International Journal of Multilingualism

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rmjm20

Third language English performance at the onset of schooling: effects of bilingualism and exposure in Catalan-Spanish bilingual children

Adriana Soto-Corominas, Marta Segura, Helena Roquet, Noelia Navarro & Yağmur Elif Met

To cite this article: Adriana Soto-Corominas, Marta Segura, Helena Roquet, Noelia Navarro & Yağmur Elif Met (12 May 2024): Third language English performance at the onset of schooling: effects of bilingualism and exposure in Catalan-Spanish bilingual children, International Journal of Multilingualism, DOI: 10.1080/14790718.2024.2348714

To link to this article: https://doi.org/10.1080/14790718.2024.2348714

Published online: 12 May 2024.

INTERNATIONAL JOURNAL OF MULTILINGUALISM
https://doi.org/10.1080/14790718.2024.2348714

Third language English performance at the onset of schooling: effects of bilingualism and exposure in Catalan-Spanish bilingual children

Adriana Soto-Corominas[a], Marta Segura[b], Helena Roquet[b], Noelia Navarro[b] and Yağmur Elif Met[b]

[a]Department of English and German Studies, Universitat Autònoma de Barcelona, Bellaterra, Spain; [b]Department of Applied Linguistics, Universitat Internacional de Catalunya, Barcelona, Spain

ABSTRACT

Research on the effects of bilingualism on third language (L3) development set in bilingual communities is scarce, outdated, and has provided mixed results. This study investigated the effects of exposure and age of onset of acquisition (AOA) of L3-English, as well as first and second language (L1/L2) skills and use, in the development of L3-English lexical and grammatical receptive abilities in Catalan-Spanish bilingual learners in Catalonia. The study followed a longitudinal design with three data collection times: Times 1 and 2 took place at the onset and the end of Grade 1, respectively, and Time 3 happened at the end of Grade 2. Results showed an overall growth of vocabulary and grammar over the first two years of primary schooling. In addition to testing time, L3-English exposure outside of school predicted higher receptive vocabulary and grammar skills, whereas an older English AOA predicted lower skills. The L1/L2 variables showed different associations to the vocabulary and grammar scores. However, the minoritized language, Catalan, bore a stronger association with the two L3 outcomes than the majority language, Spanish.

ARTICLE HISTORY
Received 4 September 2023
Accepted 23 April 2024

KEYWORDS
L3 receptive abilities; bilingualism effects; exposure; AOA; social bilingualism

Introduction

The last few decades have seen a dramatic increase in the number of English as a foreign language (FL) learners. In cases of regional/social bilingualism, where the majority of the population is a simultaneous or early sequential bilingual, L3-English acquisition (or L_n-English acquisition more generally) through schooling is widespread. This is the case of Catalonia (Spain), where this study takes place.

When a FL is acquired in a bilingual context, it is the first language that is learned non-natively. Therefore, the term *L3* in these cases entails a chronological difference with respect to the L1 and L2, as it is acquired later, but also a qualitative difference, as it is

CONTACT Adriana Soto-Corominas Adriana.Soto@uab.cat Department of English and German Studies, Universitat Autònoma de Barcelona, Plaça Cívica, 08193 Bellaterra, Barcelona, Spain

the first language learned solely in an instructed, non-naturalistic setting (Sánchez, 2020). Learners' trajectories in L3-English acquisition show great variation associated not only with L3-specific factors such as exposure in the language or its age of onset of acquisition (AOA), but also with use and proficiency in their L1 and L2.

Studies investigating the effect of bilingualism in L3 development set in bilingual communities fall under two main research lines. The first main research line includes studies comparing bilingual (e.g. Catalan-Spanish or Basque-Spanish) learners to their monolingual (e.g. Spanish) counterparts in the development and outcomes in FL-English (e.g. Lasagabaster, 2000; Sanz, 2000). Differently from studies set in largely monolingual communities where bilingual learners come from migrant backgrounds (e.g. Edele et al., 2018; Hopp et al., 2019), which show mixed results, studies comparing monolinguals to bilinguals living in bilingual communities have generally found advantages for the latter group. This bilingual advantage has been attributed to enhanced metalinguistic awareness and cognitive benefits derived from early bilingualism (Lasagabaster, 2000; Sanz, 2000). The second main line of research has narrowed in on bilinguals from bilingual communities to determine whether learners' L1/L2 proficiency and patterns of exposure modulate L3 development and outcomes, as we do in the present study. We provide a more detailed overview of these studies below.

The present study investigates the role of L1/L2 use and knowledge, together with other factors known to influence FL learning (e.g. language exposure, AOA), in the development of lexical and grammatical receptive skills in L3-English by bilingual learners at the onset of schooling, ages 6–8, using a longitudinal design that spanned the beginning of Grade 1 to the end of Grade 2. These bilingual learners were schooled in Catalonia (Spain).

Catalan, Spanish, and English in Catalonia

Catalonia is a province in Spain where Catalan and Spanish share official status. Catalan is the minoritized language, with around 35% of the population reporting it as their initial language and/or as their language of most frequent day-to-day use (IDESCAT, 2018). Catalan and Spanish are closely-related Romance languages, and share a high degree of cross-linguistic similarity in terms of their lexicon and grammar. Though empirical studies are lacking in this respect, Catalan and Spanish can be argued to have a similar degree of lexical and grammatical relatedness to English, and therefore, there is no evidence that either language is closer to English in these dimensions.

The vast majority of students in Catalonia are schooled in public and charter schools. Charter schools are semi-private schools funded by the government but also through school fees, which are most often paid for by parents. In public and charter schools in Catalonia, Catalan is the language of schooling, and Spanish and FL-English are part of the curriculum (Generalitat de Catalunya, 2009). Even though Catalan is the vehicular language of schooling, schools have some freedom to teach content courses in English and Spanish. In fact, many schools are turning to a Content and Language Integrated Learning (CLIL) approach, by which certain content courses are taught in English to boost the exposure and learning opportunities for the FL (see Soto-Corominas et al., 2023 for more details on the implementation of CLIL in the same sample of participants as in this study). Similarly, schools can choose to introduce FL-English at or before Grade 1

(Decret 175/2022, 2022), with current Catalan educational policy guidelines encouraging its introduction in kindergarten 'when possible given the sociolinguistic context at each school' (Decret 21/2023, 2023, Article 8.7).

Outside of tourist areas, English is not part of the community in Catalonia. Most parents that wish to boost their children's English language exposure have different options. First, they can resort to afterschool English language programs (Mencía & Samper, 2023). Second, they can encourage the consumption of English media in the home, such as book reading or TV watching in English (Muñoz & Lindgren, 2011). In addition, parents may also hire a babysitter or au pair to carry out their child rearing activities in English.

Bilingualism effects in L3 acquisition in bilingual communities in Spain

We focus our discussion of bilingualism effects on research conducted in two bilingual provinces in Spain, namely the Basque Country and Catalonia, given their contextual similarities. Both provinces currently have strong policies protecting Basque and Catalan, respectively, which extend to schooling. It should be noted, however, that the sociolinguistic realities of schools in these two communities are not identical. One of the most notable differences is the fact that Catalan is the main language of schooling in Catalan public and charter schools. On the other hand, in the Basque Country, Basque is the language of instruction only in some types of schooling, albeit the most popular ones (Eustat, 2024).

Four main studies have investigated bilingualism effects in L3-English vocabulary and grammar in Catalonia and the Basque Country. Cenoz and Valencia (1994) compared two groups of bilingual adolescents in the Basque Country. One group attended schools where instruction was in Basque, the minority language, whereas the other group attended schools where instruction happened in Spanish, and Basque was taught as an L2 for 3–4 hours a week. Participants completed a battery of English tests, including receptive grammar and vocabulary. The authors found that the students schooled in Basque-language schools outperformed students in Spanish programs once other sources of individual differences were controlled for.

Since the Basque Country is heavily Spanish-dominant, Basque-medium schools are argued to foster balanced bilingualism (Lasagabaster, 2000). Sanz (2008) precisely investigated the notion of whether balanced bilingualism would also be an advantage for Catalan-Spanish adolescents learning L3-English. Participants self-rated their Catalan and Spanish oral and literacy skills in terms of balance (i.e. whether their writing/reading/speaking/listening skills were similar across languages or unbalanced). The author found that those that self-rated their writing and reading skills as balanced outperformed unbalanced bilinguals in an English receptive grammar test, though no differences obtained in the lexical test. In addition, balance in speaking and listening language skills did not predict performance in vocabulary and grammar. Participants in Sanz (2008) were also asked to describe their language use with different interlocutors. The author found that increased use of Catalan with mothers and friends was positively correlated with L3 achievement, even though these associations disappeared when other variables were controlled for.

Differently from the two studies mentioned above, Muñoz (2000) and Sagasta (2003) directly assessed participants' L1/L2 proficiency instead of asking participants to judge

their own abilities. Muñoz (2000) assessed three groups of learners (ages 10, 12, and 17) in the three languages, (Catalan, Spanish, and English) using dictation and cloze tests. In L3-English, participants also completed a multiple-choice grammar test. Muñoz (2000) found that participants' home language was not a predictor of L3 performance in the tests. That is, whether participants used Catalan or Spanish most often did not predict L3 performance. However, scores were correlated across languages, meaning that higher performance in Catalan/Spanish tests was associated with better performance in L3-English.

These results contrast with Sagasta's (2003). The author tested Basque-Spanish bilingual high schoolers in their two social languages and L3-English. Among other outcomes, they measured lexical and grammatical complexity and accuracy in written production. Even though all bilinguals were exposed to Basque as their main language of schooling, those with more exposure to Basque outside of school, especially with friends, significantly outperformed those with less exposure in all the measures for Basque proficiency and either significantly or marginally outperformed them in L3-English measures. On the other hand, no differences were observed in terms of Spanish proficiency.

In summary, the body of research carried out in Spain investigating how L1/L2 can affect L3 acquisition has mainly focused on high school students and presents mixed results, likely due to discrepancies in methodology (e.g. direct vs. self-rated assessment, receptive vs. productive language skills) However, a commonality across studies is that exposure to or schooling in Basque and Catalan (the non-majority languages) does not appear to hinder the development of the L3.

An important caveat to consider when interpreting these results is that the aforementioned studies took place over 20 years ago. Multilingual education in Spanish schools, and L3-English education most especially, has experienced important changes since then (Campillo et al., 2019). The introduction of L3-English has been pushed to progressively earlier ages and its presence in the Catalan classrooms has grown, in part due to the growing popularity of CLIL approaches (Codó, 2022; European Commission, 2023). This increased exposure at school, coupled with the increased access to multilingual content readily available to children and youth through the internet and digital platforms (Cassany, 2022), could affect the influences that bilingualism may exert on early L3 development. We turn to a discussion of these factors next.

Exposure and AOA effects on L3/FL learning

As stated in the introduction, L3-English is the first language that is learned as a true FL in the Catalan context. Hence, a review of the variables that may impinge on FL-English development is pertinent.

One of the driving forces of L3/FL performance is exposure. Exposure can be conceptualized cumulatively, in terms of the accumulated experience in the language. Unsurprisingly, studies that have accounted for the number of classroom hours of exposure to the FL have generally found strong and positive associations in terms of lexical and grammatical abilities (e.g. Sanz, 2008; Soto-Corominas et al., 2023; Unsworth et al., 2015). An important component of FL exposure that is more difficult to quantify than classroom hours is extramural FL experience. In general, studies have found higher performance in FL-English vocabulary and grammar in young, primary- and secondary-school-age learners, who engage in reading (De Wilde et al., 2020, 2022; Peters, 2018), TV viewing or

videogame playing (De Wilde et al., 2020; Lindgren & Muñoz, 2013; Peters, 2018; Sundqvist, 2009), and other extramural activities in FL-English (Sanz, 2008). However, the richness of extramural English environments may depend on learners' age. Very young children (ages 6–8), like our current sample, may display impoverished FL extramural environments altogether (e.g. Unsworth et al., 2015). Children this age may still lack the literacy skills that allow them to engage in reading activities independently and may not feel comfortable enough to watch TV or play videogames in the FL.

Other than exposure, another prominent variable that is considered in the development and outcomes of a FL is its AOA. It is widely believed that an early start is beneficial for FL learning (see, for example, European Commission, 2011). It is precisely this belief, often rooted in findings for naturalistic L2 acquisition, that has promoted the early introduction of a FL in schools. However, studies that have looked into the development of FLs in instructed settings have often found an advantage in the rate of learning for later AOAs, which can neutralize possible advantages of early beginners over time (see discussion in Muñoz, 2011). A previous study including the same sample of participants from this study found that AOA was not predictive of gains between the beginning and end of Grade 1 in receptive or productive abilities in L3-English, meaning that participants developed their abilities at a similar pace regardless of when their first exposure to English took place (Soto-Corominas et al., 2023).

Present study

Our knowledge of how the L1 and L2 may affect early L3 development in situations of social bilingualism is largely rooted in a handful of studies that are currently outdated, given the European and Catalan push to foster proficiency in an L3 from an earlier age. Using a longitudinal design with three times of data collection, we investigated the lexical and grammatical skills in L3-English in a sample of Catalan-Spanish bilingual children between the onset of Grade 1 and the end of Grade 2. We asked the following questions:

(1) What is the effect of exposure to L3-English on the early development of lexical and grammatical L3-English abilities in Catalan-Spanish bilingual children?In our longitudinal study, we operationalize exposure in two ways. First, we capture the accumulated school exposure to the L3 with the variable *Time* (Time 1, 2, and 3), representing each wave of data collection. Second, we account for the extramural exposure to L3-English by grouping participants according to whether they have frequent exposure to the L3 outside of school or not. This grouping allows us to capitalize on the relatively infrequent exposure children this age have to the L3 (Unsworth et al., 2015).Our hypothesis was that *Time* would be predictive of participants' abilities, as we would see growth in lexical/grammatical skills over the span of two years. In terms of participants' extramural exposure, we expected that by grouping participants according to their frequency, whether high or low, we may be able to detect advantages for increased extramural exposure.

(2) Once L3 exposure is controlled for, does English AOA predict any further variance in L3 skills?Given that participants were at the onset of obligatory schooling, we predicted that an earlier AOA of L3-English may still predict better performance in the two

abilities. If, on the other hand, no advantages for an early AOA were detected at this stage in schooling, it would be a potent finding against current trends in hastening the early introduction of the FL in classrooms.

(3) Once L3 exposure and AOA are all controlled for, is there any additional variance in participants' abilities explained by L1/L2 skills and use?Given the paucity of studies on bilingualism effects on L3 acquisition in children from bilingual communities, we hypothesized that Catalan language use would either bear no relation to L3 abilities or would be positively associated with it (Cenoz & Valencia, 1994; Muñoz, 2000; Sagasta, 2003). In terms of proficiency, we expected a positive relationship between L1/L2 and L3 skills overall (Muñoz, 2000).

Method

Project design

The present study followed a longitudinal design with three data collection times. Time 1 (T1) took place in October/November 2021, when participants were at the onset of primary schooling, namely the beginning of Grade 1. Time 2 (T2) took place in May/June 2022, at the end of the Grade 1. Time 3 (T3) took place between March and June 2023, at the end of Grade 2. At T1 and T3 participants were tested in the three languages (Catalan, Spanish and English), and at T2, only the English tests were administered. In addition, participants' parents gave informed consent and filled in a background questionnaire at T1 and T3. The study received ethics approval by the research ethics committee at Universitat Internacional de Catalunya, protocol #MUL-2021-01.

Participants

A total of 190 participants took part in T1 of data collection. From this sample, we do not consider the data of participants whose parents reported speaking a heritage language in the home ($N = 22$), or of participants who had a history of language delay or impairment, or a diagnosis of Autism Spectrum Disorder ($N = 5$). In addition, we do not consider the data of one participant who emigrated to Spain from Switzerland at the age of 6.

Our final sample at T1 consisted of 162 participants who were an average age of 6;4 ($SD = 0;4$). At T2, one of the 14 schools was not available for testing, which reduced the sample to 130 participants. Finally, the sample we are considering for T3 is 150 participants. The 12-participant attrition between T1 and T3 is due to parents not returning the consent form at T3 ($N = 9$), or participants leaving the school ($N = 3$). Not all participants took all the language tests described below, as some participants were absent on the day of testing. The number of observations on which each statistical analysis was run is included in the Appendices.

Schools

A total of 14 schools, both public ($N = 7$) and charter ($N = 7$), in the province of Barcelona participated in this study. There was a diversity in terms of the areas where schools were located. Six of them were located in strongly Spanish-dominant cities of the metropolitan

area of Barcelona (e.g. Cornellà del Llobregat), four were in more Catalan-dominant areas in central Catalonia (e.g. Santpedor) and four were located in areas with a relatively balanced level of bilingualism (e.g. Sabadell). In order to determine the linguistic context of each area, census data corroborated by the schools' teachers were taken into account.

Schools also differed in the number of hours of weekly instruction in English, ranging from 2 to 14.5. Regarding instructional approaches, all schools offered English as a foreign language and some of them also offered CLIL classes (7 in Grade 1 and 8 in Grade 2). Within the schools that implemented CLIL, most had between 1 and 3 hours per week of CLIL instruction, with the exception of one charter school, which offered 11.5 hours per week.

Materials

Background questionnaire

Parents were given the option to complete the background questionnaire online, via a survey platform, or to do so over the phone. The questionnaire at T1 collected information about participants' demographic and linguistic background, whereas the T3 questionnaire only asked questions about the linguistic environment that could have changed over time. At T1 and T3, parents were asked to indicate the language that children used to communicate with their mother, father, older/younger siblings, maternal/paternal grandparents, uncles, cousins, friends at school and friends outside of school. Because no participants in this sample used a language other than Catalan or Spanish in the home, they used a relative scale between 1 (*Catalan always, Spanish never*) and 5 (*Catalan never, Spanish always*) to indicate their answers. In addition, for each of the three languages, parents were asked to indicate the average number of hours per week that participants spent (1) reading in the home (including joint reading with caregivers and doing homework), (2) watching TV or playing videogames, and (3) participating in extracurricular activities. At T1 only, parents were asked to indicate the participant's AOA of the three languages and the setting where participants first became exposed to each language: at home/in the family, in the community, or at school.

Vocabulary tests

In order to test receptive vocabulary abilities, we employed three tests that had the same mechanics. Participants were presented with an array of four pictures and were asked to point to the picture that best matched the word provided by the experimenter. Tests were discontinued when a given ceiling criteria was reached. For English we used the *Peabody Picture Vocabulary Task – 5th edition* (Dunn, 2019). For Spanish we used the *Test de vocabulario en imágenes PPVT-III* (Dunn et al., 2010), which had been normed on Peninsular Spanish (the variety spoken by the participants in the study). Since Catalan does not have a standardized vocabulary test, we adapted the *Test de Vocabulario en Imágenes Peabody* (TVIP; Dunn et al., 1986), which was originally developed for Latin American Spanish. The Cronbach's alpha of internal consistency for the English test was excellent (T1 = 0.98, T2 = 0.97, T3 = 0.97). The same was true for the Spanish test (T1 = 0.93, T3 = 0.93) and for Catalan (T1 = 0.92, T3 = 0.92).

Grammar tests

To assess receptive grammatical abilities, we again chose three tests that shared the same mechanics, which were similar to those of the vocabulary tests. Participants heard a sentence and chose which picture out of an array of four best matched the sentence. For English we used the *Test for Reception of Grammar 2* (Bishop, 2003) and for Spanish we used the *Test de Comprensión de Estructuras Gramaticales* (Mendoza et al., 2005).

These tests assess the comprehension of 20 different grammatical structures with each structure being tested by 4 items, for a total of 80 items. In all languages, the tested structures ranged in complexity, from short and simple constructions (*The scarf is yellow*) to complex ones (*The scarf the book is on is blue*). In order to shorten the task, we reduced the items to 40 by randomly selecting two items for each structure. Catalan, again, has no test of receptive grammar. As such, we constructed the Catalan test using the 40 items that had been eliminated from the Spanish test, translating them into Catalan, without the need for further adaptation.

Since these tests inevitably rely on vocabulary skills, participants were given a targeted vocabulary test to determine whether they knew the nouns, verbs, and adjectives included in the 40 grammar items. Those that participants did not know prior to the test were presented and reviewed twice with the help of the experimenter. The Cronbach's alpha for the English test was excellent (T1 = 0.94, T2 = 0.91, T3 = 0.92). Internal consistency was somewhat lower in Spanish (T1 = 0.64, T3 = 0.68) and Catalan (T1 = 0.71, T3 = 0.70).

Data analysis

All descriptive and inferential statistics were conducted in R (R Core Team, 2023). Preliminary data exploration of English vocabulary and grammar scores (Figure 1A and B) showed that the former were positively skewed, meaning that while most scores appeared in the lower range (with scores between 0 and 50), many values were distributed along the possible range.

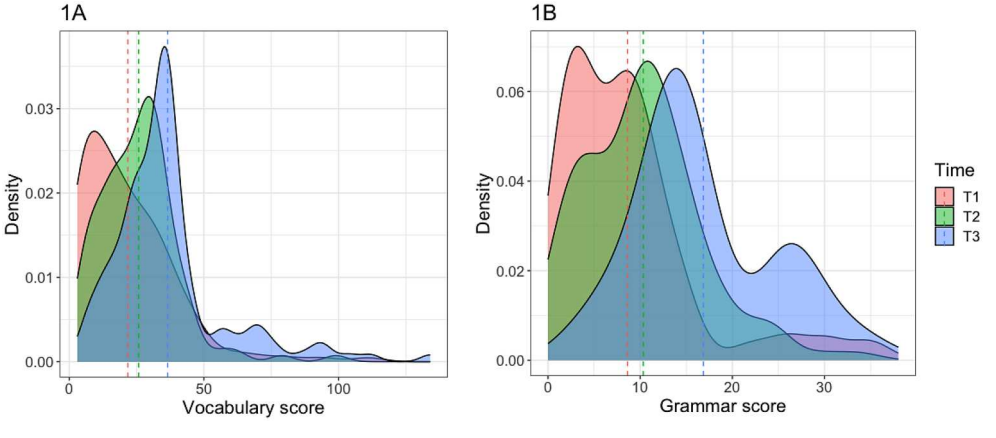

Figure 1. Density plot for English vocabulary (1A) and grammar scores (1B) for the three times. Vertical dashed lines indicate the group mean by time.

To address all research questions at once, we employed two sets of mixed-effects hierarchical regressions with a Poisson distribution: one for vocabulary and one for grammar. The outcome variable of the regressions was the total score on each test. We included a random intercept for *Participant*, which was nested within *School*. The model could not support any random slopes. Each set of regressions included three steps. Step 1 included *Time* (factor with three levels: T1, T2, T3) and *Extramural English* grouping (factor with two levels: Low vs. High). In the model for grammar, where the interaction between *Time* and *Extramural English* was significant, Step 1 included the interaction between the two factors. Step 2 added *English AOA*, which was entered as a numerical variable. Step 3 further added the L1/L2 variables of interest: *Language use with family* and *with friends*, which were both averaged from ordinal variables between 1 and 5, and the *Catalan* and *Spanish scores* in the respective tests. All numerical predictors were scaled and centered. Using hierarchical regression allowed us to see how the newly added variables show a significant improvement in the proportion of variance explained, which we calculated using pseudo-R squared with the package MuMIn (Bartoń, 2023).

It should be noted that we did not include the hours of English instruction at school in the model because this variable triggered collinearity issues with the *Time* factor (i.e. T2 entails more hours than T1, and T3 more than T2). To avoid the misinterpretation of either predictor in the presence of collinearity, and because this study does not focus on the effect of the number of hours of instruction, we include here the model with *Time* as a predictor. We note, however, that results did not change in any significant manner when *Time* was replaced by the hours of English instruction. The data, together with the scripts that generated the descriptive and inferential statistics reported in this manuscript are available at OSF (*see Data availability* below). In the script, we also share the alternative models using the variable of hours of English instruction.

A further word on the factor *Extramural English* (Low vs. High) is also necessary, as this variable was not collected as such through the questionnaires and was instead calculated during the posterior data analysis. To obtain this factor, we summed the weekly hours of English extracurricular classes, English TV/videogames, and English reading (see Table 1 for a breakdown). After creating this compound numerical variable, we binarized it using the median value. As such, at T1/T2 and T3 participants with less than 2 hours of extramural exposure, the sample median at both times, were classified as *Low,* and participants with more than 2 hours of extramural exposure were classified as *High*. The reasons why we chose to combine three numerical variables (i.e. extracurriculars, TV/videogames, reading) into one categorical variable are mainly two. First, in the three cases exposure was relatively small, with frequencies at or below 1 hour per week but with notable spread (see Table 1), making the estimation of the effect of these individual variables difficult. Second, extramural exposure is much more variable than classroom exposure (i.e. it can easily change week to week and over time). Obtaining just two measures over the two-year period likely oversimplified the calculation of this variable. As such, we binned the variable into two categories because our interest was not in the effect of small fluctuations (e.g. *1 hour* vs. *1 hour and 15 minutes*) but in a broader construct (*rich* vs. *poor extramural exposure*).

All statistical models were initially fit on *lme4* (Bates et al., 2015). Model diagnostics included inspection of the residuals with the package *DHARMa* (Hartig, 2022). We ensured that there was lack of collinearity with all variance-inflation factors under 5, as

Table 1. Participant characteristics at T1, T2, and T3. Characteristics that remain stable over time do not appear for each time.

	T1 ($N = 162$)		T2 ($N = 130$)		T3 ($N = 150$)	
	M	SD	M	SD	M	SD
Age in months	76.02	3.34	82.08	3.21	94.31	3.49
Catalan AOA in months	6.96	12.87	–	–	–	–
Spanish AOA in months	3.99	9.69	–	–	–	–
English AOA in months	29.19	16.47	–	–	–	–
Accumulated hours of English instruction at school	27.22	11.69	112.90	58.73	270.30	143.99
Weekly English extracurriculars hours	0.69	0.77	–	–	0.83	0.93
Weekly English reading hours	0.79	1.56	–	–	0.40	0.72
Weekly English TV/videogames hours	1.45	3.02	–	–	0.90	1.31
Weekly Extramural English hours[a]	2.91	3.69	–	–	2.12	1.95
Weekly Extramural Catalan hours[a]	6.42	4.61	–	–	6.50	4.73
Weekly Extramural Spanish hours[a]	8.16	6.87	–	–	8.66	5.59
Catalan vocabulary score (out of 125)	57.88	12.00	–	–	76.47	11.41
Spanish vocabulary score (out of 192)	73.63	14.58	–	–	93.95	14.66
Catalan grammar score (out of 40)	27.86	4.28	–	–	32.60	3.11
Spanish grammar score (out of 40)	28.92	3.75	–	–	33.16	2.88
Family language use[b,c]	2.98	1.49	–	–	2.96	1.53
Friends language use[b,d]	3.22	1.20	–	–	3.19	1.28

[a]: Combination of hours of reading, TV/videogames, and extracurricular activities for each of the languages. [b]: Averaged from ordinal variable between 1 and 5 (1 = Catalan always (100%), Spanish never (0%), 2 = Catalan most of the time (75%), Spanish hardly ever (25%), 3 = Both languages more or less equally (50–50%), 4 = Catalan hardly ever (25%)-Spanish most of the time (75%), 5 = Spanish always (100%)-Catalan never (0%)). [c]: Averaged across relatives (mother, father, siblings, uncles, cousins, maternal and paternal grandparents). [d]: Averaged across friends in school and friends outside of school.

assessed by the *car* package (Fox & Weisberg, 2019), and lack of overdispersion, assessed by the *performance* package (Lüdecke et al., 2021). In fact, the initial model for vocabulary was overdispersed, which was expected given the results in Figure 1A. As such, negative binomial models were fit for the vocabulary test using the package *glmmTMB* (Brooks et al., 2017). Finally, post-hoc contrasts were run with the package *emmeans* (Lenth, 2023), which implements a Tukey adjustment for multiple comparisons.

Results

Participant characteristics

A characterization of the sample at the three time points appears in Table 1. As shown by participants' AOA in the three languages, the average participant had been exposed to the three languages before age 3, with the onset for English happening comparatively later than for the other two languages. In terms of the context in which participants first became exposed to English, 90% of parents chose *at school* (i.e. school or daycare/kindergarten), whereas this was the case for only 33% of the participants for Catalan, and 12% for Spanish.

English extramural exposure, which appears broken down in Table 1 into extracurricular activities, reading and TV/videogames, was relatively low at both grades when compared to the Catalan and Spanish exposure. That is, participants engaged more frequently in extracurricular, reading, and TV/videogame activities in the two social languages. As stated in the Data analysis section above, we used this variable to break down participants into two subgroups: *High* and *Low Extramural* experience.

Language use with family and *with friends* is presented as an average of the ordinal scale for different family relatives (*Family language use*) and for friends at school and outside of school (*Friends language use*), which parents rated using a 1–5 scale. Because 5 indicates that only Spanish was used for interaction, the means, over 2.5, indicate that the sample skewed Spanish-dominant overall, which is unsurprising given the majority language status of Spanish.

Regressions for vocabulary

Table 2 includes the summary of the three regression models predicting vocabulary scores. Full model outputs appear in Appendix 1.

Model 1 shows that by including only two variables, *Time* and *Extramural grouping*, around 15% of the variance (i.e. pseudo-$R^2 = .15$) in vocabulary scores was explained. Specifically, for the effect of *Time*, the pairwise contrasts showed that scores were significantly higher at T2 than at T1 and at T3 than at T2 (both contrasts $p < .001$). In terms of the effect of *Extramural* grouping, the coefficient shows that those classified in the *Low* group obtained lower scores.

The addition of English AOA in Model 2 resulted in a significantly better fit ($\chi^2(7,8) = 6.114$, $p = .013$). *English AOA* was a negative predictor, meaning that an older English AOA predicted a lower vocabulary score. The addition of the four L1/L2 predictors in Model 3 also increased the variance explained and produced a significantly better model ($\chi^2(8,12) = 26.012$, $p < .001$). However, only one predictor was significantly associated with the lexical score, and that was *Catalan vocabulary score*, which was a positive predictor. That is, participants with higher Catalan vocabulary scores were predicted to score higher in English, as well.

Regressions for grammar

Table 3 shows the summary of the three models for grammar, and their full output appears in Appendix 2. In the grammar model, the interaction between *Time* and *Extramural grouping* was significant. The post-hoc tests indicate that both extramural groups performed better at T2 than at T1, and at T3 than at T2 and that the difference in performance between the *High* and *Low Extramural* groups was the largest at T1 and decreased over time (see Figure 2).

Table 2. Coefficient estimates for the three regression models predicting English vocabulary scores, together with *p*-values. All numerical predictors are scaled and centered.

	Model 1	Model 2	Model 3
Time_T2	0.342***	0.336***	0.334***
Time_T3	0.607***	0.604***	0.384***
Extramural_Low	−0.145*	−0.126*	−0.126*
English AOA		−0.099***	−0.120***
Catalan vocabulary			0.153***
Spanish vocabulary			0.024
Family language			−0.084†
Friends language			−0.028
Marginal pseudo-R^2	.15	.17	.27

Note: ***: *p*-value < .001; **: *p*-value < .01; *: *p*-value < .05; †: *p*-value < .10.

Table 3. Coefficient estimates for the three regression models predicting English grammar scores, together with *p*-values. All numerical predictors are scaled and centered.

	Model 1	Model 2	Model 3
Time_T2	0.238***	0.223***	0.222***
Time_T3	0.588***	0.583***	0.299***
Extramural_Low	−0.298***	−0.291***	−0.351***
Time_T2:Extramural_Low	0.158†	0.172*	0.223*
Time_T3:Extramural_Low	0.201*	0.205*	0.300**
English AOA		−0.087*	−0.107**
Catalan grammar			0.096**
Spanish grammar			0.132***
Family language			−0.150***
Friends language			0.005
Marginal pseudo-R^2	.19	.21	.37

Note: ***: *p*-value < .001; **: *p*-value < .01; *: *p*-value < .05; †: *p*-value < .10.

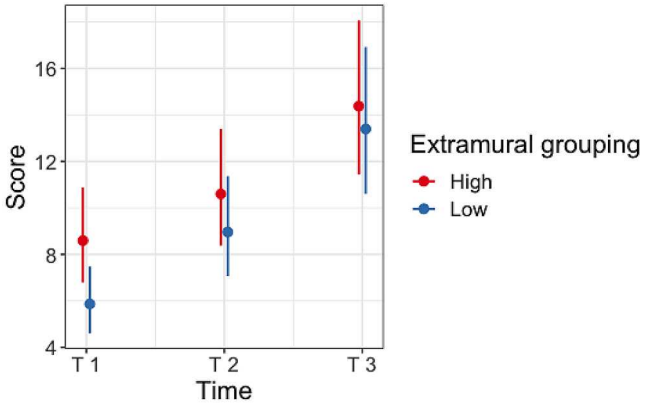

Figure 2. Visualization of the conditional effects for the interaction between Time and Extramural grouping in the model predicting English grammar scores.

The addition of *English AOA* yielded a marginally better model ($\chi^2(8,9) = 3.315$, $p = .069$). AOA was a significant and negative predictor, meaning that a later onset of English acquisition was associated with lower grammar scores overall. With the addition of the L1/L2 variables, a lot more variance was explained, producing a significantly better model ($\chi^2(9,13) = 50.418$, $p < .001$). Both the *Catalan* and *Spanish grammar* scores were associated with the *English grammar* score positively (i.e. higher Catalan/Spanish grammar scores predicted higher English scores). Finally, the variable *Family language* was a negative predictor. Because a higher value in this predictor variable means more Spanish use, the negative polarity indicates that more Spanish use within the family was associated with lower grammar scores for English.

Discussion

Given the current push in education to boost the learning opportunities in FL learning from increasingly earlier ages, more information is needed on the early stages of L3 acquisition in contexts of social bilingualism. The present study set out to investigate the bilingualism effects, together with the effects of L3 exposure and AOA, at play in the early development

of L3-English receptive vocabulary and grammatical skills in a group of participants at the onset of obligatory schooling, from the beginning of Grade 1 until the end of Grade 2. To this end, we employed a longitudinal design with three times of data collection.

Initial descriptive analyses of our sample showed that trilingualism is an early phenomenon in Catalonia, with the average AOA of English for our participants happening before age 3. This contrasts with studies that included similar demographics from just two decades ago, where participants were not typically exposed to L3-English before age 8 (Muñoz, 2000; Sagasta, 2003; Sanz, 2008). Despite participants' early start to English acquisition, English is a true FL in the Catalan context. It is learned later than the two social languages and most participants' first exposure to the language occurs at school or daycare/kindergarten. In addition, a comparison of the extramural environments in the three languages showed that participants in this sample had relatively low exposure to English outside of school, as found in other studies with similar age ranges (Unsworth et al., 2015).

Exposure effects in L3-English development

Exposure to the L3 was operationalized in two ways. First, we employed the variable *Time,* which was a factor with three levels representing each of the three data collections: T1 (beginning of Grade 1), T2 (end of Grade 1), and T3 (end of Grade 2). The variable *Time* was a coarse measure of cumulative English exposure at school, as participants had had more exposure to English at T3 than at T2, and at T2 than at T1. The second exposure variable was *Extramural English Exposure*, which was a factor with two levels: *high* and *low exposure*. Those in the *high exposure* group engaged in English extramural activities (i.e. reading, TV/videogames, extracurricular activities) with frequencies at or above the sample median.

As hypothesized, *Time* was a significant and positive predictor for both vocabulary and grammar skills. Therefore, participants showed growth in the two abilities over the two-year period. In terms of *Extramural English Exposure*, a pattern emerged whereby participants that had overall less exposure to the L3 outside of school underperformed those with more exposure in the two abilities. While this difference remained stable at the three testing times for vocabulary, the difference between the *High* and *Low* groups decreased significantly over time for grammatical abilities.

Altogether, our results demonstrate that formal classroom exposure results in gains in receptive language skills over a period of two years, which was expected given previous research (e.g. Unsworth et al., 2015). In addition, increased extramural exposure was also associated with better performance in the two abilities (e.g. De Wilde et al., 2020, 2022; Sanz, 2008). Importantly, this variable had been discretized in the present study. As such, it may not be the case that small fluctuations in extramural exposure yield meaningfully different results in receptive skills. However, those participants that engage in more English activities outside of school may have an advantage in the language. Given the trends in the results, these advantages are more short-lived in grammar than in vocabulary.

AOA effects

Research on the long-term outcomes of English has shown that AOA effects are negligible in the long run when the language is learned in situations of minimal input, as a FL

(Huang, 2016; Muñoz, 2011). However, since participants in this study were at the very early stages of L3 acquisition, we speculated that differences due to AOA may not have been neutralized yet. To this end, we included AOA in the regression models. Indeed, models indicated that an earlier exposure to the L3 predicted better performance in vocabulary and grammar, even after controlling for exposure.

In an analysis of the gains in receptive and productive abilities between T1 and T2 in the same sample of participants, Soto-Corominas et al. (2023) found no effects of English AOA. That is, while an earlier AOA conferred participants in the sample an advantage in receptive lexical and grammatical L3 abilities, their gains were not predicted by AOA, meaning that the gap between early and late acquirers is not expected to widen over time. In fact, given previous research, this gap is expected to close in the long term. A longitudinal study that follows participants for a longer period of time past the initial stages of acquisition may be able to establish when the gap ultimately closes and under which input conditions.

Bilingualism effects

Given the mixed results from previous studies, the main interest of this study was determining whether the use and/or proficiency in the two social languages, Catalan and Spanish, predicted L3 performance above and beyond the exposure and AOA factors mentioned above. Language use was captured by the variables *Language use with family* and *Language use with friends,* where higher scores indicated more Spanish use. Catalan/Spanish proficiency was included in the model by means of the respective test scores in each language.

Whereas the model for vocabulary only found the Catalan score to be significantly and positively associated with the English score, the grammar model found a positive association between the English score and both the Catalan and Spanish scores. In this regard, our results partly resemble those in Muñoz (2000), who found positive correlations between the three languages. In terms of language use, the model for grammar found that more Spanish use with family was associated with lower English scores, while the model for vocabulary only found a trend in the same direction. Sanz (2008) and Sagasta (2003) found similar trends whereby better L3 performance was associated with increased minority language use by participants.

However, unlike in Sagasta (2003), there was no evidence that language use with friends was associated with L3 performance in our study. Importantly, the sample in this study was much younger than the adolescents sampled in Sagasta (2003). Children ages 6–8 are likely to spend less time interacting with members outside of the household than teenagers, and therefore language practices outside of the home may be less likely to impact L3 acquisition.

Altogether, our results have found clear bilingualism effects in the L3 development of receptive skills. Specifically, evidence of an advantage was found for those with higher skills in Catalan, and, to a certain extent, for those with increased Catalan use in the home. These advantages are unlikely to stem from any direct transfer between languages, as Catalan cannot be argued to be closer to English in its lexicon or grammar than Spanish.

Two alternative, not mutually exclusive, explanations for this finding are possible. The first one relates to the idea of balanced bilingualism. Because Spanish is the majority

language in Catalonia, more Catalan use and stronger Catalan abilities may indicate more balanced bilingualism altogether. Previous studies have found an advantage in L3-English by balanced bilinguals, as more balanced bilinguals may attain higher bilingual abilities overall, which may in turn benefit L3 learning (Cenoz & Valencia, 1994; Lasagabaster, 2000; Sanz, 2008). The grammar model of the present study, which found both the Catalan and Spanish scores to be positively associated with English performance, strongly supports this possibility, whereas the vocabulary model does not. A second explanation for the results favoring Catalan proficiency and use is that because Catalan is the main language of instruction in Catalonia, the observed results are due to participants being more proficient in the language of schooling (Cenoz & Gorter, 2022). That is, L3 learning may be faster and easier for learners who have a strong command of the language learning occurs in. This hypothesis receives support by studies that have investigated bilingualism effects in monolingual regions, where bilinguals with higher levels in the language of instruction (their L2) tend to have better L3 outcomes (e.g. Edele et al., 2018).

Conclusions and limitations

This is one of the first studies to investigate L3-English development in a situation of social bilingualism in the current context, where efforts are made to push the start of the L3 to earlier and earlier ages. The main finding of this study is that bilingualism effects are present, even after controlling for L3 exposure and AOA. Higher proficiency in Catalan (the minoritized language) and, to a lesser extent, greater/more Catalan use within the family were predictors associated with higher receptive abilities in the L3. That is, our results prove that in bilingual societies, the linguistic abilities attained in the L1/L2 and the patterns of language use may subsequently affect L3 development. Specifically for our bilingual context, students whose proficiency in Catalan is relatively weak may not be able to take advantage of their bilingualism to the same extent as their peers with stronger abilities and may thus be at risk of falling behind in L3 development. While the precise mechanics underlying the connection between minoritized/minority language and L3 development are unclear, our findings suggest that fostering proficiency in the lesser spoken language of a given bilingual community may be beneficial for L3 development.

This study is not without limitations. One of them is that we could not narrow in on the effect of extramural exposure and chose to binarize a variable that was originally collected as continuous. As explained in the *Data analysis* section, this decision was made given the lack of fine-grained detail of the numerical variable. As such, future studies that want to examine the effects of extramural exposure may need to collect more detailed data from parents, on a weekly or monthly basis. Second, a larger sample size would have allowed for the inclusion of random slopes in the model, which would have produced more conservative results. Finally, and in an effort to not overfit our models, this study did not include other variables that may also affect L3 development at such an early age, such as the quality of the FL instruction or the intensity of the FL exposure at school (e.g. Weitz et al., 2010). We would like to note, however, that detailed information on these and other variables was collected as part of this study and is accessible through the OSF repository. The limitations of the study are nevertheless outweighed by its strengths, as this study tested participants' three languages over a two-year period of time at a stage

in development, early primary schooling, that is not featured prominently in empirical studies.

Acknowledgements

We would like to thank all participants and their parents, as well as all schools and teachers for making this study possible.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by H2020 Marie Skłodowska-Curie Actions: [Grant Number MSCA-IF-EF-RI 101027137].

Declaration of interest statement

The authors have no conflict of interest.

Data availability statement

The data and scripts for this study are openly available on the Open Science Framework (OSF) platform at [https://doi.org/10.17605/OSF.IO/AYM98].

References

Bartoń, K. (2023). _MuMIn: Multi-Model Inference. R package version 1.47.5, https://CRAN.R-project.org/package = MuMIn

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bishop, D. V. M. (2003). *The test for reception of grammar, version 2 (TROG-2)*. Harcourt Assessment.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility Among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/10.32614/RJ-2017-066

Campillo, J. M., Sánchez, R., & Miralles, P. (2019). Primary teachers' perceptions of CLIL implementation in Spain. *English Language Teaching*, *12*(4), 149–156. https://doi.org/10.5539/elt.v12n4p149

Cassany, D. (2022). Aprender ELE (y otras L2) en contextos informales. *Cuadernos Canela*, *33*(0), 5–23. https://doi.org/10.2107/canela.33.0_5

Cenoz, J., & Gorter, D. (2022). Enhancing bilingual resources in third language acquisition: Towards pedagogical translanguaging. *Educational Linguistics*, *1*(2), 338–357.

Cenoz, J., & Valencia, J. F. (1994). Additive trilingualism: Evidence from the Basque country. *Applied Psycholinguistics*, *15*(2), 195–207. https://doi.org/10.1017/S0142716400005324

Codó, E. (2022). The dilemmas of experimental CLIL in Catalonia. *Journal of Multilingual and Multicultural Development*, *43*(4), 341–357. https://doi.org/10.1080/01434632.2020.1725525

Decret 175/2022. (2022). El decret d'Educació bàsica. Accessed at https://projectes.xtec.cat/nou-curriculum/educacio-basica/decret-educacio-basica/Decret 21/2023. (2023). Ordenació dels

ensenyaments de l'educació infantil. Accessed at https://portaljuridic.gencat.cat/ca/document-del-pjur/?documentId = 951431

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition*, *23*(1), 87–111. https://doi.org/10.1017/S1366728918001062

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2022). Formal versus informal L2 learning: How do individual differences and word-related variables influence French and English L2 vocabulary learning in Dutch-speaking children? *Studies in Second Language Acquisition*, *44*(1), 87–111. https://doi.org/10.1017/S0272263121000097

Dunn, L. M. (2019). *The peabody picture vocabulary test* (5th ed.). NCS Pearson.

Dunn, L. M., Dunn, L. M., & Arribas, D. (2010). *Test de vocabulario en imágenes PPVT-III* (3rd ed.). TEA Ediciones.

Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *Manual del examinador para el test de vocabulario en imágenes peabody*. American Guidance Service.

Edele, A., Kempert, S., & Schotte, K. (2018). Does competent bilingualism entail advantages for the third language learning of immigrant students? *Learning and Instruction*, *58*, 232–244. https://doi.org/10.1016/j.learninstruc.2018.07.002

European Commission. (2011). Commission staff working paper - Language learning at pre-primary level: making it efficient and sustainable – A policy handbook. https://education.ec.europa.eu/document/commission-staff-working-paper-language-learning-at-pre-primary-level-making-it-efficient-and-sustainable-a-policy-handbook

European Commission. (2023). *Key data on teaching languages at school in Europe (2023 edition, eurydice report)*. Publications Office of the European Union. https://data.europa.eu/doi/10.2797529032

EUSTAT. (2024). Alumnado matriculado en Enseñanzas de Régimen General no universitarias. Accessed at https://www.eustat.eus/elementos/ele0002400/alumnado-matriculado-en-ensenanzas-de-regimen-general-no-universitarias-en-la-ca-de-euskadi-por-territorio-historico-y-nivel-de-ensenanza-segun-titularidad-del-centro-y-modelo-linguistico-avance-de-datos-202223/tbl0002427_c.html

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage Publications Inc.

Generalitat de Catalunya. (2009). Llei 12/2009, del 10 de juliol, d'educació. Diari Oficial de la Generalitat de Catalunya, 5422.

Hartig, F. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.6. https://CRAN.R-project.org/package = DHARMa

Hopp, H., Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, *61*, 99–110. https://doi.org/10.1016/j.learninstruc.2019.02.001

Huang, B. H. (2016). A synthesis of empirical research on the linguistic outcomes of early foreign language instruction. *International Journal of Multilingualism*, *13*(3), 257–273. https://doi.org/10.1080/14790718.2015.1066792

IDESCAT. (2018). Usos lingüístics de la població. Llengua inicial, d'identificació i habitual. Accessed August 2023. https://www.idescat.cat/indicadors/?id = anuals&n = 10364

Lasagabaster, D. (2000). Three languages and three linguistic models in the basque educational system. In J. Cenoz, & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 179–197). Multilingual Matters.

Lenth, R. (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.6, https://CRAN.R-project.org/package = emmeans

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1), 105–129. https://doi.org/10.1080/14790718.2012.679275

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source*

Software, *6*(60), 3139. https://doi.org/10.21105/joss.03139. https://pure.mpg.de/rest/items/item_3316363/component/file_3316364/content

Mencía, L., & Samper, S. (2023). *La cara i la creu de la participació extraescolar. Oportunitats i desigualtats educatives*. Fundació Jaume Bofill.

Mendoza, E., Carballo, G., Muñoz, J., & Fresneda, M. D. (2005). *Test de comprensión de estructuras gramaticales*. TEA Ediciones.

Muñoz, C. (2000). Bilingualism and trilingualism in school students in Catalonia. In J. Cenoz, & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 157–178). Multilingual Matters.

Muñoz, C. (2011). Input and long-term effects of starting age in foreign language learning. *IRAL - International Review of Applied Linguistics in Language Teaching*, *49*(2), 113–133. https://doi.org/10.1515/iral.2011.006

Muñoz, C., & Lindgren, E. (2011). Out-of-school factors. In J. Enever (Ed.), *ELLie: Early language learning in Europe* (pp. 103–124). British Council.

Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, *169*(1), 142–168. https://doi.org/10.1075/itl.00010.pet

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Sagasta, M. P. (2003). Acquiring writing skills in a third language: The positive effects of bilingualism. *International Journal of Bilingualism*, *7*(1), 27–42. https://doi.org/10.1177/13670069030070010301

Sánchez, L. (2020). Multilingualism from a language acquisition perspective. In C. Bardel, & L. Sánchez (Eds.), *Third language acquisition: Age, proficiency and multilingualism (eurosla studies 3)* (pp. 15–41). Language Science Press.

Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics*, *21*(1), 23–44. https://doi.org/10.1017/S0142716400001028

Sanz, C. (2008). Predicting enhanced L3 learning in bilingual contexts: The role of biliteracy. In C. Pérez-Vidal, M. Juan-Garau, & A. Bel (Eds.), *A portrait of the young in the new multilingual Spain* (pp. 220–240). Multilingual Matters. https://doi.org/10.21832/9781847690241-014

Soto-Corominas, A., Roquet, H., & Segura, M. (2023). The effects of CLIL and sources of individual differences on receptive and productive EFL skills at the onset of primary school. *Applied Linguistics*. https://doi.org/10.1093/applin/amad031

Sundqvist, P. (2009). Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary. [Doctoral dissertation, Karlstad University].

Unsworth, S., Persson, L., Prins, T., & De Bot, K. (2015). An investigation of factors affecting early foreign language learning in The Netherlands. *Applied Linguistics*, *36*(5), 527–548. https://doi.org/10.1093/applin/amt052

Weitz, M., Pahl, S., Flyman Mattsson, A., Buyl, A., & Kalbe, E. (2010). The input quality observation scheme (IQOS): The nature of L2 input and its influence on L2 development in bilingual preschools. In K. Kersten, A. Rohde, C. Schelletter, & A. K. Steinlen (Eds.), *Bilingual preschools. Vol. 1. Learning and development* (pp. 5–44). WVT Wissenschaftlicher Verlag.

Appendices

1. Output of negative binomial models predicting English vocabulary scores

Table A1. Coefficient table showing conditional effects for negative binomial model predicting vocabulary – Model 1.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.933	0.109	26.94	<.001***
Time_T2	0.342	0.053	6.51	<.001***
Time_T3	0.607	0.050	12.17	<.001***
Extramural_Low	−0.145	0.059	−2.46	.014*

Note: ***: *p*-value <.001; *: *p*-value <.05.

Table A2. Random effects for vocabulary – Model 1.

Groups	Name	Variance	Std. Dev.
Participant:School	(Intercept)	0.127	0.356
School	(Intercept)	0.115	0.339

Number of obs: 418, groups: Participant: School, 162; School, 14.

Table A3. Post-hoc contrasts between testing times – Model 1.

Contrast	Estimate	SE	z.ratio	*p*.value
T1 – T2	−0.337	0.050	−6.888	<.001***
T1 – T3	−0.616	0.048	−12.861	<.001***
T2 – T3	−0.280	0.049	−5.701	<.001***

Note: ***: *p*-value <.001.

Table A4. Coefficient table showing conditional effects for negative binomial model predicting vocabulary – Model 2.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.932	0.103	28.400	<.001***
Time_T2	0.336	0.052	6.405	<.001***
Time_T3	0.604	0.050	12.112	<.001***
Extramural_Low	−0.126	0.059	−2.137	.033*
English AOA	−0.099	0.038	−2.598	.009**

Note: ***: *p*-value <.001; **: *p*-value <.01; *: *p*-value <.05.

Table A5. Random effects for vocabulary – Model 2.

Groups	Name	Variance	Std. Dev.
Participant:School	(Intercept)	0.122	0.349
School	(Intercept)	0.098	0.313

Number of obs: 415, groups: Participant:School, 161; School, 14.

Table A6. Coefficient table showing conditional effects for negative binomial model predicting vocabulary – Model 3.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.030	0.099	30.686	<.001***
Time_T2	0.334	0.055	6.073	<.001***
Time_T3	0.384	0.076	5.051	<.001***
Extramural_Low	−0.126	0.058	−2.169	.030*
English AOA	−0.120	0.035	−3.382	<.001***
Catalan vocabulary	0.153	0.044	3.449	<.001***
Spanish vocabulary	0.024	0.038	0.636	.525
Family language	−0.084	0.045	−1.880	.060†
Friends language	−0.028	0.054	−0.517	.605

Note: ***: *p*-value <.001; *: *p*-value <.05; †: *p*-value <.10.

Table A7. Random effects for vocabulary – Model 3.

Groups	Name	Variance	Std. Dev.
Participant: School	(Intercept)	0.082	0.286
School	(Intercept)	0.082	0.286

Number of obs: 385, groups: Participant:School, 158; School, 14.

2. Output of Poisson models predicting English grammar scores

Table A8. Coefficient table showing conditional effects for Poisson model predicting grammar – Model 1.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.089	0.119	17.510	<.001***
Time_T2	0.238	0.056	4.263	<.001***
Time_T3	0.588	0.049	11.899	<.001***
Extramural_Low	−0.298	0.075	−3.974	<.001***
Time_T2:Extramural_Low	0.158	0.085	1.849	.064†
Time_T3:Extramural_Low	0.201	0.084	2.382	.017*

Note: ***: *p*-value <.001; *: *p*-value <.05; †: *p*-value <.10.

Table A9. Random effects for grammar – Model 1.

Groups	Name	Variance	Std. Dev.
Participant:School	(Intercept)	0.157	0.396
School	(Intercept)	0.144	0.379

Number of obs: 393, groups: Participant:School, 161; School, 14.

Table A10. Post-hoc contrasts between Extramural groups conditioned on Time – Model 1.

Time	Contrast	Estimate	SE	z.ratio	*p*.value
T1	High – Low	0.381	0.078	4.868	<.001***
T2	High – Low	0.168	0.075	2.220	.027*
T3	High – Low	0.071	0.067	1.062	.288

Note: ***: *p*-value <.001; *: *p*-value <.05.

Table A11. Post-hoc contrasts between Times conditioned on Extramural groups – Model 1.

Group	Contrast	Estimate	SE	z.ratio	*p*.value
High	T1 – T2	−0.238	0.056	−4.263	.001**
High	T1 – T3	−0.588	0.049	−11.899	<.001***
High	T2 – T3	−0.350	0.050	−7.028	<.001***
Low	T1 – T2	−0.396	0.065	−6.093	<.001***
Low	T1 – T3	−0.789	0.064	−12.233	<.001***
Low	T2 – T3	−0.393	0.063	−6.257	<.001***

Note: ***: *p*-value <.001; *: *p*-value <.05.

Table A12. Coefficient table showing conditional effects for Poisson model predicting grammar – Model 2.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.090	0.114	18.399	<.001***
Time_T2	0.223	0.056	3.971	<.001***
Time_T3	0.583	0.050	11.750	<.001***
Extramural_Low	−0.291	0.075	−3.881	<.001***
Time_T2:Extramural_Low	0.172	0.086	2.006	.045*
Time_T3:Extramural_Low	0.205	0.084	2.434	.015*
English AOA	−0.087	0.040	−2.171	.030*

Note: ***: *p*-value <.001; *: *p*-value <.05.

Table A13. Random effects for grammar – Model 2

Groups	Name	Variance	Std. Dev.
Participant:School	(Intercept)	0.161	0.401
School	(Intercept)	0.126	0.355

Number of obs: 390, groups: Participant:School, 160; School, 14.

Table A14. Coefficient table showing conditional effects for Poisson model predicting grammar – Model 3.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.202	0.100	21.927	<.001***
Time_T2	0.222	0.059	3.745	<.001***
Time_T3	0.299	0.063	4.755	<.001***
Extramural_Low	−0.351	0.076	−4.627	<.001***
Time_T2:Extramural_Low	0.223	0.090	2.485	.013*
Time_T3:Extramural_Low	0.300	0.088	3.391	.001**
English AOA	−0.107	0.034	−3.162	.002**
Catalan grammar	0.096	0.035	2.752	.006**
Spanish grammar	0.132	0.032	4.138	<.001***
Family language	−0.150	0.040	−3.707	<.001***
Friends language	0.005	0.045	0.119	0.906

Note: ***: *p*-value <.001; **: *p*-value <.01; *: *p*-value <.05.

Table A15. Random effects for grammar – Model 3.

Groups	Name	Variance	Std. Dev.
Participant:School	(Intercept)	0.090	0.301
School	(Intercept)	0.093	0.306

Number of obs: 356, groups: Participant:School, 156; School, 14.