot et al., 1985; Heisler, 1996; Mosegaard Hansen, 1997; Maury-Rouan, 2001) and Spanish (Koike, 1996; Martín Zorraquino, 1998; Portolés, 1998; Travis, 2005). According to Fraser (1999:938), DMs “impose a relationship between some aspect of the discourse segment they are part of, call it S2, and some aspect of a prior discourse segment, call it S1. In other words, they function like a two-place relation, one argument lying in the segment they introduce, the other lying in the prior discourse.” The relationship between DMs and context is often referred to in specialised research dealing with these expressions. Although it is still subject to debate, there seems to be general agreement when it comes to describing DMs as context-dependent: thus, in Schiffrin’s view, markers “can gain their function through discourse” (2001:60), and their more specific interpretation, as Fraser points out, is “negotiated by the context, both linguistic and conceptual” (1999:950).

The quantitative analysis of the DMs described below was carried out by the lecturers, who are also the authors of this study, but, as we shall see, the results were not initially presented to the student teachers, who were asked to discover for themselves how these words were used. In the quantitative analysis it immediately became clear that not all

<table>
<thead>
<tr>
<th>Item</th>
<th>Class 1 (Ireland) 55 mins</th>
<th></th>
<th>Class 2 (France) 40 mins</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Occurrences</td>
<td>DMs</td>
<td>Occurrences</td>
<td>DMs</td>
</tr>
<tr>
<td>OK [OK]</td>
<td>88</td>
<td>48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alors [so/ then]</td>
<td>29</td>
<td>23</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Donc [so/ therefore]</td>
<td>21</td>
<td>14</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>Bon [good/ well]</td>
<td>20</td>
<td>16</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Allez [come on]</td>
<td>18</td>
<td>9</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>D’accord [all right]</td>
<td>10</td>
<td>5</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>Bien [well]</td>
<td>35</td>
<td>1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>C’est-à-dire [that is to say]</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Enfin/’fin [I mean]</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

6 DMs are printed in bold in examples throughout this article.
occurrences of these items were DMs, but that they could also fulfil quite different functions. In the extract below, for example, *vamos a ver*, which elsewhere occurs as a DM, literally means “let us see”.

Prof. ...porque:: | qué trabajo nos ha costado [ríen todos] **Bueno, venga**, vamos a ver el cinco. En la misma página, página diecisiete...

[Teacher: ...because what a lot of work that was [they all laugh] **Ok, come on**, let's see number five. On the same page, page seventeen...]

Tables 1 and 2 provide details on the total numbers of occurrences of each item and the number of times they are used as DMs.

In a small number of cases it was not easy to decide whether an occurrence of an item was a DM or fulfilled some other function. In the extract below, for example, it is unclear whether *donc* is used as a DM, clarifying what has just been said or whether it expresses a consequence of what has just been said, in which case it is not a DM.
It was hesitantly decided after listening to the tape several times that donc was a discourse marker rather than expressing consequence and fulfilling what Schiffrin (2001:58) describes as its core meaning. Table 1 reveals that the French native speaker teaching nonnative speaker pupils in Ireland uses a lot more DMs than her counterpart in France, and uses the English borrowing OK 87 times in all (one occurrence is by a pupil) and 8 times as a DM, while it is not used at all in France. There are two possible explanations for these differences. Firstly, the much higher number of DMs in the class in Ireland could be explained by the fact that the pupils are nonnative speakers and thus require more direction from the teacher. Secondly, the occurrences of OK may have been influenced by the English-speaking context. These hypotheses would, of course, have to be tested in a much larger corpus. Also clear from Tables 1 and 2 is the fact that there is a closer correlation between the occurrences of the items and their use as DMs in Spanish than in French, but once again, as only four classes are involved here, it is impossible to generalise on the basis of these data. The variation nonetheless provides an interesting resource for the student teachers, provided that they are advised that generalisations cannot be drawn from such a small collection of data.

Table 3. Discourse markers used by teachers and pupils in the French classes
(*a) Used during pupil to pupil interaction in short turns of speech)

<table>
<thead>
<tr>
<th>Item (as DM)</th>
<th>Teacher Class 1 (Irl)</th>
<th>Teacher Class 2 (Fr)</th>
<th>Pupils Class 1 (Irl)</th>
<th>Pupils Class 2 (Fr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alors</td>
<td>23</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Donc</td>
<td>14</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bon</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>2(a)</td>
</tr>
<tr>
<td>Allez</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D’accord</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bien</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C’est-à-dire</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Enfin (‘fin)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Once the DMs were identified, they were divided into two categories, according to whether they were used by teachers or pupils (see Tables 3 and 4).

Two findings emerge clearly from these tables. Firstly, the fact that the pupils use hardly any discourse markers confirms the comments by Macaro (1997:56), De Fina (1997:339) and others concerning the dominant role of the teacher in the classroom. The second finding could provide encouragement to the student teachers, in that the tables reveal that the four native speaker teachers use a relatively limited number of DMs (9, 4, 10 and 8), thus implying that the acquisition of their use is not an impossibly challenging task.

The functions of the discourse markers are summarised in Table 5. Examples of these functions are given below.7

### Table 4. Discourse markers used by teachers and pupils in the Spanish classes

<table>
<thead>
<tr>
<th>Item (as DM)</th>
<th>Teacher</th>
<th>Class 1 (Irl)</th>
<th>Teacher</th>
<th>Class 2 (Sp)</th>
<th>Pupils</th>
<th>Class 1 (Irl)</th>
<th>Pupils</th>
<th>Class 2 (Sp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vale</td>
<td>23</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venga</td>
<td>21</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claro</td>
<td>20</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entonces</td>
<td>17</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bueno</td>
<td>13</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vamos a ver</td>
<td>2</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O sea</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Va</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mira</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anda</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oye</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Once the DMs were identified, they were divided into two categories, according to whether they were used by teachers or pupils (see Tables 3 and 4).

Two findings emerge clearly from these tables. Firstly, the fact that the pupils use hardly any discourse markers confirms the comments by Macaro (1997:56), De Fina (1997:339) and others concerning the dominant role of the teacher in the classroom. The second finding could provide encouragement to the student teachers, in that the tables reveal that the four native speaker teachers use a relatively limited number of DMs (9, 4, 10 and 8), thus implying that the acquisition of their use is not an impossibly challenging task.

The functions of the discourse markers are summarised in Table 5. Examples of these functions are given below.7

1. **Introduction of new topic, activity, or question.**
   a) **French**
   Prof: Inquiet?
   Élève: Avoir soucis.

---

7 An English translation of these examples is provided in Appendix 1.
Table 5. Functions of discourse markers

<table>
<thead>
<tr>
<th>Function</th>
<th>Discourse markers</th>
</tr>
</thead>
</table>
| 1. Introduction of new topic, activity, or question. | French: alors, bon, OK, bien, donc, allez  
Spanish: vamos a ver, bueno, vale |
| 2. To call the pupils’ or teacher’s attention. | French: d’accord, bon  
Spanish: vamos a ver, a ver, mira, oye |
| 3. To recap what has been said or offer clarification. | French: alors, donc, d’accord  
Spanish: o sea, vamos a ver, entonces |
| 4. To motivate or encourage the pupils. | French: allez, alors  
Spanish: anda, venga, va |
| 5. To correct oneself or rephrase what has been said. | French: c’est-à-dire, enfin, ‘fin  
Spanish: o sea, bueno |

Prof: Oui être inquiet ça veut dire avoir des soucis. Très bien oui. Être inquiet, avoir des soucis. OK? **Bon.** D’autres questions? .. OK. Alors dernière question c’était …

**b) Spanish**
Prof: ... minutos, se nos ha ido la clase volando ac (Ininteligible)
**Vamos a ver,** otra cosa ¿os estáis leyendo el libro de ...  

2. To call the pupils’ or teacher’s attention

**a) French**
Prof: Ça peut être aussi quelqu’un de  
Élève: [tousse]  
Prof: de froid qui est indifférente.  
Élève: Ouais.  
Prof: On peut dire [des élèves bavardent] **D’accord.** Donc [des élèves bavardent] vous avez la démarche …

**b) Spanish**
Prof: Pero necesitas poner “primero”: **A ver, mira,** tú tienes aquí la frase [señala la pizarra].

3. To recap what has been said or offer clarification

**a) French**  
Élève: Avec une barbe.  
Prof: Avec une barbe. Qui avec une barbe?  
Élève: Un homme.

b) Spanish
Alumno 2: y cuatro
Alumno 3: y cuatro y cinco, sí.
Prof: O sea, que esos son de la página anterior. Página dieci

4. To motivate or encourage the pupils.

a) French

b) Spanish
Alumno 3: ...tengo que cambiarlos todos ahora que he visto la pregunta.
Prof: Venga, ¡ánimo, pues cámbialo y dilo bien.

5. To correct oneself or rephrase what has been said

a) French
Prof: Quelqu’un qui a des pouvoirs, c’est quelqu’un qui est puissant. Ouais. C’est-à-dire c’est quelqu’un qui peut faire ce qu’il veut. OK. Donc il peut faire la plupart des choses qu’il veut faire. OK? […] Pourquoi vous riez c’est vrai. C’est normal. ‘fin [enfin] c’est pas normal mais c’est c’est la bonne traduction ouais [Petits rires dans la classe].

b) Spanish
Prof: ...hemos dicho: “Brasil, país, primero” que lo tenéis en letra, o sea, que lo tenéis en número, no lo tenéis en número...

These and other examples from the corpus illustrate a number of aspects of the nature and function of DMs. Firstly it is clear that they can be multifunctional. An item such as alors in French can have at least three different functions, acting as a transition between two segments (Category 1), recapping what has been said (Category 2), and motivating pupils (category 4). Vamos a ver can be used to introduce a topic, call attention, or offer clarification. Only very few items have only one function. Secondly, while at times the function is clear, it can occasionally be difficult to classify the function of the DM. In extract 4a) above, for example, alors on y va can be interpreted as a way of calling pupils’attention (Category 2) or of motivating the pupils (Category 4). In both languages the presence of clusters of DMs is evident, such as OK. Alors (5 occurrences) and OK. Donc (3 occurrences) in French, and Bueno. Venga (3 occurrences) and Bueno. A ver (3 occurrences) in Spanish. Despite this element of complexity, however, the analysis of these four classes reveals that, just as a limited number of DMs are used, so also the number of functions which they fulfil is limited, once again providing data for the student teachers which imply an
achievable goal. A more detailed analysis of the use of DMs in these classes is beyond the scope of this study, which focuses only on the aspects which were included in the classes with the student teachers.

5 Classroom application, assessment and student teachers’ evaluation

5.1 Classroom application

An inductive approach was used in the classes devoted to DMs. Taking the words *alors* in French and *bueno* in Spanish as starting points, the student teachers were given examples of the occurrences in extracts from the corpus (sufficiently long to make the context clear, not concordance lines) and asked to examine how they were used. This was done both in class and in the student teachers’ own time. They discovered that these items often did not have their core grammatical meaning and were thus introduced to DMs as a semantic-pragmatic concept and not a grammatical one. Concordance printouts were then used to illustrate the various functions of the words. The following extract from a concordance of *alors*, for example, reveals a variety of different uses, including uses as DMs and with other grammatical and pragmatic functions.


[ … lung specialist yes? Ok. So. Eh … Question number two … ruins » it is in … the subjunctive of course. So yes, it looks like the present. But … [the pupils agree] Yes ? *There we go then*. Question number five] … well no we are doing tape work tomorrow so you bring your books ! … no ? … Well I’ll give it to you then. *La lutte* can be replaced … Teacher : Yeah. In French then. Pupil1 : It’s something…]

After listening to excerpts from the tape-recordings, the student teachers became aware of the link between the different linguistic and pragmatic functions of *alors*, with differences in intonation and pitch expressing encouragement, exasperation, moving to a new topic, and other functions. In the Spanish class *bueno* was analysed, revealing that in addition to its core meaning of “good” it could be used to correct or reformulate what the speaker has said, to introduce a conclusion, to indicate agreement (replacing *sí*), to structure information, change or introduce a new topic, etc. In the very first line of the Spanish class recorded in Ireland, for example, the teacher says: “*Bueno, entonces* ¿qué teníamos para hoy?…” [*Ok then*, what did we have for today?...] Here, as the student teachers were able to point out, the use of *bueno* indicates that the initial greetings are finished and that she wants to start the class. Other DMs were also studied in both languages, using a combination of longer extracts and concordance printouts.
5.2 Assessment

As part of the assessment of the module on Language Studies, of which these classes formed part, student teachers were asked to write a 400–500 word essay in French/Spanish on teaching through the target language, in which they were asked to:

- identify an aspect of language use in the corpus which they considered relevant to them as student teachers;
- analyse the use of that aspect in one or two short extracts from the transcripts or through printouts of concordances (provided to them on demand);
- evaluate to what extent the recordings and the transcripts were of use to them in preparing to teach through the target language.

They were provided with access to the recording, the transcription, and a frequency list. At their request concordances were emailed to them of the items which they had selected, presented in the order of their occurrence in the corpus. In comparison to other studies of learners using corpora (see, for example, Bernardini, 2000; Chambers and O’Sullivan, 2004; Kennedy and Miceli, 2001), these concordance printouts are very limited, in that the student does not have the opportunity to re-sort the concordances according to the words which precede or follow the search word. However, as explained earlier, the intensive nature of the programme of study did not allow time for training in corpus consultation and analysis. As the numbers of occurrences were limited, ranging from 10 to 108 (many were in the 21–38 range), this did not present an insuperable obstacle, as the student teachers could sort the occurrences themselves using the Word version of the concordance, and consult the corpus as necessary.

Of the ten essays submitted, four were devoted to DMs, one in French and three in Spanish. The student teacher studying French studied *d’accord, OK, bon* and *bien*. Her first finding concerns the high number of occurrences of *OK* (88, of which 87 were uttered by the teacher). This triggers surprise, as she had been taught to use *d’accord* and not *OK*, which was considered an Anglicism. She then compares the use of *OK* and *d’accord* in the corpus, observing that *OK* is used exclusively in the French class recorded in Ireland. She suggests that the native speaker French teacher may have been influenced by her use of English outside class, and also within the class when having recourse to L1 (English). Indeed, as the student notes, *OK* precedes explanations in English or occurs within them 25 times. The student then identifies the main function of *OK* and *d’accord* in the corpus, noting that both are used to check pupils’ understanding when followed by a rising intonation (question mark in the transcription), and also as DMs and response tokens, to validate an answer or contribution by a pupil. She also notes that *OK* can be an interjection. In her study of *bon* and *bien*, the student identifies the main uses of both words. First she notes their grammatical use as adjective and adverb, where they cannot be considered as DMs. She notices that *bien* is often used in clusters such as *très bien* [very good] and *super bien* [excellent] and is mostly used as positive reinforcement. She also notes that *bon* is used only by the teacher in Ireland, but by both teacher and pupils in France, thus suggesting that there is a need for teachers

---

8 As only one student in the cohort was male, all are referred to as ‘she’ to preserve anonymity.
to pass on their use of ‘little words’ (Traverso, 1999; Maury-Rouan, 2001) such as bon to their pupils.

Of the three student teachers who wrote essays on DMs in Spanish, only one will be included here as an example. One student concentrates on the use of venga, which she defines as “un marcador de control de contacto”, following Portolés’s (1998) classification. Analysing the class recorded in Ireland, she notes that the first use of venga, employed by the teacher three times, is to signal her intention to start the lesson. She indicates that venga, apart from being used to encourage pupils’ participation, indicates continuation to the next exercise. She also notes that in some instances this item has a note of impatience which seems to indicate that the teacher wants to speed up the class, which the student teacher considers to have a negative effect. One could alternatively argue that it adds dynamism to the class, allowing for more interaction between teacher and pupils. The other two students each analysed a number of frequently occurring DMs: bueno, vale, venga, a ver, and vamos a ver in one case; and bueno, vamos a ver, o sea, entonces and claro in the other. Although the level of analysis in these essays varied considerably, it was nonetheless evident that the student teachers had understood and benefited from the study of DMs in the corpus as a resource for analysis and reflection.

5.3 Student teachers’ evaluation

In addition to the assessment, a semi-structured interview was held after the classes, and the student teachers were asked to complete a questionnaire evaluating these classes. Seven of the eight students attended the semi-structured interview, but only five completed the questionnaire, including five of the seven students of French and two of the three studying Spanish. The interview was based on seven questions, which were not supplied in advance. First, they were asked for their reactions to the types of recordings made available to them, namely native speaker teachers teaching classes in Ireland and in the target country. All the students appreciated having these recordings as they involved examples of native speaker teacher and pupil discourse, but they unanimously agreed that it would be very useful to include “the Irish setting”, namely English-speaking teachers teaching through the target language in Ireland. It is important to note, however, that one student teacher commented that the recording of the native speaker teacher in Spain was particularly useful for examining the use of DMs. The second question concerned the acceptability of audio recordings, but even before the question was asked the student teachers introduced the topic, once again unanimously agreeing that video recordings would be much better resources.

Claire: To listen to a tape I think you lose about 50% of what’s going on in there because you do not see the interaction.

Helen: More than 50%.

As we shall see, this strong preference for video recordings will also emerge from the

9 Aliases are used to preserve anonymity.
questionnaires. This is particularly relevant for the study of discourse markers, as researchers have highlighted the importance of the visual dimension linked to their use, noting in particular that non-verbal gestures may in themselves be considered as discourse markers (Schiffrin 1987; Maury-Rouan 2001; De Fina 1997). The issue of video recordings will be discussed further in the conclusion.

The student teachers were then asked to compare the usefulness of the corpora with the books which they had used earlier in the course to prepare them for teaching through the target language. Books of two types had been used, namely advice on teaching through the target language (Guest and Pachler, 2001) and books providing examples of useful phrases. The response of the student teachers varied in the course of their discussion. Claire began by commenting that she found the books (citing Guest and Pachler, 2001) more valuable as they familiarised her with the theory underpinning her practice, adding that she also benefited from the recordings. Resource books giving useful phrases were then mentioned, leading to the comment below:

Unidentified voice: I’d just like to say that the list that we got of classroom vocabulary, you know like open your books and go to page, sit down … you know all the different commands and all that. I think we can use that in conjunction with what we learnt this semester [from the corpora]. Before I thought that was very useful, that would stand alone. This semester what I learnt was ways to make that sound more natural …

Others: mhm

Same unidentified voice: … in the classroom. And to encourage my students to sound more natural. So I think it very useful maybe if you could combine what you’re doing in this project into those lists of classroom commands, that would make those classroom commands a lot more …

The point was then made that in the French recordings only one of the many recommended phrases was used, namely “Excusez-moi”. This led to a number of reactions, with Helen commenting that after listening to the recordings the list...

... just seemed totally inappropriate. I don’t know whether it’s a good or a bad thing but … they are obviously all correct and we use them and they are great to have them in the class but …

Suzanne: I wouldn’t say that inappropriate.

Claire: It’s designed for us NNS … in a way.

Pat: Exactly. [All agree] […]

Unidentified voice: The stuff we were given for the students to use in the class did not correspond at all to what was in the recordings.

Claire: But it could be two lists would be valuable. Of course the one we worked with last semester but the authentic language as well.

In these exchanges the student teachers appear to be unwittingly addressing a major issue facing linguists, namely should linguistic research and, in this case, resource material, be based on data of naturally occurring discourse or on the intuition of the linguist, in this case the author of the resource book deciding what phrases a teacher
would be likely to use. The concluding comment that both are appropriate calls to mind Macaro’s (2001:56) comment, quoted earlier, on the strangeness of the language classroom as a type of discourse. The other questions, relating to pupil talk, the assessment, the classes and the running of the project provide no information that is not also present in the questionnaire responses and will therefore not be analysed in detail here.

The questionnaire began with a number of questions focussing on personal details (name, age, native language etc.). These were followed by three questions on each of the three aspects of classroom discourse studied (DMs, corrective feedback, and pupil discourse). Only the responses relating to DMs will be included here. First the student teachers were asked to rate the usefulness of the classes on DMs on a five-point scale, ranging from “very useful” to “of no use at all”10. Three of the five rated the classes as very useful, one useful, and one a little useful, a slightly higher rating than was given to the other two aspects of classroom discourse studied. They were then asked what they had learned from the classes, producing factual accounts of the content of the classes. More interesting comments were provided in response to the following question: “In what way do you think you will use what you learned from these classes in your teaching?” Responses included positive comments such as the following:

More aware of discourse issues, previously unaware.
I intend to make a conscious effort to use a variety of different discourse markers in the classroom. I feel that my spoken French has slightly improved as the language I speak now sounds a bit more authentic and fluent.
The classes gave me a great insight into the authentic way of speaking Spanish.

These comments reveal that the classes had fulfilled their aim of using the recordings as stimuli to encourage the student teachers to reflect on their use of the target language in the classroom.

Four questions related to the resources provided (transcripts, tapes, concordance printouts and frequency lists). Respondents were asked how much time they had spent consulting the transcript, and given a choice of replies: more than three hours, 2-3 hours, 1-2 hours, less than one hour. All five students had consulted the transcripts for 2-3 hours, with two volunteering the information that they had done this when preparing the assessed essay. The same question was posed concerning the tapes, eliciting that three respondents had consulted them for 2-3 hours and two for less than one hour. Comments added to these responses revealed that only one respondent had found the experience of listening to the tape and following the transcript at the same time interesting and beneficial. Two found the tape difficult to follow [despite good quality recordings], with one commenting that even when consulting the transcript and listening, it was difficult to follow what was going on because of overlapping teacher and student voices. One response is clearly critical of the choice of audiotape: “It was quite disconcerting to just listen to the tape as you can’t imagine it in a classroom

---

10 Although researchers such as Dörnyei (2003: 37–38) prefer an even number of response options to omit the ‘undecided’ category, the results of this questionnaire, which used a five-point scale, did not produce responses in the middle, neutral category.
context”. Respondents were then asked to rate the usefulness of the concordance printouts on a five-point scale: very useful, useful, a little useful, not very useful, of no use at all. Three respondents found them very useful, one useful, and one not very useful. The respondent who found them not very useful commented that “it would be better to see the whole sentence or context behind the word”. Other studies of learner reaction to concordances (Yoon and Hirvela, 2004; Chambers, 2005) have produced similar results, with a majority of learners having no problem with truncated concordance lines, while a small minority sees the absence of context as problematic, one out of four in Yoon and Hirvela’s focus group (2004: 276), and one out of fourteen in Chambers’s (2005: 120) course. The same question was then posed for the frequency lists, which three respondents found very useful, and one useful (one gave no response). One respondent commented that the information on frequency gave a “clear indication of the usability and frequency of use by native speakers, which we can learn from”.

Respondents were then asked to list, in order of importance, five ways in which the resources could be improved. Two issues were listed by three respondents, namely the provision of video recordings and of recordings of nonnative speaker teachers. Other responses, all cited once, included recommendations for recordings of the student teachers themselves to be used as resources, and for the classes recorded to deal with various aspects of language teaching, including grammar, the four skills, games, songs and films. Finally the student teachers were asked if the amount of time allocated to studying classroom discourse was too little, too much or about right, eliciting the unanimous response that too little time had been allocated. This echoed the unanimity on this issue in the group interview. One respondent even commented in the questionnaire that one hour per week for a semester should be devoted to the study of classroom discourse.

6 Conclusion

According to Maury-Rouan (2001:171-172), in the context of two native speakers conversing with each other, not being able to recognise exactly the functions of vague DMs (she cites as examples alors and bon ben) does not prevent communication. The student teachers in this study, however, clearly appreciated that DMs would make both their language use and that of their pupils sound more authentic. Furthermore, the analysis of the use of DMs in the classroom by two native speaker language teachers reveals that a relatively small number is used by each individual teacher, and that they fulfil a limited number of functions, thus making mastery of their use an achievable goal for the nonnative speaker teacher.

In the broader context of the integration of corpora of classroom discourse in LTE, however, the situation is more complex. On the one hand the advantages are clear, as the study of naturally occurring classroom discourse provides student teachers with a valuable resource, adding a new dimension to existing resources such as classroom observation, books on language pedagogy, and books with “useful phrases”. Including nonnative speaker recordings in a corpus of classroom discourse, which was envisaged from the outset in this project, is clearly considered of the utmost importance by the student teachers. The strong preference for a video-based corpus, however, requires further attention. The student teachers’ preference for video recordings is clear and
justifiable, as it would clearly be much easier to understand what was going on in the classroom when watching a video-recording, and also as gestures and facial expressions have an important role to play in the interaction. Video-recording the classes would, however, present the researcher wishing to create such a corpus with a considerable challenge in practical, financial and academic terms. Firstly, providing a video recording of a language classroom with a large cast of characters requires expertise of quite a different order than an audio recording, with corresponding financial implications. Secondly, teachers who are willing to turn on a small number of tape recorders might not be so willing to have their class disrupted by the process of filming it. Finally, the filming process also alters the nature of the research project. The observer’s paradox, namely the fact that naturally occurring discourse becomes less natural when those involved are aware that they are being recorded, is already present in the audio-based corpus. How much greater would it be if teacher and pupils were subjected to the intrusion of several cameras and a crew? For all these reasons the researchers have decided to accept the loss of the visual dimension and to create an audio-based corpus. This does not mean, however, that the student teachers’ preference for video-based resources must be ignored. The solution can be found in separating to a certain extent the research instrument, the corpus, from the teaching resource. In other words, video clips of language classes could be used for analysis and independent study, and complemented by consultation of frequency lists and concordances based on an audio-based corpus.

Other issues involving the integration of the corpus in LTE are more difficult to solve. Firstly, resources in the form of corpora of classroom discourse are not yet easily available. Secondly, it is difficult to envisage finding time in the programme of study for training in corpus consultation and analysis. While it would seem more feasible to include this in undergraduate programmes of study in modern languages, there is no evidence to suggest that this is likely in the short or medium term. In the meantime, in-service courses might fill the gap, and concordancers could be developed to meet the needs of language teachers rather than researchers. This could make it possible to integrate corpus consultation in LTE in a broader context than that of mastering classroom discourse in the target language. Finally, further research in the form of empirical studies, in particular longitudinal studies, would serve to underpin developments in this area.

References

The case of discourse markers


Appendix 1  English translations of examples of functions of DMs

1. Introduction of new topic, activity, or question.
   (a) French
   Teacher: Inquiet?
   Pupil: Have worries.
   Teacher: Yes être inquiet means to have worries. Very good yes. To be worried, to have worries. Ok? **Well then.** Other questions?.. Ok. So last question was…
   b) Spanish
   Teacher: …minutes, the class flew (unintelligible). Let’s see, one more thing, are you already reading the book…?

2. To call the pupils’ or teacher’s attention
   (a) French
   Teacher: It can also be somebody
   Pupil: [coughs]
   Teacher: cold who is indifferent.
   Pupil: Yeah.
   Teacher: One can say [some pupils are chatting] **All right.** So [some pupils are chatting] you have the demeanour …
   b) Spanish
   Teacher: But you need to put in “first”: Let’s see, **look**, you have the sentence here.

3. To recap what has been said or offer clarification
   (a) French
   Pupil: With a beard.
   Teacher: With a beard. Who with a beard?
   Pupil: A man.
   Teacher: A man. Yeah a bearded man is a man with a beard. Yes? **So** [drawing on the board] here is a man. Here he is bearded .. Yes? A bearded man is a man who has a beard.
   b) Spanish
   Pupil 2: and four
   Pupil 3: and four and five, yes.
   Teacher: **That is**, those are in the previous page. Page...

4. To motivate or encourage the pupils.
   (a) French
   Teacher: But we did the second part of question four? We answered the question “Why did Antoine want to see Daniel?” [The pupils agree] yes? Ok. **There we go then.** Question number 5.
   b) Spanish
   Pupil 3: …I have to change all of them now that I have seen the question.
Teacher: **OK, go on**, change it and say it correctly then.

5. **To correct oneself or rephrase what has been said**

(a) French

Teacher: Someone who has powers, it’s someone who is powerful. Yeah. **That is to say** that is someone who can do what he wants. Ok. So he can do most of the things that he wants to do. All right? […] Why are you laughing it’s true. It’s normal. **Well** it’s not normal but it’s the right translation yeah. [giggles in the classroom]

b) Spanish

Teacher: …we have said: “Brazil, country, first”, which you have in words, **I mean**, which you have in numbers, you have it in numbers.