THIS IS AN EARLIER VERSION OF THE MANUSCRIPT. FOR THE FINAL VERSION, PLEASE CHECK THE JOURNAL WEBSITE: Serrano, R. (2011). The time factor in EFL practice. *Language Learning, 61*, 117-145.

The time factor in EFL classroom practice[1]

This paper analyzes whether the distribution of the hours of classroom practice has any effect on students' foreign language gains, by comparing two types of EFL (English as a foreign language) programs: one in which the hours of instruction are distributed in long sessions over a short period (intensive course), and another in which the students attend short sessions over a long period of time (regular course). Data from 152 participants at two proficiency levels were gathered. Learners' grammar and vocabulary knowledge, as well as listening, writing and speaking skills were examined through a variety of tasks. The results of the analyses performed indicate that intermediate-level students tend to make more language gains in intensive programs than in regular programs, whereas advanced EFL students do not seem to benefit from intensive classroom practice as much as intermediate students do.

Keywords: time distribution; English as a second language instruction; intensive language courses; EFL classroom practice

The maxim that "practice makes perfect" reflects a common belief that practice is directly related to improvement in performance. This observation has been described in cognitive psychology by the "power law of practice" (Newell & Rosenbloom, 1981), according to which time devoted to practicing a skill is a determining factor in automatization, as reflected in a more accurate and faster performance. Similarly, the second language acquisition (SLA) literature has

emphasized that, all things being equal, the more time students devote to learning a second language, the higher their level of proficiency will be[2] (Carroll, 1967; Stern, 1985).

Even though this particular study will briefly examine the effect of practice (understood as learning opportunities through time) in skill acquisition in general and second language acquisition in particular, the main focus of this article is to analyze how the time devoted to practice (more specifically foreign language practice in the classroom) should be distributed for the creation of optimal learning conditions. In cognitive psychology, several researchers have investigated the effect of time distribution (spaced or massed) on learning. Most studies in this field have claimed that items that are presented in spaced sequences, with time or other items intervening between the presentations of the target forms, are better learned and remembered than material in massed presentations (in which the items to be learned appear in continuous sequences). This phenomenon has been referred to as the *spacing effect* (Dempster, 1996). In the SLA literature, the role of time distribution in second language learning has not received much attention. While some proposals have been presented concerning the intervals at which repetitions of new words should occur in order for them to be better learned and retained (Pimsleur, 1967), not much research has been conducted in relation to how the hours of practice or exposure to a second language should be optimally distributed. The few studies that have analyzed time distribution in L2 (second language) classroom learning have reported advantages for learners in intensive/compact courses as compared to those following regular foreign language classes not only for children (Collins, Halter, Lightbown, & Spada, 1999; Lightbown & Spada, 1994; Spada & Lightbown, 1989; White & Turner, 2005) but also for adults (Serrano & Muñoz, 2007).

This study attempts to shed some light on the effect of the distribution of instructional hours on L2 learning in a classroom context while considering the students' initial proficiency level. Before presenting the empirical study, two areas of current literature will be reviewed. Firstly, the effect of

practice on the acquisition of cognitive skills in general and second/foreign language skills in particular will be discussed. Secondly, the effects of time distribution of practice will be examined, considering findings from both cognitive psychology and SLA literature.

Review of the Literature

Practice in Cognitive Psychology and in SLA

Most researchers in cognitive psychology agree that in order for a skill to become automatized, a massive amount of practice is required; however, they differ in terms of which mechanisms can account for automatic behavior. DeKeyser (2001) and Segalowitz (2003) provide comprehensive reviews of the different criteria that have been considered in order to define automaticity in the acquisition of complex cognitive skills. Following the studies examined by these two authors, automatic performance can be described as performance that involves fast, ballistic, effortless, unconscious processing, for which no attention is required and which is not hindered by disruption of interfering events.

According to Anderson's ACT-R skill acquisition theory (Anderson, 1993; Anderson, Bothell, Byrne, Douglas, Lebiere, & Qin, 2004), before performance in a given skill becomes automatized, the knowledge required to perform such skill must progress through several stages. Most knowledge begins with a declarative stage, which—in the ACT-R model—is characterized by the learning of chunks. Such knowledge is then incorporated into production rules, which are the units in which a skill can be divided. Increased practice of procedural knowledge will finally lead to automatization. The role of practice in Anderson's model is highly important, since it is through extensive practice that performance becomes faster and more reliable, or, in other words,

automatized. As was mentioned before, the effect of practice is represented mathematically by a power function, according to which, speed increases and errors decrease as the number of trials increase. Also, the power law of practice describes the rapid initial improvement and the slower subsequent change in skill acquisition until improvement in performance nears an asymptote (Anderson, 1993; Newell & Rosenbloom, 1981).

Although Anderson and his associates have not systematically applied their model to the study of foreign language learning[3], the model has proven useful in explaining many issues related to second language learning (DeKeyser, 1997; 2007a; 2007b; Johnson, 1996; O'Malley & Chamot, 1990; Towell & Hawkins, 1994). The adoption of Anderson's model to explain child (first or second) language acquisition or naturalistic L2 acquisition would be controversial; however, the application of such a model to traditional classroom-based adult L2 learning, which is the context under study here, is appropriate in the author's view. If it is assumed that adult second language learning involves cognitive mechanisms that are similar to those required for the acquisition of other complex cognitive skills, it can be said that learners have a first access to the second language material through declarative knowledge of lexical items, grammar rules or collocations. Then, such declarative knowledge is proceduralized through practice, until performance becomes automatic and no longer needs to refer to the originally learned declarative knowledge (DeKeyser, 1997; 2007a). Or, in terms of the commonly adopted terminology for SLA, traditional classroom-based L2 acquisition could be said to *generally* involve a great deal of explicit, conscious, rule-based knowledge that can become implicit (or more unconscious or automatized) through extensive practice (DeKeyser & Juffs, 2005; DeKeyser, 2007b; DeKeyser, 2009).

This language learning model has been challenged by a few authors who claim that second language acquisition takes place mainly on the basis of implicit knowledge (Ellis, 2002; 2005) and that explicit (or declarative) knowledge cannot turn into implicit (or procedural) knowledge

(Hulstijn, 2002; Krashen, 1981; 1985). Paradis (1994; 2004; 2009) has also claimed that explicit and implicit knowledge are located in different areas of the brain, and that practice of knowledge acquired explicitly can never become part of implicit competence. Furthermore, Paradis suggests that performance based on explicit knowledge that is apparently fluent is just a "speeded-up" controlled process and should not be considered evidence of implicit knowledge. The development of immersion L2 programs or intensive French as a second language (FSL) in Canada has been based on this view of language learning. These programs, however, are highly different from the type of L2 learning that takes place in many other contexts around the world, which cannot devote the same number of hours of instruction to the L2. In traditional adult foreign language classes (as is the case of the context under study), the development of automatized performance (whether it is part of procedural/implicit knowledge or just a speeding-up of declarative knowledge) is usually facilitated by declarative knowledge (DeKeyser, 2009).

The claim that increases in practice time (or more practice opportunities) lead to higher language proficiency levels has been made by several authors, although there are varying perspectives on the way in which the additional hours of instruction should be distributed. Some programs have increased the hours of foreign language instruction by offering more concentrated language exposure without adding extra weeks or extra years to the planned period of foreign language learning (e.g., intensive French or intensive English classes in Canadian schools, which will be reviewed later). Another alternative is the introduction of the foreign language at earlier grades in school, thus increasing the period of language teaching while reducing the amount of contact with the foreign language per week (Muñoz, 2006). Most research studies have concluded that time concentration usually leads to better results in learners' performance than spreading the instructional hours over longer periods of time (Collins et al., 1999). The next section will examine

the effects of time distribution on learning in general and more specifically on foreign language learning.

Time Distribution in Cognitive Psychology and in SLA

In the cognitive psychology literature, study conditions in which repetitions of items to be learned appear in spaced or distributed sequences have been found to be more favorable for subsequent retrieval and retention than presentations in which repetitions occur instantly (Braun & Rubin, 1998; Cuddy & Jacoby, 1982; Dempster, 1987; Greene, 1989; Hintzman, 1976; Russo, Mammarella & Avons, 2002; Seabrook, Brown & Solity 2005; Toppino & Bloom, 2002). This phenomenon has been called the *spacing effect* and indicates that including intervals between learning episodes (distributed practice) is more effective for subsequent learning and retention than presenting the material in "massed" training episodes, in which learning takes place in a concentrated period of time.

Although the spacing effect is found consistently in most research contexts within the cognitive psychology literature, there is not much agreement as to why immediate repetitions of an item are not as effectively recalled as spaced repetitions. Several theories have been proposed, most of them falling into two main groups: encoding variability theories and deficient processing theories.

Encoding variability theories emphasize that spaced items are better recalled than massed because each presentation of the items is encoded differently, thus providing more retrieval cues. The storage of new semantic or contextual information in the case of spaced items facilitates subsequent retrieval (Glenberg, 1979; Melton, 1970). *Deficient processing* theories (Hintzman, 1976; Jacoby, 1978), on the other hand, claim that, within a presentation, the second repetition of massed items does not receive enough processing because the previous presentation is still too

recent. In contrast, when an item is presented after some time has elapsed and after some intervening items have been shown, full processing will be necessary, as the previous presentation will not be as readily available as in the case of massed presentations (Challis, 1993; Cuddy & Jacoby, 1982; Jacoby, 1978).

Other accounts of the spacing effect include *study-phase retrieval*, according to which retrieval of the first presentation when the second one appears is essential for subsequent recall and permanence in long-term memory (Braun & Rubin, 1998; Russo, Mammarella & Avons, 2002; Thios & D'Agostino, 1976; Toppino & Bloom, 2002). Braun & Rubin (1998) suggest that spaced items tend to be better recalled than massed items because the first presentation stays in working memory for longer. Nevertheless, since it is highly important that the first presentation is retrieved at the time of the second, there is a limit to the spacing that there can be between different presentations of the same item. As Bahrick & Phelps (1987) claim, "the optimum interval is likely to be the longest interval that avoids retrieval failures" (p. 349).

There is a dearth of research in the SLA literature concerning the effect of the distribution of instruction hours on students' learning. When making decisions about time allocation of instructional hours in foreign language teaching, the factors usually taken into consideration are purely practical (urgency of achieving a certain proficiency level in a limited time, time availability, costs, etc.) and not psychological or pedagogical (Stern, 1985). Most typically, foreign language classes meet for short periods of time each week over many years, since scheduling programs in this way is less complicated than organizing concentrated courses, which often require restructuring the whole curriculum. Nonetheless, there are also foreign language courses that only last a few months, in which the students are in contact with the target language for long stretches of time per session. Intensive foreign language classes have been offered since the beginning of the 20th century (during World War II intensive L2 courses were very popular for soldiers in the US);

however, it is only quite recently that the effects of concentrated L2 instruction have started to be tested empirically and compared to courses that spread the hours of instruction through longer periods of time. In the case of adults, Serrano & Muñoz (2007) examined the performance of students in three EFL program types that offered the same amount of hours of instruction (extensive, semi-intensive, and intensive) at the intermediate level. The results of their analyses suggest that the students who made the most language gains (as reflected in a proficiency test) at the end of their respective course were those enrolled in the intensive program, while the students who demonstrated the least progress were the ones registered in the extensive EFL course.

In the case of children in primary or secondary school, several researchers have highlighted the fact that traditional second/foreign language programs that provide limited hours of instruction a week in a non-concentrated time distribution have been shown not to be effective in the acquisition of a second language (Netten & Germain, 2004; Muñoz, 2006; Spada & Lightbown, 1989). In Canada, different models have been designed in order to enhance L2 learning of both French and English, all of them increasing and/or concentrating the hours of instruction. Intensive English and intensive French, providing more hours of instruction than regular L2 classes (around 400 hours in one school year, usually fifth or sixth grade), and French block scheduling (which concentrates the hours of French instruction, but does not increase them) have been more successful in fostering the students' L2 skills than regular, "drip-feed" courses (Collins et al., 1999; Lapkin, et al., 1998; Netten & Germain, 2004; Spada & Lightbown, 1989; White & Turner, 2005). The success of intensive French has also been attributed to the change in methodology and teaching strategies, however (Netten & Germain, 2005).

The majority of the studies investigating time distribution have examined the performance of students at the beginning or intermediate proficiency level; studies comparing advanced learners—the equivalent to B2/C1 in the Common Reference Levels from the Common European Framework

of Reference for Languages—in intensive and regular foreign language classes are practically non-existent. However, Gardner, Smythe, and Brunet (1977) examined language gains in intensive courses for learners who started at different language proficiency levels. These authors reported that L2 performance improved more in the case of students whose initial proficiency level was lower.

The results from experiments in cognitive psychology (in which spaced presentations are better remembered than massed) and in the SLA literature (in which intensive language courses seem to have more positive effects for L2 learning than non-intensive programs) may lead one to believe that the evidence they provide is contradictory. However, the different nature of the experiments, as well as the different type of knowledge that is supposed to be acquired in each context can explain the dissimilarity in results. In the cognitive psychology experiments, to-be-learned items in massed conditions are usually presented right after each other, or a few seconds apart. Language teaching, as it is implemented in most classrooms today, never involves constant and continuous repetitions of the same patterns or words (which would correspond to massed practice in the cognitive psychology experiments). There are repetitions of grammar structures, vocabulary items, or collocations in the input that foreign language learners receive, yet such repetitions always appear in distributed sequences, with other material intervening in the presentations of the target items. As a consequence, language learning in both the intensive and the regular programs can be said to be distributed in cognitive psychology terms. In intensive programs, though, following DeKeyser's (2007a) model of SLA, the repetitions of the declarative knowledge that the students are supposed to acquire occur in less spaced sequences, which makes such declarative knowledge more available for retrieval and thus for proceduralization—as the study-phase retrieval account of the spacing effect would suggest. When the duration of foreign language lessons is limited, and long periods of

time elapse between sessions, the students may have problems retrieving the knowledge previously acquired, which would make proceduralization harder.

Alternatively, and taking into account a framework that emphasizes the role of implicit mechanisms in language learning, "intensity" can be considered a fundamental condition for implicit competence to develop. Second language learners, just as is the case with children acquiring their first language, would need constant and long exposure to the L2 in order for them to "acquire" (as opposed to "learn") the language system (Krashen, 1982; 1985). Paradis (1994; 2004; 2009) emphasizes that continuous practice in authentic communicative situations is necessary for the acquisition of the target language mechanisms in a subconscious way (or implicitly). Regular L2 courses rarely provide learners with such amount of practice, and the practice they offer (distributed) does not facilitate implicit learning. Intensive instruction, on the other hand, constitutes a more conducive environment for implicit learning to occur (Germain & Netten, 2005).

Research Questions and Hypotheses

The purpose of this study is to examine whether the distribution of L2 hours of instruction is related to students' L2 outcomes in different language areas and whether the students' initial proficiency level has any effect on the students' language gains in different types of programs (intensive vs. regular). In other words, the two research questions that guide the present study are the following:

1. When the number of hours of instruction is held constant, does the distribution of such hours (manifested in regular and intensive English courses) have any effect on the acquisition of English as a foreign language, as reflected in the students' gains in listening, grammar, vocabulary, reading, writing, and speaking skills?

2. Does time distribution have a different effect on EFL students at the intermediate versus advanced proficiency level in terms of the language gains experienced by those students at the end of their respective course?

Considering previous research on the effectiveness of concentrating the hours of instruction of a second language for students at early stages in the acquisition process (Collins et al., 1999; Netten & Germain, 2004; Serrano & Muñoz, 2007; Spada & Lightbown, 1989; White & Turner, 2005), we might anticipate finding differences in language acquisition that favor learning in the intensive program. Language gains at the advanced level for learners attending intensive as opposed to regular courses have not been previously analyzed. However, in light of the results reported for advanced students as compared to their peers with a lower initial level of proficiency in intensive programs (Gardner et al., 1977), fewer positive effects might be expected for the advanced students in the intensive program under analysis than for the intermediate students in the same program.

Method

Programs and Participants

Two different programs were chosen for this study: regular and intensive. The regular course offers 110 hours of instruction throughout the whole academic year (approximately seven months), distributed over 2-hour sessions that meet twice a week. In the intensive program, the students receive the same number of hours of instruction (110) in four and a half weeks during the summer, distributed over 5-hour sessions that meet five days a week. The methodology and teaching techniques used in the two program types are highly similar. The same textbooks are used in regular and intensive courses and the students have the same type of exam at the end of the instructional period. Both regular and intensive programs aim to develop the four language skills

through a mixed approach, which includes communicative activities, as well as some more grammatically-oriented tasks. Nevertheless, these courses can be considered quite traditional in the sense that interactive activities tend to be designed for students to practice specific grammar points or vocabulary items, and there is a strong focus on forms. In general, all the different skills receive the same attention in intensive and regular programs. However, due to the long sessions in the intensive summer program and in order to keep the students' attention, teachers sometimes use more audiovisual materials (TV shows or excerpts of films) and songs than in regular programs both for intermediate and advanced learners. These two program types were offered in the same institution, which is the language school of the University of Barcelona (EIM, or *Escola d'Idiomes Moderns*).

Students at two different proficiency levels were analyzed: intermediate and advanced[4]. Most of the participants included in this research were university students whose ages ranged from 18 to 23 years old, who were taking English classes in order to obtain elective credits. All the students were comparable in terms of motivation and previous experience with English.

With respect to the instructors, practically every group included in this study had a different teacher. Since it was not possible to obtain intensive and regular programs taught by the same person, arrangements were made so as to obtain a wide variety of instructors. Such variation may neutralize the influence a specific teacher might have on the results of the study, which could be confounding factor when analyzing time distribution.

The number of participants distributed by level and program type, together with the number of teachers for each group is presented in the following table:

[Table 1 here]

Procedure and Instruments

The same data collection procedure was followed for the two different types of programs and for the two proficiency levels under study. Data from the students were collected using three different types of tasks: a proficiency test, a written task, and an oral narrative. The students performed these tasks twice, once at the beginning of the course (approximately 20 hours after the classes had started) and the second time towards the end of the course, more or less after 90 hours of instruction (approximately 80 hours after the pretest[5]). The students took both the pre- and the posttest during class time, either at the beginning or at the end of the class.

The proficiency test was different for intermediate and advanced learners, yet in both cases the exercises included in the test came from a short version of the practice exam used in the language school, which had the same format as the final exam for each of the levels under study. The selection of this practice exam was motivated principally by its ability to encourage both teachers and learners to participate in this research project given that the activities included in the test were assumed to help the students prepare for the final exam. Additionally, the exercises were clearly related to the content of the course; consequently, the test was considered an appropriate measure of the kind of learning that had been taking place in the class. The activities included in the test were the following: one listening comprehension exercise, one sentence conversion exercise, one open cloze activity, and one reading activity.

After finishing the activities included in the proficiency test, the learners at the two levels wrote a composition on the topic “My best friend” in the pretest, and “Someone I admire” in the posttest. The students were given approximately 15 minutes to write the composition and were asked to use approximately 150 words.

Finally, a group of students were chosen randomly from each group to perform an oral activity. The researcher called the students one by one and asked them to leave the class (where the instructor continued teaching) and follow her into a separate area in order to perform the activity. The students were recorded while they told a narrative on the basis of a series of pictures about two children going on a picnic (Heaton, 1966). This oral narrative was extensively used in the "Barcelona Age Factor Project" (see Muñoz, 2006), and it has been applied in a variety of projects including learners with different L1's and L2's (Tavakoli & Foster, 2008). See Appendix A for the pictures.

Measures

The proficiency tests had a maximum score of 26 points for the students at the intermediate level and 19 for those at the advanced level. To complete the listening comprehension exercise, the students listened to a speaker twice and were asked to complete six blanks in the case of the intermediate learners, or nine blanks in the case of the advanced learners. Each blank was worth 1 point. The sentence conversion exercise included five sentences (each worth 1 point) that had to be rewritten using the grammar dealt with in each level. For example, the students read, "I can't buy that computer because I don't have enough money." Their task was to rewrite that sentence starting with "If I ...". The open cloze activity had ten gaps (each worth 0.5 points), which, in the case of the intermediate learners targeted different grammatical structures, and in the case of the advanced both grammar and vocabulary. The reading task, which was performed only by the intermediate-level students, consisted of two activities: one that assessed comprehension (Reading A), which was worth 5 points, and one that assessed vocabulary (Reading B), which was also worth 5 points.

The students' writing samples were analyzed in terms of fluency, complexity (both syntactic and lexical) and accuracy. Four analytic measures were chosen in order to evaluate these aspects of the students' writing samples. All four measures are considered among the most reliable measures to analyze students' written—and oral—production (Wolfe-Quintero, Inagaki, & Kim, 1998).

Fluency was examined in terms of words per T-unit (W/T), which has been commonly used in L2 writing research. Moreover, according to several studies, this measure is highly adequate to describe development in second language writing (Larsen-Freeman, 2006; Larsen-Freeman & Strom, 1977; Wolfe-Quintero et al., 1998).

In order to analyze syntactic complexity the clauses per T-unit (C/T) ratio has been adopted in this study. Both finite and non-finite clauses are considered in this ratio. Wolfe-Quintero et al. (1998) claim that the majority of the studies reviewed by them "do support the usefulness of the clauses per T-units measure" (p. 86), despite the fact that some studies did not find correlations between proficiency and the C/T ratio. In this particular research study, chunks and fragments that did not include verbs (finite or non-finite) were not considered clauses; they were classified as "fragments" and excluded from the analysis of syntactic complexity. Since the data were gathered from adult learners with some knowledge of the English language and experience writing (both in their L1 and L2), it is unsurprising that there were very few instances of fragments in the students' production.

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 Guiraud's Index is one of the most appropriate measures to analyze lexical richness in L2 learners' productions (Hout & Vermeer, 2007; Vermeer, 2000).

Errors per T-unit (Err/T) was adopted in this study as a measure of learners' accuracy. Even though this measure (together with the percentage of error free T-units over the total number of T-

units) has been widely used in order to analyze accuracy, sometimes Err/T has been reported not to capture short-term change within intact classes (Wolfe-Quintero et al., 1998).

The students' oral production was also analyzed in terms of fluency, complexity and accuracy. The same measures reported for the written task were used for the oral task with the exception of the fluency measure, which was syllables per minute (SPM) in the case of the oral production task and not words per T-unit, since W/T has been claimed to be inexact for oral data (Griffiths, 1991). When considering the use of the SPM measure, the researcher needs to make decisions concerning the syllables that should be included in the count, and whether false starts, repetitions, self-corrections, or words in the student's L1 should be taken into account in calculating the total number of syllables. Two speech rates have been proposed: Rate A, which includes all the syllables uttered by the student; and Rate B, which does not count the syllables in false starts, repetitions, self-corrections, or words in the L1 (Gilabert, 2005; Yuan & Ellis, 2003). Both of these measures were initially considered in this study; however, the differences between groups across time using the two measures were similar. Therefore, only Rate B will be reported, first for the sake of simplicity, and second, because it was assumed that this rate better defined the students' L2 fluency, as it only considers meaningful syllables.

The author collected the data for approximately 80% of all participants; the remaining 20% was collected by four research assistants, who followed the same procedure for data collection. I was also in charge of coding all the tasks; however, a research assistant coded a random 10% of the data, and the inter-rater agreement reached 93%. Intra-rater reliability reached 95%. See Appendix B and Appendix C for examples of coded writing and oral samples.

Results

In order to analyze differences in language gains between the intermediate and advanced learners in regular and intensive courses, several statistical analyses were performed. Since the data were normally distributed for all the measures according to the Kolgomorov-Smirnov Test, parametric tests were chosen for all the analyses. Differences in language gains between program types were examined through a variety of statistical procedures. It was considered highly important to evaluate language gains while controlling for pretest scores; therefore, *t*-tests were performed on the residuals of the regression of posttest on pretest scores. In more detail, the following steps were followed separately for the intermediate learners and the advanced learners. First, data were collapsed across regular and intensive programs for each of the measures. Then, posttest scores were regressed on pretest scores. The residuals from such regressions were saved, and finally *t*-tests were performed in order to find differences between the residuals from the two program types. When the differences in these *t*-tests that included the residuals of regressing posttest on pretest scores were significant at alpha = 0.05, it was concluded that there were differences in language gains between program types, while controlling for pretest scores.

In the following sections, the descriptive statistics as well as the results of the parametric tests will be presented for the intermediate and the advanced learners. The scores of the proficiency test will be analyzed first; then those of the written production task; and finally, the findings about gains in oral production will be reported.

Intermediate Learners

The total number of participants who took the proficiency test was 87 (49 in the regular group and 38 in the intensive). Table 2 presents the scores for all the exercises, including the mean and

the standard deviation (SD). Statistically significant differences regarding gains from pre to posttest between learners in the regular and in the intensive group are indicated by asterisks.

[Table 2 here]

The *t*-tests performed on the residuals from regressing posttest on pretest scores suggest that the language gains experienced by the learners in the intensive program were significantly greater than those demonstrated by their peers in the regular program in the listening [$t(82.37) = -2.66$, $p = .009$], sentence conversion [$t(85) = -2.58$, $p = .012$], reading B [$t(85) = -2.98$, $p = .004$], reading total [$t(85) = -2.44$, $p = .017$], and total scores of the proficiency test [$t(85) = -3.31$, $p = .001$]. The effect sizes[6] of these differences were medium: listening ($r = .28$), sentence conversion ($r = .27$), reading B ($r = .31$), reading total ($r = .25$), total scores ($r = .34$). No differences existed in the cloze exercise [$t(85) = -1.31$, $p = .194$] or reading A [$t(85) = -.597$, $p = .552$].

Regarding written production, samples from 83 learners[7] were collected and analyzed (45 in the regular group and 38 in the intensive). Table 3 presents the mean scores, standard deviations, and the statistically significant differences in gains from pre- to posttest between the learners in the regular and the intensive group.

[Table 3 here]

The results of the *t*-tests performed on the residuals obtained after regressing posttest on pretest scores demonstrate that the gains experienced by the learners in the two program types were highly comparable in terms of fluency, syntactic complexity, and accuracy [W/T: $t(81) = -.014$, $p = .989$; C/T: $t(81) = .190$, $p = .849$; Err/T: $t(81) = 1.28$, $p = .204$]. However, the learners in the intensive

group experienced significantly greater gains in lexical complexity [Guiraud's Index: $t(81) = -6.30$, $p < .001$], and the effect size of this difference was large ($r = .57$).

In the case of oral production, data were collected from 22 students in each program type. The descriptive statistics are presented in Table 4.

[Table 4 here]

As can be seen in the table, no statistically significant differences were found in oral production gains between the learners in the regular and intensive program regarding fluency, syntactic complexity, or accuracy [SPM: $t(42) = -.493$, $p = .624$; C/T: $t(42) = .923$, $p = .361$; Err/T: $t(42) = 1.84$, $p = .073$]. Although the effect sizes were small in the case of fluency and syntactic complexity (SPM: $r = .07$; C/T: $r = .14$), there was a medium effect size in the case of accuracy (Err/T: $r = .27$). In terms of lexical complexity, the students following intensive instruction experienced greater gains than their peers in the regular course, as shown by the marginally significant difference between the two program types in the scores of Guiraud's Index [$t(42) = -1.97$, $p = .055$]. The effect size of this difference was medium ($r = .29$).

Advanced Learners

The same analyses were performed for the proficiency test at the advanced level. A total of 65 students took the test (34 in the regular and 31 in the intensive group). Table 5 presents the descriptive statistics.

[Table 5 here]

The results of the *t*-tests on the residuals of regressing posttest on pretest scores demonstrated that there were no significant differences in the gains experienced by the advanced learners in the proficiency test [listening: $t(63) = -.795$, $p = .430$; sentence conversion: $t(63) = .212$, $p = .833$; cloze: $t(63) = -.283$, $p = .778$; total: $t(63) = -.337$, $p = .738$].

With respect to written production, the descriptive statistics for the different measures appear in Table 6. The number of students who wrote the composition was 32 in the case of the regular group and 31 in the intensive group.

[Table 6 here]

The gains experienced by the advanced learners in terms of written production were comparable for the two program types, as demonstrated by the results of the *t*-tests: W/T: $t(61) = -.398$, $p = .692$; C/T: $t(61) = .569$, $p = .571$; Guiraud's Index: $t(61) = -.541$, $p = .590$; Err/T: $t(61) = -.092$, $p = .927$).

There were a total of 45 students who performed the oral task, 23 from the regular group and 22 from the intensive. The descriptive statistics are presented in Table 7.

[Table 7 here]

The results of the *t*-tests performed on the residuals of regressing posttest on pretest scores demonstrated that there were no statistically significant differences between the gains experienced in oral production by the advanced learners in the regular and the intensive programs in fluency

(SPM: $t(43) = -1.02$, $p = .313$), syntactic complexity (C/T: $t(43) = -.526$, $p = .602$), lexical complexity (Guiraud's Index: $t(43) = -1.71$, $p = .094$), or accuracy (Err/T: $t(43) = -.539$, $p = .593$). All the effect sizes were small, except for Guiraud's Index, in which there was a medium effect size ($r = .25$).

Discussion and Conclusion

The results presented in the previous sections seem to suggest that, at the intermediate level, those students enrolled in an intensive program made more language gains in a 110-hour course than those in a regular program. Such superior gains are clear for most items included in the proficiency test (those which assessed learners' listening skills, grammar, and vocabulary knowledge). Additionally, the language gains experienced by the learners in the intensive group were significantly higher in terms of lexical complexity (in written production as well as—with marginal significance—oral production). On the other hand, the advanced learners in the intensive group did not show significantly greater gains in any of the measures used to assess these learners' performance. The gains experienced by the advanced students in the regular and in the intensive group were comparable in terms of listening skills, grammar and vocabulary knowledge, writing skills, and oral skills.

The results reported in this paper for learners at the intermediate level are in line with previous studies examining time distribution in SLA, which have found greater gains in most language skills for the learners following intensive instruction. These gains are especially significant in programs that not only concentrate the hours of instruction but also increase them. This would be the case of intensive English or intensive French in Canadian schools (from 350 to 400 hours in fifth or sixth grade) as compared to regular ESL and regular FSL classes, which offer approximately 100 hours

in an academic year (Germain, Netten, & Movassat, 2004; Germain, Netten, & Séguin, 2004; Spada & Lightbown, 1989; White & Turner, 2005). In these programs the learners following intensive instruction have concentrated practice, but also *more* practice than their peers receiving regular instruction. Thus, the fact that the learners in these intensive courses demonstrate more skillful performance is in line with what different models of skill acquisition or second language acquisition have suggested: increased practice leads to greater automatization (Anderson et al., 2004; DeKeyser, 2007a), or to the development of implicit competence (Paradis, 2004; 2009).

However, some studies in the Canadian context have also shown that learners in fifth or sixth grade in intensive English courses outperform learners in grade nine who have been exposed to approximately the same number of hours of instruction (or have practiced the language for a comparable number of hours) but distributed over a longer period of time (Spada & Lightbown, 1989)[8]. Such findings would provide evidence for the positive effect of *time concentration* without any time increase. Other studies that suggest that time concentration of L2 classroom practice (or instructional hours) is beneficial include those comparing different types of intensive English programs in Canada. Research in this area has demonstrated that L2 results are slightly more positive in programs that concentrate 350-400 hours of instruction over 5 months than in programs that deliver a comparable number of hours of instruction over 10 months (Collins et al., 1999), although the differences were minimal and could also be attributed to the fact that the learners in the concentrated program ended up receiving a few more hours of instruction. Similarly, Lapkin et al. (1998) provide some evidence for the positive outcomes of learners following concentrated French instruction (or block-scheduling) in Canada. The programs analyzed by Lapkin and her colleagues lasted for 120 hours, which were differently distributed: a 10-week period in which there is French instruction for half of the day, a 5-month period in which there is 80 minutes of French a day, or a traditional year-long course which provides 40 minutes of French instruction a

day. The learners in the most concentrated program obtained the greatest gains; however, the language gains reported in this study for learners in concentrated L2 courses are clearly less significant than those reported for intensive English or intensive French. One reason for this difference could be that the number of hours of instruction analyzed in the two contexts is highly different (120 vs. 400). As is evident from these findings, even if some advantages have been reported for concentrated L2 courses as opposed to regular courses when the programs differ in time concentration alone, those advantages are less robust than the ones demonstrated by learners attending programs that not only concentrate but also increase the hours of L2 classroom practice.

In general, then, it can be claimed that the results reported in this paper for intermediate learners are not so different from previous findings from studies analyzing the effect of the time distribution of classroom practice on learners' L2 skills, since they point in the same direction: when advantages are found, they always occur for learners receiving intensive instruction. The lack of significant differences in the gains made by the intermediate learners in regular versus intensive programs for some measures reported in this study could be explained by two issues related to the time factor. First, the effect of few hours of instruction was examined (as was the case in Lapkin et al., 1998). Had the programs been longer, more differences may have appeared. Additionally, both programs under analysis here (regular and intensive) offer the same number of hours of instruction, and greater gains have always been reported for students in intensive programs that not only concentrate but also increase the hours of classroom practice (Spada & Lightbown, 1989; White & Turner, 2005). Therefore, an important conclusion that can be reached from comparing the findings of this study and previous studies on the time factor is that time increase is generally beneficial (which is in line with the claim that "practice makes perfect"), whereas time concentration alone usually leads to more modest gains. Still, those gains are superior in many L2 areas to those

obtained by learners attending programs that distribute the hours of classroom practice over long periods of time.

Another issue that could explain why few differences were found in the students' written and oral production relates to the measures selected in this study: it is possible that some of these measures could not capture progress in a short period of time. As was mentioned before, Wolfe-Quintero et al. (1998) report that some typically used analytic measures do not always capture differences between adjacent proficiency levels or between pre- and posttest when they occur within a short period of time.

Although there is little literature on the issue, the fact that the intensive courses under examination here seem to be more beneficial for learners at the intermediate level than for those at the advanced level is also consistent with the few studies that have analyzed learners with advanced proficiency (Gardner et al., 1977). Why, then, does concentrated SLA instruction seem to lead to greater language gains for learners whose proficiency level is not advanced?

One possible explanation, following DeKeyser's (2007a) model of L2 skill acquisition and study-phase retrieval theories of the spacing effect (Toppino & Bloom, 2002), could lie in the different types of knowledge that learners at the intermediate and advanced proficiency levels need to acquire. In adult classroom EFL, the knowledge that learners need to acquire in their classes has to be practiced (or retrieved) regularly in order for such knowledge to proceduralize. When the interval between sessions is too wide, learners may have problems retrieving previously learned concepts/structures, making proceduralization and automatization harder for them. On the other hand, when there is not much spacing between practice sessions, the knowledge acquired is more readily available for retrieval, and thus also for proceduralization. According to this claim, intensive courses should lead to greater gains. However, the type of knowledge that advanced learners need to acquire is different from the learning that takes place at lower proficiency levels.

Advanced students need to acquire infrequent collocations or complex grammatical structures whose presence in intensive courses is also scarce (as compared to the most frequently used patterns that the intermediate students still need to learn). These structures include, for example, the passive voice, third conditional, different uses of *I wish* (referring to present, past or future situations), and subordinate clauses introduced by conjunctions or prepositions such as *although, despite, in spite of,* etc. Consequently, it may be more difficult to retrieve those structures for all advanced learners, since they do not occur frequently in classroom practice, regardless of program type.

Alternatively, the different language gains experienced by intermediate and advanced learners in intensive courses may be explained by the claim that intensity favors proceduralization (DeKeyser, 2007a) or the development of implicit competence (Paradis, 2009), which would roughly take place at the intermediate level. When such processes have been completed (as would be true of advanced learners), intensity is no longer as influential.

Another possible explanation for the differences in results obtained for intermediate and advanced learners is that certain structures may have stabilized in the students' interlanguage, and thus no matter how much classroom practice (or type of classroom instruction) were to be received, little progress on those structures is likely to be observed. Moreover, ceiling effects for some of the measures that were used in this study to analyze written and oral production may have occurred: samples from native speakers should be collected to check for such effects. Additionally, some of the tasks (especially the writing and oral production tasks) included in this research may not have been challenging enough for advanced learners, and their performance may already have been quite automatized in the pretest for some of the language areas under investigation here. More data should be collected in order to isolate the most plausible explanation/s for the lack of significant differences at the advanced level between learners attending regular and intensive courses.

In sum, it can be said that the findings reported in this paper contribute to the literature on time distribution of L2 classroom practice. The present study provides valuable information about typical intensive courses (1-month courses), which have not received much attention in the SLA literature. Moreover, this study also controls for the time increase variable—which is a common intervening variable in many studies examining intensive instruction—and it also includes participants with different proficiency levels. More research should be carried out, however, including a larger number of learners, who are exposed to more than 110 hours of instruction, ideally with the same teacher, and with a wider range of proficiency levels in order to check whether, in fact, there is a correlation between the language gains experienced by the students in intensive courses and their initial proficiency. Furthermore, future studies on time distribution should examine retention apart from performance at the end of the course (which was the object of analysis in this study and in previous studies on time distribution and SLA), since what learners retain after some time has elapsed may be different after receiving regular or concentrated L2 instruction. Another interesting issue that is left for further research is how the length of the L2 course—or length of the L2 classroom sessions—interact with intensity in predicting L2 gains: should intensive courses have a minimum duration (e.g., 1 month) for intensity to have a positive influence? Should there be a minimum (and a maximum) number of hours per session for time concentration to have a positive effect? The present study has investigated the effect of time distribution for different levels of proficiency. Future studies should investigate the relationship between the time factor and other variables, such as teaching techniques, motivation, age, etc. Once more research is conducted and a clearer understanding of time distribution and SLA is achieved, recommendations could be made to program administrators so that the time and practice classroom L2 learners receive is optimally distributed.

Notes

[1] A short version of this paper was presented at AILA 2008 (August 24-29, Essen, Germany).

[2] "All things being equal" should be highlighted, since research has shown that age is an important factor determining the amount of time necessary for students to reach a certain proficiency level (Muñoz, 2006; Swain, 1981; Turnbull, Lapkin, Hart, & Swain, 1998). Likewise, time available for learning does not need to correspond to the time which is actually used, as learners may not employ such available time *actively* to learn (Baddeley, 1990).

[3] This author has, however, considered language learning on several occasions (Anderson, 1980, p. 224; Anderson, 2000, p. 14).

[4] The equivalent levels in the Common European Framework of Reference would be B1 for intermediate students; and B2/C1 for advanced students (Council of Europe, 2001).

[5] The pretest could not take place earlier, since the groups are quite unstable at the beginning of each course, with students changing groups or joining in some time after the classes have started. Similarly, the posttest could not be performed right at the end of the course, because such time is necessary for some teachers to finish the syllabus and for others to review the material presented throughout the course.

[6] The effect sizes were calculated using the following formula: $r_{Y\lambda} = \sqrt{(t^2 / (t^2 + df))}$

[7] There were four students who took the proficiency test but could not take part in the composition task.

[8] The findings from such studies are less conclusive because these learners are different not only in terms of the concentration of instruction hours they have received, but also regarding age and the L2 teaching approach they have been exposed to.

References

Anderson, J. R. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (2000). *Learning and memory* (2nd ed.). New York: Wiley.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglas, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111,* 1036-1060.

Baddeley, A. (1990). *Human memory: Theory and practice.* Needham Heights, MA: Allyn and Bacon.

Bahrick H. P., & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 344-349.

Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval and time in working memory: Evidence from once-presented words. *Memory, 6,* 37-65.

Carroll, J. B. (1967). Foreign Language Proficiency Levels Attained by Language Majors near Graduation from College. *Foreign Language Annals, 1*, 131–51.

Challis, B. H. (1993). Spacing effects on cued-memory tests depend on level of processing. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 19*, 389-396.

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

Council of Europe (2001). *Common European framework of reference for languages.* Cambridge: CUP.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior, 21*, 451-457.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*, 195-221.

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). Cambridge: CUP.

DeKeyser, R. M. (2007a). 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. (2007b). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97-113). Mahwah, NJ: Lawrence Erlbaum Associates.

DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. Long, & C. Doughty (Eds.), *Handbook of second language teaching* (pp. 119-138). Oxford: Blackwell.

DeKeyser, R. M., & Juffs, A. (2005). Cognitive considerations in L2 learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 437-454). London: Routledge.

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology*, *79*, 162-170.

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In E. Ligon & R. Bjork (Eds.), *Memory* (pp. 317-344). San Diego: Academic Press.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition, 24,* 143-188.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27,* 305-352.

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

Germain, C., & Netten, J. (2005). Pedagogy and second language learning: Lessons learned from intensive French. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée , 8*, 183-210.

Germain, C., Netten, J., & Movassat, P. (2004). L'évaluation de la production orale en français intensif: Critères et résultats. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 60*, 309-332.

Germain, C., Netten, J., & Séguin, S. P. (2004). L'évaluation de la production écrite en français intensif: Critères et resultats. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 60*, 333-354.

Gilabert, R. (2005). *Task complexity and L2 narrative oral production*. Unpublished doctoral dissertation, Universitat de Barcelona. Retrieved June 16, 2007, from Tesis Doctorals en Xarxa.

Glenberg, A. (1979). Component-levels theory of the effects of spacing. *Memory & Cognition, 7*, 95-112.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology, Memory, and Cognition, 15*, 371-377.

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

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17*, 649-667.

Johnson, K. (1996). *Language learning and skill learning*. Oxford: Blackwell.

Hintzman, D. L. (1976). Repetition and memory. In: G. H. Bower (Ed.), *The psychology of learning and memory* (pp. 47-91). New York: Academic Press.

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

Hujlstijn, J. (2002). Towards a unified account of the representation, acquisition, and automatization of second-language production. *Studies in Second Language Acquisition, 24*, 347-386.

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Krashen, S. D. (1985). *The input hypothesis.* London: Longman.

Lapkin, S., Hart, D., & Harley, B. (1998). Case study of compact core French models: Attitudes and achievement. In S. Lapkin (Ed.), *French second language education in Canada: Empirical studies* (pp. 3-31). Toronto, ON: University of Toronto Press.

Lightbown, P., & Spada, N. (1994). An innovative program for primary ESL students in Quebec. *TESOL Quarterly, 28*, 563-579.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior, 9*, 596-606.

Muñoz, C. (2006). The effects of age on foreign language learning: The BAF project. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 1-40). Clevedon: Multilingual Matters.

Netten, J., & Germain, C. (2004). Theoretical and research foundations of Intensive French. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 60*, 275-294.

Netten, J., & Germain, C. (2005). Pedagogy and second language learning: Lessons learned from intensive French. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée , 8,* 183-210.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge: CUP.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism. In N. Ellis (Ed.), *Implicit and explicit learning of second languages* (pp. 393-419). London: Academic Press.

Paradis, M. (2004). *The Neurolinguistic theory of bilingualism*. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Paradis, M. (2009). *Declarative and procedural determinants of second languages.* Amsterdam/Philadelphia: John Benjamins Publishing Co.

Pimsleur, P. (1967). A memory schedule. *The Modern Language Journal*, *51*, 73-75.

Russo, N., Mammarella, N., & Avons, S. E. (2002). Toward a unified account of spacing effects in explicit cued-memory tasks. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 28*, 819-829.

Seabrook, R., Brown, G. D., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology*, *19*, 107-122.

Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition*. (pp. 382-408) Oxford: Blackwell.

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

Spada, N., & Lightbown, P. M. (1989). Intensive ESL programmes in Quebec primary schools. *TESL Canada Journal, 7,* 11-32.

Stern, H. H. (1985). The time factor and compact course development. *TESL Canada Journal, 3,* 12-27.

Swain, M. (1981). Time and timing in bilingual education. *Language Learning, 31,* 1-15.

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58,* 439-473.

Thios S. J, & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior, 15*, 529-536.

Toppino T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 437-444.

Towell, R., & Hawkins, R. (1994). *Approaches to second language acquisition.* Clevedon: Multilingual Matters.

Turnbull, M., Lapkin, S., Hart, D., & Swain, M. (1998). Time on task and immersion graduates' French proficiency. In S. Lapkin (Ed.), *French second language education in Canada* (pp. 31-55). Toronto: University of Toronto Press.

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

White, J., & Turner, C. E. (2005). Comparing children's oral ability in two ESL programs. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 61,* 491-517.

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.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity, and accuracy in L2 monologic production. *Applied Linguistics, 24*, 1-27.

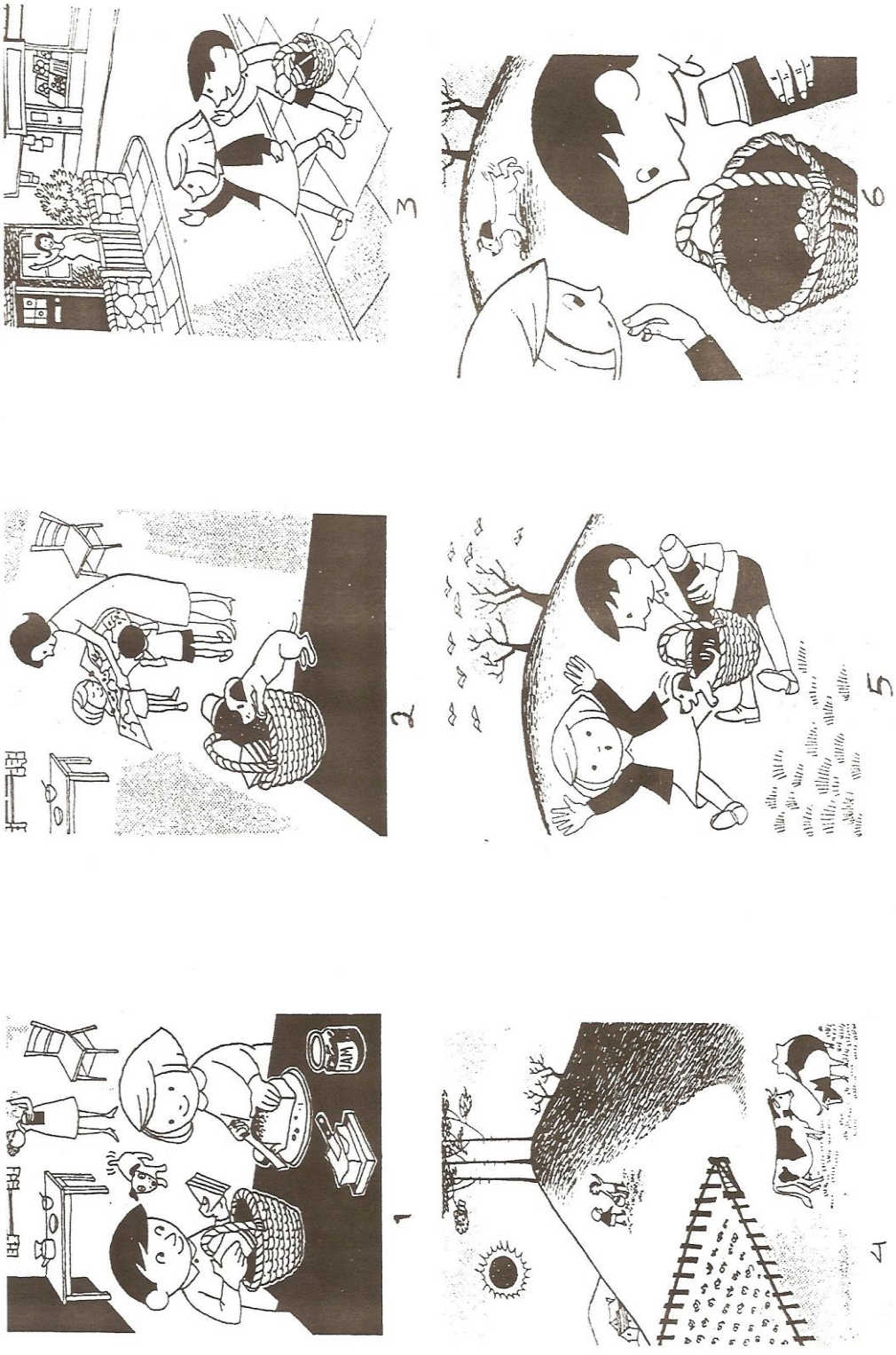
Author note

This research was supported by Grant HUM2007-64302. I would like to thank Carme Muñoz for her invaluable advice during all the different steps involved in this study, and for her feedback on an earlier draft of this paper. I also thank Robert DeKeyser and three anonymous reviewers for their insightful comments and suggestions.

Correspondence concerning this article should be addressed to Raquel Serrano, Departament de Filologia Anglesa i Alemanya, Universitat de Barcelona, Gran Via de les Corts Catalanes, 585, 08007, Barcelona, Spain. Internet: raquelserrano@ub.edu

Appendix

Appendix A Oral Task



Appendix B Example of coded writing sample (Pretest Intermediate Intensive)

*STU: My best friend is Laura [T] [C] .

*STU: I know [err] her since we were three years old [T] [C] [C] .

*STU: We went at [err] the same school [T] [C] .

*STU: and furthermore both of us [err] mothers were friends [T] [C]

*STU: Everydays [err] we played in the park with others friends' school [err] [T] [C] .

*STU: We played football, basketball [T] [C] .

*STU: For some years we broke up [err] our relationship because I changed of [err] school [T] [C] [C] .

*STU: It was very tired [err] [T] [C] !

*STU: But now we are friends other [err] time, the [err] best friends [T] [C] !

*STU: We go out together [T] [C] .

*STU: we go shopping [T] [C] [C] .

*STU: we go to the cinema, to the beach [T] [C] .

*STU: Now I know Laura and me will be together always [err] because our relationship is wonderful [err] [T] [C] [C] [C] !

Appendix C Example of coded oral sample (Posttest Advanced Regular)

STU: I guess that these two friends in the first picture [#] they are preparing uh their meal to [x2] [#] go to [x2] [#] the [#] picnic [T] [C] [C] [C] .

*STU: and well [#] the mother of one of them or maybe their mother is helping them to [#] [T] [C] +...

*STU: I don't know [T] [C] .

*RES: It’s a map .

*STU: Ok . She well I guess she's showing where is [err] the place they are going to or something like that [T] [C] [C] [C] [C] .

*STU: and then they leave in the third picture [T] [C] .

*STU: and [#] in the fourth picture they are maybe deciding the place to [#] have lunch [T] [C] [C] .

*STU: and then then realize that [false start] +...

*STU: well when they stop and see the basket they [x2] realize that the dog that was staying [err] with them has eat [err] everything [T] [C] [C] [C] [C] [C] .

* [#] indicates pause

Tables

Table 1

Participants

	Students	Groups	Teachers
Intermediate Regular	49	4	4
Intermediate Intensive	38	3	3
Advanced Regular	34	4	3
Advanced Intensive	31	3	3
Total EFL students	152	14	13

Table 2

Proficiency test results for intermediate learners

	Pretest			
	Regular (N = 49)		Intensive (N = 38)	
	Mean	SD	Mean	SD
Listening /6	4.06	1.42	4.22	1.30
Sentence conversion /5	.694	.865	.829	.983
Cloze /5	2.42	1.06	2.71	.984
Reading A (comprehension) /5	3.17	.733	3.33	.700
Reading B (vocabulary) /5	1.45	1.12	1.41	1.16
Reading Total /10	4.62	1.42	4.74	1.47
Total /26	11.80	3.23	12.50	3.48
	Posttest			
	Regular (N = 49)		Intensive (N = 38)	
	Mean	SD	Mean	SD
Listening /6	4.52	1.39	5.15**	.968
Sentence conversion /5	1.06	1.10	1.72*	1.34
Cloze /5	2.94	1.13	3.35	.937
Reading A (comprehension) /5	3.52	.721	3.63	.665
Reading B (vocabulary) /5	1.98	1.05	2.60**	1.08
Reading Total /10	5.50	1.43	6.24*	1.37
Total /26	14.03	3.66	16.46***	3.44

** $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$*

Table 3

Written production scores for intermediate learners

	Pretest			
	Regular (N = 45)		Intensive (N = 38)	
	Mean	SD	Mean	SD
Fluency (W/T)	8.85	1.60	8.60	1.69
Syntactic Complexity (C/T)	1.58	.333	1.54	.289
Lexical Complexity (Guiraud's Index)	7.26	.772	7.05	.757
Accuracy (Err/T)	.762	.389	.692	.426
	Posttest			
	Regular (N = 45)		Intensive (N = 38)	
	Mean	SD	Mean	SD
Fluency (W/T)	11.15	2.97	11.13	3.13
Syntactic Complexity (C/T)	2.02	.574	2.00	.578
Lexical Complexity (Guiraud's Index)	6.56	.686	7.56***	.892
Accuracy (Err/T)	.948	.394	.810	.468

** $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$*

Table 4

Oral production scores for intermediate learners

	Pretest			
	Regular (N = 22)		Intensive (N = 22)	
	Mean	SD	Mean	SD
Fluency (SPM)	62.17	12.71	74.69	23.20
Syntactic Complexity (C/T)	1.47	.407	1.66	.353
Lexical Complexity (Guiraud's Index)	4.97	.563	4.93	.558
Accuracy (Err/T)	1.07	.453	1.30	.606
	Posttest			
	Regular (N = 22)		Intensive (N = 22)	
	Mean	SD	Mean	SD
Fluency (SPM)	73.67	16.69	83.20	18.33
Syntactic Complexity (C/T)	1.68	.363	1.65	.394
Lexical Complexity (Guiraud's Index)	4.91	.609	5.21	.500
Accuracy (Err/T)	1.23	.550	1.04	.593

$* p \leq .05, ** p \leq .01, *** p \leq .001$

Table 5

Proficiency test scores for advanced learners

	Pretest			
	Regular (N = 34)		Intensive (N = 31)	
	Mean	SD	Mean	SD
Listening /9	3.95	1.38	4.37	1.60
Sentence conversion /5	2.21	1.24	2.47	1.27
Cloze /5	2.43	1.13	2.74	1.05
Total /19	8.43	2.67	9.51	2.67
	Posttest			
	Regular (N = 34)		Intensive (N = 31)	
	Mean	SD	Mean	SD
Listening /9	4.72	1.38	5.23	1.58
Sentence conversion /5	2.40	1.35	2.51	1.25
Cloze /5	3.12	.869	3.29	.834
Total /19	10.15	2.63	11.03	2.54

** $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$*

Table 6

Written production scores for advanced learners

	Pretest			
	Regular (N = 32)		Intensive (N = 31)	
	Mean	SD	Mean	SD
Fluency (W/T)	11.31	3.03	13.12	3.67
Syntactic Complexity (C/T)	1.98	.512	2.19	.671
Lexical Complexity (Guiraud's Index)	7.69	.735	7.82	.607
Accuracy (Err/T)	.678	.353	.836	.431
	Posttest			
	Regular (N = 32)		Intensive (N = 31)	
	Mean	SD	Mean	SD
Fluency (W/T)	13.56	3.56	14.22	3.24
Syntactic Complexity (C/T)	2.47	.639	2.43	.566
Lexical Complexity (Guiraud's Index)	7.46	.719	7.60	.652
Accuracy (Err/T)	.732	.478	.815	.486

$* p \leq .05$, $** p \leq .01$, $*** p \leq .001$

Table 7

Oral production scores for advanced learners

	Pretest			
	Regular (N = 23)		Intensive (N = 22)	
	Mean	SD	Mean	SD
Fluency (SPM)	108.8	22.42	118.1	21.91
Syntactic Complexity (C/T)	2.01	.550	1.83	.365
Lexical Complexity (Guiraud's Index)	5.63	.506	5.80	.674
Accuracy (Err/T)	.666	.297	.802	.434
	Posttest			
	Regular (N = 23)		Intensive (N = 22)	
	Mean	SD	Mean	SD
Fluency (SPM)	117.2	27.03	131.7	28.73
Syntactic Complexity (C/T)	1.90	.334	1.89	.484
Lexical Complexity (Guiraud's Index)	5.54	.552	5.90	.666
Accuracy (Err/T)	.552	.270	.657	.411

*$* p \leq .05$, $** p \leq .01$, $*** p \leq .001$