

Long-term effects of an earlier start

An analysis of EFL written production*

Teresa Navés, M. Rosa Torras and M. Luz Celaya
Universitat de Barcelona

The aim of this paper is to analyse the development of the written production of six groups of learners (N = 520) in a formal instruction setting and to investigate the effects of onset and age on the levels of attainment in writing performance. The written production was measured after 200, 416 and 726 hours of instruction. Both intragroup and intergroup analyses were carried out (a) to analyse the long-term effects of an earlier start in second language writing in a school setting, (b) to analyse the patterns of development between the four writing component measures depending on learners' age, and (c) to test whether the relationships found between the measures in the four areas of writing differ depending on learners' age group. This study concludes that (a) at the end of the instructional period Early Starters still have not surpassed Late Starters, (b) there seem to be two patterns of writing performance depending on learners' age, and (c) the correlations found between the writing component indicators change depending on learners' age group.

Introduction

During the past decade many countries in Europe and other parts of the world have decided to lower the starting age of foreign language instruction in school, on the grounds that learning at a younger age would result in better outcomes. This is the case of Spain, where in 1990 the introduction of English as a foreign language was lowered from grade 6 (age 11) to grade 3 (age 8). This new situation raises many questions concerning the influence of age on the development and outcomes of second language learning. Although there is a large body of theory and research on this topic, most studies have focused on naturalistic contexts and relatively few have been concerned with school settings where

exposure to the language is limited to the school classroom and to just a few hours per week.

The present study forms part of a larger project on the effects of age on the acquisition of English as a foreign language in a school context. The specific aim of this paper is to address the question of the effects of age and the long-term effects of an earlier start, with special reference to written production as an indicator of language development. With this in mind, we analysed the development of the written production of two groups of EFL learners who started their English instruction at different ages (8 and 11, respectively). Data collection took place after 200, 416 and 726 hours of instruction.

Age and second/foreign language acquisition

Despite extensive research into the relationship between age and second language acquisition, the conclusions remain controversial (Birdsong 1999, Collier 1988, DeKeyser 2000, Harley 1986, Harley and Wang 1997, Hyltenstam and Abrahamson 2001, Long 1990, Marinova-Todd, Marshall and Snow 2000, Singleton 1989, 1997, 2001). Studies conducted in untutored learning environments suggest that older learners may present initial short-term advantages in some features of SLA (for example, aspects of the lexicon and morphosyntax (Olsen and Samuels 1973, Snow and Hoefnagel-Höhle 1978)), whereas children may be better in other areas, for example pronunciation (Scovel 2000) or in what Cummins (1981) has labelled basic interpersonal communicative skills (BICS). However, the weight of evidence suggests that in the long run, learners whose exposure begins in childhood ultimately surpass those learners who started their exposure as adults, in both pronunciation (Asher and Garcia 1969, Munro, Flege and Mackay 1996, Oyama 1976, Tahta, Wood and Loewenthal 1981) and syntax (Harley 1986, Harley and Hart 1997, Johnson and Newport 1989, 1991, Patkowski 1980, Schachter 1990). According to these studies the general trend is to accept the view that 'younger is better', although it is important to note that children's claimed advantage does not lie in their rate of acquisition but in their ultimate attainment.

Despite the wealth of studies supporting the 'younger is better' position in untutored contexts, Slavoff and Johnson (1995) found no age differences concerning initial morphosyntactic intuition in a study carried out with 107 subjects who had arrived in the United States between the ages of 7 and 12 and whose native languages were typologically different from English. This responds to the claim by Snow and Hoefnagel-Höhle (1978) that the 'older = better'

results found in some studies carried out with learners of a second language that is typologically related to their L1 may be explained by the older learners' more sophisticated pattern recognition skills and their ability to transfer them from one language to the other, and not by their greater learning abilities.

With regard to the influence of age in formal L2 learning settings, specifically in school environments, the evidence favours the hypothesis that older learners are more successful than younger ones. Most studies carried out in this context sought to analyse the results of introducing the teaching of a foreign language into the primary school curriculum. A number of these studies analysed school learners with different starting ages but with the same number of hours of instruction. The results have consistently indicated that rate of acquisition increases with age. If the amount of exposure is held constant, older learners learn faster than younger ones especially in grammar (Bland and Keislar 1966, Cenoz 2002, Ekstrand 1978, Stankowski Gratton 1980). These better results generally obtained by older learners have also been confirmed by Canadian immersion programs. As Harley (1986) reports, older immersion learners had as much success as younger learners in short time periods.

Results obtained by the current authors in the Barcelona Age Factor Project (BAF), a large project (of which the current paper forms a part) on the effects of age on the acquisition of English as a foreign language, have also confirmed these results. In this project we collected data from over 1,700 participants who began English instruction at the age of 8 (Primary school) or 11 (Secondary school). Using a battery of both oral and written tests, data were collected after 200, 416 and 726 hours of instruction for both starting age groups. Results after 200 and 416 hours of instruction showed that learners who had started instruction at the age of 11 obtained significantly higher scores on the tests examined than participants who began at the age of 8. This was the case of those tests that may be considered more literacy-related or academically oriented, such as a cloze-test, a grammar test, a dictation, and a written composition (Muñoz 1999, Pérez-Vidal, Torras and Celaya 2000) and also an oral narrative in which a story was elicited from a series of pictures (Muñoz 2000). However, the results obtained from the listening comprehension test did not yield such large differences between the two groups (Muñoz 2003).

Other studies have compared school learners with different starting ages, but with different total hours of instruction, i.e. groups of pupils who received instruction at primary school with others who did not (see, for example, Blondin, Candelier, Edenlenbos, Johnstone, Kubanek-German and Taeschner 1998, Burstall, Jamieson, Cohen and Hargreaves 1974, Justman and Nass 1956,

Oller and Nagato 1974). In none of the cases did early starters with more hours of instruction show an advantage over late starters. However, preliminary results from two studies in the BAF project may have reached different conclusions. In an exploratory study (N= 115) Navés and Miralpeix (2002) found that at Time 1 (after 200 hours of instruction) Late Starters aged 12.9 no longer outperformed those early starters aged 10.9 who had received 100 additional instruction hours in most writing areas. Preliminary results from Navés (in progress) suggest that at Time 3 (after 726 hours of instruction), Late Starters aged 17.9 no longer outperform those Early Starters aged 16.9 who had received an average of 200 additional hours of instruction.

These studies show that the advantage of young beginners in the long run in naturalistic contexts did not appear in the data from school contexts. In this regard, Singleton (1989: 121) points out that before rushing to the conclusion that formal instructional conditions lead to different age-related results from natural exposure conditions in terms of eventual second language attainment, the sense attached to 'eventual' should be considered. As Singleton (1995: 3) suggests,

the eventual benefits of early second language learning in a formal instructional environment might be expected to show up only in rather longer-term studies than have to date been attempted.

There is a need for studies that explore whether it is feasible for younger school beginners to catch up to older formal beginners. The amount of exposure to the foreign language of instructed learners is a crucial factor,

maybe as crucial as the age at which initial exposure takes place, that is, the age at which pupils begin their instruction in the foreign language.
(Muñoz 1999: 21)

In addition to these long-term studies, there is also the need to explore how, throughout the school years, the age factor affects the acquisition of particular language modalities (for example, oral/written language), whether the development of different linguistic features is sensitive to it and what role amount and intensity of exposure play.

Foreign language writing

Foreign language classrooms have always had writing as a relevant activity for learning since, due to the characteristics of the context (the high number of students per class, no real communicative needs, etc.), writing has been a useful

tool to assess interlanguage development (for a review see Hillocks 1986). Besides, since writing is the only language skill that can be examined independently of the physical presence of the L2 learner (Leki 2000), it allows researchers to analyse large samples. Moreover, as Weissberg (2000) states, writing seems to reflect better than speech the emergence of new morphosyntactic forms and the development of grammar.

The best way to gauge development, however, is still controversial for two main reasons. First, the analysis of interlanguage has often suffered from the comparison with native-like competence (Cook 2002) and so, unfortunately, the same measures as in L1 written language have often been used for the analysis of L2 written products (Gass and Selloff Magnan 1993, Lumley 2002). Moreover, as Cumming (1998) points out, writing in a second language has been evaluated in an imprecise way or through scales with little validation to be able to provide indicators of achievement. Secondly, the analysis of written production in the specific context of the foreign language classroom, especially at beginner levels, has received little attention (Leki 1996). Therefore, more data are needed on written development in the school curriculum, as one of the indicators of interlanguage development.

Researchers have proposed a wide variety of measures that Wolfe-Quintero, Inagaki and Kim (1998) classify according to four major categories corresponding to different aspects of development: fluency, accuracy, syntactic complexity and lexical complexity. Several studies have used such measures to analyse both oral and written data in order to, for example, compare learners of different levels and different ages, and to study the effect of pedagogical treatment (Bardovi-Harlig and Bofman 1989, Carlisle 1989, Foster and Skehan 1996, Frantzen 1995, Kepner 1991).

One question that is receiving increasing attention is whether Wolfe-Quintero et al.'s four aspects of language development progress at the same rate, with more proficient second language learners being more fluent, accurate and complex than less proficient learners; or whether one of the components may progress at the expense of the others, as other researchers have suggested (Mackay 1982, Tedick 1990, 1995, Torras and Celaya 2001). Wolfe-Quintero et al. (1998) note that accuracy does not always proceed uniformly in relation to proficiency levels, whereas global complexity and fluency measures do. They found that accuracy measures were most often significantly related to holistic judgements, and were consistently linear for holistic judgements and short-term change in intact classes, whereas fluency, syntactic and lexical complexity measures were most often significantly related to program or school levels:

These results suggest that fluency and complexity measures may be related to the construct 'development' but that accuracy measures may be related to a different construct 'error'. (Wolfe-Quintero et al. 1998: 118)

Researchers have also indicated that there may be trade-offs between accuracy and complexity (Foster and Skehan 1996, Tedick 1990) or between accuracy and fluency (Hamp-Lyons 1991, MacKay 1982).

(...) syntactic complexity metrics would be misapplied if they were to be used as absolute developmental indices and as direct indices of language ability. (...) On the other hand, 'more complex' does not necessarily mean 'better' regarding communicative competence and speaking or writing ability. (Ortega 2000: 18)

Cumming, Kantor and Powers (2002) argue that particular types of writing and even specific writing tasks probably need unique scoring criteria. In a similar vein, in Celaya, Pérez-Vidal and Torras (2001) it is suggested that the specific features of the data in the BAF project (different ages, different levels of proficiency, especially beginner levels, and a school context) imply taking careful decisions on how to measure written production as an indicator of interlanguage development. We therefore designed a set of measures drawing on previous research but also creating new measures which could be used effectively in this context.

As mentioned above, this study forms part of a set of studies collectively known as the BAF Project. The project seeks to analyse the written production in EFL by bilingual Catalan-Spanish children. In an earlier publication, the written production of both early and late starters (henceforth ES and LS) was analysed after 200 and 416 hours of instruction applying a set of 17 measures (Celaya, Torras and Pérez-Vidal 2001, Pérez-Vidal, Torras and Celaya 2000, Torras and Celaya 2001). The study presented below adds data from written productions collected at a third time, after 726 hours of instruction when the ES are 16 years old and LS are 17. In addition, the set of measures used in the previous studies has been enlarged to include 23 new measures in order to make possible the analysis of the more developed interlanguage in the written compositions produced by the older learners with more hours of instruction.

Three research questions guided the present study:

1. Will learners who started instruction earlier (8 years old) (ES) catch up to those who started later (11 years old) (LS) at Time 3 of data collection (after 726 hours of instruction)?

2. Will the measures in the areas of fluency, lexical complexity, syntactic complexity and accuracy present similar patterns of development depending on learners' age?
3. Will the relationship between the measures in the four areas of writing differ depending on learners' age group?

Method

Participants

A total of 520 students participated in this study. The participants in this study are learners of English from different state schools in Barcelona. They are all Catalan-Spanish bilinguals. None of these subjects had followed extra curricular classes in or outside the school, nor did they have additional exposure to the language, apart from the usual channels of music, TV and Internet, which are almost impossible to control for. In identifying the subjects we have benefited from the fact that two different educational systems have coexisted for several years. We have therefore been able to include in our sample students who followed the new curriculum and began learning English at the age of eight (Elementary/Primary School) within the new state education system and also students that started English at the age of eleven within the former state education system in 6th grade (Secondary School). Table 1 presents the distribution of the sample into six groups according to age and number of hours of instruction.

Table 1. Subjects

	T1 200 h.	T2 416 h.	T3 726 h.
	G1	G3	G5
ES	10;9 N=129	12;9 N=107	16;9 N=55
	G2	G4	G6
LS	12;10 N=111	14;9 N=68	17;9 N=50

G1, G3 and G5 (odd-group numbers) represent the early starters (ES) groups of learners, i.e., those who started learning English when they were 8. The average age of G1 is 10;9 with 200 hours of instruction; the average age of G3 is 12;10 with 416 hours of instruction and the average age of G5 is 16;9 with 726 hours of instruction.

G2, G4 and G6 (even-group numbers) represent the late starters (LS) groups of learners who started English when they were 11. The average age of G2 is 12;10 years with 200 hours of instruction; the average age of G4 is 14;9 with 416 hours of instruction and the average age of G6 is 17;9 with 726 hours of instruction.

Procedure

For the present study the data come from a written composition, which was administered to participants in their own classroom by an external researcher. Both teachers and researchers made it clear to the students that the task would not be assessed as an exam. All the participants were given 15 minutes to write on the topic 'Introduce yourself: Me, my past life and my future'. In this way, both time and topic constraints were controlled for so as to make results comparable (Wolfe-Quintero et al. 1998).

Table 2 shows the 39 measures used to analyse the written production of the six groups of learners. The writing measures have been classified into four areas following Wolfe-Quintero et al. (1998) and previous studies from the authors of the present study: fluency, accuracy, syntactic complexity and lexical complexity (a detailed description of the measures and their rationale appears in Celaya, Pérez-Vidal and Torras 2001).

Analysis

The following statistical tests were used for the analysis of the data by means of the SPSS 10.0. First, in order to compare ES and LS at Time 3 (726 hours), G5 and G6, respectively, a t-test (independent samples) was performed in order to compare the differences of means between G5 and G6. Normality and homogeneity of variance was checked using the Z Kolmogorov-Smirnov test and the Levene test, respectively. The two groups of learners were normally distributed for each of the variables except for six measures, for which a Mann-Whitney test was performed.

Second, in order to find the developmental patterns of writing measures

Table 2. Measures for the analysis of EFL writing

Accuracy	Fluency	Syntactic Complexity	Lexical Complexity
1.* Error-free sentences (EFS)	4. Total number of words (TW)	12. Number of subordinated clauses (Sub)	28.* Noun tokens (TkNoun)
2. Error-free sentences in % (EFS%)	5.* Total number of words in English (TL2W)	13. Number of coordinated clauses (Coord)	29. Noun types (TyNoun)
3.* Number of rejected units (RejU)	6. Total number of nodes (Nod)	14.* Number of combined clauses (Comb)	30.* Adjective tokens (TkAdj)
	7. Total number of clauses (Cl)	15.* Coordination Index (CI)	31. Adjective types (TyAdj)
	8. Total number of sentences (S)	16. Number of non-finite nodes (NFNod)	32.* Adverb tokens (TkAdv)
	9. Words per Clause	17.* Ratio of nonfinite nodes per clause (NFNodxCl)	33. Adverb types (TyAdv)
	10. Words per sentence (WxS)	18.* Ratio of nonfinite nodes per sentence (NFNodxS)	34.* Lexical verb tokens (TkVlex)
	11.* (WxCl) (only for G5andG6)	19.* Ratio of clauses per sentence (ClxS)	35. Lexical verb types (TyVlex)
		20.* Ratio of subordinated clauses per clause (SubxCl)	36. Primary verb types (Prim)
		21.* Ratio of coordinated clauses per clause (CoordxCl)	37.* Open class words (OCW)
		22.* Ratio of combined clauses per clause (CombxC)	38.* Lexical Density (LD)
		23.* Ratio of subordinated clauses per sentence (SubxS)	39.* Total number of words in L1 (TWL1)
		24.* Ratio of Coordinated clauses per sentence (CoordxS)	40.* Total number of invented words (TWL12)
		25.* Ratio of combined clauses per sentence (CombXS)	
		26. Ratio of nodes per sentence (NodxS)	
		27. Auxiliary modal verbs (Aux/Mod)	

The ‘*’ indicates new measures presented in this study.

amongst the groups, that is, to compare the difference of means amongst the six groups of learners, Z Kolmogorov-Smirnov and Levene tests were first computed on each variable to test normality and the homogeneity of variances. The results from both tests show that normality could not be assumed. For this reason non-parametric analyses were conducted despite our large population.

The Kruskal-Wallis test, the non-parametric One-Way ANOVA, was performed to check whether there were significant differences between the six groups. Significant differences ($p < 0.05$) were found for each of the measures, which allowed us then to perform 15 series of Mann-Whitney tests.

Finally, two Pearson bivariate correlations were conducted to see the extent to which measures correlated with one another for two groups of learners: G1, G2 and G3, that is, learners below 12;10 years old, on the one hand, and G4, G5 and G6, that is, above 12;10, on the other. These two large groups were formed based on the results obtained from the series of Mann-Whitney tests, as discussed in the next section.

Results

The first goal of the present study was to examine whether ES (G5) would have caught up to LS (G6) in the long run within the school system (Time 3: 726 hours of instruction).

Table 3 indicates that LS (G6) significantly outperform ES (G5) in three areas of writing (accuracy, fluency and lexical complexity) and in most measurements of syntactic complexity. LS's (G6) accuracy in writing as measured by the percentage of error-free sentences (EFS %) is significantly higher than that of younger learners (G5). G6 also produces a higher total number of EFS, although this difference does not reach significance. G6 also significantly outperforms G5 in sentence length measured as the total number of words per sentence (WxS), and in lexical complexity, G6 significantly outperforms G5 in three domains. First, G6 significantly uses more adverbs (tokens) than G5. Second, G5 still significantly relies on the use of L1 words when compared to G6. G6, on the other hand, takes significantly more risks when attempting to use target-like words and produces a significantly higher number of invented words. Syntactic complexity measures give mixed results. On the one hand, LS (G6) significantly outperform ES (G5) in eight of the syntactic complexity indicators (see Table 3), while ES (G5) get significantly higher scores in four of the measures. A more careful look at the measures involving non-finite nodes might account for this finding and is discussed in the next section.

To sum up, out of the 39 measures, the mean of 31 of them obtained by the LS was higher than the mean obtained by ES learners. Statistically significant differences between the two groups of learners were found in sixteen measures of writing ability: for 12 measures, LS outperforms ES; and in four measures ES outperform LS. LS learners' accuracy, fluency, and lexical complexity are significantly better than ES. LS learner's syntactic complexity is significantly higher as measured by means of subordination, CI, and other ratios involving subordination, but significantly lower in non-finite nodes measures. In other

Table 3. T-Test results from G5 and G6 (ES and LS) (Time 3)

	Group	Mean	SD	P		Group	Mean	SD	p
S	5	7.67	2.65	0.070	NodxS	5	2.29	0.75	0.099
	6	6.68	2.87			6	2.54	0.77	
EFS	5	2.57	1.73	0.232	TWL1	5	0.95	1.84	0.008
	6	3.02	1.81			6	0.22	0.68	
EFS%	5	29.41	17.23	0.001	TWL12	5	0.02	0.13	0.047
	6	43.67	23.64			6	0.14	0.40	
Clauses	5	13.73	5.70	0.704	TkNoun5	5	14.78	6.33	0.241
	6	14.18	6.39			6	13.30	6.51	
Sub	5	2.44	2.09	0.000	TyNoun	5	12.44	5.33	0.092
	6	4.24	2.61			6	10.74	4.89	
Coord	5	3.82	2.40	0.657	TkAdj	5	4.02	2.88	0.293
	6	3.62	2.16			6	4.68	3.48	
Comb	5	6.05	4.09	0.069	TyAdj	5	3.42	2.18	0.240
	6	7.52	4.08			6	4.00	2.78	
CI	5	66.93	29.93	0.004	TkAdv	5	4.05	3.21	0.041
	6	50.51	26.88			6	5.50	3.89	
NFNod	5	3.31	2.50	0.000	TyAdv	5	3.20	2.59	0.059
	6	1.82	1.66			6	4.20	2.76	
NFNod	5	0.31	0.38	0.402	TkVLex	5	5.09	2.86	0.955
	6	0.23	0.57			6	5.06	2.77	
xCl	5	0.51	0.51	0.320	TyVLex	5	7.33	3.41	0.398
	6	0.40	0.62			6	7.90	3.49	
ClxS	5	1.80	0.51	0.002	TkPrim	5	6.56	3.20	0.934
	6	2.15	0.57			6	6.62	3.73	
SubxCl	5	15.93	10.57	0.000	TyPrim	5	4.05	1.53	0.687
	6	28.38	12.39			6	3.92	1.85	
Coordx	5	26.19	11.63	0.638	TyAux/	5	4.85	1.82	0.150
	6	25.08	12.36			6	5.48	2.50	
Cl	5	42.12	15.30	0.000	TOCW	5	27.95	12.31	0.821
	6	53.46	16.57			6	28.54	14.31	
SubxS	5	0.31	0.26	0.000	TW	5	90.58	33.44	0.731
	6	0.65	0.36			6	93.14	41.62	
Coordx	5	0.51	0.33	0.391	LD	5	30.42	5.31	0.826
	6	0.57	0.36			6	30.19	5.13	
S	5	0.82	0.48	0.000	TWL2	5	89.62	33.85	0.672
	6	1.21	0.59			6	92.78	41.56	
Combxs5	5	13.58	5.74	0.628	Wxs	5	11.92	3.56	0.003
	6	14.16	6.40			6	14.08	3.60	
Tnod	5	16.89	6.48	0.474	WxCL	5	6.71	1.41	0.736
	6	15.98	6.48			6	6.63	0.89	

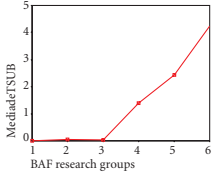
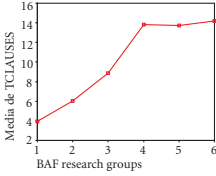
words, in the long run within the school system, at Time 3 (726 hours). LS's writing is still significantly better than ES's in the four writing areas of fluency, accuracy, lexical and syntactic complexity. (ES perform significantly better than LS in only four indicators involving non-finite nodes of syntactic complexity.)

The second major goal of this study was to find the development of writing measures within the groups. The results of the analysis of the Mann-Whitney tests performed allowed us to arrange them into two patterns.

Table 4 summarises the results obtained in the analysis of written production in each of the four areas (fluency, accuracy, syntactic and lexical complexity) for each group of learners. The first row of Table 4 shows the qualitative tendency being represented by the two graphs corresponding to the total number of subordinated clauses and the total number of clauses, respectively. These two graphic representations of the two patterns found are clearly distinct and with minor changes they represent all the measures listed in the rest of the table. All the measures included in each pattern show significant differences between either the younger (G1, G2, G3) or the older groups (G4, G5 and G6).

Pattern I shows a tendency for almost no interlanguage development as measured by writing between the first three groups of younger learners (below 12;10) at Times 1 and 2 (after 200 and 416 hours of instruction, respectively) and then a steady rise in the older groups (above 12;10) for most syntactic complexity measures and for adverbs (tokens and types) and lexical verb (types). Moreover, in Pattern I, significant differences are found amongst learners older than 12;10. That is, significant differences are found between G3, G4, G5 and G6. G6 (LS) significantly outperforms G5 (ES) in the measures of fluency, syntactic and lexical complexity listed in the table. G5 in turn outperforms G4, and G4 outperforms G3. ($G6 > G5 > G4 > G3$). The eldest group of learners is better in most syntactic complexity indicators, in Words per Sentence (WxS) and in lexical measures involving adverbs and lexical verbs (see Table 4 for details).

Table 4. Patterns of writing development

		
General characteristic	<p>The general trend shows a flat gain from G1 to G3 to then dramatically increase from G3 to G5 (or G6)</p> <p>(*) Significant gains found above 12;10 year-old learners between G3 and G5</p>	<p>The general trend shows a sharp gain from G1 to G4 to then stop growing</p> <p>(*) Significant differences found below 14;9 year-old learners G1, G2, G3 and G4</p>
Specific Features	<p>(a) Substantial gains also found between below 12;10 year-old learners although the line is much flatter. G3 > G2 > G1</p> <p>(b) Low gains from G1 to G3 to then rapidly and significantly increase from G3 to G5. No statistical differences found for the two older groups (G5 and G6)</p> <p>(c) Low gains from G1 to G3 to then rapidly and significantly increase from G3 to G5. Significant decline found between G5 and G6.</p>	<p>(d) There are no statistical differences found between the older learner groups (above 12;10 years old) G4, G5 and G6.</p> <p>(e) There is a significant drop from G4 to G5 but no significant differences found between G5 and G6</p>
Accuracy Measures		<p>Rejected Units (*) (d)</p> <p>Error-free sentences (*) (e)</p>
Fluency Measures	Words per Sentence (*) (a)	<p>Total number of Words (*) (d)</p> <p>Clauses (*) (d)</p> <p>Sentences (*) (e)</p> <p>Total N of L2 words (*) (d)</p> <p>Coordinated clauses (*) (d)</p>
Syntactic Complexity Measures	<p>Subordinated clauses (*) (a)</p> <p>Combined clauses (*) (a)</p> <p>Coordination Index (*) (except between G4 and G5) (a)</p> <p>Clauses x Sent (*) (a)</p> <p>Sub x Clause (*) (a)</p> <p>Comb x clause (*) (a)</p> <p>Sub x Sent (*) (a)</p> <p>Comb x Sent (*) (a)</p> <p>Aux Mod (*) (b)</p> <p>Coord x Sent (*) (b)</p> <p>Coordinated clauses per Clause (b)</p> <p>Non-finite nodes (*) (c)</p> <p>Non-finite nod x Sent (*) (c)</p> <p>Non-finite nodes per Clause (c)</p>	
Lexical Complexity Measures	<p>Adverb Tokens (*) (b)</p> <p>Adverb Types (*) (b)</p> <p>Lexical Verb Types (*) (b)</p>	<p>Finite verbs (tokens) (*) (d)</p> <p>Noun Tokens (*) (d)</p> <p>Noun Types (*) (d)</p> <p>Adjective Tokens (*) (d)</p> <p>Adjective Types (*) (d)</p> <p>Primary as Main (Tkns) (*) (d)</p> <p>Open class words (*) (d)</p>

Pattern II shows a steady development in writing for the first four groups of younger learners (below 14;9 years) at Times 1 and 2, after 200 and 416 hours of instruction which suddenly stops in the older groups (aged above 14;9) for accuracy, fluency and some lexical measures. Results from the series of pairs of Mann-Whitney tests show that the development of most accuracy and fluency measures follow Pattern II, that is, there is an increasing gain in these writing components for the younger groups aged below 14;9 (G1, G2, G3 and G4) whereas most syntactic complexity indicators follow Pattern I in which increasing gains are found above 12;10 year-old learner groups (G4, G5 and G6). The development of lexical complexity measures is split between the two patterns. Most lexical complexity measures follow Pattern II (total number of finite lexical verbs, nouns (both tokens and types), adjectives (both tokens and types), primary main verbs and total number of open class words), with significant differences found for all the indicators between the groups below 14;9 (G1, G2, G3 and G4). Lexical complexity indicators measured as total number of adverbs (both as types and as tokens) as well as total number of lexical verbs (types) follow Pattern I, although no significant differences were found between G5 and G6.

The third aim of this study was to find out how the different measures of writing ability correlated with each other and to check the extent to which the strength of the correlation changed depending on the learners' age group. Table 5 reports the correlations between measures which either show a high correlation between indicators of different writing components (lexical and syntactic complexity or accuracy and fluency, for example) or those whose correlation coefficients substantially change across the age-span.

Accuracy and Fluency: Error-free sentences (EFS) and total number of words (TW) are strongly correlated (.70) in the younger group but weakly correlated (.34) in the older group. In other words, the more fluent younger learners are, the more accurate their writing may also be. Accurate and fluent writing, however, do not seem to go hand in hand when learners are older.

Accuracy and Lexical Complexity: The same trend is found between accuracy and lexical complexity measures. In the younger group of learners, Error-free sentences (EFS) strongly correlate with lexical complexity as measured by tokens and types of nouns. In the older group, however, these correlations are much weaker.

Fluency and Lexical Complexity: The total number of sentences, clauses or finite verbs strongly correlates with nouns and adjectives (tokens and types); clauses

Table 5. Correlations between measures

Correlations of measures within < 12;10 year-olds and > 12;10 year-olds.		< 12;10	> 12;10	
Unless indicated, the correlation is significant ($p < .005$) and $p = .000$		group	group	
Total number of words	Error-free sentences	.707	.345	↓
Total number of words	Error-free sentences %	.263	-.173	↓
Total number of words	Aux/Modals	.150	.544	
Total number of words	Subordinated clauses	($p = .002$) .168	.594	↑↑
Total number of words	Adverbs (tokens and types)	.437	.661	
		.347	.646	↑↑
Total error-free sentences	Error-free sentences in %	.666	.699	≡
Total error-free sentences	Nouns (tokens and types)	.668	.435	
		.633	.442	↓
Total error-free sentences	Adjectives (tokens and types)	.579	.334	
		.592	.313	↓
Total number of sentences	Total of Finite verbs	.942	.754	↓
Total number of sentences	Total number of clauses	.965	.769	
Total number of sentences	Nouns (tokens and types)	.864	.771	↓
		.827	.724	
Total number of sentences	Adjectives (tokens and types)	.700	.599	↓
		.676	.555	
Total number of sentences	Lexical verbs	.596	.645	↑
Subordinated clauses	Adverbs (tokens and types)	No sig. .011	.602	↑↑
		No sig. .014	.606	
Subordinated clauses	Coordination Index	-.809	-.506	↓
Subordinated clauses	Finite verbs	No sig. .143	.630	↑↑
Subordinated clauses	Words per sentence	.213	.539	↑
Coordinated clauses	Finite verbs	.614	.729	↑
Coordinated clauses	Nouns (tokens)	.625	.543	↓
Coordinated clauses	Words per sentence	.611	.454	↓
Finite Verbs	Nouns (tokens and types)	.891	.743	↓
		.854	.655	
Finite Verbs	Adjectives (tokens and types)	.716	.669	↓
		.697	.611	
Finite Verbs	Adverbs (tokens and types)	.438	.628	↓
		.334	.601	
Finite Verbs	Words per sentence	.543	.277	↓
Clauses	Nouns (tokens and types)	.906	.759	↓
		.863	.670	
Clauses	Adjectives (tokens and types)	.728	.681	↓
		.708	.621	
Clauses	Adverbs (tokens and types)	.424	.597	↑
		.318	.661	
Nouns (tokens)	Adjectives (tokens and types)	.687	.588	↓
		.661	.529	
Nouns (tokens)	Lexical verbs (tokens and types)	.595	.598	
		.566	.554	≡
Nouns (tokens)	Words per sentence	.647	.209	↓
Lexical verbs (types)	Adverbs (tokens and types)	.284	.619	↑
		.187	.652	
Words per sentence	Lexical Density	.486	$p = 0.60$ -.143	↓
Words per sentence	Aux Modal	$P = 0.12$.135	.585	↓

and finite verbs also strongly correlate with adverbs regardless of learners' age group. The relationship between words per sentence (WxS) and lexical complexity measures such as the total number of nouns (tokens) and lexical density (LD) is also very strong within the younger group, but gets weaker as learners get older. In contrast, the total number of words and adverbs (tokens and types) and classes of words and adverbs (tokens and types), respectively, are weakly correlated within the younger group, but as learners get older this correlation turns into a rather strong one.

Fluency and Syntactic Complexity: In the younger group, there is a weak correlation between sentence length and total number of words and the total number of modal auxiliary verbs and subordinate clauses, respectively; the same measures correlate strongly in the older group of learners. In other words, there are more chances that, for older learners, the more fluent they are, the more syntactically complex their writing. The same does not apply, or does not apply so strongly, for younger learners.

Syntactic Complexity and Lexical Complexity: In younger learners, we did not find a significant correlation between number of subordinated clauses and adverbs, on the one hand, and between finite verbs and subordinate clauses, on the other. However, as learners get older, the two metrics are not only significantly but also strongly correlated.

Discussion

In this study we investigated English language development as reflected in the writing of six groups of learners and among the older learners with greater amount of instruction in particular. The comparison between the oldest two groups of learners (G5 and G6) allows us, first, to see long-term effects of an earlier start and, secondly, to identify how the four indicators of writing proficiency (fluency, accuracy, syntactic and lexical complexity) develop in the long run within a school-setting.

Previously published parts of this study had found that ES did not catch up to LS either after 200 or after 416 hours of instruction (Celaya, Torras and Pérez-Vidal 2001). The question was, therefore, whether they would do so after 726 hours.

Table 6 summarises the results found after 726 hours of instruction. As we have seen, starting age was found to have a clear influence on the scores

Table 6. Results from ES and LS at Time 3 (G5 and G6)

Summary of results for RQ1	G6 significantly outperforms G5	G5 significantly outperforms G6
Accuracy	EFS % *(0.00)	
Fluency	WxS *(0.00)	
Syntactic Complexity	Sub *(0.00)	NFNod (0.00)
	Comb *(0.04)	NFNxCl * (p=0.01)
	CI *(0.00)	NFNxS *(0.04)
	ClxS * (0.00)	Combxs (0.00)
	SubxCl * (0.00)	
	CombxCi* (0.00)	
	SubxS * (0.00)	
Lexical Complexity	TWL1 *(0.00)	
	TWL12 *(0.04)	

obtained for the written composition. At Time 3 ES (G5) have not surpassed LS (G6), even if some differences in favour of LS found in previous studies by the authors may be fading, as shown by the indicators for which there are now no significant differences found between the two groups. At Time 3, accuracy, lexical complexity and syntactic complexity show clear advantages for LS over ES. Fluency as measured by total number of words shows no significant differences between the two groups; whereas words per sentence show an advantage for LS. Sentence length measured as words per sentence (WxS) has traditionally been regarded by researchers as a syntactic complexity measure. Wolfe-Quintero et al. (1998) however, after examining 39 empirical studies concluded that WxS was best seen as a fluency measure. The authors of the present study have therefore adopted the same stance, both in this study and in their previous studies with younger learners, despite the fact that recent studies on writing (Li 2000a, 2000b, Ortega 2000, Sotillo 2000) still question whether WxS should be regarded as a fluency or a complexity measure.

If WxS is regarded as a measure of fluency, LS (G6) outperform ES (G5) in the long run within a school setting (Time 3: 726 hours) in the four areas of writing ability. If, on the other hand, mean length of a sentence is regarded as a measure of syntactic complexity, then at Time 3 LS outperform ES only in three of the writing components but not in the area of fluency.

The fact that ES have not surpassed LS after 726 hours does not come as a surprise. Independent measures of learners' proficiency (Muñoz 1999) obtained

Table 7. Summary of correlation results

	Fluency		Syntactic complexity		Lexical complexity	
	<12;10	>12;10	<12;10	>12;10	<12;10	>12;10
Accur.	ESF – TW Strong	EFS – TW Weak	None	None	EFS – Noun Strong	EFS – Noun Weak
					EFS – Adv Weak	EFS – Adv Weaker
Fluency			TW – Sub/AuxM Weak	TW – Sub/AuxM Strong	S – N/Adj Stronger	S – N- Adj Strong
					S – Vlex Strong	S – Vlex Stronger
					Cl – FinV/N/ Adj Stronger	Cl – FinV/N/Adj Strong
			WxS – Sub/ AuxM Weak	WxS – Sub/AuxM Strong	TW/Cl – Adv Weak	TW/Cl – Adv Strong
					WxS – N Strong	WxS – N Weak
					WxS – LD Strong	WxS – LD Weak
Syntac Compl			Cl – Sub Strong	Cl – Sub Weak	Sub – Adv Non signif.	Sub – Adv Strong
					Sub – FinV Non signif.	Sub – FinV Strong

both through a standardised multiple-choice grammar test and a cloze test show that the late starters are significantly more proficient than earlier starters in this same population of learners. This finding is also consistent with the research findings on writing competence with a similar population of Euskera-Spanish bilingual children conducted in the Basque Country. Cenoz (2002) and Doiz and Lasagabaster (2001) also found that LS significantly outperformed ES at the three times of data collection. It remains to be seen whether the same relationship between earlier and late starters will still remain by the time the learners reach the end of compulsory schooling.

Alternative explanations to the higher written proficiency of late starters may be sought in the fact that Time 3 is the end of the teaching period for the older

group of learners (G6), who will be taking a University entry exam after finishing the school year. The wash-back effect of a national test may have influenced the results, since the entry exam contains a composition, and it is not uncommon for secondary school teachers to train their students, particularly during the last academic year of schooling in writing 75–100 word compositions.

Another factor that may have influenced the results is that late starters are one year older than early starters, and therefore more cognitively mature. Cummins (1980) has argued that older learners show higher mastery of L2 syntax, morphology and other literacy-related skills, such as vocabulary and reading comprehension, due to their greater cognitive maturity. Muñoz (2002) has suggested that the late starters' faster rate in this study is due not only to their superior cognitive development but also to their more advanced communication skills.

Finally, the intensity of the instruction learners received, and the way instruction was distributed in the two groups may also account for the differences found. The methodological changes that the introduction of a new educational system may have involved are still unclear. The intensity of instruction during the last period of schooling and, therefore, of data collection was greater for late starters than for early starters. The amount of instruction received by early starters is far more spread out within the three collection times than that of the late starters.

With respect to those measures where the early starters outperformed the late starters, measures involving coordination and non-finite nodes, these seem to mark an early stage of syntactic complexity (Allen 1972, Doughty 1991, Pienemann 1998) and are typical of early development in both second and first language development (Ferraris 2001). It seems as if the early starters can cope with *-ing* and *to + inf* verb phrases (nonfinite subordination) but not full finite subordinate clauses. Whether this is due to a developmental constraint, lack of the structures in the input or because 'non-finite nodes' are in fact less complex than subordinated clauses is something further research needs to investigate.

This study confirms the poor results obtained by young learners in educational settings in previous studies (Burstall et al. 1974, Cenoz 2002, Doiz and Lasagabaster 2001, Ekstrand 1976, Muñoz 2000, Oller and Nagato 1974). The lower scores obtained by the younger group could be due to differences in cognitive maturity, linguistic development, and test-taking strategies. The differences observed could also be related to type of input or instructional style (Harley 1986, Harley and Hart 1997) with wash-back exam training explaining the higher ultimate attainment in written production of the older learners.

While the underlying assumption of much second language writing research is that fluency, accuracy, syntactic and lexical complexity progress in tandem (Wolfe-Quintero 1998), other researchers have suggested that these four components of writing might develop at different rates:

For reasons that at present are unclear, ESL writers often acquire different components of written control at different rates. (...) We also sometimes see writers who have mastered a wide vocabulary but markedly less syntactic control (...) It may be that as ESL writers advance, the strands of competence in their writing converge until at the highest levels unity (that is, an equal measure of competence in all areas) is achieved. (Hamp-Lyons 1991:241)

Previous analyses with participants from this study from T1 and T2 had found that the four writing areas (fluency, accuracy, syntactic and lexical complexity) did not seem to develop in parallel (Torras and Celaya 2001: 116). The present study suggests that the development of writing ability follows two clear patterns as summarised in Table 4. On the one hand, syntactic complexity and lexical complexity as measured by adverbs (types and tokens) show a pattern of almost no development when learners are younger than 12;10 but a steady and significant development above that age. On the other hand, accuracy and fluency measurements along with noun and adjective measures of lexical complexity (types and tokens), follow a pattern of even and steady development up to the age of 14;9 followed by a plateauing of development after that age. These results suggest that syntactic complexity as well as the acquisition of adverbs take place once learners' written accuracy, fluency and basic lexical complexity has developed to a critical level.

One possible interpretation for the lack of continuous growth in accuracy, fluency and lexical measures of nouns and adjectives is the fact that writing is a cognitively complex skill. As they become more proficient, second and foreign language writers find it hard or impossible to improve the four major areas of writing simultaneously. Allen (1997), for example, in an intervention program designed to improve second language learners' writing found that 'whereas mastery of a majority of features improved, more errors were observed in the use of demonstrative/relative pronouns, subject-verb agreement, and punctuation of coordinated and embedded clauses' (Allen 1997:9).

Uneven development has been noted in both L1 acquisition where the lexicon is expanded by, first, using nouns, then predicates and finally closed-class items (D'Odorico, Carubbi, Salerno and Calvo 2001, Foster-Cohen 1999, Harris and Chasin 1999, Maital, Dromi, Sagi and Bornstein 2000) and L2

research where nouns and adjectives seem to be more readily learnable than verbs and adverbs (Broeder, Extra and Van Hout 1993, McShane 1991). Vocabulary researchers have also suggested that in both first and second language acquisition nouns are acquired before adjectives, and adverbs are the last to be acquired (Nation 1990). Similar evidence is provided by the current study which has suggested that the development of nouns (tokens and types) and adjectives (tokens and nouns) steadily increases among younger learners while the development of adverbs takes place much later, once adjectives and nouns have been established after an average of 416 hours of instruction and when learners are older than 12;10.

With respect to the third aim of the study, the results suggest that the relationship between accuracy, fluency, syntactic and lexical complexity changes depending on learners' age group. Fluency and accuracy measures are strongly correlated if measured in young learners' writing but are weakly correlated among old learners (above 12;10), suggesting that the more young learners write, the more accurate their writing becomes; and that writing accurately may induce longer compositions. Fluency and syntactic complexity on the other hand are only weakly correlated within the younger group but strongly correlated within the older group, possibly because of the increased, and late, use of complex sentences.

The relationship between lexical complexity and fluency follows the same trend as that between accuracy and fluency, being strongly related in the group of younger learners but only weakly related in the group of older learners. Finally, a relationship between syntactic and lexical complexity is non-existent among the younger learners whereas it is quite strong among the older groups. One possible explanation suggested above is the late development of syntactic indicators involving subordination and lexical indicators such as adverbs. In sum, the correlations found between different measures of development in writing seem to depend on learners' age group.

Conclusions

1. Early starters do not surpass late starters after 726 hours of instruction (the end of instruction for the late starters) in a school setting. Overall, late starters significantly outperform early starters in the four areas of writing (fluency, accuracy, syntactic complexity and lexical complexity). However, there are four measures of syntactic complexity (involving ratios of non-finite nodes in

particular) for which early starters significantly outperform late starters. There are also a considerable number of measures which do not turn out to be statistically significant. In the light of these results, it cannot be claimed that early starters have caught up to late starters, although they might do with a longer period of comparison given the reduction in significance of certain measures between the two groups.

2. There seem to be two different patterns of development in EFL written production (accuracy, fluency, syntactic complexity and lexical complexity) depending on learners' age. Pattern I shows almost no interlanguage development as measured by writing between the first three groups of younger learners (below 12;10) at Times 1 and 2 and then a steady increase in the older groups (above 12;10) for most syntactic complexity measures and for adverbs. Pattern II shows a steady development in the first four groups of younger learners (below 14;9 years) at Times 1 and 2. This development stops in the older groups (aged above 14;9) for accuracy, fluency and some lexical measures.

3. Accuracy, fluency, syntactic and lexical complexity do not develop in tandem, but correlate differently depending on the learners' age group and the strength of the relationship between the measures in the four components.

Note

* The authors acknowledge the support of the Ministerio de Educación y Cultura (Spain) through projects PB-94-0944 and PB-97-090. We are grateful to Maribel Peró (Facultat de Psicologia, Universitat de Barcelona) for her advice on statistics. We are indebted to Carme Muñoz, Joe Hilferty and Imma Miralpeix (Facultat de Filologia, Universitat de Barcelona) for their insights on earlier versions of this paper. Responsibility for the final version, however, rests with the authors.

References

- Allen, A. E. 1972. "Frequency and control of selected syntactic structures as related to quality rating of written expression". *DAI* 32: 4477A.
- Allen, R. and Rochecouste, J. 1997. "Intervention in the tertiary writing process". *TESOL in Context* 7 (2): 9-13.
- Asher, J. and Garcia, R. 1969. "The optimal age to learn a foreign language". *Modern Language Journal* 53: 334-341.
- Bardovi-Harlig, K. and Bofman, T. 1989. "Attainment of syntactic and morphological accuracy by advanced language learners". *Studies in Second Language Acquisition* 11 (1): 17-34.

- Birdsong, D. 1999. "Introduction: Whys and why nots of the critical period hypothesis for second language acquisition". In *Second Language Acquisition and the Critical Period Hypothesis*, D. Birdsong (ed.), 1–22. Mahwah, NJ: Lawrence Erlbaum.
- Bland, M. and Keislar, E. 1966. "A self-controlled audio-lingual program for children". *French Review* 40: 266–76.
- Blondin, C., Candelier, M., Edenlenbos, P., Johnstone, R., Kubanek-German, A. and Taeschner, T. 1998. *Foreign Languages in Primary and Pre-School Education. A Review of Recent Research within the European Union*. London: CinterlanguageT.
- Broeder, P., Extra, G. and Van Hout, R. 1993. "Richness and variety in the developing lexicon". In *Adult Language Acquisition: Field Methods*, C. Perdue (ed.), 145–163. Cambridge: Cambridge University Press.
- Burstall, C., Jamieson, M., Cohen, S. and Hargreaves, M. 1974. *Primary French in the Balance*. Windsor: NFER.
- Carlisle, R.S. 1989. "The writing of Anglo and Hispanic elementary school students on bilingual, submersion and regular programs". *Studies in Second Language Acquisition* 11 (3): 257–280.
- Celaya, M.L., Pérez-Vidal, C. and Torras, M.R. 2001. "Matriz de criterios de medición para la determinación del perfil de competencia lingüística escrita en inglés". *Revista Española de Lingüística Aplicada* 14: 87–98.
- Celaya, M.L., Torras, M.R. and Pérez-Vidal, C. 2001. "Short and mid-term effects of an earlier start: an analysis of EFL written production". In *EUROSLA Yearbook*, S. Foster-Cohen and A. Nizegorodcew (eds), 195–209. Amsterdam: John Benjamins.
- Cenoz, J. 2002. "Age Differences in Foreign Language Learning". *International Review of Applied Linguistics* 135–136: 125–142.
- Collier, V.P. 1988. "The Effect of Age on Acquisition of a Second Language for School". *New Focus. Occasional Papers in Bilingual Education* 2: <http://www.ncela.gwu.edu/ncbebpubs/classics/focus/02aage.htm>
- Cook, V. 2002. "Background to the L2 user". In *Portraits of the L2 User*, V. Cook (ed.), 1–27. Clevedon: Multilingual Matters.
- Cumming, A. 1998. "Theoretical perspectives on writing". *Annual Review of Applied Linguistics* 18: 61–78.
- Cumming, A., Kantor, R. and Powers, D.E. 2002. "Decision making while rating ESL/EFL writing tasks: A descriptive framework". *The Modern Language Journal* 86: 67–96.
- Cummins, J. 1980. "The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue". *TESOL Quarterly* 14 (2): 175–187.
- Cummins, J. 1981. "Age on arrival and immigrant second language learning in Canada: A reassessment". *Applied Linguistics* 2: 132–149.
- DeKeyser, R.M. 2000. "The robustness of critical period effects in second language acquisition". *Studies in Second Language Acquisition* 22: 499–533.
- D’Odorico, L., Carubbi, S., Salerni, N. and Calvo, V. 2001. "Vocabulary development in Italian children: a longitudinal evaluation of quantitative and qualitative research". *Journal of Child Language* 28 (2): 351–372.
- Doiz, A. and Lasagabaster, D. 2001. "El efecto del factor edad en la producción escrita en inglés". In *La Lingüística Aplicada a Finales del Siglo XX. Ensayos y Propuestas*, I. De la

- Cruz, C. Santamaría, C. Tejedor and C. Valero (eds), 63–68. Alcalá de Henares: Universidad Alcalá de Henares.
- Doughty, C. 1991. “Second language instruction does make a difference: Evidence from an empirical study of ESL relativization”. *Studies in Second Language Acquisition* 13 (3): 431–469.
- Ekstrand, L. 1978. “English without a book revisited: The effect of age on second language acquisition in a formal setting”. In *Didakometry 60. Department of Educational and Psychological Research*. School of Education, Malmö.
- Ferraris, S. 2001. “Text Organisation in Italian L2 Learner Varieties”. In *EUROSLA Yearbook*, S. Foster-Cohen and A. Nizegorodcew (eds), 225–237. Amsterdam: John Benjamins.
- Foster, P. and Skehan, P. 1996. “The influence of planning and task type on second language performance”. *Studies in Second Language Acquisition* 18 (3): 299–323.
- Foster-Cohen, S. 1999. *An Introduction to Child Language Development*. London: Longman.
- Frantzen, D. 1995. “The effects of grammar supplementation on written accuracy in an intermediate Spanish content course”. *Modern Language Journal* 79 (3): 329–355.
- Gass, S. and Sieloff Magnan, S. 1993. “Second-language production: SLA research in speaking and writing”. In *Research in Language Learning. Principles, Processes, and Prospects*, A. Omaggio Hadley (ed.), 156–196. Illinois: National Textbook Company.
- Hamp-Lyons, L. 1991. “Scoring procedures for ESL contexts”. In *Assessing Second Language Writing in Academic Contexts*, L. Hamp-Lyons (ed.) 241–276. Westport, Conn: Ablex.
- Harley, B. 1986. *Age in Second Language Acquisition*. Clevedon Avon: Multilingual Matters.
- Harley, B. and Hart, D. 1997. “Language aptitude and second language proficiency in classroom learners of different starting ages”. *Studies in Second Language Acquisition* 19: 379–400.
- Harley, B. and Wang, W. 1997. “The Critical Period Hypothesis: Where are we now?”. In *Tutorials in Bilingualism: Psycholinguistic Perspectives*, A.M.B. de Groot and J.F. Kroll (eds), 19–51. Mahwah, N.J.: Erlbaum.
- Harris, M. and Chasin, J. 1999. “Developments in early lexical comprehension: a comparison of parental report and controlled testing”. *Journal of Child Language* 26: 453–460.
- Hillocks, G. 1986. “Research on written composition. New directions for teaching”. *ERIC Clearinghouse on Reading and Communication Skills*. National Institute of Education. <http://www.ncte.org/rte/Hillocks.pdf>
- Hyltenstam, K. and Abrahamsson, N. 2001. “Age and L2 learning: The hazards of matching practical ‘implications’ with theoretical facts”. *TESOL Quarterly* 35 (1): 151–170.
- Johnson, J.S. and Newport, E.L. 1991. “Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language” *Cognition* 38: 215–258.
- Johnson, J.S. and Newport, E.L. 1989. “Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language”. *Cognitive Psychology* 21: 60–99.
- Justman, J. and Nass, M. 1956. “The high school achievement of pupils who were and were not introduced to a foreign language in elementary school”. *Modern Language Journal* 40: 120–23.

- Kepner, C.G. 1991. "An experiment in the relationship of types of written feedback to the development of second language writing skills". *Modern Language Journal* 75 (3): 305–313.
- Laufer, B. and Nation, P. 1995. "Vocabulary size and use: Lexical richness in L2 written production". *Applied Linguistics* 16 (3): 307–322.
- Leki, I. 1996. "L2 composing: strategies and perceptions". In *Writing in a Second Language*, B. Leeds (ed.), 27–37. London: Longman.
- Leki, I. 2000. "Writing, literacy, and applied linguistics". *Annual Review of Applied Linguistics* 20: 99–115.
- Li, Y. 2000. "Assessing second language writing: The relationship between computerized analysis and rater evaluation". *International Review of Applied Linguistics* 127–128: 37–52.
- Li, Y. 2000. "Linguistic characteristics of ESL writing in task-based e-mail activities". *System* 28 (2): 229–245.
- Long, M. 1990. "Maturational constraints on language development". *Studies in Second Language Acquisition* 12: 257–280.
- Lumley, T. 2002. "Assessment criteria in a large-scale writing test: what do they really mean to the raters?" *Language Testing* 19 (3): 246–276.
- MacKay, G.D. 1982. "The problems of flexibility, fluency, and speed-accuracy tradeoff in skilled behavior". *Psychological Review* 89: 483–506.
- Maital, S., Dromi, E., Sagi, A. and Bornstein, M.H. 2000. "The Hebrew communicative development inventory: language specific properties and cross-linguistic generalizations". *Journal of Child Language* 27: 43–67.
- Marinova-Todd, S.H., Marshall, D. and Snow, C.E. 2000. "Three misconceptions about age and L2 learning". *TESOL Quarterly* 34 (1): 9–34.
- McShane, J. 1991. *Cognitive Development: An Information Processing Account*. Oxford: Blackwell.
- Munro, M.J., Flege, J.E. and Mackay, I.R.A. 1996. "The effects of age of second language learning on the production of English vowels". *Applied Psycholinguistics* 17: 313–324.
- Muñoz, C. 1999. "The effects of age on instructed foreign language acquisition". In *Essays in English Language Teaching. A Review of the Communicative Approach*, S. González Fernández-Coruguedo, R.A. Valdeón, D. García, A. Ojanguren, M. Urdiales and A. Antón (eds), 1–21. Oviedo: Servicio de Publicaciones Universidad de Oviedo.
- Muñoz, C. 2000. "Bilingualism and trilingualism in school students in Catalonia". In *English in Europe: the acquisition of a third language*, J. Cenoz and U. Jessner (eds), 157–178. Clevedon: Multilingual Matters.
- Muñoz, C., Pérez-Vidal, C., Celaya, M.L., Navés, T., Torras, M.R., Tragant, E. and Victori, M. 2002. "En torno a los efectos de la edad en el aprendizaje escolar de una lengua extranjera". *ICE Universitat de Barcelona*.
- Muñoz, C. in press. "Variation in oral skills development and age of onset". In *Age and the Acquisition of English as a Foreign Language*, M.P. García-Mayo and M.L. García Lecumberri (eds). Clevedon: Multilingual Matters.
- Nation, P. 1990. *Teaching and Learning Vocabulary*. New York: Newbury House.

- Navés, T. and Miralpeix, I. 2002. "Can early children starters catch up later children starters in writing in the short run?" Paper presented at the XXVI AEDEAN International Conference. Universidade de Santiago.
- Navés, T. and Pérez-Vidal, C. 1999. "Optimal non-native written competence in EFL". In *Contribuciones al Estudio de la Lingüística Aplicada*, J. Cuevas and D. Falsa (eds), 93–98. Logroño: Asociación Española de Lingüística Aplicada (AESLA).
- Oller, J. and Nagato, N. 1974. "The long-term of FLES: An experiment". *Modern Language Journal* 58: 15–19.
- Olsen, L. and Samuels, S. 1973. "The relationship between age and accuracy of foreign language pronunciation". *Journal of Educational Research* 66: 263–267.
- Ortega, L. 2000. Understanding syntactic complexity: the measurement of change in the syntax of instructed L2 Spanish learners. Unpublished Ph.D. dissertation, University of Hawai'i at Manoa, Honolulu.
- Oyama, S. 1976. "A sensitive period for the acquisition of a non-native phonological system". *Journal of Psycholinguistic Research* 5: 261–284.
- Patkowski, M. 1980. "The sensitive period for the acquisition of syntax in a second language". *Language Learning* 30 (2): 449–472.
- Pérez-Vidal, C., Torras, M. R. and Celaya, M. L. 2000. "Age and EFL written performance by Catalan/Spanish bilinguals". *Spanish Applied Linguistics* 4 (2): 267–290.
- Schachter, J. 1990. "On the issue of completeness in second language acquisition". *Second Language Research* 6: 93–124.
- Scovel, T. 2000. "A critical review of the critical period research". *Annual Review of Applied Linguistics* 20: 213–223.
- Singleton, D. 1989. *Language Acquisition. The Age Factor*. Clevedon, Avon: Multilingual Matters.
- Singleton, D. 1995. "A critical look at the Critical Period Hypothesis in second language acquisition research". In *The Age Factor in Second Language Acquisition*, D. Singleton and Z. Lengyel (eds), 1–29. Clevedon: Multilingual Matters.
- Singleton, D. 1997. "Age and second language learning". In *Encyclopedia of Language and Education. Volume 4: Second Language Education*, R. Tucker and D. Corson (eds), 43–50. Dordrecht: Kluwer.
- Singleton, D. 2001. "Age and second language acquisition". *Annual Review of Applied Linguistics* 21: 77–89.
- Slavoff, G. R. and Johnson, J. S. 1995. "The effects of age and the rate of learning a second language". *Studies in Second Language Acquisition* 17 (1): 1–16.
- Snow, C. and Hoefnagel-Höhle, M. 1978. "The critical period for language acquisition: evidence from second language". *Child Development* 49: 1114–1128.
- Sotillo, S. M. 2000. "Discourse functions and syntactic complexity in synchronous and asynchronous communication". *Language Learning and Technology* 4 (1): 82–119.
- Stankowski Gratton, R. 1980. "Una ricerca sperimentale sull'insegnamento del tedesco dalla prima classe elementare". *Rassegna Italiana di Lingüística Applicata* 128 (3): 119–41.
- Tahta, S., Wood, M. and Loewenthal, K. 1981. "Foreign accents: Factors relating to transfer of accent from the first language to a second language". *Language and Speech* 24: 265–72.
- Tedick, D. J. 1990. "ESL writing assessment: Subject-matter knowledge and its impact on performance". *Studies in Second Language Acquisition* 9: 123–143.

- Tedick, D.J. and Mathison, M.A. 1995. "Holistic scoring in ESL writing assessment: What does an analysis of rhetorical features reveal?" In *Academic Writing in a Second Language. Essays on Research and Pedagogy*, D. Belcher and G. Braine (eds), 205–230. Norwood, New Jersey: Ablex.
- Torras, M.R. and Celaya, M.L. 2001. "Age-related differences in the development of written production. An empirical study of EFL school learners". *International Journal of English Studies* 1 (2): 103–126.
- Weissberg, B. 2000. "Developmental relationships in the acquisition of English syntax: writing vs. speech". *Learning and Instruction* 10: 37–53.
- Wolfe-Quintero, K., Inagaki, S. and Kim, H. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Hawai'i: University of Hawai'i at Manoa.