THE EFFECT OF PROGRAM TYPE AND PROFICIENCY LEVEL ON LEARNERS' WRITTEN PRODUCTION

RAQUEL SERRANO SERRANO[*]
Universitat de Barcelona

ABSTRACT. *The aim of this study is to analyze the effects of proficiency level and time distribution of instructional hours on adult second language learners' written production. Two English as a foreign language programs were considered: "regular" and "intensive". A total of 292 writing samples from 83 intermediate-level learners and 63 advanced learners were analyzed by means of different measures of fluency, complexity, and accuracy. Additionally, a group of native English speakers (N = 29) was included to provide baseline data. The results of the statistical analyses suggest positive effects of time concentration for the development of fluency and lexical complexity at the intermediate level. The results of the students at the advanced level are discussed on the basis of their performance with respect to the native English speakers.*

KEY WORDS. *Time distribution, intensive instruction, L2 proficiency, L2 writing, fluency, complexity, accuracy.*

RESUMEN. *El objetivo de este estudio es analizar el efecto del nivel de competencia lingüística y la distribución de las horas de instrucción en la producción escrita de estudiantes adultos de segundas lenguas. Se han considerado dos programas de inglés como segunda lengua: "regular" e "intensivo". Se han analizado 292 redacciones de 83 estudiantes de nivel intermedio y 63 de nivel avanzado utilizando diferentes medidas de fluidez, complejidad y corrección. Además, se ha incluido un grupo de nativos anglófonos (N = 29) como referencia. Los resultados de los análisis estadísticos sugieren un efecto positivo para la intensidad en fluidez y riqueza léxica para los estudiantes de nivel intermedio. Los resultados obtenidos por los estudiantes de nivel avanzado se discutirán en relación a la producción escrita de los nativos.*

PALABRAS CLAVE. *Distribución del tiempo, instrucción intensiva, competencia lingüística en L2, escritura en L2, fluidez, complejidad; corrección.*

1. INTRODUCCIÓN

This study attempts to shed some light on how different types of instruction in terms of time distribution ("regular" vs. "intensive") affect the development of English

proficiency as reflected in writing for learners at an intermediate and an advanced level. In order to examine this development, different measures of fluency, complexity, and accuracy have been adopted, which have been used and examined by many researchers in the field as the main object of analysis (Wolfe-Quintero, Inagaki, and Kim 1998; Larsen-Freeman 2009; Norris and Ortega 2009; Skehan 2009; Pallotti 2009), as well as in terms of how they interact with other factors both in written (Torras, Navés, Celaya, and Pérez-Vidal 2006; Ishikawa 2007; Kuiken and Vedder 2008; Celaya and Navés 2009) and oral production (Skehan 1998; 2009; Skehan and Foster 1999; Robinson 2001, 2003; Gilabert 2007; Tavakoli and Foster 2008; Robinson, Cadierno and Shirai 2009). The special issue edited by Housen and Kuiken (2009) in *Applied Linguistics* provides evidence for the interest in these aspects of learners' production, as well as for the need to conduct more studies on the topic. The present study contributes to this second line of research in providing more evidence on how these three aspects of L2 learners' written performance develop in relation to other factors, namely the distribution of the instructional hours they receive (regular/intensive), as well as learners' initial proficiency level (intermediate/advanced). The literature on the effect of time distribution of L2 instructional hours on learners' L2 acquisition is certainly scarce, and even more so are studies that analyze learners' written performance, or learners of different initial proficiency levels. Serrano (2011) compared the language *gains* experienced by learners in intensive and regular programs in a variety of language skills and outlined a possible cognitive explanation for why, for certain learners, intensive contact with the L2 could be advantageous for overall L2 development. The present study focuses exclusively on writing and can therefore examine in depth learners' written production in relation to program type and proficiency level. Moreover, the present study includes a control group of native English speakers with the aim of providing baseline data that could be useful, especially in examining the performance of the learners at the advanced level, which could not be explored in Serrano (2011).

Research on time distribution and L2 learning has been performed mainly in the Canadian context and in the case of primary school children. The learners in intensive programs (which include around 400 hours of L2 instruction in one year) have been reported to make more progress than their peers in regular, "drip-feed" courses (Spada and Lightbown 1989; Lightbown and Spada 1994; Lapkin, Hart, and Harley 1998; Netten and Germain 2004; White and Turner 2005). Additionally, analyses of time concentration within intensive English courses have suggested that the learners in the more concentrated or "massed" program tend to make more L2 gains, but the differences between more or less concentrated intensive courses are not always significant (Collins *et al*. 1999; Collins and White 2011). Some of these studies in the Canadian context have focused especially on writing with similarly encouraging findings for intensive instruction (Lightbown and Spada 1997; Lapkin *et al*. 1998; Germain, Netten, and Séguin 2004). The few studies investigating the effects of time distribution on foreign language learning by adults also tend to report advantages in the

case of more concentrated programs in different skills, including writing (McKee 1983; Hinger 2006; Serrano and Muñoz 2007).

When considering the learners' proficiency level, some studies have examined language gains in intensive courses for learners who started at different language proficiency levels (Gardner, Smythe, and Brunet 1977). The evidence from these studies suggests that intensive foreign language instruction has a more positive effect for students with lower initial skill in the foreign language.

In sum, this study aims to fill the gap in different areas of SLA and L2 writing which have not been commonly investigated. First, L2 development as manifested in writing will be examined for intermediate and advanced learners attending different types of English as a foreign language (EFL) programs (regular vs. intensive). Studies of this kind are certainly relevant in foreign language contexts in order to determine the type of instruction that enhances learners' development of L2 proficiency as represented in written production in situations in which the exposure to the L2 is limited. Moreover, both the population examined in this study (adult EFL learners), and the type of intensive program under investigation (one-month intensive course) have not received much attention in the literature on intensive L2 learning. However, research on this area is necessary, considering that the intensive programs that adults tend to follow –mostly because of time constraints– are typically one-month long (and not five or ten months, as is the case of the Canadian programs for school children previously mentioned). Furthermore, the possible effects of initial proficiency level will be considered on L2 learners' written production, which has indeed been analyzed in the L2 literature on writing (Larsen-Freeman 1983; Wolfe-Quintero *et al*. 1998; Ortega 2003), but not with respect to other variables, such as type of L2 program. An additional aim of this study is to analyze how the written performance of advanced learners of English compares to the performance of native English speakers. By analyzing the advanced learners' production in relation to a baseline of native English speakers we could have a better insight on how the proposed measures of written production work at the advanced level and whether there are ceiling effects.

More specifically, the present study will try to provide answers to the following research questions:

1. Are there interaction effects that relate time distribution and proficiency level in explaining EFL learners' development of written fluency, complexity, and accuracy?

2. Regardless of learners' proficiency level, do intensive courses enhance EFL learners' development of written fluency, complexity, and accuracy?

3. How does the written production of advanced English learners in terms of the proposed measures of fluency, complexity, and accuracy compare to the performance of native English speakers?

2. METHOD

2.1. *Programs*

Two different EFL programs were considered for this study (regular and intensive), which were offered at the same language school in Barcelona, Spain. The two programs differed in terms of how the hours of instruction were distributed, but not in the total amount of hours they offered. The regular course offered 110 hours of instruction throughout the whole academic year (26 weeks), distributed in two 2-hour sessions a week. The intensive program consisted of the same number of hours of instruction (110) but distributed in 5-hour sessions, five days a week during four and a half weeks in the summer. The methodology used in the two program types was highly similar (same syllabus, same textbook, same exams, etc.). In general, the two types of courses concentrated on the four different skills through a combination of language-centered exercises and other more communicative activities. However, the main focus of these classes was on language forms. The courses under study gave a certain prominence to writing skills (both program types equally), and the students' writings (the ones performed throughout the course plus the one included in the final exam) constituted approximately 25% of their total final grade for the course.

2.2. *Participants*

The written performance of students at two different proficiency levels was analyzed: intermediate (N = 83), which would correspond to level B1 from the Common European Framework of Reference; and advanced (N = 63), which would correspond to level B2/C1 (Council of Europe, 2001). These learners belonged to 14 different groups/classes, 8 classes for the regular program and 6 for the intensive (see Table 1 for details). With respect to the instructors, each of the groups included in this study had a different teacher with the exception of the "Advanced Regular", in which one teacher taught two different classes. Since it was not possible to obtain intensive and regular classes taught by the same teacher, arrangements were made so as to obtain a wide variety of instructors. This variation may neutralize the influence a specific instructor might have. Proficiency was defined by "level" within the language school under study.

	Students	Groups	Teachers
Intermediate Regular	45	4	4
Intermediate Intensive	38	3	3
Advanced Regular	32	4	3
Advanced Intensive	31	3	3
Total EFL students	146	14	13
Native English Speakers (NESs)	29	–	–

Table 1. *Participants*.

Most of the participants included in this research were university students whose ages ranged from 18 to 23 years old. They were, for the most part, bilingual speakers of Spanish and Catalan. The majority of them were taking English classes in order to obtain elective credits, although there were a few young professionals who needed to learn English for their job. All the students were comparable in terms of motivation and previous experience with English.

A further group of participants was considered in order to provide a baseline for comparison and to assess ceiling effects at the advanced level. This group included 29 native English speakers (NESs), who were college students from two American universities. These students' age, as well as fields of study, was comparable to the EFL learners considered for this particular research. It must be emphasized that the role of the NESs in this study is not to provide a model that EFL learners should imitate, but rather to gauge ceiling effects in the particular measures under study. As Pallotti (2009: 589) suggests,

> "native speakers' baseline data are crucial, not because learners' aim is necessarily to behave like native speakers, but because looking at what native speakers do may overcome the researchers' bias toward seeing learners as defective language users, who always need to 'do more'".

Consequently, the English data coming from these NESs will be especially useful to examine the performance of the advanced learners in this study. It could certainly be the case that the advanced learners do not show much progress from pre- to posttest because their performance in the pretest in some areas could already be "native-like", and thus these learners do not have much room for progress.

2.3. *Tasks and Procedure*

The same data collection procedure was followed for the two different types of programs and for the two proficiency levels under study. The author of the present study as well as some research assistants were in charge of collecting the learners' samples. First of all, towards the beginning of the course (approximately 20 hours after the classes had started) the students were asked to write a composition of approximately 150 words with the title "My best friend". Then, 80 hours after the pretest, the students wrote another 150-word composition, this time with the title "Someone I admire". These topics were considered appropriate as students of different proficiency levels could write on them as the content is quite accessible. Moreover, they are similar to the typical descriptive essays that are required in language courses. The students wrote these two compositions in 15 minutes during class time, either at the beginning or at the end of the class. Descriptive writing tasks on familiar topics were preferred because they were more accessible to intermediate learners, even if such an easy task could be a problem for the assessment of advanced learners' writings (see discussion). Additionally, time constraints were imposed because the main focus of the present study is on English proficiency as evidenced in written production, and learners are supposed to demonstrate such proficiency in a

relatively constrained task. The present study does not focus on the writing process itself or on writing skills, for which more time would be necessary in order to give learners the opportunity to plan, revise, edit, etc. More time, together with a more complex topic, would be more appropriate if the focus of the study were on the writing process or writing skills and not proficiency as measured in writing. These two pieces of writing were performed just as part of this research and not as part of the students' coursework. The same tasks were adopted for the two proficiency levels to facilitate comparison.

The NESs only wrote one composition and they were given a choice of topics "My favorite city", and "Someone I admire". The NESs were asked to produce the same amount of words (approximately 150) in 15 minutes (although most students finished in 10 minutes). The reason why "My best friend" was replaced by "My favorite city" was that it was assumed that for NESs writing in their L1, the former topic would be more "childish" than for L2 learners, who are more used to writing on similar topics in the L2. Out of the 29 NESs who produced writing samples, 22 chose the topic "Someone I admire" and 7 chose "My favorite city".

2.4. *Data Analysis*

For this particular study five objective measures have been chosen in order to evaluate the students' writing samples. All of these measures have been claimed to be among the most reliable objective measures to analyze students' L2 written production in terms of fluency, complexity and accuracy (Wolfe-Quintero *et al*. 1998). In almost all of the measures the T-unit has been adopted as the production unit.

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

Fluency was examined in terms of words per T-unit (W/T), which has been a highly 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 to describe development in second language writing, which is the focus of this study (Larsen-Freeman and Strom 1977; Wolfe-Quintero *et al*. 1998; Larsen-Freeman 2006). It must be indicated that W/T has sometimes been assumed to measure grammar complexity more than fluency (Ortega 2003; Norris and Ortega 2009). Nevertheless, as Cooper (1976) and Wolfe-Quintero *et al*. (1998) suggest, longer does not necessarily mean more complex. Longer T-units can incorporate a higher number of words and phrases, yet such words do not have to be included necessarily in complex grammatical clauses. Nevertheless, since some concerns have been raised about whether W/T assesses fluency or syntactic complexity, another commonly used measure has been adopted to examine learners' fluency, which is the total number of words in the written sample (which will be referred to as W).

216

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 reviewed by them "do support the usefulness of the clauses per T-units measure", despite the fact that some studies did not find correlations between proficiency and C/T (see also Ortega 2003).

Lexical complexity 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 to analyze lexical complexity in L2 learners' productions (Vermeer 2000; Van Hout and Vermeer 2007). 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. Even though some studies have found Guiraud's Index to be a reliable measure of lexical complexity, this index has also received some criticism (Jarvis 2002; McCarthy and Jarvis 2007).

The measure errors per T-unit (Err/T) was adopted in this research 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 errors were not taken into account. In terms of the development of accuracy in SLA, it seems reasonable to expect that, as learners make progress in their second/foreign language knowledge, their language production (written, but also oral) should be more accurate. Nevertheless, this expectation is not always met, and sometimes L2 learners' accuracy has been reported to decrease when their complexity increases (Foster and Skehan 1996; Skehan and Foster 1997), although other studies maintain that complexity and accuracy do not need to be in competition (Robinson 2003, for oral production; or Ishikawa 2007; Kuiken and Vedder 2008, for written performance). The relationship between fluency, complexity, and accuracy is thus far from being clear.

The CLAN program (MacWhinney 2000) and the Statistical Package for the Social Sciences (SPSS 16.0) were used for the coding and analyses of the writing samples. The assessment of two different coders was compared for 15% of the total number of compositions, and inter-rater agreement (by means of percentage agreement) reached 100% in number of T-units, 97.4% in number of clauses, and 83.2% in number of errors. Decisions on what constituted an error were taken by one coder and some guidelines were provided to the second coder, including the following general criteria: ignore mechanical errors (spelling, punctuation, capitalization, paragraph formation); focus on morphology, syntax, and vocabulary (within vocabulary do not consider problems with register or expressions that are technically correct, but for which a more "commonly used" term could have been provided). Then, when specific discrepancies appeared, the case was discussed and an agreement was reached. Intra-rater reliability reached 100%

in number of T-units, 99.5% in number of clauses, and 90.5% in number of errors. The number of types and tokens was computed automatically through CLAN (MacWhinney 2000). See Appendix 1 for an example of a coded writing.

The data were submitted to different statistical analyses to examine the relations between the three independent variables and the five dependent variables. The independent variables were the following:

– Time-of-test: pretest/posttest

– Proficiency level: intermediate/advanced

– Program-type: regular/intensive

The dependent variables and the measures used to examine them were the following:

– Written fluency: W, W/T

– Written syntactic complexity: C/T

– Written lexical complexity: Guiraud's Index

– Written accuracy: Err/T

In order to answer the first research question (Are there interaction effects that relate time distribution and proficiency level in explaining EFL learners' development of written fluency, complexity, and accuracy?), a three-way mixed ANOVA was performed to examine the interaction between the three independent variables for each dependent variable. Then, two-way mixed ANOVAs analyzing the interaction of time-of-test and program-type were performed separately for intermediate and advanced learners so as to answer the second research question (Regardless of learners' proficiency level, do intensive courses enhance EFL learners' development of written fluency, complexity, and accuracy?) Finally, independent-sample *t*-tests were carried out to compare the performance of the advanced learners and the NESs in the pretest to control for ceiling effects. More details about these tests will be provided in the next section.

3. RESULTS

The first research question that was proposed in this study aimed at analyzing the interaction between time-of-test, proficiency level, and program type when examining learners' written fluency, complexity, and accuracy. A three-way mixed ANOVA was thus performed with the previously-mentioned variables as independent variables, and W, W/T, C/T, Guiraud's Index, Err/T as dependent variables. Table 2 represents the descriptive statistics for all the groups included in this study. It should be kept in mind that the scores in the Err/T measure should be interpreted differently from the scores in the rest of the measures: a higher score in Err/T would indicate *less* accuracy (involving

a higher ratio of errors per T-unit), while a higher score in the other measures would indicate *more* fluency, syntactic and lexical complexity. Also, the scores from the NESs have been presented twice (next to the scores from the advanced learners in both program types) in order to facilitate comparisons.

	REGULAR EFL				NATIVES
	INTERMEDIATE (N = 45)		ADVANCED (N = 32)		NES (N = 29)
	Pretest Mean (SD)	Posttest Mean (SD)	Pretest Mean (SD)	Posttest Mean (SD)	Test Mean (SD)
Fluency (W/T)	8.85 (1.60)	11.15 (2.97)	11.31 (3.03)	13.56 (3.56)	14.06 (4.14)
Fluency (Total Words)	144.62 (42.81)	131.87 (36.27)	156.31 (40.42)	164.03 (33.94)	168.31 (37.01)
Syntactic Comp. (C/T)	1.58 (.333)	2.02 (.574)	1.98 (.512)	2.47 (.639)	2.22 (.631)
Lex. Comp. (Guiraud's I.)	7.26 (.772)	6.56 (.686)	7.69 (.735)	7.46 (.719)	8.10 (.736)
Accuracy (Err/T)	.762 (.389)	.948 (.394)	.678 (.353)	.732 (.478)	.031 (.041)
	INTENSIVE EFL				NATIVES
	INTERMEDIATE (N = 38)		ADVANCED (N = 31)		NES (N = 29)
	Pretest Mean (SD)	Posttest Mean (SD)	Pretest Mean (SD)	Pretest Mean (SD)	Test Mean (SD)
Fluency (W/T)	8.60 (1.69)	11.13 (3.13)	13.12 (3.67)	14.22 (3.24)	14.06 (4.14)
Fluency (Total Words)	136.36 (29.28)	188.41 (59.32)	147.37 (29.68)	158.93 (28.59)	168.31 (37.01)
Syntactic Comp. (C/T)	1.54 (.289)	2.00 (.578)	2.19 (.671)	2.43 (.566)	2.22 (.631)
Lex. Comp. (Guiraud's I.)	7.05 (.757)	7.56 (.892)	7.82 (.607)	7.60 (.652)	8.10 (.736)
Accuracy (Err/T)	.692 (.426)	.810 (.468)	.836 (.431)	.815 (.486)	.031 (.041)

Table 2. *Descriptive statistics.*

The results of the three-way mixed ANOVA showed that the interaction between time-of-test, proficiency-level, and program-type was significant in the case of fluency as measured by W ($F(1,142) = 15.08$, $p < .001$; partial *eta squared* = .096), and lexical com-plexity as measured by Guiraud's Index ($F(1,142) = 18.20$, $p < .001$; partial *eta squared* = .114). For the other measures, no significant interactions were found between the three variables. Answering the first research question, these results suggest that intensity affected intermediate and advanced learners differentially in terms of fluency and lexical complexity.

In order to answer the second research question and examine whether intensity was positive regardless of program type, two two-way mixed ANOVAs were performed, one including the intermediate learners in the regular and the intensive program, and another one including the advanced learners in both program types. Through this test, interactions were examined between time-of-test and program-type for each level independently. The results of the first test, including the *intermediate learners*, indicate that there were interaction effects between time-of-test and program-type in fluency as measured by W ($F(1,81) = 30.10$, $p < .001$; partial *eta squared* = .271), and lexical complexity (Guiraud's Index: $F(1,81) = 36.26$, $p < .001$; partial *eta squared* = .309). In both cases, the students in the intensive program significantly outperformed their peers in the regular program. No other significant interactions were found for the other dependent variables. The two-way mixed ANOVA including the *advanced students* yielded no significant interaction effects between time-of-test and program-type for any of the dependent variables. These results suggest that intensity was especially favorable for the intermediate learners in terms of fluency and lexical complexity.

Additional tests were performed in order to compare the advanced classroom learners with the NESs in order to control for ceiling effects at the pretest. As expected, the performance of the NESs in this study was more fluent, complex, and accurate than the performance of the advanced learners in the regular and in the intensive program, according to the descriptive statistics reported in Table 2. The results of the independent-sample *t*-test comparing the advanced students in the *regular program* and the NESs demonstrate that significant differences existed in all the areas in favor of the NESs (W/T: $t(50.98) = -2.94$, $p = .005$; Guiraud's Index: $t(59) = -2.20$, $p = .032$; Err/T: $t(31.95) = 10.30$, $p < .001$), except in the case of fluency as measured by W ($t(59) = -1.20$, $p = .233$), and syntactic complexity ($t(59) = -1.63$, $p = .108$). When examining the performance of the advanced learners in the *intensive course* in comparison to the NESs, it was found that no significant differences existed between the two groups in any of the measures (W: $t(58) = -4.67$, $p = .642$; W/T: $t(58) = -.935$, $p = .354$; C/T: $t(58) = -.140$ $p = .889$; Guiraud's Index: $t(58) = -1.66$, $p = .102$), with the exception of written accuracy (Err/T: $t(30.60) = 10.34$, $p < .001$). These results demonstrate that the performance of the advanced learners in the intensive group was more "native-like" at the time of the pretest, and that this group did not have as much room for improvement from pre- to posttest.

4. DISCUSSION

The results obtained in this study suggest that there are interaction effects between time-of-test, proficiency level, and program-type in the development of written fluency, and lexical complexity in favor of the intermediate learners in the intensive program. Moreover, there seem to be interaction effects between time-of-test and program-type (in favor of the intensive program) for the intermediate learners in those same areas, but not for the advanced learners. This means that intensity does not always enhance the development of written fluency, complexity and accuracy. More specifically, intensity was shown to enhance the development of two aspects of written performance (fluency and lexical complexity) and only in the case of intermediate-level students.

It has to be acknowledged, then, that the results found in the present study (as is the case for all the studies analyzing fluency, complexity and accuracy) rely on the validity of the measures used to examine the three areas. Nevertheless, these results are quite consistent with what other studies investigating the effect of time distribution on L2 performance have found.

The literature examining intensive L2 programs has mostly been concerned with areas other than writing; however, the few studies analyzing written production have demonstrated that concentrated instruction leads to more advanced written performance (McKee 1983; Lightbown and Spada 1997; Lapkin *et al.* 1998; Germain *et al.* 2004). The studies examining intensive L2 courses included participants that were not advanced L2 learners; consequently, the results of the present study are in line with those findings. According to this, it can be concluded that, at the intermediate proficiency level, intensive instruction tends to foster a more optimal development of written production than a more distributed type of instruction, which offers an equal number of classroom hours over a longer period of time. Through a more prolonged daily exposure to the L2, the learners become more accustomed to the target language and are able to acquire more fluency. Some studies have indeed reported advantages for the learners receiving intensive instruction in oral fluency (Spada and Lightbown 1989). Similarly, daily practice with the L2 lexicon probably promotes the learning of the different vocabulary items, and the retrieval of such vocabulary may become easier because of the more concentrated practice, which prevents forgetting (in fact, Collins *et al.* 1999 also report a positive effect of concentrated language courses on the acquisition of vocabulary).

Nonetheless, practice effects may have favored the intensive learners, as these students took the pre- and posttest only a few weeks apart as opposed to several months in the case of the learners in the regular program. This is a common problem in studies examining the effect of time distribution on classroom L2 learning. Studies analyzing retention, apart from performance at the end of the course, would be helpful in order to examine whether the positive effects of intensity last. Retention data are not available for these students; therefore, the possibility of practice effects cannot be completely dismissed.

The effect of time distribution on the advanced learners' written performance does not seem to be as evident as in the case of the intermediate learners, according to the

results of the statistical analyses performed in this particular study. In fact, the few studies that have examined the effect of proficiency level on language gains in intensive programs have reported that the learners with lower initial proficiency are the ones that seem to obtain the most benefit in this type of program (Gardner *et al*. 1977). Nevertheless, it must be borne in mind that, when comparing the advanced students' performance with that of a group of NESs in the pretest, it was observed that the EFL learners in the intensive program were only different from the NESs in accuracy. On the other hand, the written production of the students in the regular program was significantly different from the production of the NESs in terms of written fluency (W/T), lexical complexity (Guiraud's Index) and accuracy (Err/T). As a consequence, the fact that the learners in the intensive program did not demonstrate significant gains from pre- to posttest in fluency, syntactic or lexical complexity could be attributed to ceiling effects and not to the fact that intensity of exposure is not beneficial for advanced learners. In this respect, these findings complement the ones reported by Serrano (2011). The reason why these learners' production was not significantly more accurate in the posttest still remains to be explained and could be due to the type of instruction received, which targets complex grammatical structures and not so much more basic aspects that are dealt with in lower levels, and seemed to be the source of many errors in the advanced learners' written productions (confusion between *do/make*, inappropriate use of *-ing/to infinitive,* inverted word order in embedded *wh*-clauses as in *I know who is he*, lack of third person singular *–s* in simple present, etc.)

Alternatively, the lack of improvement in accuracy at the advanced level could be tentatively explained by the type of task that was chosen for this study. Writing two descriptions about the students' best friend or someone they admire might not have been challenging (or complex) enough for the advanced learners, and, as some researchers have suggested, more complex tasks might lead to more complex and accurate production, due to their potential need for greater attentional resources (Robinson 1995, 2001; Kuiken and Vedder 2008). This same fact could also explain why the number of errors per T-unit that the advanced learners made at the pretest was not significantly different from the number of errors made by the learners at the intermediate level.

On the other hand, when examining the descriptive statistics at each data collection time for each group (see Table 2), it also becomes apparent that, as syntactic complexity increases, accuracy decreases. With values lower than 2 C/T, accuracy values tend to stay lower than 0.77 Err/T (e.g., regular intermediate pretest, C/T: 1.58, Err/T: .762; regular advanced pretest, C/T: 1.98, Err/T: .678; or intensive intermediate pretest, C/T: 1.54, Err/T: .692), but when syntactic complexity values are higher than 2 C/T, accuracy tends to decrease, and more errors per T-unit are made (e.g., regular intermediate posttest, C/T: 2.02; Err/T: .948; or intensive advanced posttest, C/T: 2.43, Err/T: .815). Therefore, another possible explanation for these results could be provided by Skehan's Trade-off Hypothesis (Foster and Skehan 1996; Skehan 2009), according to which, there are certain tensions between fluency, complexity and accuracy, and often improvement in one area could lead to lower performance in other areas.

5. Conclusion

In conclusion, the findings from the present study support the positive effects that have been reported for concentrated L2 instruction at proficiency levels below advanced. Nevertheless, the results for the advanced learners are quite inconclusive but they highlight the need to incorporate baseline data when analyzing L2 learners' production. More research should be performed using other measures of written production at the advanced level as well as other language skills so as to find more empirical evidence for the type of time distribution that is the most effective in developing L2 learners' writing in particular and L2 proficiency in general.

Acknowledgments

Notes

* Correspondence to: Raquel Serrano Serrano. Universitat de Barcelona. Gran Vía de les Corts Catalanes, 585. 08007 Barcelona. E-mail: raquelserrano@ub.edu.

References

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

Celaya, M. L., and T. Navés. 2009. "Age-related differences and associated factors in foreign language writing. Implications for L2 writing theory and school curricula". *Writing in Foreign Language Contexts*. Ed. R.M. Manchón. Clevedon: Multilingual Matters. 130-155.

Collins, L., R. H. Halter, P. M, Lightbown and N. Spada. 1999. "Time and the distribution of time in L2 instruction". *TESOL Quarterly* 33: 655-680.

Collins, L. and J. White. 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.

Council of Europe. 2001. *Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press.

Foster, P. and P. Skehan. 1996. "The influence of planning and task type on second language performance". *Studies in Second Language Acquisition* 18: 299-323.

Gardner, R. C., P. C. Smythe, and G. R. Brunet. 1977. "Intensive second language study: Effects on attitudes, motivation and French achievement". *Language Learning* 27: 243-261.

Germain, C., J. Netten, and S. P. Séguin. 2004. "L'évaluation de la production écrite en français intensif: Critères et resultats". *The Canadian Modern Language Review* 60: 333-354.

Gilabert, R. 2007. "The simultaneous manipulation of task complexity along planning time and (+/- here and now): Effects on L2 oral production". *Investigating Tasks in Formal Language Learning*. Ed. M.P. García Mayo. Clevedon: Multilingual Matters. 44-68.

Hinger, B. 2006. "The distribution of instructional time and its effect on group cohesion in the foreign language classroom: A comparison of intensive and standard format courses". *System* 34: 97-118.

Housen, A., and F. Kuiken, eds. 2009. *Complexity, Accuracy and Fluency (CAF) in Second Language Acquisition Research*. *Special issue in Applied Linguistics* 30: 461-473.

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.

Ishikawa, T. 2007. "The effect of manipulating task complexity along the (+/- here and now) dimension on L2 written narrative discourse". *Investigating Tasks in Formal Language Learning*. Ed. M. P. García Mayo. Clevedon: Multilingual Matters. 136-156.

Jarvis, S. 2002. "Short texts, best-fitting curves and new measures of lexical diversity". *Language Testing* 19: 57-84.

Kuiken, F. and Vedder, I. 2008. "Cognitive task complexity and written output in Italian and French as a foreign language". *Journal of Second Language Writing* 17: 48-60.

Lapkin, S., D. Hart, and B. Harley. 1998. "Case study of compact core French models: Attitudes and achievement". *French Second Language Education in Canada: Empirical Studies*. Ed. S. Lapkin. Toronto, ON: University of Toronto Press. 3-31.

Larsen-Freeman, D. 1983. "Assessing global second language proficiency". *Classroom-Oriented Research in Second Language Acquisition*. Eds. H.W. Seliger and M. Long. Rowley, MA: Newbury House. 287-304.

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. 2009. "Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition". *Applied Linguistics* 30: 579-589.

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

Lightbown, P. M., and N. Spada. 1994. "An innovative program for primary ESL students in Quebec". *TESOL Quarterly* 28: 563-579.

Lightbown, P. M. and N. Spada. 1997. "Learning English as a second language in a special school in Quebec". *The Canadian Modern Language Review* 53: 315-355.

MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. *Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

McCarthy, P. M., and S. Jarvis. 2007. "*vocd*: A theoretical and empirical evaluation". *Language Testing* 24: 459-488.

McKee, E. 1983. "The effects of intensive language instruction on student performance in beginning college French". Paper presented at the Central States Conference on the Teaching of Foreign Languages. St. Louis, Missouri.

Netten, J. and C. Germain. 2004. "Theoretical and research foundations of Intensive French". *Canadian Modern Language Review* 60: 275-294.

Norris, J. M. and L. Ortega. 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.

Pallotti, G. 2009. "CAF: Defining, refining, and differentiating constructs". *Applied Linguistics* 30: 590-601.

Polio, C., C. Fleck, and N. Leder. 1998. "'If I only had more time': ESL learners changes in linguistic accuracy on essay revisions". *Journal of Second Language Writing* 7: 43-68.

Robinson, P. 1995. "Task complexity and second language narrative discourse". *Language Learning* 45: 99-140.

Robinson, P. 2001. "Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA". *Cognition and Second Language Instruction*. Ed. P. Robinson. Cambridge: CUP. 287-318.

Robinson, P. 2003. "The cognition hypothesis, task design, and adult task-based language learning". *Second Language Studies* 21: 45-105.

Robinson, P., T. Cadierno, and Y. Shirai. 2009. "Time and motion: Measuring the effect of the conceptual of tasks on second language speech production". *Applied Linguistics* 30: 533-554.

Serrano, R. 2011. "The time factor in EFL classroom practice". *Language Learning* 61: 117-143.

Serrano, R. and C. Muñoz. 2007. "Same hours, different time distribution: Any difference in EFL?" *System* 35: 305-321.

Skehan, P. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. 2009. "Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis". *Applied Linguistics* 30: 510-532.

Skehan, P. and P. Foster. 1997. "Task type and task processing conditions as influences on foreign language performance". *Language Teaching Research* 1: 185-211.

Spada, N. and P. M. Lightbown. 1989. "Intensive ESL programmes in Quebec primary schools". *TESL Canada Journal* 7: 11-32.

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

Torras, M. R., T. Navés, M. L. Celaya, and C. Pérez-Vidal. 2006. "Age and IL development in writing". *Age and the Rate of Foreign Language Learning*. Ed. C. Muñoz. Clevedon: Multilingual Matters. 156-182.

Van Hout, R. and A. Vermeer. 2007. "Comparing measures of lexical richness". *Modelling and Assessing Vocabulary*. Eds. H. Daller, J. Milton, and J. Treffers-Daller. Cambridge: CUP. 93-115.

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

White, J. and C. E. Turner. 2005. "Comparing children's oral ability in two ESL programs". *The Canadian Modern Language Review* 61: 491-517.

Wolfe-Quintero, K., S. Inagaki, and H. Y. Kim. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity. Technical Report 17*. Manoa, Hawai'i: University of Hawai'i Press.

APPENDIX 1

Example of a coded writing sample:

*ELS: I want [err] speak about someone I admire [T] [C] [C] [C].
*ELS: I like the [err] solidary [err] people [T] [C].
*ELS: When a person explain [err] [err] me your [err] projects which [err] thought to help someone I like it [T] [C] [C] [C] [C].
*ELS: I think many times we should look around [T] [C] [C].
*ELS: many people haven't a [err] little things [err] very importants [err] in our life [T] [C].
*ELS: We are in a good position [T] [C].
*ELS: We don't think in [err] the poor people [T] [C].
*ELS: They a lot of times are [err] in the street [T] [C].
*ELS: They sleep in the [err] cash-point [T] [C].
*ELS: When we find [err] in the street we turn our face and we don't look [err] us [err] [T] [C] [C] [C].
*ELS: We prefer [err] don't see the situation [T] [C] [C].
*ELS: We prefer window-shop [err] or another [err] things for [err] we don't feel bad [T] [C] [C] [C].
*ELS: When I see someone who worried [err] a little bit in [err] another person I feel very good [T] [C] [C] [C].