

Article

The Emergence and Development of Syntactic Patterns in EFL Writing in a School Context: A Longitudinal Study

M. Luz Celaya 

Department of Modern Languages and Literatures and English Studies, University of Barcelona, 08007 Barcelona, Spain; mluzcelaya@ub.edu

Received: 21 March 2019; Accepted: 18 June 2019; Published: 19 June 2019



Abstract: The present longitudinal study analyses the emergence and development of syntactic patterns in English as a Foreign Language (EFL) written production. Data were elicited by means of a paper and pencil task from sixteen school learners at three different times over a time span of 6 years. Studies in the area of Second Language Writing (SLW) have mainly focused on English as a second language and very few longitudinal studies have been carried out with low proficiency learners in a school context. To further contribute to the field, we have tried out a new measure of analysis, namely, the production of syntactic patterns. The results show that most learners produce both a wider variety and a higher number of patterns from one time to another, although statistically significant differences vary across patterns in relation to the times of data collection; secondly, it was found that the behaviour of two of the learners differed from that of the rest of the participants. It is concluded that in an instructional setting, the development of syntactic patterns in EFL writing as proficiency increases does not always show progression towards complexity and that it may be learner dependent.

Keywords: English as a foreign language; individual variation; syntactic patterns; Second Language Writing

1. Introduction

A large number of studies in the acquisition of second and foreign languages have for long focused on the analysis of second language writing (SLW) (see [Silva and Matsuda 2001](#); [Manchón 2011](#); [Manchón and Matsuda 2016](#); among others). However, research in the area mainly analyses English as a second language (ESL), as opposed to English as a foreign language (EFL) and other second (L2)/foreign languages (FL); moreover, few studies have focused on longitudinal data in EFL written production over a long period of time with more than one learner, especially at low levels of proficiency in EFL. The present study tries to contribute to fill these gaps in the field by analysing EFL writing from a longitudinal perspective. The study also addresses individual variation, an issue which, to the best of my knowledge, has not yet been dealt with in EFL writing, contrary to the case of other languages (see [Vyatkina 2012](#) for German). More specifically, the research reported here aims at (a) analysing the emergence and development of syntactic patterns in EFL writing in a longitudinal sample of school learners, and (b) investigating whether the same behaviour is presented by all learners in the cohort or, alternatively, whether there exists individual variation.

To this end, the present paper addresses two main lines of research in the review of the literature that follows, namely, the issue of measuring writing in an L2/FL and the relationship between the areas of complexity, accuracy and fluency and writing development with a special focus on low levels of proficiency and on syntactic complexity (which is the area dealt with in the present study) and, second,

individual variation in L2 writing. Due to the scarce research on both EFL longitudinal studies in writing and on individual variation in EFL writing, ESL and other second languages are also included in the review below. After the review of previous studies, the two research questions in the present study are presented. The context, participants, instrument and procedure are then described, followed by the sections on the results and discussion. The conclusions and limitations are dealt with in the last section.

1.1. Measuring L2 Writing Development at Low Levels of Proficiency

The topic of measurement still remains an unresolved issue in L2 writing at low levels of proficiency, which is the most common profile presented by young learners in FL school contexts, as in the case of the present study. Two main trends stand out in the literature; on the one hand, the issue of the best measures to gauge L2 written production at such low levels in an effective way, and, on the other, the relationship between the areas of complexity, accuracy and fluency (widely known in the literature as CAF, see [Housen et al. 2012](#)) in their development over time. In this sense, since the present study focuses on the development of syntactic patterns as indicators of complexity, special emphasis is given below to the review of syntactic complexity in writing.

Overall findings in the analysis of FL writing at low levels of proficiency show that written production in such cases is not always sensitive to the same measures that have been frequently used to analyse writing in second language learning contexts, on the one hand, or more advanced learners, on the other. A case in point is the study by [Celaya et al. \(2000\)](#) in the same EFL context as in the present study where the researchers used seventeen CAF measures that had been applied to L2 writing in [Wolfe-Quintero et al. \(1998\)](#) to analyse the written data coming from 289 EFL participants at grades 5 and 7. The *Principal Components Analysis* performed on the data showed that the amount of variance was explained differently depending on the school grade. Thus, at grade 5 with the lowest level of proficiency, four components, namely, all measures of fluency, one measure of accuracy (*number of rejected units*), and two complexity measures (*number of non-finite verbs* and *types of auxiliary verbs*) explained 75% of the variance, whereas in the case of Grade 7 learners, who were slightly more proficient than Grade 5 learners, one more component of complexity appeared as responsible for the 69% of the variance and the measures themselves had changed. Similar findings appear for German as a foreign language with older learners at a university level but still at low levels of proficiency. [Vyatkina \(2012\)](#) study of linguistic complexity in the writing of beginning learners over four semesters shows that some of the measures (e.g., clause-type units and subordinating conjunctions) turned out to be reliable indices to tap progress, whereas others such as *clause-length* did not discriminate as proficiency increased. In the case of ESL learners, [Evans et al. \(2014\)](#), for instance, compared three metrics (*weighted clause ratio*, *error-free T-unit ratio* and *error-free clause ratio*) in a sample of 350 writings and concluded that only one of the measures (*weighted clause ratio*) proved effective to assess low levels of proficiency.

As for the relationship between the three areas (CAF) in writing development, findings in FL and SL contexts do not differ much. In the same EFL context as above, [Torras and Celaya \(2001\)](#) carried out a study with the production from a total of 63 EFL low proficiency learners after 200 and 416 h of instruction in English, which covered a total period of four years for each group. A set of already existing measurements as well as a number of measures that had to be designed ad hoc were applied. These new measures were deemed necessary due to the scarcity of studies carried out with low proficiency young learners and, hence, the low number of measures to analyse such levels. The areas under analysis showed not to progress in parallel. In other words, some areas developed at the expense of the others along time. In other words, although the four areas analysed in the study—lexical and syntactic complexity, fluency and accuracy—presented higher scores at the second time of data collection, the results from normalized scores showed that fluency developed faster and achieved higher levels than both lexical and syntactic complexity and accuracy. Similar findings as to the different pace in the development of subcomponents of writing are reported in three other studies

in different contexts. Thus, [Verspoor and Smiskova \(2012\)](#) analysed the EFL written production by 22 Dutch students in their first and third year of high school in two types of programs and concluded that some measures discriminated best at certain stages than at others. [Vyatkina et al. \(2015\)](#) analysis of a corpus of learners of German as a FL at a US university shows that, although the range of the modification system did not significantly change from the beginning to the end of the period analysed, there were continuous changes which reflected the non-linear growth of modifier categories in the learners' writing. In an ESL context, [Bulté and Housen \(2014\)](#) studied a corpus of essays by adult learners after the first and last months of a four-month English course and concluded that the measures used to tap the development of L2 English showed an increase in the mean scores from the first to the second essay, but that different subcomponents of complexity in L2 writing developed at a different pace.

Writing development has often been analysed in relation to changes in syntactic complexity. Although the notion of "linguistic complexity" is not without controversy and still in need of definition in SLA research (see [Pallotti 2015](#)), the studies reviewed below apply the concept of complexity to those measures used to describe L2 written production and to assess proficiency along time. [Mellow \(2008\)](#) carried out a case study with a 12-year-old girl with Spanish as her L1 while spending nine months in an English-speaking context (US). The learner was found to gradually produce several different constructions and increase the complexity of her production. Similarly, [Crossley and McNamara \(2014\)](#) analysed syntactic complexity in 57 learners of ESL at a US university by means of a written essay and found significant growth in most of the features of syntactic complexity (e.g., phrasal and clausal complexity) from the beginning to the end of the semester.

Some other studies, however, do not always yield clear-cut evidence in favour of the growth of syntactic complexity in L2 writing as proficiency increases, as in the studies carried out in the Australian context. Thus, in her analysis of 25 learners with several L1 Asian languages, [Storch \(2009\)](#) found no improvement in complexity after one semester at an Australian university. In the same line, both [Knoch et al. \(2014\)](#) and [Knoch et al. \(2015\)](#) analysed ESL writing by university learners with different L1s after one year and three years of immersion in an English-medium university (101 and 31 participants, respectively). The results showed no significant improvement in grammatical complexity (measured by *clauses per T-unit*, *words per clauses* and *ratio of dependent clauses per clauses*) or in accuracy at the second time of data collection, despite the improvement in fluency. In both studies, the researchers claim that the lack of improvement in complexity might be due to the absence of language instruction, since the participants were enrolled in several university degrees and not in an English course.

Because of these controversial findings on the parallel growth between syntactic complexity and proficiency, it seems reasonable to establish a link with yet another feature intrinsic to L2 writing, namely, writing quality to try to establish constructs that may be applied. Although research had generally shown that syntactic growth does not necessarily coincide with the assessment of quality (see [Bulté and Housen 2014](#); [Crossley and McNamara 2014](#)), a recent study with first year ESL writers ([Casal and Lee 2019](#)) provides a different picture when combining the use of holistic and fine-grained measures of complexity, since only one of the measures (clausal complexity) does not show a relationship with writing quality. We can, therefore, point to the use of different measures as one of the reasons why findings are not always comparable. Besides, the fact that studies take place at different contexts (a writing course in the latter study) and with different learners (age and proficiency), as has also been seen in this section, seems to suggest the need for further research in the area, especially at low levels of proficiency in the L2.

1.2. Individual Variation in L2 Writing

A few recent longitudinal studies have analysed individual variation in writing; however, to the best of my knowledge, the focus on EFL is still missing. In her study of five high school Swedish learners of L2 French as a foreign language over a period of 30 months, [Gunnarsson \(2012\)](#) analysed

the development of fluency, complexity and accuracy in their written production, and the relationship between these areas in the different participants. Gunnarsson claims that individual variation stands out as the most evident result in her study; some learners were found to focus on accuracy at the expense of fluency and complexity, whereas other learners focused on fluency at the expense of accuracy. Therefore, one of the main conclusions in the study is the need to analyse individual patterns of development. With German as a foreign language, [Vyatkina \(2012\)](#) and [Vyatkina et al. \(2015\)](#) also address the issue of individual variation. The former study describes the individual developmental paths in the development of writing complexity of two college-learners of German as a foreign language whose data were compared to that in the larger cross-sectional cohort to which they belonged. The specific progress of one of the learners showed a higher level of syntactic complexification than her classmates or the other focal learner; on his part, this second focal learner produced longer sentences than the rest of the participants. [Vyatkina et al. \(2015\)](#) analysed the development of L2 writing syntactic complexity in a longitudinal corpus of writing in L2 German. The researchers report different trends for five learners in the use of attributive adjectives, as opposed to the other seven learners in the study. The authors claim that the coexistence of such high variability together with the uniform group trends in the study is a finding that shows the need to investigate individuals, as also seen in [Gunnarsson \(2012\)](#) above.

In light of this previous research, the first objective in the present study is to analyse the emergence and development of syntactic patterns in the written production of sixteen longitudinal Spanish/Catalan bilingual learners of EFL. The second objective is to analyse individual behaviour in the production of syntactic patterns to observe possible differences from the overall behaviour of the cohort, an approach that allows researchers (and teachers, ultimately) to get a more detailed picture of how second and foreign languages are learnt by focusing on outliers. The following research questions are therefore posed:

1. When do syntactic patterns emerge and how do they develop in EFL writing? In other words, how does syntactic complexity—operationalized as types of syntactic patterns—change in the learners' written production over time?
2. Do all learners conform to the behaviour of the cohort, or is there any individual variation?

2. Materials and Methods

2.1. Context

The present study analyses a subsample of data from the Barcelona Age Factor Project (BAF) (see www.ubgral.com and [Celaya and Navés 2009](#)). The study was carried out in several state-funded schools in a bilingual community in Spain, which shared a similar socioeconomic background and similar teaching methodologies. The contents provided in class followed the official curriculum and were controlled for in all cases. Data for the present study are from three primary schools and the corresponding four high schools that the participants attended later.

2.2. Participants

The participants are sixteen Spanish/Catalan bilingual learners of EFL whose data were collected longitudinally at three times that were set by the criteria in the larger project: Time 1 (T1) when they were 10 years old, Time 2 (T2) when they were 12 years old, and Time 3 (T3) at 16 years, after 200, 416 and 726 h of instruction as part of their school curriculum, respectively; the age of onset was the same for all learners. The very strict learner selection is the reason for the small number of participants; because the large study aimed at controlling exposure to the target language, learners in the sample could not have spent any time abroad in an English-speaking country or have attended out-of-school English classes. As learners grew older, these conditions could not always be met, and so the number of potential longitudinal participants gradually decreased.

All subjects gave their informed consent for inclusion before they participated in the study. Since the subjects were under age and data collection took place in the school premises, the informed consent was given by school directors, teachers and parents. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Projects PB94-0944, HUM2004-05167 and FFI2010-21478.

2.3. Instruments and Procedure

The data are from a paper and pencil task, a topic (“Introduce yourself: Me, my past life and my future”) and time-constrained (15 min) written essay, which was administered to participants in their own classroom. Both topic and time limit were kept constant across learners and times of data collection following Wolfe-Quintero et al. (1998) claim that such condition guarantees the comparability of results. It was made clear to the students that the task would not be assessed as an exam; instructions were given in English and in the learners’ L1s to avoid misunderstandings. The procedure followed a repeated-task design since, as mentioned above, data were collected from each of the sixteen learners at three different times, as specified above. Thus, a total of 48 compositions were available for analysis in the present study, three from each learner.

A standardized proficiency test (*Oxford Placement Test* (OPT), Allan 1992) was also administered at each of the three times of data collection to assess the learners’ level of proficiency in English and thus obtain individual scores independently of their school year.

Contrary to previous studies that have used well-tried measures of syntactic complexity (see discussion on measures to capture complexification in Norris and Ortega (2009), and several perspectives on complexity theory and language development in Ortega and Han (2017)) and in view of the very low level of the learners’ productions that did not allow us to use some of the most popular measures, the decision was taken to focus on syntactic patterns as the measure to gauge syntactic complexity and development, as in Klein and Perdue (1997) longitudinal study for the analysis of the acquisition of second languages in Europe. In spite of the differences vis-à-vis the present study both in relation to context (second vs foreign), mode (oral vs written) and participants (adults vs. children and teenagers), the patterns produced by our participants closely resembled those analysed in the *Basic Variety* (Klein and Perdue 1997). However, as the learners’ production complexified, there was a need to draw on a descriptive grammar of English (Huddleston and Pullum 2005) for the coding of the data, since the patterns in the *Basic Variety* could not be used to account for all instances.

A syntactic pattern, therefore, is understood here from a descriptive perspective of language, that is, as a structure where units are ordered to create meaningful sentences. Furthermore, it is claimed that the patterns in the present study present different degrees of complexity on the basis of the number and types of constituents, an approach that, as Pallotti (2009, p. 123) states, “has been widely used in SLA research”. Thus, as shown in Table 1 below, Pattern 1 (to be), Pattern 2 (to like) and Pattern 3 (to have) imply a similar degree of syntactic complexity in that the verb is followed by only one compulsory constituent, whereas Pattern 4 includes the use of any verb and, hence, a wider variety and number of constituents in the clause. Pattern 5 and Pattern 6, coordination and subordination, respectively, are considered in the present study as more syntactically complex than the other four patterns not only in terms of their syntactic organization (inter-clausal) but also because they require a higher cognitive effort by the learner, who must focus on meaning more than in the previous patterns. It goes without saying that the relationship between syntax and morphology and lexis in the acquisition of English in a school context should be borne in mind, but such issue is out of the scope of the present study (in this line, see Agustin-Llach and Agustin-Llach and Jimenez-Catalan (2018) on lexical profiles).

The first step in the analysis of the data consisted of identifying syntactic patterns (P henceforth) in the learners’ writing and classifying them according to the only six types that learners produced (see Table 1 below for the description of patterns together with unedited examples from the data). Patterns 5 and 6 were counted twice, that is, as one unit representing coordination or subordination, respectively, and also as each of the types represented by each of the clauses. In line with previous

analyses of data from the same learners (see mainly Pérez-Vidal et al. 2000), and bearing in mind the objectives of the present study, the decision was taken to accept errors of any type (spelling, morphological or syntactical) because the focus of the study was not on accuracy.

Table 1. Typology of syntactic patterns used in the analysis and their corresponding unedited examples.

| | |
|------------------------------------|---|
| P1. S + V to be + Cs/A | <i>Her mascot (= pet, from Spanish mascota) is a cat</i> <i>My favourite eat is meat and potatos</i> <i>I am in the classroom</i> |
| P2. (A) + S + V to like + Od + (A) | <i>I like football very much</i> |
| P3. (A) + S + V/V to have + Od (A) | <i>In my free time I play football</i> <i>Has got a one sister and one bird</i> |
| P4. S + V + A | <i>I study at school x</i> |
| P5. Coordination | <i>My sister is called Silvia and she have ten years old</i> |
| P6. Subordination | <i>When I had 4 years I start the school</i> |

Note: S = Subject; V = Verb; Cs = Subject Complement; A = Adverbial; Od = Direct Object.

Inter-rater reliability for the analysis of the units was carried out with another researcher and a research assistant who both had many years of experience in the analysis of EFL writing. The agreement reached was 98%; disagreements were finally resolved by the author of the present study. Data were submitted to statistical analysis run with RStudio 3.0.2. To answer the first research question, regression analysis was conducted with the non-parametric Cochran's Q test and the Wald Test to test for heterogeneity, and for the second research question the semiparametric Pearson Residual was used to test for goodness of fit (see further explanation in the Results section).

3. Results

The first research question inquired into the emergence and development of syntactic patterns in EFL writing and how the data changed in the written production each time in terms of syntactic complexity (operationalized as types of syntactic patterns, as described above). To answer this question, first a descriptive analysis was carried out in which tokens, that is, syntactic patterns, were counted and classified. The total number of tokens under analysis is 488, which corresponds to the total number of tokens that the learners produced at the three times of data collection; that is, 52 tokens at T1, 185 tokens at T2 and, finally, 251 at T3. It must be mentioned here that no data produced by the participants were left out from the analysis; that is, even if numbers and types of patterns may seem low when compared to those studies that analyse production at higher levels of proficiency, learners did not produce any other types of patterns and so the analysis covers the whole amount of the learners' production. As can be inferred from the description of the results below, the type of pattern used by the learners is related to the contents of the classes, which, as mentioned above, was the same for all learners. The learners had been exposed and instructed in all the syntactic patterns, except for P6, already at the first time of data collection. However, their frequency along time is precisely taken here as the key element to answer the first research question.

P1 (verb *to be*) presents the highest total number of tokens (198); P1 stands for the 67.3% of tokens when measured against the total number of tokens at T1, for the 53.51% at T2 and for the 25.49% at T3. P3 follows as the most widely used pattern with 93 tokens, distributed into 5.76% at T1, 13.51% at T2 and 25.89% at T3. Then follows pattern P5, with a total of 67 tokens and the following percentages: 9.6%, 15.13% and 13.54% at T1, T2 and T3, respectively. P2 yields a total of 50 instances, although the percentages present a different distribution since at T1 P2 accounts for 11.5% of the total number of tokens, at T2 the percentage decreases to 9.72%, and, finally, it rises again at T3 with 10.35%. P4, with a total of 47 tokens, is distributed into 5.76%, 7.56% and 11.95% for each of the three times of data collection. The pattern with the fewest number of tokens (33) is P6 (subordination), which appears in the data at T2 for the first time (0.54%), and then the percentage rise to 12.74% at T3.

These results can be summarized by focusing on P1, P5 and P6. Except for P6, the rest of the patterns emerge at T1 after 200 h of instruction; P6 emerges at T3 after 726 h of instruction, except for an occasional token produced by one of the participants at T2, which will be discussed below. However, both P1 and P5 present a decrease from T2 to T3, in contrast to the behaviour of the rest of the patterns, which start growing from T2 to T3. Such a picture implies that syntactic complexity in the learners' written production shows a gradual growth as indicated by two factors, namely, by the increase in the number of tokens of P2, P3, and P4 from one time to the next and the sharp decrease in P1 after T1, and by the emergence of P6 (subordination) at T3.

After the descriptive analysis of the types of patterns and their distribution over time, the results were analysed from a second perspective, namely, in relation to the number of participants who produced the patterns at each time. This approach was deemed necessary to be able to give a more accurate description of the participants' behaviour over time. In this sense, the results show an increase in the number of participants who produce each of the patterns from one time to the next. This runs parallel to the increase in the number of tokens of five of the patterns (except for P1 from T2 to T3), as described above. Figure 1 below presents the percentages of participants who produced each pattern from T1 to T3.

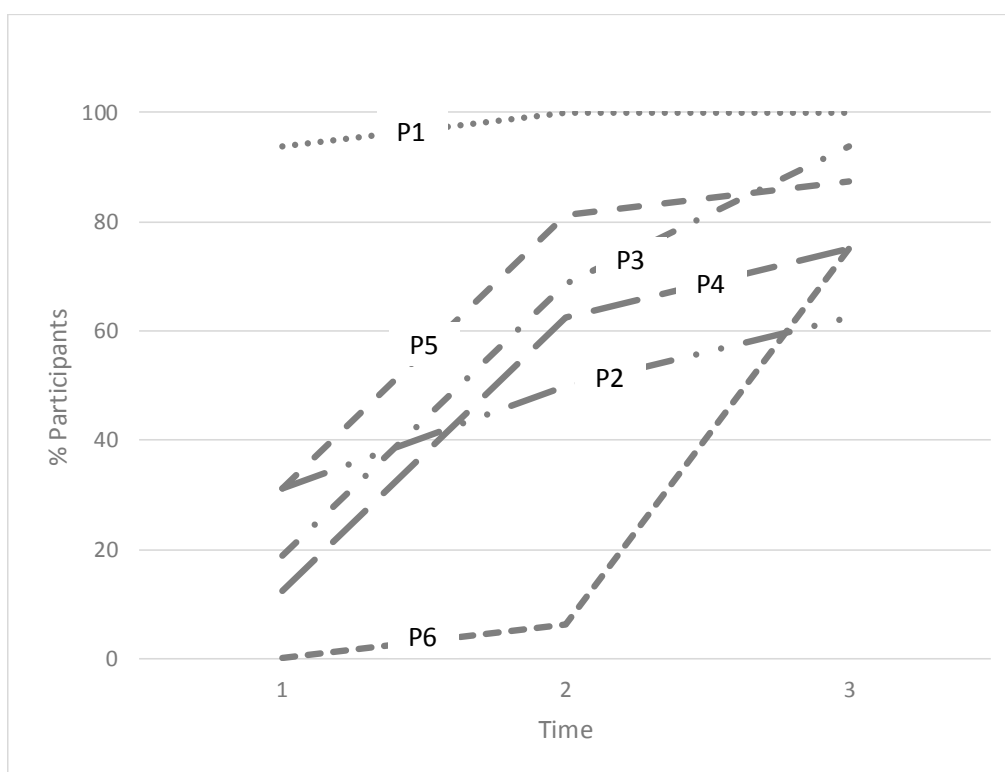


Figure 1. Participants (in percentages) who produced each pattern from each data collection time, Table 1. to T2 to T3.

Each pattern presents a different type of emergence and development in relation to their use by the learners. P1 is used by almost all the learners (93.75%) at T1, whereas all the participants (100%) produce it at both T2 and T3. Both P2 and P5 show a steady increase from between time points, although with notable differences in terms of numbers between the two patterns, especially from T1 to T2; 31.25% of the participants produce P2 at T1, a number that rises to 50% at T2 and, finally, to 62.50% at T3. Similarly, P5 is produced by the same percentage of participants as P2 at T1 (31.25%), but the number then rises to 81.25% at T2 and 87.50% at T3. As for P3 and P4, they are produced by a much lower percentage of learners at T1 (18.75% and 12.50%, respectively), but then they rise spectacularly to above 60% at T2. At T3, P3 is produced by 93.75% of the learners, close to the percentage for P1.

Finally, P6 presents a very different distribution; no learners produce it at T1 and only one learner (6.25%) uses it at T2; however, as specified above, at T3, it is produced by 75% of the learners.

After the descriptive analysis of the data to delve into the relationship between EFL written production and syntactic complexity at the three times of data collection, Cochran’s Q was first applied to the data (the significance level was set at < 0.01). This non-parametric test analyses whether *k* samples (the three data collection times in this case) have identical effects on the dependent variable (each of the syntactic patterns in our study). The analysis revealed statistically significant differences for four of the patterns, namely, P3 ($p = 0.001$), P4 (0.005), P5 ($p = 0.001$) and P6 ($p = 0.000$), but not for P1 (the statistical test could not be performed on the data, because of a ceiling effect) or P2 ($p = 0.258$). The individual test results for the patterns that yield an effect of data collection time present statistically significant differences between different times in each of the patterns analysed, as shown by the *Wald Chi-Square test*. As seen in Table 2 below, P3, P4 and P5 present the same behaviour in relation to data collection times: the differences between T1 and T2 are statistically significant ($p = 0.002$; $p = 0.000$; $p = 0.000$, respectively), but those between T2 and T3 are not.

Table 2. Syntactic patterns and times of data collection.

| Pattern | Time Repeated Contrast | Contrast Estimate | Std. Error | Wald Chi-Square | df | Sig. |
|---------|------------------------|-------------------|------------|-----------------|----|-------|
| P3 | T3 vs. T2 | 0.25 | 0.108 | 5.333 | 1 | 0.021 |
| | T2 vs. T1 | 0.49 | 0.155 | 9.864 | 1 | 0.002 |
| P4 | T3 vs. T2 | 0.13 | 0.174 | 0.516 | 1 | 0.472 |
| | T2 vs. T1 | 0.49 | 0.130 | 14.392 | 1 | 0.000 |
| P5 | T3 vs. T2 | 0.06 | 0.107 | 0.340 | 1 | 0.560 |
| | T2 vs. T1 | 0.48 | 0.127 | 14.347 | 1 | 0.000 |

In contrast, P6 presents statistically significant differences between T2 and T3 ($p = 0.000$) but not between T1 and T2 ($p = 0.302$) (see Table 3 below).

Table 3. Pattern 6 and times of data collection.

| Pattern | Time Repeated Contrast | Contrast Estimate | Std. Error | Wald Chi-Square | df | Sig. |
|---------|------------------------|-------------------|------------|-----------------|----|-------|
| P6 | T3 vs. T2 | 0.69 | 0.116 | 35.200 | 1 | 0.000 |
| | T2 vs. T1 | 0.06 | 0.061 | 1.067 | 1 | 0.302 |

To explore the second research question of the study (Do all learners conform to the behaviour of the cohort, or is there individual variation?), first we calculated the means of tokens for each pattern that learners produced at each data collection time in relation to the number of learners who produced the tokens (see Table 4 below). Except for P6, the results show that the lowest means appear at T1 in both P3 (1.00) and P5 (1.00), followed by P2 (1.20); the highest means appear in P1 at T2 (6.19), followed by P3 at T3 (4.33). Four out of the six patterns present a steady increase in their means from T1 to T2 to T3. The exceptions appear in P4, with a mean that decreases slightly from T1 to T2 (1.50 vs. 1.40), and in P1, with a mean that decreases from T2 to T3 (6.19 and 4.00, respectively), although the number of participants who produce P1 is the same at both times (16 participants), that is, all the members in the cohort.

Table 4. Means of tokens for each pattern in relation to the number of learners who produced them at each data collection time (T1, T2, T3).

| Times of Data Collection | | P1 | P2 | P3 | P4 | P5 | P6 |
|--------------------------|------|-------|-------|-------|-------|-------|-------|
| T1 | Mean | 2.33 | 1.20 | 1.00 | 1.50 | 1.00 | |
| | N | 15 | 5 | 3 | 2 | 5 | |
| | SD | 0.900 | 0.447 | 0.000 | 0.707 | 0.000 | |
| T2 | Mean | 6.19 | 2.25 | 2.27 | 1.40 | 2.15 | 1.00 |
| | N | 16 | 8 | 11 | 10 | 13 | 1 |
| | SD | 2.588 | 0.886 | 1.348 | 0.516 | 1.214 | - |
| T3 | Mean | 4.00 | 2.60 | 4.33 | 2.50 | 2.43 | 2.67 |
| | N | 16 | 10 | 15 | 12 | 14 | 12 |
| | SD | 1.713 | 2.547 | 2.350 | 1.382 | 1.505 | 1.371 |

These means were the basis for the statistical analysis to check for goodness of fit at each time of data collection by means of the Pearson Residual. With this kind of semiparametric regression, the means between the standardized residuals of items in each pattern were calculated to detect anomalous cases based on deviations from the norms of their cluster groups. In relation to the behaviour of the cohort just described, the statistical results show the existence of individual variation in two of the learners, but for different reasons. One out of the sixteen learners (Learner A) presented an anomalous index of 1.930; this learner differed from the rest of the cohort because he never produced P3, P5 or P6. The other anomalous case was Learner V, who presented an anomalous index of 1.493; this learner's writing differed from that of the rest of the cohort in that at T3 he produced 6 tokens of P6, whereas the mean of the group at T3 for the pattern was 2.67, as seen in Table 4 above.

4. Discussion

The current study set out to investigate EFL writing in relation to the emergence and development of syntactic patterns in a cohort of sixteen school learners. The two objectives in the study were (a) to analyse the emergence and development of syntactic patterns as indicators of syntactic complexity, and (b) possible individual variation among the participants.

Interesting trends appear when we consider the frequency of use of each pattern at the three times of data collection. To begin with, and as described above, our results show that at T1 (after 200 h of instruction), five out of the six patterns analysed have already emerged, whereas P6 emerges at T3 (except for one token produced by one learner at T2, as explained above). On the other hand, P1, probably used as a chunk by learners at low levels of proficiency, presents the highest percentage of use at the three times of data collection but its use gradually decreases from time to time. On the contrary, the production of P3 and P4 follows a gradual increase from one data collection time to the next. The fact that the frequency of use of three of the patterns changes from T1 to T2 is in line with the idea of syntactic complexification as proficiency increases (P3, P4 and P5 yielded statistically significant differences between T1 and T2). These results confirm previous findings in ESL contexts, as in Mellow (2008), Bulté and Housen (2014) and Crossley and McNamara (2014).

Secondly, the production of P2 and P5 is also in line with findings in previous studies. The percentage of use of P2 decreases from T1 to T2; in fact, P2 appears as the least frequently used pattern at T3, thus suggesting that the production of this pattern has decreased at the expense of more complex ones such as subordination. However, in contrast to the rest of the patterns, no statistically significant differences were found for P2 between the three times of data collection, as reported above. Similarly, the production of P5, which is the second most frequently used at T2, decreases from T2 to T3. Therefore, the production of P2 and P5 seems to confirm the non-parallel development between written production with increasing levels of proficiency along time (each of the times of data collection) and syntactic complexity, as in Navés et al. (2003), who analysed learners in the same context, and in longitudinal studies such as Storch (2009) and Knoch et al. (2014) in ESL

contexts, which did not find any improvement in syntactic complexity. Furthermore, such finding can also be interpreted as showing a ceiling effect in the development of syntactic complexity, in line with studies such as Rifkin (2005) where it is suggested that, although necessary, traditional classroom instruction might not be enough to boost proficiency (as compared to other learning contexts such as immersion abroad or at home). Alternatively, it might also be the case that it is the type of text required that did not favour the production of complex patterns, as in the study with oral data by Lázaro-Ibarrola and Lázaro-Ibarrola and Hidalgo (2017) where accuracy, but not syntactic complexity, slightly improved at the third time of task repetition.

In relation to our second research question, that is, the appearance of individual variation in a cohort of homogeneous learners, two of the participants have been seen to differ from the other fourteen learners even though their features (context, ages, sociolinguistic background and proficiency) were similar. Although, as stated in the review of the literature, no previous studies have focused on individual variation and EFL in a school setting, our findings are in line with studies on other foreign languages such as Gunnarsson (2012), Vyatkina (2012), and Vyatkina et al. (2015). It is claimed here, therefore, that there is a need to approach data from a distinct perspective to be able to explain individual behaviour in groups, since as stated by Ellis and Larsen-Freeman (2006, p. 564) “what generalizations exist at the group level often fail at the individual level”. Research on inter-individual variation in groups of language learners seems then foremost nowadays, since it may have relevant pedagogical implications not only to create tailor-made tasks but also for assessment purposes, especially at low levels of proficiency so that individual needs can be addressed from the very beginning of the learning process.

Three main differences between the present study and previous ones should be acknowledged here to further discuss our findings. To begin with, the fact that the operationalisation of syntactic complexity differs between studies might make comparisons difficult. Crossley and McNamara (2014), for instance, used Coh-Metrix indices involving phrasal, clausal and sentential levels; in the present study, we had to be more conservative, due to the low level of proficiency in EFL of our participants, and so the decision was made to try out the measure “syntactic patterns” based on the description of ILs in Klein and Perdue (1997) and on descriptive grammars of English, as mentioned above, in contrast to previous studies that have used widely accepted measures to analyse writing. Another relevant feature in studies on L2 writing development appears in relation to the different lengths of time that are considered necessary for syntax to develop. Thus, in longitudinal studies, although the measure of length of time between tests is used to mark milestones in writing development, such periods differ from study to study. Crossley and McNamara (2014) claimed that an eight-week period was not enough for syntactic growth to appear in their study, whereas Knoch et al. (2014) concluded that even three years of degree study might not have been enough for complexity to emerge. Although with only three times of data collection, the present study covers a time span of six years, a larger time span than in most longitudinal studies. Finally, both the different ages and contexts of participants are issues that must be borne in mind when comparing results in SLA research. The present study deals with primary and secondary school learners in a foreign language learning context; in a school context, especially if there is no explicit focus on writing in the school curriculum, the emergence of certain patterns in writing may be delayed as compared to ESL contexts where learners are exposed to the L2 in a non-gradual way on a daily basis.

5. Conclusions and Limitations

This study has analysed EFL writing development and individual behaviour in a cohort of sixteen bilingual learners of EFL in a school context over a period of six years, with three times of data collection and school-only instruction in EFL. The first research question inquired into the changes in syntactic complexity—operationalized as types of syntactic patterns—in the learners’ written production over time. The results show a similar emergence and production for three of the patterns analysed (P3, P4 and P5), with statistically significant differences between T1 and T2, and a different picture in the

other three patterns. The second research question focused on the behaviour of each of the learners to identify possible differences among participants. Two main conclusions can be drawn from our findings. The first is that the parallel relationship between written production and syntactic complexity along time seems to differ from one time point to another (T1, T2 and T3). As a line for further research it may be advisable to make use of holistic measures (as in Ruiz de Zarobe and Celaya 2011) together with CAF measures to obtain a more complete picture of the development of the second or foreign language.

The second conclusion is that there exists individual variation, which is in line with the few studies that have addressed this issue, although not in EFL, as seen in the review of the literature. It is suggested, therefore, that individual variation needs to be addressed in the future, especially in the area of foreign language teaching in school contexts and more specifically in EFL at low levels of proficiency. As pointed out above, such a line of research may yield important pedagogical implications for instructed contexts, which are, however, out of the scope of the present study.

At this point, it is necessary to acknowledge a number of limitations of the present study. First, although it is a longitudinal study over a time span of 6 years, a larger amount of time than most studies, data were collected only at three times. The reason for this lies in the design of the large project (see Section 3 above), which aimed at comparing learners who had two different ages of onset as marked by two different curricula in the teaching of English in Spain. Therefore, another limitation to be acknowledged is that the amount of data may be scarce if compared to other studies with more points of data collection. However, it is very relevant to bear in mind that EFL written production by young and low proficiency learners in a school context differs from the production that we come across in studies that work with ESL adult proficient learners.

Finally, the fact that the topic chosen for the writing task may have influenced the types of syntactic patterns produced by the learners is another limitation in the present study. A different task or a different topic might have yielded other syntactic patterns. However, as explained above, maintaining the same task and topic at the three times of data collection is the only way to ensure a reliable comparison of all the productions over time. Mollet et al. (2010) discussion of the nature of different measurements of language leads us to conclude that any decision regarding the measures chosen for a particular study will always be subjective. Future research, then, may adapt the typology used here to analyse L2 writing on a variety of other topics.

Funding: This research belongs to the projects funded by grant numbers PB94-0944, HUM2004-05167, FFI2010-21478 and 2014SGR1089.

Acknowledgments: The author acknowledges support through grants FFI2010-21478 and FFI2016-80564-R, and through 2014SGR1089 to the GRAL research group. Ferran Carrascosa provided invaluable help with statistics and Mike Maudsley carried out a thorough language revision. I am very grateful to both of them. My special thanks to Júlia Barón, Rosa M. Manchón and Carmen Muñoz for their support throughout the process of writing this paper. My most sincere gratitude to M. Teresa Navés for her insightful ideas throughout all the years we have been working together. Finally, I am deeply grateful to the two anonymous reviewers for their valuable feedback. Any mistakes remain my own.

Conflicts of Interest: The author declares no conflict of interest.

References

- Agustin-Llach, M. Pilar, and M. Rosa Jimenez-Catalan. 2018. Teasing out the role of age and exposure in EFL learners' lexical profiles: A comparison of children and adults. *International Review of Applied Linguistics* 56: 25–43. [CrossRef]
- Allan, Dave. 1992. *Oxford Placement Test*. Oxford: Oxford University Press.
- Bulté, Bram, and Alex Housen. 2014. Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing* 26: 42–65. [CrossRef]
- Casal, J. Elliott, and Joseph J. Lee. 2019. Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing* 44: 51–62. [CrossRef]

- Celaya, M. Luz, and M. Teresa Navés. 2009. Written production in English as a Foreign Language: age-related differences and associated factors. In *Teaching and Researching Writing in Foreign Language Learning Contexts. Learning, Teaching, and Research*. Edited by Rosa M. Manchón. Clevedon: Multilingual Matters, pp. 130–55.
- Celaya, M. Luz, Carmen Pérez-Vidal, and M. Rosa Torras. 2000. Matriz de criterios de medición para la determinación del perfil de competencia lingüística escrita en inglés. *Revista Española de Lingüística Aplicada (RESLA)* 14: 87–98.
- Crossley, Scott A., and Danielle S. McNamara. 2014. Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing* 26: 66–79. [[CrossRef](#)]
- Ellis, Nick C., and Diane Larsen-Freeman. 2006. Language emergence: Implications for Applied Linguistics—Introduction to the special issue. *Applied Linguistics* 27: 558–89. [[CrossRef](#)]
- Evans, Norman W., K. James Hartshorn, Troy L. Cox, and Teresa Martin de Jel. 2014. Measuring written linguistic accuracy with weighted clause ratios: A question of validity. *Journal of Second Language Writing* 24: 33–50. [[CrossRef](#)]
- Gunnarsson, Cecilia. 2012. The development of complexity, accuracy, and fluency in the written production of L2 French. In *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy, and Fluency in SLA*. Edited by Alex Housen, Folkert Kuiken and Ineke Vedder. Amsterdam and Philadelphia: John Benjamins, pp. 247–76.
- Housen, Alex, Folkert Kuiken, and Ineke Vedder, eds. 2012. *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy, and Fluency in SLA*. Amsterdam and Philadelphia: John Benjamins.
- Huddleston, Rodney, and Geoffrey K. Pullum. 2005. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Klein, Wolfgang, and Clive Perdue. 1997. The basic variety (Or couldn't natural languages be much simpler?). *Second Language Research* 13: 301–47. [[CrossRef](#)]
- Knoch, Ute, Amir Rouhshad, and Neomy Storch. 2014. Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? *Assessing Writing* 21: 1–17. [[CrossRef](#)]
- Knoch, Ute, Amir Rouhshad, Sue Ping Oon, and Neomy Storch. 2015. What happens to ESL students' writing after three years of study at an English medium university? *Journal of Second Language Writing* 28: 39–52. [[CrossRef](#)]
- Lázaro-Ibarrola, Amparo, and M. Angeles Hidalgo. 2017. Procedural repetition in task-based interaction among young EFL learners. Does it make a difference? *ITL-International Journal of Applied Linguistics* 168: 183–202. [[CrossRef](#)]
- Manchón, Rosa M., ed. 2011. *Learning to Write and Writing to Learn in an Additional Language*. Amsterdam and Philadelphia: John Benjamins.
- Manchón, Rosa M., and Paul Kei Matsuda, eds. 2016. *Handbook of Second and Foreign Language Writing*. Berlin: De Gruyter Mouton.
- Mellow, J. Dean. 2008. The emergence of complex syntax: A longitudinal case study of the ESL development of dependency resolution. *Lingua* 118: 499–521. [[CrossRef](#)]
- Mollet, Eugène, Alison Wray, Tess Fitzpatrick, Naomi R. Wray, and Margaret J. Wright. 2010. Choosing the best tools for comparative analyses of texts. *International Journal of Corpus Linguistics* 15: 429–73. [[CrossRef](#)]
- Navés, M. Teresa, M. Rosa Torras, and M. Luz Celaya. 2003. Long-term effects of an earlier start. An analysis of EFL written production. In *EUROSLA Yearbook*. Edited by Susan Foster-Cohen and Simona Pekarek. Amsterdam: John Benjamins, pp. 103–29.
- Norris, John M., and Lourdes Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics* 30: 555–78. [[CrossRef](#)]
- Ortega, Lourdes, and Zhaohong Han, eds. 2017. Complexity Theory and Language Development. In *Celebration of Diane Larsen-Freeman*. Amsterdam: John Benjamins.
- Pallotti, Gabriele. 2009. CAF: Defining, refining and differentiating constructs. *Applied Linguistics* 30: 590–601. [[CrossRef](#)]
- Pallotti, Gabriele. 2015. A simple view of linguistic complexity. *Second Language Research* 31: 117–34. [[CrossRef](#)]
- Pérez-Vidal, Carmen, M. Rosa Torras, and M. Luz Celaya. 2000. Age and EFL written performance by Catalan-Spanish bilinguals. *Spanish Applied Linguistics* 4: 267–90.
- Rifkin, Benjamin. 2005. A ceiling effect in traditional classroom foreign language instruction: Data from Russian. *The Modern Language Journal* 89: 3–18. [[CrossRef](#)]

- Ruiz de Zarobe, Yolanda, and M. Luz Celaya. 2011. CLIL across languages: Research outcomes in two bilingual communities. *Teoría de la Educación: Educación y Cultura en la Sociedad de la Información (TESI)* 1: 200–14.
- Silva, Tony, and Paul Kei Matsuda, eds. 2001. *On Second Language Writing*. Mahwah: Erlbaum.
- Storch, N. 2009. The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing* 18: 103–18. [[CrossRef](#)]
- Torras, M. Rosa, and M. Luz Celaya. 2001. Age-related differences in the development of written production. An empirical study of EFL learners. *International Journal of English Studies (IJES)* 1: 103–26.
- Verspoor, Marjolijn, and Hana Smiskova. 2012. Foreign language writing development from a dynamic usage-based perspective. In *L2 Writing Development: Multiple Perspectives*. Edited by Rosa M. Manchón. Boston: De Gruyter Mouton, pp. 17–46.
- Vyatkina, Nina. 2012. The Development of Second Language Writing Complexity in Groups and Individuals: A Longitudinal Learner Corpus Study. *The Modern Language Journal* 96: 576–98. [[CrossRef](#)]
- Vyatkina, Nina, agen, irschmann, and Felix Golcher. 2015. Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing* 29: 28–50. [[CrossRef](#)]
- Wolfe-Quintero, Kate, Shunji Inagaki, and Hae-Young Kim. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity. Technical Report 17*. Manoa: University of Hawai'i Press.



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).