A longitudinal analysis of the effects of one year abroad

Abstract

The purpose of this study is to analyze the progress of 14 Spanish-speaking learners of English during a period abroad from a longitudinal perspective. Oral and written data were collected three times during an academic year at a British university. These samples were analyzed in terms of fluency, syntactic complexity, lexical richness and accuracy. The results of the statistical analyses indicate that, while a few months abroad might be sufficient for some gains in oral performance to occur, improvement in written production is slower and does not seem to take place until the students have spent more than one semester abroad. Additionally, it was observed that the type of interaction experienced while abroad as well as some attitudinal features can partly explain language development in some areas.

Literature review

Learning context has been an important focus of SLA research during the past decade (Collentine, 2009; Freed, 1995, 1998; Freed, Segalowitz, & Dewey, 2004; Freed, So, & Lazaar, 2003; Llanes, 2011). This interest in learning context has grown in tandem with

the popularity of study abroad (SA) experiences. According to the International Institute of Education and the European Commission for Higher Education, which report data regarding the SA participation in the U.S.A. and Europe respectively, the number of students studying abroad has increased dramatically during the past decade (see Figures 1 and 2). Similarly, Canadian universities are becoming more and more interested in sending their students abroad, and an increasing number of these students (up to 17,850 in 2006, three times more than in the year 2000) are taking advantage of this opportunity, according to the Association of Universities and Colleges of Canada (AUCC, 2007).

[Figure 1 here]

[Figure 2 here]

It has been documented that SA has an impact on several areas of second language development, both in non-linguistic aspects such as motivation (Allen, 2010), affective and cultural factors (Ismail, Morgan, & Hayes, 2005) and in linguistic ones. The effects of SA on the participants' linguistic development in particular have been widely documented. The most investigated domain in relation to learning context is oral production, especially oral fluency, as it is believed to be the most sensitive to learning context (Freed, 1995; Freed, Segalowitz & Dewey, 2004; Lennon, 1990; Llanes & Muñoz, 2009; Segalowitz & Freed, 2004; Yager, 1998). Vocabulary development is another important domain that has been reported to be different in SA and instructional settings.

The studies by Lennon (1990) and Llanes and Muñoz (2009) analyzed the oral fluency of a group of learners who spent time abroad and found that time abroad was crucial for the improvement of second language (L2) fluency. Freed's (1995) study also examined oral production using both objective measures and rating scales through which native speakers judged the native-likeness of learners' speech samples. The results obtained on both sets of measures revealed that the SA context led to greater gains than the AH context. Similarly, Yager (1998) found that after the SA experience, participants were perceived to be more fluent. Further evidence for the benefits of SA on participants' oral production comes from Segalowitz and Freed (2004), who examined oral fluency using a series of objective measures and attempted to relate the gains that participants achieved to their cognitive abilities. Segalowitz and Freed concluded that cognitive abilities also play a role in the oral fluency improvement that participants experienced.

Vocabulary acquisition is another commonly researched aspect of learning context studies. Dewey (2008) examined the receptive vocabulary of American undergraduates learning Japanese in three learning contexts: AH, IM (domestic immersion) and SA. Dewey found that participants in the SA group scored higher than participants in the IM group, who in turn scored higher than participants in the AH group. Foster (2009) compared the L2 vocabulary of learners in different contexts, AH and SA, and also included data from native speakers of the L2. She found that SA participants' L2 use was closer to the native speakers' than that of the AH group. Other studies such as Ife, Vives, and Meara (2000), Llanes and Muñoz (2009), and Milton and Meara (1995) examined the L2 vocabulary development of participants who spent some time abroad, but they did not offer a comparison group. Although they analyzed vocabulary use in different ways, they

have all concluded that stays abroad are beneficial for the participants' lexical development.

Although not as commonly investigated as oral fluency and vocabulary, other language skills have also been the subject of learning context studies. Some evidence has been provided in the literature for the advantage of the SA context for the development of listening skills (Cubillos, Chieffo, & Fan, 2008; Dyson 1988; Llanes & Muñoz, 2009), reading comprehension skills (Dewey, 2004; Lapkin, Hart, Swain, 1995), writing skills (Sasaki, 2004, 2007, 2009), and also grammar (Guntermann 1995; Howard, 2005; 2006). Interestingly, no benefits have been reported for the SA context in terms of pronunciation (Díaz-Campos, 2004; Mora, 2008). In relation to sociolinguistic appropriateness, Regan (1995:261) claims that 'the effect of the year abroad is very striking in the acquisition of the vernacular grammar and sociolinguistic competence'. Regan corroborated her own claim in her 2005 study, in which she examined the *ne* 'no', 'not' deletion in French (L2) by a group of five Irish undergraduates who spent an academic year in France. The author collected data at three different points (pretest, posttest and a delayed posttest) and found that, between the pre- and the posttest, participants deleted the particle *ne* significantly more frequently (i.e. showing more native-like sociolinguistic patterns) and that the *ne* deletion rates attained after their year abroad were still maintained one year after their return from France.

It must be borne in mind, however, that not all the empirical evidence in terms of context of learning suggests significant differences in favor of students going abroad over students receiving classroom instruction. Some studies have reported no differences between learning contexts, or no significant improvement after a period abroad

(Collentine, 2004; DeKeyser, 1991; Dewey, 2004; Díaz-Campos, 2004; Mora, 2008). Similarly, studies examining different linguistic areas do not necessarily find advantages in all of these areas for the SA context (Freed et al., 2003).

The fact that the SA context has not been found uniformly more beneficial for language development than classroom L2 learning –despite the popular belief that the best (or even the only) way to learn a language is by spending time abroad– can be due to the different factors that determine whether students will take advantage of the opportunities they supposedly have abroad. One such factor is the length of stay abroad. It has been shown that, in general, students who stay abroad for a whole academic year tend to show significantly greater gains than those who stay for only one semester (Dwyer, 2004; Ife et al., 2000). Other variables that can have an impact on the type and rate of L2 development that occurs abroad include initial proficiency level (Brecht, Davidson, & Ginsberg, 1995; Freed, 1990; 1995; Ife et al., 2000; Milton & Meara, 1995), language contact while abroad (Freed et al., 2004; Juan-Garau & Pérez-Vidal, 2007; Segalowitz & Freed, 2004), personality (Kinginger, 2008), or even gender (Brecht et al., 1995; Polanyi, 1995; Regan et al., 2009).

The present study aims to investigate language gains in a study abroad context from a longitudinal perspective, which is not a commonly adopted design in the literature (with a few exceptions; e.g., Regan, 2005). The same group of participants was followed at different time points during their stay abroad and their oral and written production was assessed throughout time. The learners' progress was examined through different data collection points after a few months abroad and after a whole academic year. This design allows us to analyze, first of all, whether L2 development in an SA context is linear in the

different areas of oral and written production under examination, and whether oral and written production develop in tandem. Additionally, a longitudinal design also facilitates the possibility of examining whether there are some L2 areas that develop more quickly than others. In addition to the longitudinal development of students' oral and written production, we have also considered two factors for analysis that we thought might affect language progress in the SA context: attitudes towards the L2 and its speakers, which has been claimed to affect second language acquisition in general (Masgoret & Gardner, 2003), and language interaction while abroad.

More specifically, this study aims to answer the following research questions:

1. Does L2 proficiency in oral and written production develop at the same pace while abroad, or is improvement in one modality faster than in the other?

2. Can learners' individual variables, such as attitudes or chances to interact abroad, explain certain aspects of language development in oral and written production?

Method

Participants

The participants in this study are 14 Spanish-speaking students who were enrolled at the University of Southampton (United Kingdom) for one year as part of the Erasmus European exchange program, which is the most popular program for college students to study abroad in Europe. The participants were all young adults between the ages of 20 and 24, with a mean age of 22. There were nine females and five males. The participants

studied different majors in Spain, seven of them related to English studies or translation, six in scientific fields, and one in history. The students also differed in terms of their academic year at their universities of origin: two participants were in their second year, eight were in their third or fourth, and four were writing their undergraduate thesis. For all of the students the study abroad period was optional and for six, or roughly half, this was their first study-abroad experience.

In terms of their experience with the English language in formal settings, all of them had received instruction at school beginning between the ages of six and ten. Apart from this, eight students had also taken 'extra-curricular' courses in language schools, while six students never had this experience. When they were asked about their perception of their English proficiency in reading, writing, speaking and listening, they all rated themselves between lower intermediate and advanced, and this rating corresponds with students' actual proficiency level as determined by their performance in the pretest.

Instruments

The instruments that were used in this study were designed to examine students' oral and written production on the one hand, and students' background information referring to language attitude and language use on the other.

Students' oral production was elicited by means of an oral narrative ('The picnic story', Heaton, 1966). To the authors' knowledge, this task was first used for research purposes by the 'Barcelona Age Factor Project' (see Muñoz, 2006), and since then it has been used in a variety of studies (Collins & White, 2011; Llanes & Muñoz, 2009; Serrano,

2011; Tavakoli & Foster, 2008). The participants were shown a series of pictures representing some children preparing a picnic with their mother first, and then spending some time in the country with their dog, who gets in their picnic basket while they were preparing their sandwiches (see Appendix A). The interviewer allowed the students to become familiar with the story before they were asked to narrate it.

In order to assess students' written production, a descriptive essay was elicited. The students were asked each time to write a description of a person, who, in their first essay was 'their best friend', in their second essay 'someone they admired', and in their third essay 'their best friend in the study abroad context'. The students were given 15 minutes and were asked to write approximately 150 words.

Self-reported data, in the form of a written questionnaire, was used to obtain bio data including information about their language learning history as well as attitudinal data and information about different aspects of their stay abroad. The present study will focus on the questions that elicited information about students' attitudes towards English people and the English language, as well as language contact while abroad (see Appendix B for some examples of key questions).

Procedure

The data collection took place 'in situ', that is, in the study abroad context (University of Southampton). Most studies analyzing gains after an experience abroad tend to assess students' competence when they have returned to their home country. We believe that analyzing students' language production while they are still abroad provides a better

reflection of the actual language gains that take place abroad compared to examining their skills once they have returned back home. First of all, the students are still in contact with the L2 and should have less interference from the L1 and more automatic production of the L2 than when they are in a setting in which the L1 is dominant. Also, depending on how long students are in their home country again before their language production is examined, some of the gains that occurred while abroad might not be as apparent as in the SA context (especially those referring to procedural knowledge, using DeKeyser's (2007) terminology).

Longitudinal data were collected at three time points. The pretest (Time 1) took place towards the beginning of the stay abroad (last week of September). The data collection at Time 2 occurred in December, before the students returned to their home country for the Christmas holidays. Finally, the data at Time 3 were collected in the month of May. Even though the time lapse is longer from Time 2 to Time 3 than from Time 1 to Time 2, it should be borne in mind that the Easter break occurs between Time 2 and Time 3, and most students travel during that break, often to their home country.

The same procedure was followed for all three data collection points by the same researcher (one of the authors of this study). The researcher met with the students individually or in pairs at the university premises where they first completed the oral task, which was recorded in a quiet room with the presence of the researcher only. The students then performed the written task, and the questionnaire was completed at the end of the session. The students spent an average of 20-25 minutes to finish all the tasks.

Analysis

Measures

Oral and written production

The same measures were adopted to analyze oral and written production, except for the case of fluency, for which syllables-per-minute (SPM) was adopted for oral fluency, while words per T-unit (W/T) was used for written fluency. The T-unit has been adopted as the production unit except as otherwise noted.

The T-unit, defined as 'one main clause with all subordinate clauses attached to it' (Hunt, 1965: 20), was developed by Hunt (1965) as an alternative to the sentence, the latter being subject to the learner's knowledge and command of the punctuation system of a specific language. This unit was considered appropriate for this study for the same reason.

Fluency was examined in terms of words per T-unit (W/T), which has been a very frequently used ratio. The total number of words in a sample was divided by the total number of T-units. Several studies have claimed that W/T is a good measure of development in second language writing (Larsen-Freeman, 2006; Larsen-Freeman & Strom, 1977; Wolfe-Quintero, Inagaki, & Kim, 1998). It must be indicated that W/T has sometimes been assumed to measure grammar complexity more than fluency (Norris & Ortega, 2009; Ortega, 2003). Nevertheless, as Cooper (1976) and Wolfe-Quintero et al. (1998) suggest, longer does not necessarily mean more complex. Some evidence for the fact that longer T-units do not need to include more complex clauses is found in

Casanave (1994), who observed that many of her students produced longer and more accurate T-units after some hours of instruction, but they were less complex. Fluency in oral production was examined by means of syllables per minute (SPM), since this measure is generally considered more appropriate for oral fluency than W/T (Griffiths, 1991). For our study, the syllable count did not include false starts, repetitions, self-corrections, unfinished sentences or words in a language other than English.

In order to analyze syntactic complexity, the T-unit complexity ratio (clauses per T-unit, or C/T) has been adopted in this study, and within the term 'clauses', both finite and non-finite clauses were considered. The total number of clauses in a sample was divided by the total number of T-units. Wolfe-Quintero et al. (1998:86) claimed that the majority of the studies they reviewed 'do support the usefulness of the clauses per T-units measure'.

Lexical richness was examined using Guiraud's Index of Lexical Richness: word types divided by the square root of the word tokens (Types/√Tokens). Some studies have shown that this measure is one of the most adequate in analyzing lexical richness in L2 learners' productions (Van Hout & Vermeer, 2007; Vermeer, 2000). In her review of the most commonly used measures of lexical richness in spontaneous speech data, Vermeer (2000) concludes that Guiraud's Index is highly reliable, while the traditionally used Type/Token ratio lacks validity and reliability due to its dependence on text length.

The measure errors per T-unit (Err/T) was adopted in this study in order to examine learners' accuracy. Err/T was obtained by dividing the total number of errors by the total number of T-units. The errors that were considered included lexical, morphological and syntactic errors. Mechanical or pronunciation errors were not taken into account.

It should be borne in mind that the accuracy scale works in the opposite direction from the other measures described above. While a higher number of W/T, SPM, C/T or a higher Guiraud's Index would indicate improvement over time, in the case of Err/T the opposite pattern occurs: fewer errors would indicate more accurate performance over time.

The CLAN program (MacWhinney, 2000) and the Statistical Package for the Social Sciences (SPSS, 2007) were used for the coding and analyses of the writing samples. Three different researchers (the three authors of this study) coded the data for the more objective measures (W/T, C/T, SPM). Inter-rater reliability was calculated for the division of the oral and written samples in T-units, clauses, as well as for errors. In the first two cases, percentage agreement reached 100% (on 15% of the data, coded by all three researchers). For accuracy, which is usually more problematic, two researchers were in charge of the coding. Inter-rater reliability was calculated on 30% of the data, reaching 95% agreement. After all the samples were coded, analyses were performed using SPSS.

Questionnaire

Attitudinal data included six items related to attitudes towards English people and four items related to attitudes towards the English language. All the items used semantic differential five-level scales. The bipolar adjectives regarding English people included: sociable/unsociable, friendly/unfriendly, open/narrow minded, humble/snob, honest/false,

reliable/unreliable. The adjectives regarding the English language included the following: simple/complex, beautiful/ugly, well/badly-sounding, easy/difficult to learn.

Regarding language contact, students were asked to state the type of accommodation they had chosen as well as to state a maximum of four people they had most contact with in their place of residence while abroad (either in their residence hall or apartment/house), a variable that is referred to as 'interaction' in this article. Students were asked to indicate the language of communication and nationality of each person. Students who were living with British families were excluded from the analysis since they were too few (n=2). Students were also asked if there was someone from Spain whom they spent considerable time with while in England. For this 'interaction' variable, two values were calculated, one for the number of reported Spanish-speaking roommates and one for the number of English-speaking roommates. The data obtained through the questionnaire were also analyzed with SPSS.

Statistical analyses

In order to analyze the language progress from Time 1 to Time 2 and Time 2 to Time 3, Wilcoxon Signed Rank Tests were performed, with the different measures of fluency, complexity and accuracy as dependent variables, first for the oral production task and then for the written production task. Non-parametric tests were preferred due to the low number of participants (n=14 in written production; n=13 in oral production).

In the analysis of the self-reported data, the Mann-Whitney U test, also a non-parametric test, was used. Due to the small sample, the 'exact sig.' value (instead of the

'asymp.sig (2-tailed)' value), was used to determine the level of significance of the results, as recommended by Field (2005). Independent variables with more than two levels in the original questionnaire were transformed into two levels due to the small size of the sample.

Results

The results of the different statistical analyses will be presented first for the oral production data, followed by the written production data, and lastly the results of the self-reported data.

Oral production data

The descriptive statistics for the mean scores obtained by the participants in the oral production task appear in Table 1. This table also contains information about the standard deviations in parentheses and the median scores.

[Table 1]

The results of the Wilcoxon Signed Rank Tests performed for each of the measures comparing Time 1–2; Time 2–3; and Time 1–3 appear in Table 2. After the significance value, we also include Cohen's *d* for effect size.

[Table 2]

In view of these results, it appears that one semester abroad was enough for significant progress to occur in certain areas of oral production, namely, fluency and lexical richness. The effect size of these differences is large in the case of fluency and medium-large in the case of lexical richness. In contrast, the progress that the students experienced between the end of the first semester and the end of the second semester was not significant except in accuracy. In this area, the effect size of the difference between Time 2 and Time 3 was large. Considering the whole stay, all the areas of oral production under examination experienced a significant improvement (with the effect size of the differences from Time 1 to Time 3 being large), with the exception of syntactic complexity.

Written production data

Table 3 presents the descriptive statistics, including the means, medians and standard deviations, for the scores obtained by the students at each data collection time for all the measures of written production.

[Table 3]

The results of the Wilcoxon Signed Rank Tests as well as the effect sizes appear in Table 4.

[Table 4]

Unlike the results for oral production, no significant progress was experienced by the students in terms of written production during their first semester abroad. Some significant progress begins to occur from the end of the first semester to the end of the second semester in terms of accuracy and syntactic complexity, and the effect size of these differences is large. However, the most significant development in terms of written production occurs between Time 1 and Time 3, that is when initial and final performance is compared. All four areas under analysis (fluency, syntactic complexity, lexical richness and accuracy) show significant growth and the effect size of the differences between the two time points is large (or medium-large in the case of lexical richness).

Comparing oral and written production

As can be seen from the descriptive and inferential statistics, the progress experienced by the students abroad during the first and second semester in oral and written production differs. Some significant progress in oral production was already apparent in the first semester, but significant improvement in written production did not manifest itself until the second semester. The progress in each of the areas analyzed (fluency, syntactic complexity, lexical richness and accuracy) is represented in Figures 3-6. Note that in the case of fluency (Figure 3), the scores for written fluency (W/T) have been multiplied by 10, in order to have a similar scale to SPM, which makes the relationship more apparent

in the visual representation. Also, as explained above, the accuracy measure (Err/T) is the only one in which lower scores indicate improvement (fewer errors, more accuracy).

[Figure 3]   [Figure 4]

[Figure 5]   [Figure 6]

In Figure 3, it can be observed that, even though students' development of fluency can be said to be linear in both oral and written production, the progress in oral fluency during the first semester is more significant than during the second semester. The opposite is true for written fluency, for which the second semester seems to be especially significant. In the case of syntactic complexity (Figure 4), development is apparent in the case of oral production. However, for written production, syntactic complexity declined at the end of the first semester, but improved by the end of the second semester. Figure 5 also shows that the lexical richness of oral production improved more than written production during the first semester. Finally, Figure 6 indicates that significant development in accuracy did not occur until the second semester.

Self-reported data

In order to examine the relationship between attitudes and linguistic gains, students' answers to the questionnaire at Time 3 and students' gains from Time 1 to Time 3 were examined, as it was expected that attitudes would have more impact on language gains after a longer period of time. Out of the six scales related to attitudes towards English people, significant differences were found in two of the six bipolar adjectives in the

questionnaire. Those students who rated English people as more 'sociable' than 'unsociable' made more gains as regards accuracy in their written production [U= 8, n1= 7, n2=7, Z=-2.11, p=.04]. The same is true for students who rated English people as more 'humble' than 'snob' [U=5, n1=9, n2=5, Z=-2.33, p=.02]. The effect size in both tests was large (Cohen's d = 0.59 and 0.65 respectively).

Out of the four scales related to attitudes towards the English language, significant differences were found in one of the four bipolar adjectives in the questionnaire. Those students who rated English as more 'complex' than 'simple' made more gains in the lexical measure both in their written [U=9, n1=8, n2=6, Z=-1.94, p=.05] as well as their oral production [U=6, n1=6, n2=7, Z=-2.14, p=.03]. In both tests, the effect size was large (Cohen's d=0.52 and 0.59 respectively).

In examining the relationship between language contact and learning gains, gains from Time 1-3 were used in the analyses of two variables that remained constant throughout the academic year: accommodation and contact with a close Spanish friend during the academic stay. Gains from Time 1-2, Time 2-3 as well as Time 1-3 were used in the analysis of a variable that was more liable to change from Semester 1 to Semester 2, which is the linguistic profile of the people students had more contact with in their residence hall or apartment. As regards accommodation, it was found that there were significant differences between students who were living in apartments/houses and those living in a residence hall, with the former having more gains between Time 1 and Time 3 in lexical richness (oral production) [U=3, n1=4, n2=7, Z=2.08, p=.04; Cohen's d=0.63]. Results also indicated that students who did not tend to have someone from Spain with whom they did almost everything experienced more gains in lexical richness (written

production) [U=4, n1=5, n2=8, Z=-2.34, p=.02] and accuracy (oral production) [U=5, n1=5, n2=8, Z=-2.03, p=.05]. The effect size in both tests was large (Cohen's d=0.62 and 0.72 respectively). Whether students were living with only English-speaking people or with one or more Spanish-speaking mates turned out to be significant in the lexical richness measure (written production) between Time 2-Time 3 [U=7, n1=6, n2=7, Z=-2.0, p=.05; Cohen's d=0.5], with those living with only English-speaking people experiencing more gains. No significant differences were found in the oral production measures between Time 1 and Time 2 or between Time 1 and Time 3.

Discussion

In answer to the first research question, our results seem to suggest that L2 proficiency in oral and written production while abroad develop in somewhat different ways. The longitudinal design made it possible to observe that the 14 English learners that were examined in this study made significant progress in some areas of oral production (namely fluency and lexical richness) at the end of the first semester abroad, while no parallel improvement was registered in terms of written production. It is especially interesting that the areas that seem to improve the most after one semester abroad coincide with what most studies in the literature seem to suggest as the areas for which spending time abroad could be especially beneficial, namely oral fluency (Freed, 1995; Lennon, 1990; Llanes & Muñoz, 2009; Segalowitz & Freed, 2004; Yager, 1998) and vocabulary (Ife et al., 2000; Milton & Meara, 1995). Similarly, the findings from this study concerning development from Time 1 to Time 2 in terms of writing are in line with

results reported by other researchers in which the study abroad context is not found to be particularly helpful for the development of written production (Freed et al., 2003). Indeed, most of the studies from the literature that report advantages in written production for SA students seem to analyze long periods of time (Sasaki, 2004; 2007; 2009). To our knowledge, there is only one study (Pérez-Vidal & Juan-Garau, 2009) that shows significant improvement in some aspects of written production after a relatively short experience abroad (three months). However, the improvement was observed on only two of the five measures considered to analyze fluency, complexity and accuracy.

It is from Time 2 to Time 3 that students' oral accuracy improved. It seems as if the students benefited first from the SA context in terms of fluency and lexical richness, and only later does this progress extend to accuracy. The period between Time 2 and Time 3 is also when accuracy in written production develops significantly. The implications from these findings are that for L2 accuracy to develop, longer stays might be necessary in some cases, which could also explain why some studies focusing on accuracy in short-term stays have found little or no improvement (DeKeyser, 2010). It might also be the case, as has been found in previous studies, that other areas (or 'sub-systems') need to develop before a development in accuracy can occur (Caspi, 2010).

Considering the progress experienced throughout the whole academic year (from Time 1 to Time 3), the results reported in this study are quite hopeful for the SA experience, as significant improvement occurs in almost all the areas of oral and written production under analysis. These results could imply that the reason a clear advantage has not been unanimously reported in the literature might be related (among other possible

factors, of course) to the short-term stays that tend to be analyzed (usually one semester or less).

Another objective of this study was to analyze whether some attitudinal and interactional factors were associated with the progress the students experienced abroad. In answer to the second research question we can say that several factors appear to have a certain relation with language development. We have found that some attitudes towards the L2 (English is a complex language) or the people who speak it (English people are sociable and humble rather than snob) were associated with gains in accuracy and lexical richness. The reason why the adjective 'sociable', as well as 'humble', was a key adjective could be due to the fact that those students who considered English people sociable and not snobbish might have interacted more with them, which contributed in turn to language gains. It is also interesting that those who found the English language more complex were the ones who made more gains in lexical richness. Probably, these learners paid more attention to complexity, were challenged by this feature of the language, and as a result their production was more complex in terms of vocabulary. Nevertheless, it should be emphasized that not all the attitudes towards the English language or English people under analysis were associated with language gains. This might be due to the fact that the choice of adjectives (which was done through an Internet search of stereotypes of British people by foreigners) might have not been exhaustive enough or that some adjectives included might refer to attitudes that have less effect on L2 learners' use of the language. More studies should analyze attitudes in a more detailed way in order to establish a clearer relationship between this variable and language gains abroad, as the present study has demonstrated that this is an area worth exploring in depth.

Moreover, our results also suggest that living arrangements also seemed to have a role in the progress experienced by the students. Similarly, those students who did not spend most of their time with a Spanish student improved their lexical richness more than those who did. In fact, these two situations likely lead to more possibilities for interaction in the L2, and such use/practice is probably responsible for the language improvement.

Although the current study was not designed within the Dynamic Systems Theory (DST), the results that have been reported here can be explained using some of the major tenets of this theory (de Bot, 2008; Larsen-Freeman & Cameron, 2008; Verspoor, de Bot, & Lowie, 2011). Indeed, under a DST perspective, language development is seen as the interaction between a wide variety of internal and external factors that can be grouped in different levels and sub-levels (Lowie, Verspoor, & de Bot, 2009). Among these levels, Lowie et al. highlight the social, psycholinguistic, cultural and linguistic level. In the present study, we are considering all of these levels and how they interact: we have examined how the learning context (which encompasses socio-cultural factors) may be related to language development. Additionally, we have analyzed different sub-systems within the linguistic level: fluency, syntactic complexity, lexical richness and accuracy both in oral and written mode. Our findings certainly demonstrate that there is an interaction between the different levels and sub-levels.

Furthermore, the results of the present investigation suggest that some sub-systems develop faster than others in the SA context: globally, it seems that progress occurs earlier in oral production than in written production. Then, in the case of oral production, it seems that the development of accuracy is slower, perhaps requiring other areas to develop before it (namely fluency and lexical richness). This finding is in line with

Caspi's (2010), whose study suggests that the development of both lexical and syntactic complexity precedes the development of lexical and syntactic accuracy, which is explained by the 'nestedness' and hierarchical structure of dynamic systems (van Geert, 1995). As suggested by Caspi (2010), these two characteristics (nestedness and especially hierarchical structure), which are typical of language development according to DST, can explain why, for example, vocabulary acquisition is a prerequisite for the development of syntactic complexity.

For all the above-mentioned reasons, we consider that DST offers an appropriate framework to investigate language development abroad and further studies should be conducted that examine the SA context under a DST perspective, ideally with more data collection points than the present study, as well as including more information about other variables and individual development of participants instead of focusing on group means.

Conclusion

In conclusion, considering the results reported in the present investigation, it can be said that the SA context potentially provides an advantageous situation for students to improve their L2 skills. Nevertheless, the word 'potentially' must be emphasized here, as not all learners will necessarily find such a context beneficial, as studies with larger groups of participants and different measures of socio-cultural and individual variables may reveal. According to the findings from this study, length of stay is an influential variable in terms of the progress that is to be expected for oral and written skills. More

time is necessary for measureable progress in written production to occur than it is for oral production. The findings of this study are certainly innovative in this respect since they contribute to the debate on whether the SA context is beneficial for written development or not. According to our findings, written development can occur while abroad; however, a substantial amount of time in the L2 country (in this study, two full semesters) is necessary before such development can take place. Similarly, our results also suggest that attitudes as well as type of interaction can also influence linguistic improvement to a certain extent.

In this study, we have only analyzed three factors that can contribute to the L2 development in the SA context, namely duration of the stay, attitudes, and living arrangements. We are also aware that there are many other factors that can determine whether the potential of the SA context materializes: initial proficiency level, which according to DeKeyer (2007; 2010) is crucial, aptitude (DeKeyser, 2010), and motivation (Dwyer, 2004; Isabelli-Garcia, 2006), to name a few. Future studies should concentrate on different individual factors and relate them to the kind of progress expect that occurs abroad.

More longitudinal studies like the one reported here (and ideally including more L2 samples) are also necessary in order to gain a better insight into L2 development in the SA context. As in the present study, it is important that longitudinal analyses include a variety of measures that tap different areas of language proficiency so as to understand better what L2 aspects are more likely to improve after a stay abroad experience and in which order L2 gains should be expected to appear.

Acknowledgements

References

Allen, H. W. (2010).   Language-learning motivation during short-term Study Abroad: An activity theory perspective. *Foreign Language Annals, 43,* 27-49.

Association of Universities and Colleges of Canada [AUCC]. (2007). Canadian universities and international student mobility. Retrieved January 2010, from http://www.aucc.ca/_pdf/english/publications/student_mobility_2007_e.pdf.

Brecht, R., D. Davidson, D. E. & Ginsberg, R. B. (1995). Predictors of Foreign Language Gain During Study Abroad. In B. F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 37-66). Amsterdam: John Benjamins.

Casanave, C. P. (1994). Language development in students' journals. *Journal of Second Language Writing, 3*, 179-201.

Caspi, T. (2010). A dynamic perspective on Second language development. PhD. Dissertation, University of Groningen.

Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, *26*, 227-248.

Collentine, J. (2009). Study abroad research: Findings, Implications and Future Directions. In C. Doughty & M. Long (Eds.), *Handbook of language teaching* (pp. 218-233). Malden, MA: Blackwell Publishing, Ltd.

Collins, L., & White, L. (2011). An intensive look at intensity and language learning. *TESOL Quarterly, 45,* 106-133.

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research, 69,* 176-183.

Cubillos, J. H., Chieffo, L. & Fan, C., 2008. The impact of short-term study abroad programs on L2 listening comprehension skills. *Foreign Language Annals*, *41*, 1, 157–185.

de Bot, K. (2008). Introduction: Second language development as a dynamic process. *The Modern Language Journal, 92,* 166-178.

DeKeyser, R. M. (1991). Foreign language development during a semester abroad. In B. F. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp.104-119). Lexington, Mass.: D. C. Heath.

DeKeyser, R. M. (2007). Study abroad as foreign language practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208-226). Cambridge: CUP.

DeKeyser, R. M. (2010). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals, 43,* 80-92.

Dewey, D. P. (2004). A comparison of reading development by learners of Japanese in intensive and domestic Immersion and Study Abroad contexts. *Studies in Second language Acquisition*, *26*, 303-327.

Dewey, D. P. (2008). Japanese Vocabulary Acquisition by Learners in three contexts. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *XV*, 127-148.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, *26*, 249-273.

Dwyer, M. M. (2004). More Is Better: The Impact of Study Abroad Program Duration. *The interdisciplinary Journal of Study Abroad, X*, 151-163.

Dyson, P. (1988). The year abroad. Report for the Central Bureau for Educational Visits and Exchanges. Oxford University Language Teaching Centre.

Field A. (2005) *Discovering statistics using SPSS* (2nd ed). London: Sage.

Foster, P. (2009). Lexical Diversity and Native-Like Selection: The Bonus of Studying Abroad. In B. Richards, M. Daller, D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition* (pp. 91-106). Hampshire: Palgrave Macmillan.

Freed, B. F. (1995). What makes Us Think that Students Who Study Abroad Become Fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam: John Benjamins.

Freed, B. F. (1998). An overview of issues and research in language learning in a study abroad setting. *Frontiers*: *The interdisciplinary Journal of Study Abroad*, *IV*, 31-60.

Freed, B. F., Segalowitz, N. & Dewey D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, *26*, 275-301.

Freed, B. F., So, S. & Lazar, N. A. (2003). Language learning abroad: How do gains in written fluency compare with gains in oral fluency in French as a second language? *ADFL Bulletin*, *34*, 3, 34-40.

Griffiths, R. (1991). Pausological research in an L2 context: A rationale, and review of selected studies. *Applied Linguistics, 12*, 345-364.

Guntermann, G. (1995). The Peace Corps Experience: Language Learning in Training and in the Field. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 149-170). Amsterdam: John Benjamins.

Heaton, J. (1966). *Composition through pictures*. Essex: Longman.

Howard, M. (2005). On the role of context in the development of learner language: Insights from study abroad research. *ITL International Journal of Applied Linguistics*, *148*, 1-20.

Howard, M. (2006). The expression of number and person through verb morphology in French interlanguage. *International Review of Applied Linguistics*, *44*, 1, 1-22.

Hunt, K. W. (1965). *Differences in grammatical structures written at three grade levels.* Research Report No. 3. Urbana, IL: National Council of Teachers of English.

Ife, A., Vives, G. & Meara, P. (2000). The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics*, *4*, 1, 55-84.

Isabelli-García, C. (2006). Study Abroad Social Networks, Motivation and Attitudes: Implications for Second Language Acquisition. In M. A. DuFon & E. Churchill (Eds.), *Language learners in study abroad contexts* (pp. 231-258). Clevedon: Multilingual Matters.

Ismail, B., Morgan, M. & Hayes, K. (2005). Effect of short study abroad course on student openness to diversity. *Journal of Food Science Education*, *1*, 15-18.

Juan-Garau, M. & Pérez-Vidal, C. (2007). The effect of context and contact on oral performance in students who go on a stay abroad. *VIAL*, *4*, 117-134.

Kinginger, C. (2008). Language learning in study abroad: Case histories of Americans in France. *Modern Language Journal Monograph Series*, 1.

Lapkin, S., Hart, D. & Swain, M. (1995). A Canadian Interprovincial Exchange: Evaluating the Linguistic Impact of a Three-Month Stay in Quebec. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 67-94). Amsterdam: John Benjamins.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese Learners. *Applied Linguistics, 27,* 590-619.

Larsen-Freeman, D., & Cameron, L. (2008). Research methodology on language development from a complex systems perspective. *Modern Language Journal, 92,* 200-213. Oxford: Oxford University Press.

Larsen-Freeman, D., & Strom, V. (1977). The construction of a second language acquisition index of development. *Language Learning, 27*, 123-134.

Lennon, P. (1990). Investigating Fluency in EFL: A Quantitative Approach. *Language Learning*, *40*, 387-417.

Llanes, À. (2011). The many faces of study abroad: An update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism, 8,* 189-215.

Llanes, À. & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, *37*, 3, 353-365.

Lowie, W.M., Verspoor, M.H. & de Bot, K. (2009). A dynamic view of second language development across the lifespan. In K. de Bot & R. Schrauf (Eds.), *Language development over the lifespan* (pp. 125-146). New York/London: Routledge.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Masgoret A. M. & Gardner R.C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning, 53*, 167-210.

Milton, I. & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITI Review of Applied Linguistics*, *107*, 8, 17-34.

Mora, J. C. (2008). Learning Context Effects on the Acquisition of a Second Language Phonology. In C. Pérez-Vidal (Coord.), M. Juan-Garau & A. Bel (Eds.), *A portrait of the young in the new multilingual Spain* (pp. 241-263). Clevedon: Multilingual Matters.

Muñoz, C. (Ed.) (2006). *Age and the rate of foreign language learning*. Clevedon, UK: Multilingual Matters.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30,* 555-578.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*, 492-518.

Pérez-Vidal, C. & Juan-Garau, M. (2009). The effect of study abroad (SA) on written performance. *EUROSLA Yearbook, 9,* 269-295.

Polanyi, L. (1995). Language Learning and Living Abroad: Stories from the Field. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 271-292). Amsterdam: John Benjamins.

Regan, V. (1995). The acquisition of sociolinguistic native speech norms: Effects of a year abroad on second language learners of French. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 245-268). Amsterdam: John Benjamins.

Regan, V. (2005). From speech community back to classroom: What Variation Analysis can tell us about the role of context in the acquisition of French as a foreign language. In J-M. Dewaele (Ed.), *Focus on French as a foreign language* (pp. 191-209). Clevedon: Multilingual Matters.

Regan, V., Howard, M. & Lemée, I. (Eds.) (2009). *The acquisition of sociolinguistic competence in a study abroad context.* Clevedon: Multilingual Matters.

Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, *54*, 525 - 582.

Sasaki, M. (2007). Effects of Study-Abroad experiences on EFL writers: A multiple-data analysis. *Modern Language Journal* 91, IV, 602–620.

Sasaki, M. (2009). Changes in English as a foreign language students' writing over 3.5 years: A sociocognitive account. In R. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 49-76). Clevedon: Multilingual Matters.

Segalowitz, N. & Freed, F. B. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, *26*, 173-199.

Serrano, R. (2011). The time factor in EFL classroom practice. *Language Learning, 61*, 117-145.

SPSS 16.0 for Windows [Computer software] (2007). SPSS Inc.

van Geert, P. (1995). Growth dynamics in development. In R. F. Port & T. J. Van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 313-338). Cambridge, MA: The MIT Press.

van Hout, R. & Vermeer, A. (2007). Comparing measures of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary* (pp. 93-115). Cambridge: CUP.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing, 17,* 65-83.

Verspoor, M. H, de Bot, K., & Lowie, W. (2011). *A dynamic approach to second language development: Methods and techniques*. Amsterdam: John Benjamins.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity. Technical Report 17.* Manoa, Hawai'i: University of Hawai'i Press.

Yager, K. (1998). Learning Spanish in Mexico: The effect of informal contact and student attitudes on language gain. *Hispania*, *81*, 898-913.

Appendix A: Oral narrative (Heaton, 1966)
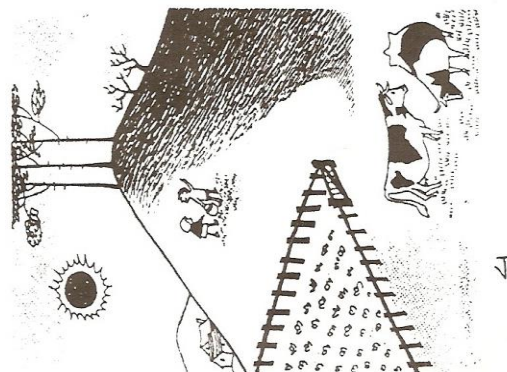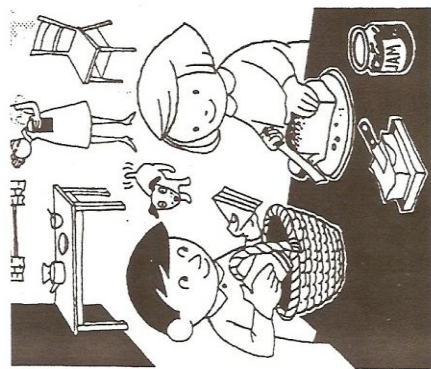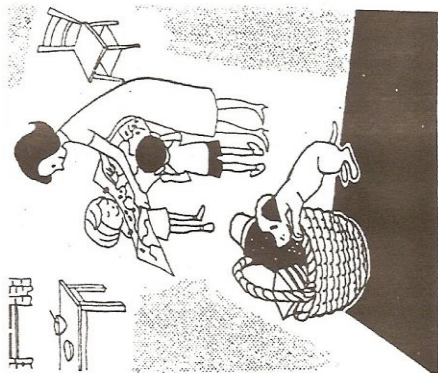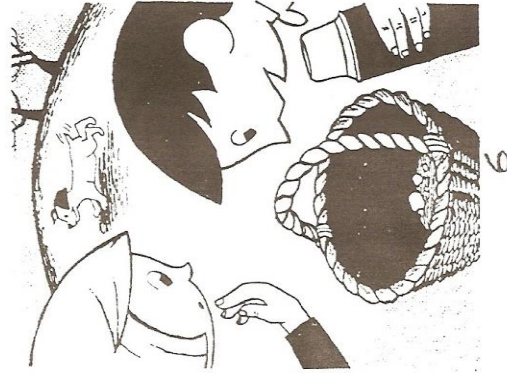
Appendix B: Sample of key questions included in the questionnaire (original version in Spanish)

Name …………………… Age………………

1. How would you classify your level of English in relation to the following skills?
   (1) elementary (2) pre-intermediate (3) intermediate (4) upper-intermediate (5) advanced

listening	speaking	reading	writing

2. Describe very briefly how you've learned English until now (starting age, stays abroad, language schools, etc.)

3. When was the last time you followed an English language course? Where was it? How long did it last?

4. In general, what do you think about English people?

sociable	____ ____ ____ ____ ____	unsociable
friendly	____ ____ ____ ____ ____	unfriendly
open-minded	____ ____ ____ ____ ____	narrow-minded
humble	____ ____ ____ ____ ____	snob
honest	____ ____ ____ ____ ____	false
reliable	____ ____ ____ ____ ____	unreliable

5. And about the English language?

simple	____ ____ ____ ____ ____	complex
beautiful	____ ____ ____ ____ ____	ugly
well-sounding	____ ____ ____ ____ ____	bad-sounding
easy to learn	____ ____ ____ ____ ____	difficult to learn

6. Where are you living now?
   - Single room in a residence hall
   - Shared room in a residence hall
   - Shared student apartment/house
   - Single room in a private house

7. Make a list of the four people you've had the most contact with while abroad (at home, university or weekends).

Relationship (classmate, friend, roommate, etc.)	How much contact have you had with this person? A little/Some/A lot	Nationality	Language used for interaction

8. Among these people, is there anyone you've done almost everything with? If the answer is 'yes', add a 'X' next to her/him.

Figures

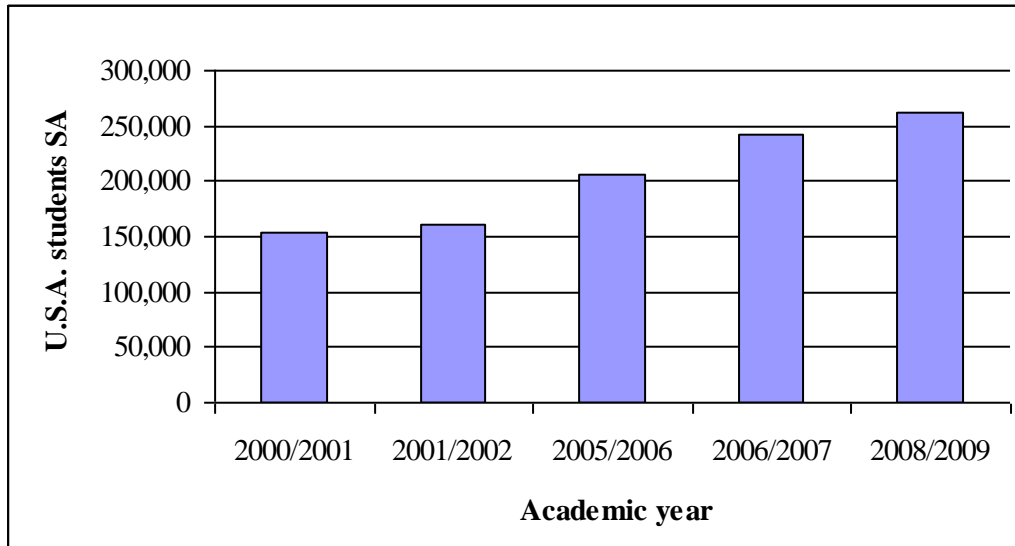Figure 1: Number of U.S.A. students staying abroad

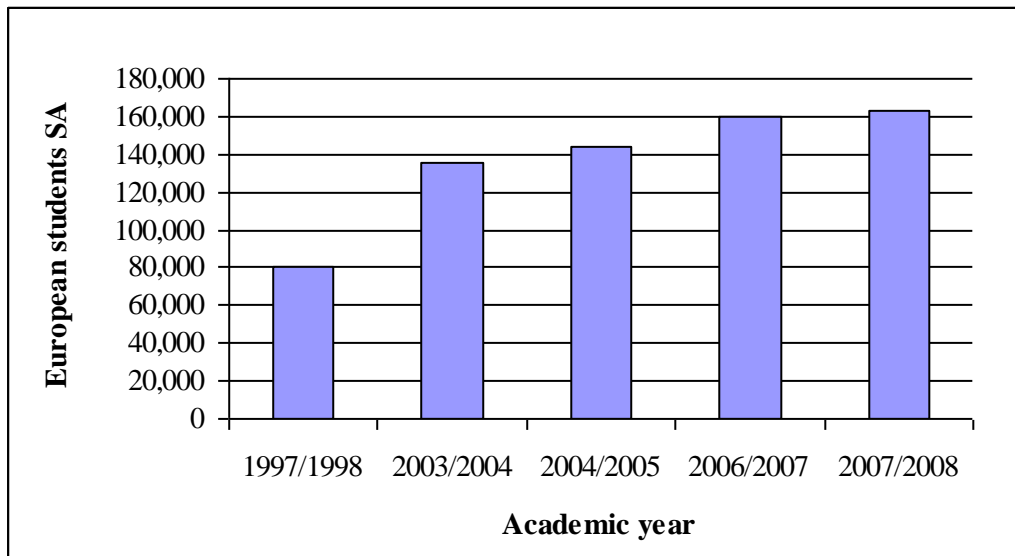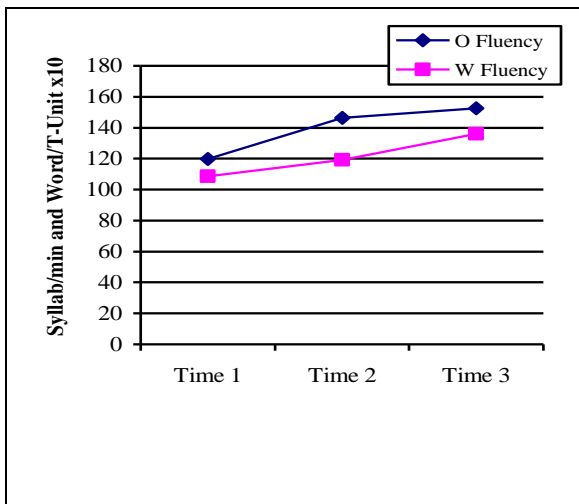Figure 2: Number of participants in Erasmus programs in Europe
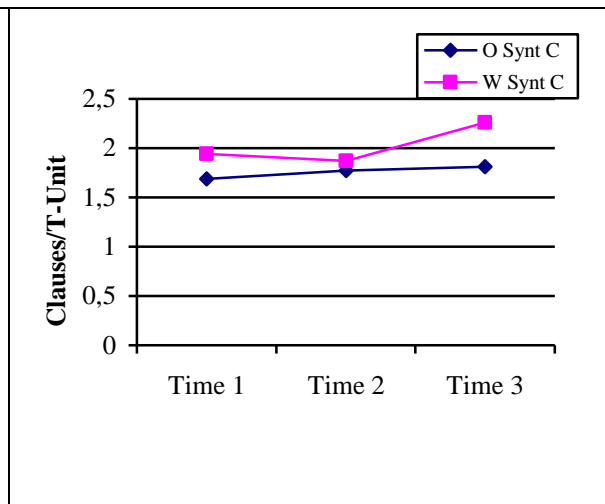
Figure 3 Fluency
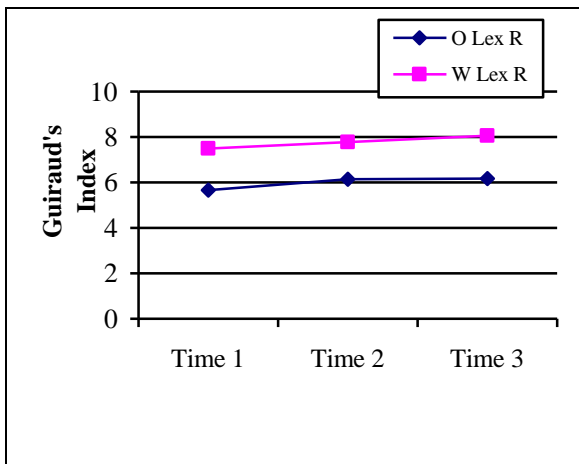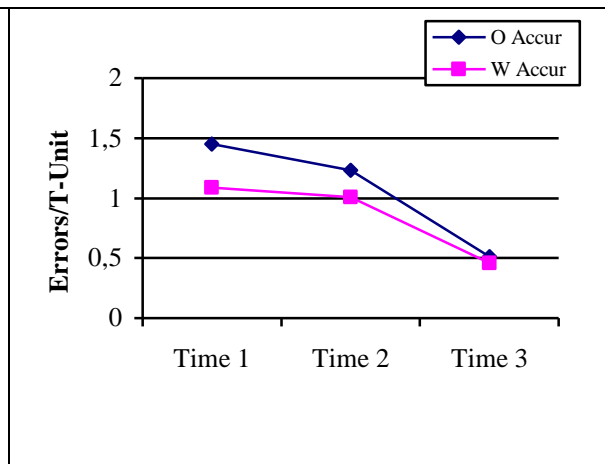
Figure 4 Syntactic complexity

Figure 5 Lexical richness

Figure 6 Accuracy

Tables

TABLE 1
Descriptive statistics oral production

	TIME 1	TIME 2	TIME 3
Fluency: Syllables/minute			
Mean (Standard deviation)	119.89 (25.81)	146.22 (30.48)	152.45 (33.57)
Median	122.42	145.61	147.25
Syntactic Comp.: Clauses/T-Unit			
Mean (Standard deviation)	1.69 (0.20)	1.77 (0.30)	1.81 (0.27)
Median	1.64	1.72	1.74
Lexical Comp.: Guiraud's Index			
Mean (Standard deviation)	5.66 (0.91)	6.13 (0.83)	6.16 (0.59)
Median	5.83	5.92	6.11
Accuracy: Errors/T-Unit			
Mean (Standard deviation)	1.45 (0.94)	1.23 (0.72)	0.51 (0.33)
Median	1.07	1.29	0.50

TABLE 2
Inferential statistics and effect size oral production

	TIME 1 –TIME 2	TIME 2 – TIME 3	TIME 1 – TIME 3
Fluency	$Z=-3.18$	$Z=-1.01$	$Z=-2.90$
(Syllables/minute)	**$p=.001$**	$p=.311$	**$p=.004$**
	Cohen's $d=0.93$	Cohen's $d=0.10$	Cohen's $d=1.08$
Syntactic Comp.	$Z=-0.80$	$Z=-0.31$	$Z=-1.64$
(Clauses/T-Unit)	$p=.424$	$p=.753$	$p=.101$
	Cohen's $d=0.31$	Cohen's $d=0.14$	Cohen's $d=0.50$
Lexical Comp.	$Z=-2.48$	$Z=-0.25$	$Z=-1.99$
(Guiraud's Index)	**$p=.013$**	$p=.807$	**$p=.046$**
	Cohen's $d=0.54$	Cohen's $d=0.04$	Cohen's $d=0.65$
Accuracy	$Z=-1.25$	$Z=-2.34$	$Z=-2.55$
(Errors/T-Unit)	$p=.209$	**$p=.019$**	**$p=.011$**
	Cohen's $d=-0.43$	Cohen's $d=-1.28$	Cohen's $d=-1.33$

TABLE 3
Descriptive statistics written production

	TIME 1	TIME 2	TIME 3
Fluency: Words/T-Unit			
Mean (Standard deviation)	10.85 (2.43)	11.91 (2.00)	13.60 (2.49)
Median	10.57	11.76	13.13
Syntactic Comp.: Clauses/T-Unit			
Mean (Standard deviation)	1.94 (0.38)	1.87 (0.31)	2.26 (0.39)
Median	1.93	1.92	2.26
Lexical Comp.: Guiraud's Index			
Mean (Standard deviation)	7.50 (0.95)	7.78 (0.84)	8.07 (1.02)
Median	7.50	7.64	8.32
Accuracy: Errors/T-Unit			
Mean (Standard deviation)	1.09 (0.72)	1.01 (0.79)	0.46 (0.28)
Median	1.08	0.68	0.44

TABLE 4
Inferential statistics and effect size written production

	TIME 1 –TIME 2	TIME 2 – TIME 3	TIME 1 – TIME 3
Fluency	Z=-1.60	Z=-1.85	Z=-2.60
(Words/T-Unit)	p=.109	p=.064	**p=.009**
	Cohen's d=0.47	Cohen's d=0.75	Cohen's d=1.11
Syntactic Comp.	Z=-0.22	Z=-2.41	Z=-1.99
(Clauses/T-Unit)	p=.826	**p=.016**	**p=.046**
	Cohen's d=-0.20	Cohen's d=1.10	Cohen's d=0.83
Lexical Comp.	Z=-0.91	Z = -1.16	Z = -2.04
(Guiraud's Index)	p=.363	p = .245	**p = .041**
	Cohen's d=0.31	Cohen's d = 0.31	Cohen's d = 0.58
Accuracy	Z=-0.53	Z=-2.04	Z=-2.16
(Errors/T-Unit)	p=.594	**p=.041**	**p=.030**
	Cohen's d=-0.10	Cohen's d=-0.92	Cohen's d=-1.15