Lexical Reiteration in EFL Young Learners' Essays: Does it Relate to the Type of Instruction?

Mª PILAR AGUSTÍN LLACH
ROSA M. JIMÉNEZ CATALÁN[*]
Universidad de la Rioja

ABSTRACT

The purpose of this study is to ascertain whether the type of instruction (English as vehicular language and English as a subject) is related to the use of reiteration ties. In the first place, we identified, classified, and counted the number and kind of reiteration ties used by two groups of EFL learners. Secondly, we examined whether reiteration as a mechanism of lexical cohesion correlates with learners' language level. Thirdly, we looked at students data in order to determine if lexical reiteration is associated with lexical variation. Results point to a similar pattern for both groups with regard to the use of reiteration ties: content and non-content students resort to word repetition rather than to other devices of lexical reiteration. Differences were found in favour of content students concerning lexical variation, language level, and use of antonyms and general nouns.

KEYWORDS: foreign language, lexical reiteration, lexical cohesion, lexical richness, lexical competence, school context, writing development, young language learners.

I. INTRODUCTION

The language learning context has received considerable attention in second language acquisition research. Most of it has been devoted to the discussion of the differences between natural and instructional settings, and less to the comparison of different language classroom settings and how they affect the learners' language acquisition and development. A particular area where studies are needed is the effectiveness of Content and Language Integrated learning (CLIL) in English as Foreign Language (EFL) classrooms. Influenced by the good results of immersion programmes in Canada, CLIL programmes have begun to be adopted in different European countries. Following this trend, a movement in favour of implementing CLIL in Spanish primary and secondary schools has been gaining ground, particularly in recent years. In the absence of official guidelines from educational policy makers, enthusiastic teachers from several bilingual and monolingual communities have begun to experiment with CLIL programmes in the belief that using the target language as a medium of instruction to learn the content of other subjects will result in the acquisition of higher target language proficiency level. Unfortunately, this belief is sustained mostly on good will rather than on empirical evidence, since research on the effectiveness of CLIL in the development of learners' target language competence is scarce, inconclusive, and very disperse.

The present study responds to the need of empirical studies on the effect of CLIL instruction in the development of different dimensions of communicative competence. It also responds to the need for research on young learners in both writing and vocabulary produced in different instructional contexts. Specifically, it aims to contribute to the understanding of how lexical cohesion is achieved in essays written by young EFL learners in two different language learning contexts: CLIL ─in which, English is used as medium of instruction in subjects of the school curriculum─ and English as a classroom subject.

Lexical and discourse competence are part of communicative competence. They both describe the ability to use language that conforms to the norms of different contexts and genres. Lexis and discourse make an indissoluble couple: on the one hand, lexis is intrinsically related to discourse type since it establishes grammatical and lexico grammatical relations that help texts be coherent and cohesive; on the other, discourse is linked to lexis through the selection of the appropriate vocabulary to deal with specific topics in specific genres. Knowing how to use the appropriate vocabulary goes hand in hand with knowing how

to handle genre rhetorical conventions as well as knowing how to produce coherent and cohesive texts.

In the last decades, a considerable amount of research has focused on the study of the cohesive markers used by second language learners. Nevertheless, most studies have looked at grammatical cohesion rather than lexical cohesion, and the majority of studies have focused on texts produced by adult language learners rather than on texts by young learners. Allard and Ulatowska (1991), in a study on the relationship between number of lexical and conjunctive ties and text quality, note the importance from an educational and research perspective of understanding how cohesion is achieved in L1 children's writing. We believe that it is also important to understand how lexical cohesion is achieved in EFL young learners' written discourse in different instructional contexts. L2 vocabulary research has provided evidence of the positive effect of vocabulary on the quality of writing (Engber, 1995; Laufer, 1994; Lee, 2003; Muncie, 2002). Research has also shown that lexical cohesion correlates with compositions scored holistically (Lieber, 1980; Meisuo, 2000; Witte and Faigley, 1981; Yang, 1989). With this in mind, in the next paragraphs we will define lexical cohesion and lexical reiteration and will review the main studies on lexical cohesion in written discourse. We will then state our objectives, pose our research questions, describe the instruments and procedures and finally present and discuss the esults of our study.

Coherence and cohesion are related concepts but by no means synonyms. Halliday and Hasan (1976) define them bearing in mind the text as a unity that hangs together thanks to inner and surface relationships. With regard to cohesion, Halliday and Hasan made a distinction between two main types of meaning relations: grammatical and lexico-grammatical. Grammar relations is comprised of the following categories: reference, substitution, ellipsis, and conjunction, whereas lexico-grammatical relations conform the category of lexical cohesion, defined by the authors as: "the cohesive effect achieved by the selection of vocabulary by means of reiteration or by collocation". (274). In Halliday and Hasan's framework, lexical reiteration is a mechanism of producing cohesion in a text by means of repetition of two or more lexical items that are observable at the surface of the text. Within this framework, lexical reiteration comprises four categories: repetition of the same word, use of a synonym, use of a superordinate, or use of a general noun. These four categories are understood as a set of options to achieve cohesion that range from most specific (use of the same word) to most general (use of a general noun). As Halliday and Hasan (1976: 278) put it: "Reiteration is a form of lexical cohesion which involves the repetition of a lexical

item, at one end of the scale; the use of a general word to refer back to a lexical item, at the other end of the scale; and a number of things in between —the use of a synonym, or superordinate".

Research on lexical cohesion in writing discourse is characterized by a great deal of fragmentation. Jiménez Catalán and Moreno Espinosa (2005) distinguish four tendencies in research: i) studies on the frequency of cohesive ties used by language learners (Anderson, 1980; Lieber, 1980; Meisuo, 2000; ii) research on the relation between coherence, frequency of cohesive ties, and quality of written production (Allard & Ulatowska, 1991; Johnson, 1992; Karasi, 1994; Khalil, 1989; Norment, 1994; Meisuo, 2000; Tierney & Mosenthal, 1983); iii) comparative studies on the frequency and variety of lexical cohesive ties employed by native speakers of and non-native speakers of different languages (Connor, 1984; Field & Yip, 1992; Guzmán, García & Alcón, 2000; Johnson, 1992; Liu & Robinet, 1990; Zanardi, 1994); iv) research oriented towards the discovery of whether genre, or topic has any influence of the different lexical ties employed by language learners (Allard & Ulatowska, 1991; Norment, 1994; Tierney & Mosenthal, 1983). In their review, Jiménez Catalán and Moreno Espinosa (2005) conclude that: a) lexical cohesion and within it, lexical reiteration, is the most frequent category among other types of cohesion; b) research is not conclusive regarding the relation between coherence and frequency of cohesive ties; c) there are significant differences in the use of lexical cohesive ties between native speakers and English learners; and the fact that d) genre has an effect on the choice of lexical ties. In their study of the lexical cohesive ties employed by a sample of 19 Business English students in a business letter and a dialogue, they reported word repetition as the most frequent lexical device used by students, in line with previous research on the effect on the type of genre on the choice of lexical ties.

Most of the studies mentioned earlier focus either on lexical cohesion in L1 children's writing or look at adult L2 language learner's lexical cohesion. There is hardly any research concerning lexical cohesion in the written discourse of young foreign language learners, let alone of learners in different instructional contexts. Yet, the study of lexical cohesion in EFL young learners' writing is of paramount importance. Manchón, Roca, and Murphy (2000) noted that most of the difficulties learners experience with writing have to do with vocabulary.

We believe that this study can be useful for English language teachers as well as for research on different areas such as vocabulary acquisition, learner corpus, and CILL. First, it may provide quantitative evidence of lexical cohesive ties used by EFL young learners in

written discourse. Second, the data elicited from this group may be useful for understanding the relationship between use of lexical cohesion ties and vocabulary richness as well and language level. Third, it may contribute to the investigation of the effectiveness of CLIL on two important dimensions of communicative competence: lexis and discourse.

The main purpose of this study is to learn whether the type of instruction (English as vehicular language or English as a curriculum subject) is related to the word types and lexical cohesive ties used by students in a written task. To that end, we aim at the following specific objectives. First, to identify the number and kind of reiteration ties used by each group as well as the number of students who use each kind of reiteration. Second, to ascertain whether reiteration as a mechanism of lexical cohesion correlates with learners' language level. Third, to find out if lexical reiteration is related to lexical richness.

Taking into account the above objectives the following research questions guided our study:

1)   What are the most frequent lexical reiteration ties used by content and non-content students?

2)   Will be there differences between the two groups within the category of lexical reiteration?

3)   Is reiteration associated with learners' language level and if it is, with what categories of reiteration?

4)   Can we observe differences concerning the number of types per 100 tokens produced by content and non-content students?

5)   Do content students outperform non-content students? If so, does lexical reiteration relate to lexical richness?

II. METHOD

II.1. Participants

Our sample consists of two groups of Spanish students (*N*= 60), learners of English as a foreign language, in 6th grade of primary education in two types of instructional contexts: content and non-content. The content group consists of *N*=30 female students from a CLIL school located in Bilbao. They have Spanish as L1, Basque as L2, and English as their L3. On its part, the non-content group consists of *N*=30 female students randomly selected out of a sample of 114 females from four Logroño primary schools. All the students have Spanish as L1 and English as L2.

In the time of eliciting the data (Spring 2006), the content group had received approximately 960 hours of English instruction at the rate of 5 hours per week. This amount refers only to English classes. In addition, they had been taught two curricular subjects by using English in several courses of primary education: Science, for two hours a week throughout 1st and 2nd grade, and Science and Art and Craft, for two hours a week throughout 3rd, 4th, 5th, and 6th grade. In contrast, the non-content group had received approximately 629 hours of English instruction at the rate of 3 hours per week. Both, the content and the non-content group were taught English following a communicative approach with emphasis on listening and speaking, although they had also been introduced to writing.

II. 2. Data collection

All participants were asked to complete an 8-point cloze test (Corporate Author Cambridge ESOL, 2004) as a reference for language level and to write a composition in English to an English family. Students were given the reason for writing this letter: they had to imagine that they were going to spend a month in Oxford with the English family. The purpose was introducing themselves to the family, writing about their own family, their own town, school, hobbies, and interests. The adequacy of the topic to students' age and language level was considered in the selection of the data elicitation instrument. Likewise, students' familiarity with writing informal letters was checked by consulting the teachers of both groups as well as students' textbooks and materials. Both groups had some experience in writing E-mails, postcards, and informal letters to friends. Oral instructions in Spanish were given to students before performing the task. Students were asked to write the composition in English in 30

minutes. Dictionaries and other kinds of help were not allowed as to elicit as much spontaneous learners' productive vocabulary as possible.

II.3. Procedure

Each composition was transcribed and encoded electronically as a separate file via hand typing. The corpus of essays was edited to remove unintelligible words and Spanish words, except names of places and people, as they can be used in texts to establish cohesive links. Spelling mistakes were corrected but punctuation and grammar errors were left intact.

II.4. Data Analysis

All compositions were subjected to the textual analysis program *WordSmith Tools* (Scott, 1996). In our analysis, words were not lemmatised but counted as separated entries. For the purpose of the present study we understand *the word*, following Sinclair (1991:41), as " a 'word' is any string of characters with a word space on either side, so boy and boys, and come, came are all different words". Likewise, following Richards and Schmidt (2002:567), in this study we use the term *type* as a class of linguistic item, *token* as an actual example or number of occurrences of a type, and *Type-Token Ratio* (TTR) as "a measure of the ratio of different words to the total number of words in a text". In corpus studies as well as in L2 vocabulary research, the TTR is regarded as a measure of lexical richness. This looks at the degree of lexical variation or lexical diversity in a text and may be used to compare lexical differences between texts (Barnbrook, 1996) and to identify learners' language level (Engber, 1995; Laufer, 1994; Read, 2000). In this sense, Read notes: "It is reasonable to expect that more proficient writers have a larger vocabulary knowledge that allows them to avoid repetition by using synonyms, superordinates and other kinds of related words" (2000:200).

As has been remarked in vocabulary research literature, TTR is text-length dependent (Baayen 2001; Laufer & Nation 1995; Malvern, Richards, Chipere, & Durán, 2004). This means that as the text increases in length, the possibilities of using new types are reduced. When analysing lexical variation in foreign language learners' compositions this is a serious limitation as researchers can come up with incongruent results. For instance, a learner who only writes 60 tokens may achieve a higher TTR than a learner who writes 120 tokens. Several solutions have been proposed to avoid this drawback. One of them is to divide the

corpus material in equally sized text samples (see for instance, Arnaud, 1984; Laufer, 1991; and Laufer and Nation, 1995). However, this solution is not without problems. As Miralpeix (2006: 92) notes: "…data are lost when cutting the texts". We agree with her although we also believe that there is still some advantage in using this method. By dividing the corpus material in equally sized text samples it is possible to carry out systematic comparisons of the vocabulary produced by EFL learners of different age, sex, grade, and/or language level. Moreover, not much is discarded when dealing with young EFL learners as in a 30 minute period they do not usually write essays that contain more than 100 tokens. Thus for the purpose of this study each composition count was stopped when the counting processor indicated 100 tokens.

A qualitative analysis was later applied to each composition in order to identify and classify instances of lexical reiteration based on lexical word types. In order to warrant reliability in the identification, classification and counting of lexical reiteration ties, the process was carried out twice independently by the two authors of this study. A close agreement was found between them. Nevertheless, a third independent researcher was consulted to revise the identification and classification of the data[1]. In the classification and counting of lexical reiteration ties we followed the taxonomy proposed by Halliday and Hasan (1976), which comprises four categories: word repetition, superordinates/hyponyms, synonyms/cuasi-synonyms, and general nouns. To this classification, we added two additional categories: meronymys and antonyms. These two are included in more recent classifications of lexical cohesion such as the ones proposed by Hoey (1991) and by Martin (1992), among others.

Following Halliday and Hasan's (1976) framework of cohesion, the letter-writing task is understood here as a text on its own that has semantic unity. That is to say, a text in which grammatical and lexical relationships are realized. Some of these relationships are manifested by means of lexical cohesion ties. According to Halliday and Hasan (1976) the term lexical tie refers to a single instance of cohesive relationship between two or more content words.

In order to find out whether the results were significant, different statistical tests were applied to the data: chi-square, Spearman correlation, and t-test[2]. The use of chi-square test was justified mainly because of the nominal level of lexical reiteration ties as they were classified according to categories within a classification. Spearman correlation test was used to ascertain the relation between use of reiteration ties and language level on one part and,

reiteration ties and lexical variation on the other. The use of this test rather than Pearson correlation test was due to the fact that the samples were not normally distributed. Finally, a t-test was applied to the data to see if there were statistical differences between content and non-content students concerning their means in types and tokens.

IV. RESULTS

IV. 1. Lexical Reiteration

Our first objective was to identify the number and kind of reiteration ties used by content and non-content students. Relevant questions resulting from this objective are: what are the most frequent lexical reiteration ties used by content and non-content students? Will be there differences between the two groups within the category of lexical reiteration?

As can be inferred from the raw frequencies and the means shown in Table 1, word repetition is the most frequent lexical reiteration tie in the compositions of both groups, followed by superordinates/ hyponyms, meronymys and synonyms. The least frequently used devices for content and non-content students are antonyms and general nouns.

	Content			Non-content		
	Fr.A	Mean	sd	Fr.A	Mean	sd
Word repetition	64	2.13	1.25	111	3.70	1.60
Superordinate /Hyponymy	41	1.37	1.03	37	1.23	1.30
Meronymy	21	.70	.79	26	0.87	1.00
Synonym / cuasi-synonym	21	.70	.83	18	0.60	.81
Antonymy	8	.27	.45	7	0.23	.50
General Noun	7	.23	.43	5	0.17	.53

Table 1. Distribution of lexical reiteration in content and
non-content students' compositions

As shown in Table 1, the distribution of the different categories of lexical reiteration is roughly the same in the two groups. However, the number of word repetitions as well as the number of meronymies is higher in the compositions of non-content students. This tendency is reversed for the other categories, in which content students surpass non-content students.

Regarding our first objective, another perspective of analysis is to look at the number of students who use each kind of reiteration. As Table 2 indicates, compared to content students, a greater number of non-content students make use of word repetition and meronymy. However, this tendency reverses in the remaining classes of reiteration, where the number of students who use any of these categories is greater in the content group than in the non-content one.

	Content	Non-content
Word Repetition	90	100
Superordinates /Hyponymys	83.3	70
Synonyms	50	43.3
General nouns	23.3	10
Meronymy	53.3	56.7
Antonymy	26.7	20

Table 2. Percentages of students who produce at least one instance of eat class of lexical reiteration

As far as lexical reiteration is concerned, the results of the chi-square test applied to the data allows us to conclude that i) use of word repetition is related to the type of instruction as non-content students produce a significant higher number of this kind of lexical reiteration than content students (chi 18.67 $p < .01$); ii). Likewise, production of general nouns is related to the type of instruction with content students using a significant higher number of this kind of lexical cohesive ties than non-content students (chi 6.82 $p < .05$); and, iii) no significant differences are found between content and non-content students concerning the use of superordinates/hyponyms (chi 3.61, $p > 0.05$), synonyms/cuasi-synonymys (chi 0.320, $p > 0.05$), meronymys (chi 1.223, $p > 0.05$), or antonyms (chi 1.773, $p > 0.05$).

IV.2. Reiteration and language level

Our second objective was to ascertain whether reiteration as a mechanism of lexical cohesion correlates with the learners' language level. Thus, the pertinent questions here are: does reiteration, as a device of lexical cohesion, relate to learners' language level and if it does, with what categories of reiteration? In order to find an answer to these questions, first, the means and standard deviations on a cloze test were calculated for each group. Table 3 reveals that content students outperform non-content students.

	Content	**Non-content**
Mean	6.53	4.47
s.d.	1.25	1.63
Max.	8	8
Min.	4	2

Table 3. Mean, standard deviation, and maximum and minimum scores in the cloze test by content and non-content students

A chi square test was performed in order to find out the significance of the scores obtained by the content and non-content groups of subjects. We found that scores on the cloze test depend on the type of instruction (chi 22.52, $p < .01$) with content students scoring higher on the test than non-content students.

As to the relation between reiteration and learners' language level, no significant correlation (Spearman test) was found between the different categories of lexical reiteration and scores on the cloze test. The values obtained were as follows: $r = -0.225$, $p > 0.05$ for Word Repetition, $r = -0.014$, $p > 0.05$ for Superordinates /hyponyms, $r = 0.231$, $p > 0.05$ for Synonyms, $r = 0.180$, $p > 0.05$ for General noun, $r = 0.107$, $p > 0.05$ for Meronymys, $r = 0.071$, $p > 0.05$ for Antonymys.

IV.3. Reiteration and lexical variation

In this section, we will address the issue of lexical variation and try to meet our third objective: to verify if lexical reiteration is associated with lexical richness. The pertinent questions related to this objective are: can we observe differences concerning the number of types per 100 tokens produced by content and non-content students? Do content students outperform non-content students? If so, does lexical reiteration relate to lexical richness? Table 4 contains the results obtained by each group:

	Content Group	**Non-content Group**
Mean	56.93	53
s.d.	4.46	5.27
Max.	67	62
Min.	47	43

Table 4. Types per 100 tokens in content and non-content students' essays

The comparison shows that the content group achieves a higher means than the non-content group. A t-test was performed in order to check if there were significant differences between the two groups. Results indicate that there are significant differences in favour of content students (t= 3.63, p < 0.001). Due to the reduction of the corpus of compositions to equal size of 100 tokens, the means and the TTR coincide. Therefore, we can conclude that, on average, content students use more different words in their compositions than do non-content students, which indicates a higher degree of lexical variation in favour of the content group's compositions.

Once found out that content students outperform non-content students on their TTRs, we are in position to examine whether lexical reiteration relates to lexical richness. Spearman test applied to the data yielded the following results: i) a negative correlation between types and word repetition (r–0.298, p < .05), meaning that as the number of lexical types decreases, the number of word repetition increases; ii) a positive but moderate to low correlation between production of lexical types and use of antonyms (r 0.274, p < .05), meaning that as the number of lexical types increases, antonyms also increase; iii) no correlation was found between lexical types and other categories of lexical reiteration: superordinates/hyponyms (r = -0.020, p > 0.05), synonymys (r = 0.225, p > 0.05), meronyms (r = -0.172, p > 0.05), and general nouns (r = 0.007, p > 0.05).

V. DISCUSSION

In this study we examined lexical reiteration in the compositions of EFL young learners in two instructional contexts. Our purpose was to compare a sample of texts written by students in a content school with a sample of texts of students in an English as a classroom subject. Results show that word repetition is the most frequent lexical cohesion device used by both content and non-content students. This finding corroborates studies on cohesion in the written discourse of non-native speakers (Bae, 2001; Connor, 1984; Khalil, 1989; Ferris, 1994; Guzman, García & Alcón 2000; Jiménez Catalán & Moreno Espinosa, 2005; Meisuo, 2000), where learners have been reported to resort to word repetition rather than to other devices of lexical cohesion such as synonyms, general nouns, or superordinates.

Although, similar patterns of use are found in content and non-content students concerning word repetition, differences are also observed. In the first place, the percentages of

students who produce at least one instance of superordinates, synonyms, antonyms, and general nouns are higher for content students; on the contrary, in the case of word repetition and meronymys the percentages of use are higher in the non-content group than in the content one. However, significant differences can be claimed only for two types of lexical reiteration: word repetition and general nouns.

Empirical evidence on the relation between language level and use of lexical cohesive ties is contradictory in its results. For instance, in Ferris (1994), it is reported that higher language level learners use a wider variety of lexical cohesion ties than lower language learners, who rely more on word repetition. In contrast, in the studies conducted by Connor (1984) and Scarcella (1984), no significant differences were reported between learners' language level and frequency of cohesive ties were found (quoted in Ehrlich, 1988). The findings of the present study add data in this sense as, although from the scores on the cloze test administered to content and non-content students we can assume content students to have a higher language level than non-content students, it still cannot be claimed that the learners' language level has an effect on any of the categories of lexical reiteration (among them, word repetition). In our study, no correlation was found between these variables.

Significant differences in favour of content students' performance were also found in their higher TTRs, which points to a higher degree of lexical variation in their compositions. Concerning lexical variation, another additional sign of the supremacy of content over non-content students is observed in the negative correlation of type of instruction and use word repetition (r–0.298, $p<$ .05), and the positive correlation of type of instruction and use of antonyms (r 0.274, $p <$ .05). These data give clear evidence of the relationship between the production of lexical types and use of two lexical reiteration ties by content and non-content students: the implications are that as content students have a higher lexical variation, they do not need to resort to word repetition as frequently as non-content students. By the same token, given content students' higher degree of lexical variation, it is not surprising to find a higher number of antonyms in their compositions.

Although results point to higher percentages of use of lexical reiteration ties by content students, we cannot obviate the fact that, except for the percentage using word repetition, a rather low the percentage of students used other classes of lexical reiteration such as supeordinates/hyponyms, synonyms, antonyms, or general nouns. Students in 6th grade seem to make a very restricted use of lexical reiteration devices to make their compositions

cohesive, regardless of whether or not they attended CLIL classrooms. It is not easy to compare our results with other studies on the effect of CLIL on learners' language acquisition. First, there are hardly studies conducted in EFL contexts and, as mentioned in the introduction, the few studies found have focused more on university or secondary students than on young learners. Second, research is very different concerning objectives and methodology and disperse on different research topics. With caution, the findings of our study could be related to the ones obtained in ESL Canadian contexts regarding content students higher performance in oral production (Spada & Lightbown, 1989; White & Turner, 2005). More related to our study are the research projects carried out by Llinares and Whittaker in the Madrid Autonomous Community, and the one conducted by the REAL and the GLAUR groups in the Basque country and La Rioja. Llinares and Whittaker (2006) report two studies: CLIL students' progress in the acquisition of register features on history as content discipline, and an analysis on the oral and written EFL young learners' production. In this second study, they found no statistical differences regarding students' capability of differentiating the features of written and spoken discourse. On their part, under the framework of a joint research project aimed at identifying the effect of type of instruction (English as vehicular language versus English as medium of instruction) on the acquisition and development of English as L2 and L3 learners' competence, Villarreal Olaizola and García Mayo (2007) report significant differences among content and non-content students regarding the use of –s and –ed morphemes: the former performing significantly better. Within the same joint research project, Jiménez Catalán and Ruiz de Zarobe (2007) investigated content and non-content student performance on a receptive vocabulary test, finding highly significant scores for content students.

VI. CONCLUSION

Lexical cohesion plays an important role in the development of students' written discourse. This study has attempted to provide evidence on the effect of the type of instruction on the use of lexical reiteration ties as an important mechanism of lexical cohesion in an English letter written by EFL young learners studying English in two different instructional contexts: CLIL instructional context in which English is used as a vehicle to learn Science and Art and Craft, and English as a school subject, in which English is taught as one subject among others. We found differences in the use of lexical reiteration ties by both groups of students: although

both groups resorted to word repetition rather than to other devices of lexical reiteration, differences in favour of content students were found regarding lexical variation, and language level. One of the most remarkable results is that non-content students use word repetition more frequently than content students, whereas content students use general nouns more frequently than non-content ones. Although the differences of use are not great (in particular regarding the latter), they are highly significant. In the light of results, it can be claimed that content instruction is positively associated with several aspects of learners' vocabulary: lexical richness, as content students produce a significantly higher number of types per composition, a higher percentage of use of superordinates, synonyms, antonyms, and a significantly higher number of general nouns.

The exploratory nature of this study imposes some limitations in terms of generalization of results. A major limitation lies in the sampling. Due to the shortage of continued student experience with CLIL in EFL classrooms, it is extremely difficult for researchers to select a sample of content students who have been instructed over time purely in English as a vehicular language to learn other languages and who have never received instruction in English as a subject in any of their primary education grades. This would be the ideal situation for conducting a comparison with students in English as a subject schools, but so far it is, if not impossible, very difficult to find. This situation places the researcher in a dilemma as to whether to investigate these incipient CLIL classrooms or wait until this approach is extended throughout all the grades in compulsory education. Still, we believe that researchers have a responsibility to investigate the effectiveness of language teaching experiences as they come, even if they cannot intervene to control all the variables. This was our option, and in our sample of content students it is difficult to separate and identify the effect of content proper, from the effect of more hours of instruction to the language. The reason is that, even if they belong to the same educational level, in addition to learning curricular content subjects in English, content students also have 350 more hours of instruction in English than non-content students. In a study on the effects of starting age on the acquisition of EFL in schools of Catalonia, Celaya, Torras, and Pérez-Vidal (2001) measured the written performance of two groups of 12-year-old secondary students who had received different amount of instruction (416 hours compared to 200 hours). Among their results, relevant for the interpretation of our study is the fact that they report significant differences in 6 measures of complexity related to lexis in favour of students with more hours of instruction. The connection with our study is clear: if there is empirical evidence on the positive effect of

exposure on learners' vocabulary, we cannot venture a strong interpretation of the data in favour of the effect of CLIL. We hope that with the gradual implementation of CLIL programmes in Spanish schools, it will possible for researchers to control the CLIL variable more tightly. Nevertheless, we believe that this exploratory study may be useful for researchers and teachers alike as it draws a profile of the frequency of use of lexical reiteration ties by content students compared to non-content ones in 6th grade in EFL in primary education. It also provides evidence of the relationship between type of instruction and students' lexical richness, language level, and use of lexical reiteration ties. The picture that emerges reveals significant differences in favour of content students regarding lexical richness and language level, as well as regarding the use of antonyms and general nouns. Nevertheless, we cannot claim the same regarding the use of other lexical reiteration categories such as supeordinates/hyponyms, synonyms, or meronymys, as both, content and non-content students hardly use any of these categories in their compositions.

Acknowledgement

NOTES

1. We are grateful to Professor Francisco Ruiz de Mendoza from University of La Rioja for his insightful comments and suggestions in the identification and classification of the examples of lexical reiteration.

2. We would like to thank Monserrat Martín for her expert advice on test selection and statistical analyses.

REFERENCES

Allard, L. & Ulatowska, H. K. (1991). Cohesion in written narrative and procedural discourse of fifth-grade children. *Linguistics and Education,* 3:1, 63-79.

Anderson, P. (1980). *Cohesion as an index for written and oral compositions of ESL learners*. U.S, ERIC Document. ED 198529.

Arnaud, P.J.L. (1984). The lexical richness of L2 written productions and the validity of vocabulary tests. In T. Culhane, C. Klein-Braley & D.K.Stevenson (Eds.), *Practice and Problems in language testing*, pp: 14-28.

Baayen, H. (2001). *Word frequency distributions*. Dordrecht: Kluwer Academic Publishers.

Bae, J. (2001). Cohesion and coherence in children's written English: immersion and English-only classes. *Issues in Applied Linguistics*, 12:1, 51-88.

Barnbrook, G. (1998). *Language and computers*. Edinburgh: Edinburgh University Press.

Celaya, M.L.; Torras, M.; & Pérez-Vidal, C. (2001). Short and mid-term effects of an earlier start. An analysis of EFL written production. In S, Foster-Cohen & Nizegorodcew (Eds.), *EUROSLA Yearbook*, vol. 1, pp: 195-209.

Connor, U. (1984). A study of cohesion and coherence in English as a second language students' writing. *Papers in Linguistics, International Journal of Human Communication,* 17:3, 301-316.

Ehrlich, S. (1988). Cohesive devices and discourse competence. *World Englishes*, 7:2, 11-118.

Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of L2 Writing*, 4:2, 138-155.

Ferris, D. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly,* 28:2, 414-420.

Field, Y. & Yip. L.M.O. (1992). A Comparison of international conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, 23:1, 15-25.

Guzmán, J; García, I; & E, Alcón. (2000). Cohesión léxica y textos narrativos: análisis cuantitativo en el contexto de la enseñanza de la L1, la L2, y la LE. In Alcántara, F.; Barreras, A.; Jiménez, R.; Moreno, M.A. (Eds.), *Panorama actual de la lingüística aplicada. Conocimiento, procesamiento y uso del lenguaje*. Logroño: Universidad de La Rioja, pp: 1761-1768.

Halliday, M.A.K & R. Hasan (1976). *Cohesion in English*. London: Longman.

Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press.

Jiménez Catalán, R. & Moreno Espinosa, S. (2005). Lexical cohesion in English L2 students' compositions. In Salazar, P., Esteve, M. J.; & Codina, V. (Eds.), *Teaching and learning the English language from a discourse perspective.* Castelló: Universitat Jaume I, pp: 73-90.

Jiménez Catalán, R.M. & Ruiz de Zarobe, Y. (2007). Does the type of instruction have any bearing on EFL learners' receptive vocabulary? *Paper presented at ELIA X*, 14-16 March, University of Seville.

Johnson, P. (1992). Cohesion and coherence in compositions in Malay and English, *RELC Journal,* 23, pp: 1-17.

Karasi, M. (1994). *Cohesive features in the expository essays of Secondary Four (Express) and Secondary Five (Normal) students in Singapore*. M.A. Dissertation, Nanyang Technological University.

Khalil, A. (1989). A study of cohesion and coherence in Arab EFL college students' writing. *System*, 17:3, 359-371.

Laufer, B. (1991). The development of L2 lexis in the expression of the advanced language learner. *Modern Language Journal*, 75:4, 440-448.

Laufer, B. (1994). The lexical profile of second language writing: does it change over time? *Regional English Language Center Journal*, 25:2, 21-33.

Laufer, B. & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16:3, 307-322.

Lee, S. (2003). ESL Learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System*, 31, 537-561.

Lieber, P. (1980). *Cohesion in ESL students' expository writing*. Ph. D. Dissertation, New York University.

Liu, CH. & Robinett, B.W. (1990). Cohesive devices in the paragraph contrastive Chinese-English use. ELT in China. In *Papers presented at the International Symposium on Teaching English in the Chinese context*, Beijing, Foreign Language Teaching and Research Papers.

Llinares, A. & Whittaker, R. (2006). Linguistic analysis of secondary school students' oral and written production in CLIL contexts: studying social sciences in English. *WIEWZ*, 15:3, 28-32.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. London: Palgrave, Macmillan.

Manchón, R.M.; Roca, J.; & L. Murphy (2000). La influencia de la variable «grado de dominio de la L2» en los procesos de composición en lengua extranjera: hallazgos recientes de la investigación. In Muñoz, C. (Ed.) (2000). *Segundas lenguas*. Adquisición en el aula Barcelona: Ariel Lingüística, pp: 277-297.

Martin, J.R. (1992). *English text: System and structure*. Amsterdam: John Benjamins.

Meisuo, Z. (2000). Cohesive features in the expository writing of undergraduates in two Chinese Universities, *RELC Journal,* 31:1, 61-95.

Miralpeix, I. (2006). Age and Vocabulary Acquisition in English as a Foreign Language (EFL). In C. Muñoz (Ed.), *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters, pp: 89-106.

Muncie, J. (2002). Process writing and vocabulary development: comparing lexical frequency profiles across drafts. *System*, 30:2, 225-235.

Norment, N. (1994). Contrastive analyses of cohesive devises in Chinese and Chinese ESL in Narrative and Expository written texts. *Chinese Language Teaching Association Journal,* 29:1, 49-81.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Richards, J. & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. London: Longman.

Scarcella, R. (1984). *Cohesion in the writing development of native and non-native English speakers. Unpublished dissertation*. University of Southern California.

Scott, M. (1996). Comparing corpora and identifying key words, collocations, and frequency distributions through the WordSmith Tools suite of computer programs, Ghadessy, M. & R. L. Roseberry (Eds.), *Small Corpus Studies and ELT,* Amsterdam, John Benjamins Publishing Co, pp: 1-17.

Sinclair, J. (1991). *Corpus, concordance, collocation*, Oxford: Oxford University Press.

Spada, Nina. & P.M. Lightbown (1989). Intensive ESL programmes in Quebec Primary schools. *The Canada Journal*, 7:1, 11-32.

Tierney, R. & J. Mosenthal (1983). Cohesion and textual coherence, *Research in the Teaching of English,* 17:3, 215-229.

Villarreal-Olaizoila, I. & García-Mayo, P. (2007). Finiteness in English as an L3: Content-Based vs. English as a School Subject Learners. *Paper given at the 6th International Symposium on Bilingualism*, May 30-June 2, University of Hanburg.

White, J. & Turner, C. (2005).Comparing Children's Oral Ability in Two ESL programs. *The Canadian Modern Language Review*, 61:4,491-517.

Witte, S.P. & Faigley, L. (1981). Coherence, cohesion and writing quality. *College Composition and Communication*, 32:2, 189-204.

Yang, W. (1989). Cohesive chains and writing quality. *Word*, 40:1-2, 235-254.

Zanardi, N. (1994). Cohesion in Italian adult learners' and native speakers' compositions: a comparison. *Australian Review of Applied Linguistics,* 12:2, 22-50.