THIS IS AN EARLIER VERSION OF THE MANUSCRIPT. FOR THE FINAL VERSION, PLEASE CHECK THE PUBLISHER'S WEBSITE: Tragant, E. & Muñoz, C. (2023). Longitudinal research: The ELLiC project . In *Ten Years of Language Learning at School* (pp. 1-27). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-32759-9.

Chapter 1

Longitudinal research: The ELLiC project

Abstract

This chapter starts with a description of the story behind the ELLiC project (English Language Learning in Catalonia): from the initial goal of examining young language learners during their first years of foreign language instruction at school to following up with some of those learners

through the end of secondary school. The next section is devoted to longitudinal research. We first introduce the reader to SLA (second language acquisition) research with a focus on change over time and highlight the suitability of this temporal focus to answer fundamental problems in SLA. Some of the challenges of longitudinal research such as 'study length' are also discussed. A description of six exceptionally long studies follows to illustrate the affordances of long-term longitudinal research. The chapter's subsequent section is devoted to the ELLiC project by first situating it in the context of Catalonia (Spain) and the region's school system and then presenting the project's design. The main methodological elements of ELLIC follow including a description of the multiple instruments that characterise the project. In sum, this first chapter provides the reader with the necessary context to get a full understanding of the next five chapters in this book.

Keywords: School learners, Longitudinal research, Learner trajectories, L2 development

"Why is longitudinal research essential to advancement of knowledge in the field of SLA? The simple but uncontestable answer is that many questions concerning second language learning are fundamentally questions of time and timing"

Ortega and Iberri-Shea (2005, p. 27)

1.1 The story behind a long journey

The ELLiC project (Early Language Learning in Catalonia) had its beginning when we joined the ELLiE project (Early Language Learning in Europe). See Enever (2011) for a report of the ELLiE project, which was conceived as a

transnational longitudinal initiative in order to give a rich description of language learning and teaching in the first years of foreign language (FL) instruction in primary schools in Europe. When ELLiE ended, after four years of close collaboration with our colleagues in Croatia, England, Italy, the Netherlands, Poland, and Sweden, we felt compelled to continue following our ELLiE learners¹ in Catalonia (Spain) until the end of primary education. We had been in their English lessons from grade 1 (from hereon G1), spoken to their English teachers, measured their progress from grade to grade and spoken to a subgroup of focal learners year after year. We resisted the idea of putting an end to our longitudinal investigation.

After G6 we were clear about not wanting to miss the opportunity to witness how those young learners were going to cope with English in their transition to high school. With that aim in mind, we decided to trace our focal learners from their last year in primary school to their first year in high school (G7). We succeeded and were able to observe them intensely during that "special" year. We did not go back to the high schools for two years until G10. We knew that G10 would be our last chance to contact our focal learners and that was when we collected data from them for the last time. Our ten-year-long "journey" with them ended then.

Our experience conducting the ELLiC project for such a long time has been both a gratifying and demanding endeavour. Like Ortega and Iberri-Shea (2005) and many other SLA (second language acquisition) researchers, like Long and Doughty (2003) or R. Ellis (2021), we also acknowledge the centrality of studying change over time and the need to conduct longitudinal research in our field. It is not uncommon for cross-sectional studies to end with a call for longitudinal research or for longitudinal studies to suggest a lengthier follow-up study.

Longitudinal research is not the norm in our field and deserves a review in section 1.2 of this chapter as a preamble to the ELLiC project. In this review we first introduce different methodological approaches that are appropriate for the study of change over time and then focus on "classical" longitudinal research with special attention to studies that run over several years in school contexts, which is what the ELLiC project is about. Section 1.3 is an introduction to the ELLiC project and provides the contextual and methodological information that the reader will need later on to understand the empirical studies in this book.

1.2 Longitudinal research

One of the first decisions SLA researchers must make when planning research is about its design. If the primary research focus is on change over time, which corresponds to the approach of this project, researchers have four main options to choose from: a simultaneous cross-sectional design, a repeated cross-sectional design, a "classic" longitudinal design and a retrospective study. In studies following a simultaneous cross-sectional design, data from learners of different age groups or proficiency levels are recorded. In studies following a repeated cross-sectional design, measures of the same population (but different participants) are recorded at different times. In studies following a longitudinal design, the same participant or participants are recorded at different points in time or continuously over a period of time. Retrospective longitudinal studies, like life history studies, draw on participants' reporting of past events.

The metaphor of a long-distance journey by train with a number of intermediate train stations may be helpful to exemplify the difference between these four designs with a focus on time. In a simultaneous cross-sectional study, researchers simultaneously collect data at different train stations (= at different times) from different passengers (usually groups of learners) travelling during the same period of time (e.g., academic year, semester etc.). In a

¹ The terms 'learner' and 'student' are used interchangeably in the book.

repeated cross-sectional design, the researcher collects data from one single train station once every x time from passengers who are different but have the same profile and make the same journey on different occasions. In a longitudinal study, on the other hand, researchers wait to get on the same train at a number of train stations to be able to trace or take record of one or more passengers as they travel. Finally, in retrospective studies the researcher will meet the passenger(s) at one of the stations or even the final destination and will ask them to rememorate the journey from its beginning or from a certain location or point in time to the present location or train station.

On some occasions, a simultaneous cross-sectional design with a focus on time is the preferred option because data can be collected more quickly and can involve large samples. If the study requires data from a large number of participants, it may involve a large team of researchers but this type of design allows us to have results much sooner than if we had opted for a repeated cross-sectional or a "classic" longitudinal design. Other advantages of simultaneous cross-sectional research with a focus on time is that it is often more likely to lead to generalisation than "classic" longitudinal research, which necessarily involves smaller samples. Cross-sectional research of this type also avoids the problem of attrition, inherent in longitudinal studies.

Repeated cross-sectional designs can be considered partially longitudinal in nature because data is collected at different rounds from the same population. This type of research is often conducted by institutions to record changes over extended periods of time (i.e., every five years) or to evaluate educational programmes. Whereas this type of research can address important questions regarding the impact of policies, historical changes, and second language (L2) programmes, it may be less attractive for individual researchers or PhD students whose work is evaluated in shorter time scales.

Longitudinal research, which by definition has a focus on time, has different advantages and challenges when compared to the two types of cross-sectional research presented above. According to Ortega and Iberri-Shea (2005), there are many questions concerning L2 learning that require a truly longitudinal approach, like for example how L2 competence develops over time, the durability of the effects of L2 instructional practices over L2 learning, developmental changes, among others. Longitudinal research may be more manageable for individual researchers or small research teams than simultaneous cross-sectional research since data collection spreads over time. The focus on smaller samples over time also allows for a more situated approach to SLA research. The research design that we adopt in this book can be classified as longitudinal research.

If we go back into the history of SLA, it is not difficult to find influential longitudinal studies. In the seventies, for example, Wode (1976) carried out systematic longitudinal observations of German-speaking children learning English as an L2 and English-speaking children learning German as an L2. He was able to question a universal developmental sequence for the acquisition of a language irrespective of whether it is acquired as a first language (L1) or an L2. A good example of longitudinal research from the mid-1980's is the comparative study conducted by Perdue and colleagues (1993a, 1993b) in the context of the European Science Foundation. The aim of the study was to investigate the spontaneous L2 acquisition in adult immigrants learning English, German, Dutch, French, and Swedish over a two-and-a-half-year period. A later illustration of an influential longitudinal study was carried out by Norton (2000) in her two-year-long study on five recently arrived adult learners in Canada, where she suggests a poststructuralist conception of identity in language learning. In the school context, the Barcelona Age Factor project (BAF) is an often-cited contribution to longitudinal research where a subset of learners who had started learning English at different ages were tracked over time (Muñoz, 2006).

Despite the suitability of longitudinal research to answer fundamental problems in SLA, it is surprising to find a dearth of SLA research that follows this approach. In 2005 Ortega and Iberri-Shea conducted a review of 20 major applied linguistics journals over a period of three years (2002-2004) and could only identify 38 longitudinal studies.

They also observed that among those 38 studies, only 20% took place in high school or middle school contexts and even fewer involved younger learners. In a later review by Barkaoui (2014) on quantitative longitudinal research, the scarce presence of this approach in published research was also made evident. Only 36 studies could be identified over a ten-year period (2003-2013) from the more than 30 journals under review that quantitatively measured the same dependent variable(s) on the same participant(s) on two or more occasions regardless of study length. Years later, Bridgewater (2020) conducted a review similar to Ortega and Iberri-Shea's work (2005) and identified studies where the authors self-labelled their work as "longitudinal" from a collection of 9 journals over three and a half years (2017-2020). The journals under review were: Applied Linguistics, Language Learning, Language Teaching Research, Second Language Research, TESOL Quarterly, The Modern Language Journal, Annual Review of Applied Linguistics, Journal of Second Language Writing, and Studies in Second Language Acquisition. Out of the roughly 640 articles published in the years under review, 97 were self-labelled as longitudinal. In comparison to Ortega and Iberri-Shea's review, this figure indicates an increase in the presence of longitudinal research in SLA journals, yet the proportion continues to be very low (roughly 15% of all published research in those journals). In fact, in a recent brief history of SLA, R. Ellis (2021, p. 196) reiterates that there still exists a lack of longitudinal studies, something that Long and Doughty had already observed in their introduction to The handbook of second *language acquisition* back in 2003.

1.2.1 Longitudinal research and study length

Study length is a characteristic feature of longitudinal research even though it is not a defining feature in definitions of the term such as in Barkaoui (2014), Loewen and Plonsky (2016) or Riazi (2016). This is because study length can vary widely depending on the nature of the study. Ortega and Iberri-Shea (2005) opt to describe typical study length based on the longitudinal studies they review. In the case of studies on L2 development, the time span is between four months and four years. For longitudinal research on L2 programmes outcomes, the study length is typically between four to six years, whereas for investigations of L2 instructional effectiveness there seems to be a preference for eight-week-long interventions and post-tests between one and three months later. No temporal references are provided in Ortega and Iberri-Shea for qualitative longitudinal SLA research but they do mention that much of this research does not document change over time even if data is collected over extended periods of time.

Variation in study length is also evidenced in Bridgewater's review of longitudinal classroom research where the shortest study ran over six days (Kunitz and Marian, 2017) and the longest ones spanned 8 years (Kibler, 2017; Kibler and Hardigree, 2017). She also observed that a majority of studies covered relatively short periods of time. Out of the 27 studies that she analysed, there were only four publications that ran over one year and most of them had a duration of a semester or a few weeks. This is especially understandable for research taking place at universities, where the basic unit of tuition (i.e., courses) are organised in semesters or quarters and courses do not tend to have continuity.

Longitudinal research taking place over more than a few weeks can be more feasible in a school context than in other contexts because the duration of compulsory education ranges over several years (from 9-15 years in Europe, according to the European Commission, 2022). It may be even more feasible in educational systems like Spain's where courses tend to have a year-long duration or in schools with fairly stable student populations, as is the case of the participating schools in the present project. But even in school contexts, long-term longitudinal research is always a challenge. For example, in Costache et al.'s (2022) three-year-long study taking place in Swiss high schools, one of their eight participating schools dropped out. Something similar happened in Cable et al.'s (2010) larger study in England, where three primary schools (out of a pool of 40) dropped out during their three-year longitudinal study. Attrition at the level of individual learners has also proved to be problematic in school-based research, especially in studies that cover the transition from primary to secondary education. De Wilde et al. (2021) could

only track 1 in 7 (that is, 13%) of the 800 learners in a study that first tested students in their last year of primary school in Flanders and a second time at the end of the second year of secondary school. Attrition was also unavoidable in the BAF project (Muñoz, 2006) in part because of the learners who were lost in the transition from primary to secondary school and also because learners who began extracurricular English classes during the period of the study or who repeated a school year were excluded from the longitudinal analyses. This meant that only 19% of the 11-year-old learners from the initial sample were retained two years later. The proportion was even lower (9%) among 13-year-old learners.

1.2.2 Long-term longitudinal research at schools

In spite of the challenges of conducting longitudinal research, studies do exist in SLA that run over exceptionally long periods of time. In this section, six pieces of research (with a minimum study duration of six years) have been selected to illustrate different affordances of long-term longitudinal research.

Tragant and Victori's work (2006) on learning strategies and age is a good illustration of long-term descriptive research in the context of the BAF project in Catalonia. In that study, a subset of the longitudinal learners (n=48) answered the same survey at three points in time over a period of six years (ages 10;9 at Time 1; 12;9 at Time 2 and 16;9 at Time 3). The survey provided evidence of progress in the number of strategies learners reported as they grew older and in the nature of the strategies reported over time. Overall the results from the longitudinal sample complemented those from the larger cross-sectional sample (n = 360) and strengthened the internal validity of the study.

A different example of long-term longitudinal research comes from a case study of three Swedish learners (Lindgren and Enever, 2017) who were followed over their six years in primary school. In the thick descriptions of these three learners, who were originally part of the ELLiE project, the reader gets a vivid depiction of how their L2 use in the home environment was not originally the same for the three learners and did not evolve in the same way over time. The reader also learns about how the home environment had a visible effect on learners' interlanguage development as they grew older. Had this case study been based on a shorter time scale, the reader would have lost depth in understanding the learners' trajectories with English.

A long-term study of a different nature was conducted by Lightbown et al. (2002) with over 800 French-speaking learners in the Canadian province of Nouveau-Brunswick who started learning English in grade 3 (age 8 to 9). The learners participated in an experimental programme based on comprehension-based learning that consisted of learners basically learning English from age-appropriate materials of their choice that they would read and listen to independently (30 minutes per day). The programme was originally evaluated after the first three years of its implementation with very promising results (Lightbown, 1992; Lightbown and Halter, 1989). Learners' attitudes towards the experimental programme were very positive and their performance in English was comparable to that of learners who received audio-lingual instruction (regular ESL programme). Such positive outcomes motivated the continuation of the comprehension-based program into secondary school and its evaluation six years later with a sample of 73 learners (age 14 to 15). This follow-up study revealed that secondary school learners' assessment of their learning was in sharp contrast to findings from these same learners in primary school: they now had different expectations of what an ESL program should be like. Their performance was still comparable to that of learners in the regular program regarding measures of comprehension and some measures of oral production but not on measures of written production. In this case, the long-term evaluation of the program was crucial in detecting the limits of the comprehension-based approach.

In Sparks et al. (2009), long-term longitudinal research was used to predict language proficiency in high school students in the US where foreign language study generally starts at the beginning of secondary education. The

study started with data collection of L1 skills and L1 academic measures from 156 students in 1st grade, continued during elementary school and carried over into secondary school. The participants were 54 students who had chosen to study a foreign language in high school in 9th grade and had done so for two years (until 10th grade) when L2 proficiency measures were administered. The availability of data from early elementary school allowed the researchers to conclude that differences in L2 proficiency (and L2 aptitude) could be related to students' lower levels of L1 skills (particularly L1 decoding skills) that had been acquired many years before L2 instruction took place, rather than to low motivation. The ten-year-long period in the design of this study was a necessary element to be able to predict the power of L1 skills in the US school system.

Heining-Boynton and Haitema's work (2007) was also conducted over a 10-year period and consisted of two studies about attitudes towards early foreign language learning. The first one was a large-scale four-year-long survey-based investigation of over 20,000 school learners of Spanish and French in elementary schools in the US (second to fifth graders). The second follow-up study was based on structured interviews to 13 high school learners (11th and 12th graders) who had answered the surveys in the first study, were enrolled in a foreign language course in high school and volunteered to be interviewed. The time interval of this investigation allowed the researchers to conclude that early foreign language instruction had had a positive long-standing impact on adolescents' perceptions of foreign language speakers, cultures and their own foreign language education.

Another piece of research with a large gap between the major study and the follow-up study is that of Lamb (2018). This study was part of a research project originally conducted when a group of 12 junior high students in Indonesia started their formal study of English. Years later Lamb contacted those same learners again twice when they were in senior high school at the age of 22-24 (between 12 and 13 years after the start of the project). The study was inspired by learners' spontaneous references to the researcher's influence on their motivation to learn English in the earlier interviews. In his last contact with these studies, Lamb asked them to talk directly about his influence in their language learning trajectory and re-analysed earlier interviews. The study is valuable in that it underscores the relationship that is established with participants when research, particularly interviews, is conducted by the same researcher over such long periods of time. In the case of Lamb, his occasional interactions with learners (fewer than 10 spread over 12-13 years in less than 30-minute interviews) had a tangible and long-lasting effect on both learners who had been originally chosen because of their high motivation as well as those who were less motivated. After each meeting the relationship between Lamb and his participants became closer, the interviews less structured, and the data more trustworthy.

The last piece of research that we would like to bring to the readers' attention is Pfenninger's study (2020) on age of onset in an immersive school context in Switzerland. Unlike the studies we have just described above, where data was often collected once a year, this study stands out because data was collected every three months. The many measurements that were gathered throughout the eight years of the study (which ranged from 16 to 32 times depending on the group) allowed the researcher to map out the height and shape of patterns of L2 development between groups. Inspection of individual learner trajectories over the multiple measurements also showed that systematic variability within learners grew with time.

In sum, longitudinal research is a logical design when the aim is to explore change over time even though it has never been a very frequent methodological approach in SLA research. This is because, among other reasons, tracing change over time is challenging, with school and participant attrition being a major drawback. There is also a tendency in many longitudinal studies for study length to not go beyond a few weeks. Nevertheless, by going through a selection of long-term longitudinal studies, we have shown that this type of exceptional research is versatile and can be used to describe patterns of change over extended periods of time, reveal individual variability with time, provide thick descriptions of language learners' trajectories, make long-term predictions of L2 performance, evaluate L2 programs in the long run and even create fruitful researcher-participant relationships. Long-term longitudinal research is especially attractive in school contexts and is the approach that the ELLiC project has adopted.

1.3 The ELLiC project

The ELLiC project took place in the Barcelona area of Catalonia, where Catalan and Spanish are widely used in the community. Spanish is the majority language in Spain and Catalan is by law the language of communication at school. The use of Catalan to teach at school guarantees that all learners, independently of their family background, end up having a full command of the two official languages (Catalan and Spanish) in Catalonia. The educational system also guarantees that learners have a good command of at least one foreign language by the end of compulsory education, which is almost always English. Compulsory education in Catalonia and the whole of Spain lasts for 10 years. The first six years take place in primary schools (from ages 6 to 7 to ages 11 to 12) while the next four years take place in high schools (from ages 12 to 13 to ages 15 to 16) and are part of lower secondary education ('ESO'). Many primary schools also provide preschool education (from ages 3 to 4 to ages 5 to 6).

Foreign language instruction is required to start no later than G1 in primary schools when students are 6-7 years old even though schools can start foreign language instruction in pre-school. Schools also have the option of starting a second foreign language in the last two years of primary education and teach one or more content subjects in the first foreign language. At the time of the ELLiC project, primary schools had certain flexibility in the amount of hours devoted to the first foreign language ranging from a minimum of 315 hours to a maximum of 420 hours (in six years). The distribution of the minimum hours of instruction was regulated and increased gradually from a total of 70 hours distributed between G1 and G2 to a total of 105 hours distributed between G3 and G4, and a total of 140 hours distributed between G5 and G6.

Foreign language instruction continues to be compulsory in lower secondary education in Catalonia and Spain but the school curriculum for foreign languages is a bit more restrictive than in primary education. At the time of the ELLIC project, learners in Catalonia received a total of 420 hours of instruction in the first foreign language which almost always was English. Foreign language instruction was distributed fairly equally during the four years of compulsory schooling in secondary education. Nevertheless, there was some variability in the instruction of a second foreign language, which could be on offer as an optional subject in some grade levels and high schools, especially in large high schools. Like in primary education, high schools could choose to teach content subjects in English.

1.3.1 The ELLiC design

The ELLiC project started as part of ELLiE and focused on five participant primary schools in the Barcelona area. One or two classes per school were followed year after year until the end of primary education (G6). The ELLiC project continued throughout secondary education and followed a smaller number of the original participants in eight different high schools. Data collection in those high schools was carried out in their first and last years of secondary education only (G7 and G10). Figure 1.1 illustrates the design of the study with black and grey dots standing for primary and secondary education respectively. The blank dots indicate no data collection. Access to primary schools was requested through the schools' principals and access to secondary schools was requested through the Department of Education (regional government).

Figure 1.1 The ELLiC design



The ELLiC project includes two samples, a general sample with data from whole groups during primary school, and a focal learner sample gathered from a few learners from each group in primary and secondary school (see sections 1.3.2 and 1.3.3 respectively). The project includes quantitative as well as qualitative studies where multiple instruments are used (see section 3.4) and data is gathered from learners of English as well as their English teachers and parents.

1.3.2 The primary schools in ELLiC and participants

Five primary schools (A, B, C, D, and E) were selected for participation in this project because they were representative of average state schools in the area of Barcelona. Their students came from a wide range of socioeconomic backgrounds. None of the schools were located in deprived neighbourhoods. The proportion of subsidised lunches was low in all five schools and ranged from 1% to 11%. Immigration was low in all schools even if it ranged from schools with hardly any newcomers to schools with 2 newcomers out of every 10 students. School A was located next to a university's teacher training school and used to be a site for student teachers to do their practicum. This school together with school C were involved in an institutional project to promote English instruction. Schools D and E were originally independent schools and had a good reputation in their neighbourhoods. School B also had a good reputation for English instruction. Except for school C, where English started in preschool (age 5), in the rest of the schools English instruction started in G1 at the beginning of the ELLIE project. Table 1.1 summarises the main features of the five participating schools. Parental permission was obtained for each participant through the school.

Table 1.1 The five primary schools in ELLiC

	School A	School B	School C	School D	School E	
School	near downtown		close to	downtown	residential	
location	downtown	Barcelona	waterfront in	Barcelona	area in mid-	
	Barcelona		Barcelona		size town (45	
					km from	
					Barcelona)	
School size*	2	2	1	2	1	
3611001 3120	Ζ.	Z	L	Ζ	1	
Family	well-	Z lower middle	Iower middle,	z middle lass	mixed	
Family profiles	well- established	lower middle class	Iower middle,	middle lass	mixed background	
Family profiles	well- established working class	lower middle class	lower middle, lower class and some	z middle lass	1 mixed background	
Family profiles	well- established working class	lower middle class	lower middle, lower class and some international	z middle lass	nixed background	

Subsidised					
lunches	11,3 %	6 %	4.6 %	5.1 %	1%
Newcomers	≤ 20 %	≤ 10 %	≤ 10 %	≤ 5 %	1%

*Number of classes per grade level

From G1 to G6, some of the data was gathered collectively from the whole group of students. In G1 these classes ranged from 20 to 27 students per group. One class participated from each school except for school A in which two groups participated. Initially there were a total of 140 students in the G1 class lists and a considerable number of them stayed in the school until the end of primary education. Comparison of class lists in G1 and G6 yields an attrition of 28%. More information about the whole group or general sample can be found in Chapter 2.

Some data was collected individually from focal learners, a subgroup of students in each school (initially 2-4 boys and 2-4 girls per school). These learners were chosen by their English teachers in G1 after telling them we were interested in a variety of learner profiles. Nevertheless, in the long run we came to realise that some of the teachers mostly selected good learners, probably to make a positive impression on us. Our initial focal learner sample in G1 included 30 learners (16 boys and 14 girls). Except for one learner who was born in Argentina, the rest of the students were born in Spain and spoke Catalan (n = 17), Catalan and Spanish (n = 9), Spanish (n = 2), or Chinese (n = 1) at home². Seventeen of the focal learners had at least one parent with a university degree. In one of the schools a student left in G3 and was substituted by another student. Two other students left their schools later (in G5 and G6) and were not substituted.

1.3.3 The secondary schools in ELLiC and participants

When the ELLiC students moved from primary to secondary education, we succeeded in tracing 28 of the focal learners in G6 in their new high schools in G7. This was possible because the majority of students (n = 25) transferred to the same five high schools, and the remaining three students who attended three other high schools were convenient for us to reach. Six of the eight high schools where we collected data from in G7 were state schools. The other two schools were charter schools, that is, independent schools that were partially funded by the regional government. The proportion of newcomers in these schools ranged from 5% to 20%. See table 1.2 for further information about the high schools in the ELLiC project.

By the time these 28 focal learners were in G10, three students were lost because they were no longer in their initial high school or we were not able to contact them. The final number of focal learners from whom we could collect data during both primary school (G1-G6) and secondary school (G7 and G10) is 25. Six of these learners obtained a low final course mark in English in G10 (\leq 5 out of a maximum 10), fourteen obtained a high mark (\geq 8) and the remaining five got either a 6 or 7.

² This information was not available from one of the focal learners.

Table 1.2The eight secondary schools in the ELLIC project (G7)

	School A	School B	School C1	School C2	School D1**	ol D1** School D2		School E2**
	(n = 6)	(n= 5)	(n = 5)	(n = 1)	(n = 5)	(n = 1)	(n = 4)	(n = 1)
School	near	residential	close to	close to	downtown	near	mid-size	village close
location	downtown	area	waterfront	waterfront		downtown	town	to mid-size
		outside city					(45 km	town
		centre					from	
							Barcelona)	
School	state	state	state	religious charter	state	charter	state	state
typology								
School	4	4	4	2	5	5	4	2
size*								
Family	mixed	mixed	mixed	mixed	mixed	middle and	mixed	mixed
profiles	background	background	background	background	background	upper middle	background	background

*Number of classes per grade level ** Students from secondary schools D1 and D2 had attended primary school D. The same applies for secondary schools E1 and E2.

1.3.4 English instruction in ELLiC schools

A total of 17 primary school teachers taught English to the ELLiC participants for one to up to four school years. Most of these teachers, but not all of them, had been trained as primary school English teachers. The number of English sessions per week and their duration varied from school to school and from year to year within the same school. However, there was a general trend in G1-G2 for English instruction to be provided for two to three 45minute-long class sessions per week. From G3 on, the trend was for English instruction to be provided in 45- or 60minute-long sessions three to four times a week. Whenever possible in some schools, English instruction was taught to half the class at a time for one of the weekly sessions. In primary schools A, C, and E this happened more often than in primary schools D and B. The methodological approach in the five primary schools varied from teacher to teacher, but there was a tendency for teachers with less training to rely on mechanical practice and a textbook and for better trained teachers to follow a more communicative approach and use a wider variety of teaching materials. Students in schools A and C learned science in English for one hour a week from G3 to G6.

A total of twelve different English teachers taught the 28 focal learners in ELLiC when they moved to secondary school. In these schools English was taught for three sessions a week and one of these sessions was taught with half the group of students. In some schools one or two optional quarterly subjects were taught in English. English teachers in these high schools faced a complex situation because of large classes (30-33 students) and because they had very little knowledge of what or how their 'new' students had been learning English previously in primary schools. This was due to a lack of communication with English teachers in feeder schools. Another source of complexity was the range of learners' levels of proficiency because of feeder schools with different levels of English and some learners having attended or attending private English language schools in addition to their school lessons. The general trend for these English teachers, who had a good level of grammar knowledge as English philologists, was to follow a form-focused approach and start teaching grammar from zero in G7 in an attempt to balance learners' levels of proficiency. For more information about the transition from primary to secondary school in the ELLiC project see Muñoz et al. (2015) and Camuñas (2013).

1.3.5 Instruments

Interviews, questionnaires and other instruments

This section describes data elicited from interviews, questionnaires and other instruments in order to capture perceptions from primary and secondary school English teachers, learners and their parents.

Teacher interviews and questionnaires

Data was elicited from primary school English teachers in G1-G6 and high school teachers in G7. In G10 it was not possible to gather any data from the focal learners' English teachers. Once every year teachers answered our questions and completed a questionnaire about their teaching methodology and the focal learners. In G7 teachers were also asked about the topic of transition from primary to secondary school in an additional interview. All teacher interviews, which were conducted in Catalan or Spanish and often lasted for 30-45 minutes, were audio recorded and transcribed.

Learner interviews and questionnaires

Focal learners were interviewed individually at least once a year in primary school, twice in G7 (in December and May/June) and one last time in G10. During these interviews, which were originally developed as part of the ELLIE project, learners were encouraged to express their perceptions of their own learning process and their attitudes towards the classroom. In the G7 and G10 interviews, the ELLIE interview script was expanded to better fit the high

school context. From G6 on a second interview was developed as part of the ELLiC project with a focus on metalinguistic and crosslinguistic awareness. All the interviews were conducted in Catalan (the school language). The length of the interviews tended to increase over time. For example, in G1-G2 most of them lasted between 5-6 minutes while in G10 they lasted from 12 to 25 minutes.

An L2 motivation/attitudes questionnaire was administered to all students in class during primary education and to focal learners only in G7 and G10. It was created as part of the ELLiE project (Enever, 2011) and included questions like: *"Do you like English this year?"; "Is English easy or difficult for you this year?"*. Information about extracurricular instruction was also obtained in this questionnaire (except in G1 and G2 when the teacher provided us with this information). Illustration 1.1 shows the answer choice in the G2-G6 questionnaire. The questionnaire was always written in Catalan and administered towards the end of each school year. In G7 it was also administered in December. Five additional items were added in G10 about parental encouragement (Brady, 2015).

Illustration 1.1

Learner questionnaire in G6. Sample items from answer sheet.

Τ.	L								
	88	8		\odot	\odot				
	Dislike a lot	Dislike a lot Dislike		Like	Like a lot				
2				·					
	88	8		\odot	\odot				
	Very difficult	Difficult	Neither easy nor difficult	Easy	Very easy				

Parents' questionnaires

Two pencil and paper questionnaires were sent home with the children in G2 and G4 as part of the ELLiE project (Enever, 2011). The questionnaires enquired about the parents' bio data, their children's out-of-school contact with English, observations of their children's learning of English, etc. A total of 73 and 175 parents' questionnaires were collected from G2 and G4 respectively.

Other instruments

In an attempt to use age-appropriate data collection methods, two instruments with a strong visual component were used in G4 and G10. This modality of data made sense given the centrality of visual images in children's and teenagers' lives. In G4 learners were asked to produce a drawing in connection with English (*"When you hear the word 'English', which image comes to your mind?"*). In G10 learners were asked to complete two word webs with the prompts "English now" and "English in the future".

Linguistic and cognitive measures

In this section we provide a general description of the following linguistic measures: two general proficiency measures, two comprehension tests and two production tasks. The comprehension tests and the oral production tasks were originally developed as part of the ELLiE project (Enever, 2011). A language aptitude test is also described at the end of this section.

General proficiency measures

Two indicators of general English proficiency were: the scores obtained from an official external examination in G6 and final course marks. The official examination assessed L2 listening and reading comprehension (14 and 24 multiple choice items respectively) and L2 written production (between 8-10 lines). The exam is considered a reliable indicator of learners' English language achievement at the end of primary school. Scores from the Catalan and Spanish tests in the official external examination were also collected. All tests in the official examination were standardised tests. See Chapter 2 for additional information about these exams. The second English measure was the final course marks for English, which were collected every year.

Listening and reading comprehension tests

The listening test included a selection of items from the test used in the BAF project (Muñoz, 2006) and it consisted of a picture selection task in which learners were expected to match oral prompts with one of three visuals (see Illustration 1.2). The test, which was administered from G1-G6, never required any reading or writing from learners but the number and difficulty of the items increased over time to adapt to learners' increasing level of proficiency. The test was administered collectively during class time at the end of the school year.

Illustration 1.2. *Listening test: a sample item from the G6 test*

Instructions: Circle the correct picture. Oral prompt 8: "Mum puts John's lunch on the table"



The reading test, which was administered from G4 to G6, consisted of a picture comic strip for children. Learners were asked to read short sentences and match them to the illustrations (see illustration 1.3). In G6 additional true/false items were added to increase the difficulty of the test. The reading test, like the listening test, never required learners to do any writing and was administered collectively during class time at the end of the school year.

Illustration 1.3 *Comic strip*³

Instructions: Write a letter in every bubble. There is an extra letter.

³ Pesquis I Baliga-CF1119 ©Cavall Fort 2009, by Viladoms – All rights reserved.



A) THE REFRIGERATOR IS EMPTY.
B) WE'VE GOT SOME CHOCOLATE SPREAD, YUM!
C) MMMMM! THIS IS GOING TO BE GOOD! ...HEY! WHERE IS THE PINEAPPLE? WHERE IS THE BREAD?

Production tasks

Learners were asked to perform a written and an oral production task towards the end of every school year. The written task was administered from G5 on while the oral task was used from G4 on. The written task was about writing a letter to a pen friend from a different country. Learners completed the task in ten minutes (maximum) by handwriting during class time in G5 and G6 and individually outside class in G7 and G10.

The oral task, which was always administered individually, included short questions about themselves, their family, and friends and a guessing game based on a picture in which many different people appeared. The scene of the picture was a classroom scene in G4, G6, and G7, a supermarket in G5 and a scene at the beach in G10 but all scenes always included a similar number of people (21-24). During the guessing game, the researcher and the learner took turns guessing a character of their choice.

Language aptitude test

Learners' language-learning aptitude was measured by means of the MLAT-E (elementary version of the Modern Language Aptitude Test) in its Catalan adaptation (Suárez, 2010). The test was administered to intact classes in G6. It consists of four parts: Hidden Words (ability to associate sounds and symbols), Matching Words (sensitivity to grammatical structure), Finding Rhymes (ability to hear speech sounds), and Number Learning (auditory alertness and ability to remember).

Lesson observations

Two types of lesson observations were conducted: group and focal learner observations. At least two group observations per teacher were carried out each year except in G7 when observations took place in only some schools. It was not possible to conduct lesson observations in G10. A template was used while the lesson was going on in which a chronological description of the stages of the lessons plus field notes were recorded. After each lesson, a post-observation sheet was completed that included structured and open questions. From G1 to G4, observations of the focal learners were also conducted in which three students were observed at a time once or twice per school year. The templates were originally developed as part of the ELLiE project (Enever, 2011) and were gradually modified during the course of the study.

Summary of instruments

Table 1.3 lists all the instruments/measures in the ELLiC project with an indication of when they were used.

Table 1.3Overview of instruments/measures

		Primary					Secondary		
	Instruments / Measures		education					education	
			G2	G3	G4	G5	G6	G7	G10
	Teacher interview and questionnaire	х	х	Х	х	х	Х	х	
s Ient:	Learner ELLiE interview	х	х	х	х	х	х	х	х
ws, naire trum	Meta- and cross ling awareness interview						х	х	х
ervie cionr r ins	Learner questionnaire	х	х	х	х	х	х	х	х
Inte uest other	Parents' questionnaire		х		х				
and e	Learners' drawings				х				
(0	Learners' word webs								х
	Official external examination scores						Х		
gnitive	Course marks (English)	x	х	x	x	x	x	х	x
d cog	Listening test	х	х	х	х	х	х		
: anc easu	Reading test				х	х	х		
Linguistic me	Written task					х	х	х	х
	Oral task				х	х	х	х	х
	Language aptitude test						х		
	Classroom observations	х	х	х	х	х	х	х	
Observ.	Focal learners' observations	x	x	x	x				

References

Barkaoui, K. (2014). Quantitative approaches for analysing longitudinal data in second language research. *Annual Review of Applied Linguistics*, 34, 65-101. <u>https://doi-org/10.1017/S0267190514000105</u>

Brady, I. K. (2015). *The ideal and ought-to L2 selves in Spanish learners of English* [Doctoral dissertation]. Universidad de Murcia. TESEO <u>https://digitum.um.es/xmlui/handle/10201/45976</u>

Bridgewater, C. (2020). *Longitudinal classroom-based research: A synthesis of L2 studies* [Master's thesis]. Universitat de Barcelona. Retrieved February 9, 2023, from <u>http://hdl.handle.net/2445/193380</u>

Cable, C., Driscoll, P., Mitchell, R., Sing, S., Cremin, T., Earl, J., et al. (2010). *Language learning at key stage 2: A longitudinal study*. London: Department for Children, Schools, and Families (DCSF Research Reports DCSF-RR198)

Camuñas, M. (2013). The provision of English as a foreign language at transition from primary to secondary school in the classrooms of *Catalonia* [Master's thesis]. Universitat de Barcelona. Retrieved February 9, 2023, from http://hdl.handle.net/2445/193307

Costache, O. Becker, E. S., & Goetz, T. (2022). Is English the culprit? Longitudinal associations between students' value beliefs in English, German, and French in multilingual Switzerland. *Modern Language Journal*, 106(2), 313-327. https://doi.org/10.1111/modl.12774

De Wilde, V., Brysbaert, M. & Eyckmans, J. (2021). Young learners' L2 English after the onset of instruction: Longitudinal development of L2 proficiency and the role of individual differences. *Bilingualism: Language and Cognition*, 24, 439–453. https://doi.org/10.1017/S1366728920000747

Ellis, R. (2021). A short history of SLA: Where have we come from and where are we going? *Language Teaching*, 54, 190–205. <u>https://doi-org/10.1017/S0261444820000038</u>

Enever, J. (Ed.) (2011). ELLIE. Early Language Learning in Europe. British Council

European Commission, European Education, and Culture Executive Agency, Motiejūnaitė-Schulmeister, A., Sicurella, A., Birch, P., (2022). The structure of the European education systems 2022/2023: Schematic diagrams. Publications Office of the European Union. https://data.europa.eu/doi/10.2797/21002.

Heining-Boynton, A. L. & Haitema T. (2007). A ten-year chronicle of student attitudes toward foreign language in the elementary school. *The Modern Language Journal*, 91(2), 149–168. http://www.jstor.org/stable/4625998

Kibler, A. K., (2017). Becoming a "Mexican feminist": A minoritized bilingual's development of disciplinary identities through writing. *Journal of Second Language Writing*, 38 (October), 26–41. <u>https://doi-org/10.1016/j.jslw.2017.10.011</u>

Kibler, A. K., & Hardigree, C. (2017). Using evidence in L2 argumentative writing: A longitudinal case study across high school and university. *Language Learning*, 67(1), 75–109. <u>https://doi-org/10.1111/lang.12198</u>

Kunitz, S., & Marian, K. S. (2017). Tracking immanent language learning behaviour over time in task-based classroom work. *TESOL Quarterly*, 51(3), 507–535. <u>https://doi-org/10.1002/tesq.389</u>

Lamb, M. (2018). When motivation research motivates: Issues in long-term empirical investigations. *Innovation in Language Learning and Teaching*, 12(4), 357-370. https://doi.org/ 10.1080/17501229.2016.1251438

Lightbown, P. M., Halter, R. H., White, J., & Horst, M. (2002). Comprehension-based learning: The limits of "do it yourself." *Canadian Modern Language Review*, 58, 427–464. <u>https://doi-org/10.3138/cmlr.58.3.427</u>

Lightbown, P. M. (1992). Can they do it themselves? A comprehension-based ESL course for young children. In R. Courchêne, J. St. John, C. Thérien, & J. Glidden (Eds.), *Comprehension-based language teaching: Current trends* (pp. 353-370). University of Ottawa Press

Lightbown, P. M. & Halter, R. (1989). *Evaluation of ESL learning regular and experimental programs in four New Brunswick school districts: 1987-88 (Vol. 2)*. Research report prepared for the Ministry of Education of New Brunswick

Lindgren E. & Enever J. (2017). Employing mixed methods for the construction of thick descriptions of early language learning. In J. Enever & E. Lindgren (Eds.), *Early language Learning. Complexity and mixed methods* (pp. 201-221). Multilingual Matters

Long, M., & Doughty, C. (2003). SLA and cognitive science. In C. Doughty, & M. Long (Eds.), *The handbook of second language acquisition* (pp. 866–870). Blackwell

Loewen S., & Plonsky L. (2016). An A-Z of applied linguistics in research methods. Palgrave

Muñoz, C. (Ed.) (2006). Age and the rate of foreign language learning. Multilingual Matters

Muñoz, C., Tragant, E., & Camuñas, M. (2015). Transition: Continuity or a fresh start? *APAC ELT Journal*, 80, 11-16. Retrieved January 17, 2023, from

https://static1.squarespace.com/static/594fcf2ca5790afdb0a78fb6/t/6050c7aaf88b82622ea045ad/161590676219 5/APAC+ELT+Journal+n.80.pdf

Norton, B. (2000). Identity and language learning: Gender, ethnicity, and educational change. Longman

Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26–45. <u>https://doi-org/10.1017/S0267190505000024</u>

Perdue, C. (Ed.) (1993a). Adult language acquisition: Cross-linguistic perspectives. Vol. I, Field Methods. Cambridge University Press

Perdue, C. (Ed.) (1993b). Adult language acquisition: Cross-linguistic perspectives. Vol. II, The Results. Cambridge University Press

Pfenninger, S. E. (2020). The dynamic multicausality of age of first bilingual language exposure: Evidence from a longitudinal content and language integrated learning study with dense time serial measurements. *The Modern Language Journal*, 104, 662–686. <u>https://doi-org/10.1111/modl.12666</u>

Riazi, A. M. (2016). The Routledge encyclopedia of research methods in applied linguistics. Quantitative, qualitative, and mixed-methods research. Routledge

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term relationships among early first language skills, second language aptitude, second language affect, and later second language proficiency. *Applied Psycholinguistics*, 30(4), 725-755. https://doi.org/10.1017/S0142716409990099

Suárez, M. (2010). Language aptitude in young learners: The elementary modern language aptitude test in Spanish and Catalan [doctoral dissertation]. University of Barcelona. Retrieved January 30, 2022, from http://www.tesisenred.net/handle/10803/38244

Tragant, E. & Victori, M. (2006). Reported strategy use and age. In C. Muñoz (Ed.), Age and the rate of foreign language learning (pp. 208-236). Multilingual Matters

Wode, H. (1976). Developmental sequences in naturalistic L2 acquisition. *Working Papers on Bilingualism*, 1(1), 1–13