

Receptive Vocabulary Acquisition in Pre-Primary Education through Soft-Content and Language Integrated Learning

Marta Segura¹, Helena Roquet¹ & Júlia Barón²

¹ Institute for Multilingualism, Department of Applied Linguistics, Universitat Internacional de Catalunya, Spain

² Department of Modern Languages and English Studies, Philology and Communication Faculty, Universitat de Barcelona, Spain

Correspondence: Marta Segura, Institute for Multilingualism, Department of Applied Linguistics, Universitat Internacional de Catalunya, Josep Trueta s/n, 08195, Sant Cugat del Vallès, Barcelona, Spain.

Received: June 25, 2021

Accepted: August 11, 2021

Online Published: September 2, 2021

doi: 10.5539/elt.v14n10p1

URL: <https://doi.org/10.5539/elt.v14n10p1>

Abstract

Content and Language Integrated Learning (CLIL) has become the focus of Foreign Language (FL) research within the last decades. CLIL provides a more complete, naturalistic, and meaningful context to FL learning, which has proven to bring many benefits to learners, such as a higher motivation and promotion of creativity, and better results in receptive skills, vocabulary, morphology and fluency. Nevertheless, most CLIL research has focused on primary and secondary level students and, thus, more research is needed with younger learners, namely, pre-primary students. The present study examines the learning of FL vocabulary in pre-primary learners following a soft-CLIL program, as compared to their same age peers following Formal Instruction (FI) of English. Over the course of six months, pre-primary students of two grades, namely 4- and 5-year-old students (N=155), took part in such program, aiming at teaching two curricular preschool units, traditionally taught in the mother tongue (L1), in English in the FL sessions. A longitudinal study was conducted, and students were administered a general vocabulary level pre-test, as well as a target words receptive vocabulary post-test after the two units had been worked on. The focus of the research was on receptive vocabulary acquisition, but age and word frequency effects were also analyzed. Results showed positive tendencies in receptive vocabulary development through soft-CLIL, although not statistically significant. A significant frequency effect was found, indicating that high-frequency words are recalled more easily than lower-frequency ones, but no significant differences were found when comparing learners from the two grades.

Keywords: Content and Language Integrated Learning, English as a Foreign Language, pre-primary education, receptive vocabulary, vocabulary acquisition

1. Introduction

Foreign language (FL) learning research has gained ground in the latest decades, during which age of onset of FL learning in schools has been brought down (Pérez-Vidal, Escobar & Roquet, 2013) and new FL teaching methodologies have been explored and implemented. That is the case of integrated teaching approaches, such as Content and Language Integrated Learning (CLIL), in which a curricular subject is taught in the FL (Lasagabaster, 2008), providing a more meaningful and natural FL learning context (Artieda, Roquet, & Nicolás-Conesa, 2020; Lasagabaster & Sierra, 2009; Pérez-Vidal, Lorenzo, & Trenchs, 2015).

Immersion has been claimed to be an ideal FL learning context, as it provides a natural learning environment with high input exposure and interaction opportunities (Pérez-Vidal, 2011). However, immersion is not always possible and FL teaching approaches such as CLIL have developed to attempt to bring to the classroom a more naturalistic and meaningful context. Research has already been conducted to analyze the effects of CLIL programs and benefits have been found in many linguistic and non-linguistic areas (Dalton-Puffer, 2008; Lasagabaster, 2008; Lorenzo, Casal & Moore, 2010; Pérez-Vidal, 2011; Roquet & Pérez-Vidal, 2017).

Although there have been many studies researching the effects of CLIL in different skills and settings, most of such research has been conducted with primary and secondary level students. Very few studies (Asensio Arjona, 2020; García Esteban, 2015) have focused on the analysis of CLIL programs in pre-primary education. Therefore,

the present research aims to shed some light into receptive vocabulary acquisition in pre-primary EFL students following a CLIL program.

1.1 Content and Language Integrated Learning (CLIL)

In light of the recent increasing interest to promote FL learning since an early age, many countries and schools have increased the quantity and quality of exposure students have to the FL within the school, by implementing CLIL programs in which “a foreign language is used as the medium of instruction to teach content subjects” (Artieda, Roquet & Nicolás-Conesa, 2020: 1). When CLIL programs are correctly implemented, both the content and the FL share the focus of attention (Dale & Tanner, 2012; Ortega, 2015), which reduces the pressure students may have to learn the target language and provides more interaction opportunities, as well as contextualized learning (Lasagabaster & Sierra, 2009; Pérez-Vidal, 2011).

While all CLIL programs share the main feature of the FL and the content being taught integrated, there are some variable characteristics that change in each program, such as their timing, the institutional support and resources, their continuation (Pérez-Vidal, 2013) and their intensity (Dale & Tanner, 2012; Pérez-Vidal, 2013). Two subtypes of CLIL can be distinguished, depending on their intensity: hard-CLIL and soft-CLIL (García Esteban, 2015). While in hard-CLIL almost half or more of the curriculum is taught in a FL, in soft-CLIL “language teachers do cross-curricular work or teach topics from the curriculum as part of a language course” (García Esteban, 2015: 30). Therefore, a soft version of CLIL provides the opportunity to language teachers to teach the FL using other curriculum contents and to bring to the classroom meaningful exposure to the target language. The fact that language teachers are in charge of the teaching helps mitigate one of the main problems that CLIL content teachers may have to face when teaching by the means of a FL, which is the lack of high FL proficiency (Mair, 2018; Pavón, 2014).

Regardless of their intensity and other variable characteristics, in all CLIL programs the curricular content and the FL should be taught in an integrated way (Lasagabaster, 2008), bringing into the classroom higher quality input and communication opportunities. Additionally, CLIL hours are usually added hours to the regular FL classes, thus increasing the time of exposure students have to the target language, which is key in many contexts where contact with the FL is very limited outside the classroom (Muñoz, 2008) and where immersion is not possible. Therefore, CLIL programs bring to the classroom an environment more similar to the immersion context, in which there is a higher quantity and quality of contextualized input, as well as more opportunities for natural interaction (Dalton-Puffer, 2008; Lasagabaster & Sierra, 2009).

According to previous research, the FL learning context that CLIL provides brings many benefits to the students, due to the increased contact with the language and the naturalistic learning environment. Higher results have been found in CLIL learners, as compared to their same age peers following Formal Instruction (FI), in some linguistic skills, such as vocabulary, morphology, and receptive skills (Dalton-Puffer, 2008; Lasagabaster, 2008; Lorenzo, Casal, & Moore, 2010; Pérez-Vidal, 2011). Some other personal skills have also been found to be enhanced through CLIL, namely, motivation, risk-taking, and self-confidence towards FL learning (Dalton-Puffer, 2008). Nevertheless, there are some other areas in which contradictory results have been found, that is the case of productive skills, syntax, pronunciation, and pragmatics (Dalton-Puffer, 2008; Pérez-Vidal, 2011).

In terms of vocabulary, several researchers have reported an advantage for students following CLIL programs. Lasagabaster (2008) found that bilingual secondary education CLIL students outperformed their non-CLIL peers in all tests, including the four main linguistic skills, grammar, and vocabulary. Those results are in line with the findings reported in Artieda, Roquet & Nicolás-Conesa (2020) in which bilingual secondary education students following FI+CLIL showed an advantage over the FI group in reading comprehension, which is closely correlated with vocabulary knowledge. Similar results were also found by Jiménez-Catalán, Ruíz de Zarobe & Cenoz (2006), who analyzed primary education students' performance in reading and writing measured through vocabulary use, and reported higher results by the CLIL group in both skills. In a longitudinal study, Pérez Cañado (2018) also found an advantage in primary and secondary education CLIL students over their non-CLIL peers in all linguistic aspects, including vocabulary.

When analyzing receptive vocabulary development in primary and secondary education CLIL students, Agustín-Llach & Canga Alonso (2014) concluded that CLIL seemed to enhance receptive vocabulary acquisition. However, they pointed out that such advantage in CLIL learners may be attributed, not only to the CLIL program, but also to the increased amount of exposure to the language. In a posterior study, Canga Alonso (2015) compared CLIL primary education students with non-CLIL secondary education students with higher amount of exposure to the FL and found that the older non-CLIL students did not show a significant difference in receptive

vocabulary knowledge, even after a larger exposure to the FL. Such results led the author to state that the CLIL context provides enough meaningful input to enhance vocabulary acquisition.

There is, thus, a significant number of studies that have found that CLIL gives an advantage to learners, in terms of vocabulary acquisition, which seem to be accentuated even more in the long term and with increased language exposure. Nevertheless, such results are contradictory with the ones reported in Admiraal, Westhoff & de Bot (2006), who did not find any differences in receptive vocabulary development between the secondary education CLIL and FI groups after four years of instruction.

Most research tapping onto the effects of CLIL programs has focused on FL learning in primary and secondary school students. To our knowledge, very few studies (Asensio Arjona, 2020; García Esteban, 2015; Mair, 2018) have analyzed how CLIL may enhance FL acquisition in very younger learners, namely pre-primary education students.

1.2 Age and Learning Context in FL Learning

In our current globalized and multilingual society, learning a FL has become an essential element of education, and schools have started to design strong FL teaching programs. Two of the main measures that have been promoted by the European Union to enhance multilingualism are the implementation of new integrated FL teaching methodologies (Pérez-Vidal, 2011), such as CLIL, and the promotion of an earlier start of FL learning in schools (Muñoz, Tragant, & Torras, 2010).

It has traditionally been believed that the sooner children start to learn a language, the better the chances are for them to reach native-like proficiency. Such belief was supported by the Critical Period Hypothesis, first formulated by Penfield (Penfield & Roberts, 1959) and later popularized by Lenneberg (1967), according to which there is a certain age after which attaining native-like proficiency in a FL becomes more difficult. This, however, is not to be generalized, since the context of acquisition plays a crucial role (Miralpeix, 2006; Muñoz, Tragant, & Torras, 2010) both in children and adults when learning a FL: while children may benefit more from naturalistic environments where they can learn implicitly, older learners seem to benefit more from FI contexts where they can profit from explicit and analytic learning, since they have developed greater problem solving skills and analytic mechanisms (DeKeyser, 2000), as well as a higher cognitive development. That is in accordance with the results reported in Artieda, Roquet, & Nicolás-Conesa (2020), which showed that biological age did indeed have an impact on language learning, since older learners following FI outperformed their younger peers following FI+CLIL after receiving the same number of hours of instruction. In a previous study, Muñoz (2006) found that older learners had a higher rate of FL acquisition at initial stages, thanks to their explicit learning mechanisms; while younger students acquired the language at a slower rate, since they mostly had to draw from their implicit learning mechanisms, which explains why younger learners seem to outperform their older peers in a natural context.

Therefore, the context of FL acquisition is an essential factor to consider when analyzing the effects of early age of onset (Miralpeix, 2006; Miralpeix, 2007; Muñoz, 2008; Muñoz, Tragant, & Torras, 2010). In a study conducted by Krashen, Long, & Scarcella (1979) it was reported that in a natural immersion context, children learning the FL ended with a higher proficiency than learners that began learning the language as adults. Nevertheless, it cannot be claimed that there is a significant advantage of an early age of onset when learning the language in a FL setting, where input time and exposure are limited (Muñoz, 2008). Thus, it cannot be stated that an early age of onset will result in higher level of attainment in any context, since there are some differences between contexts that have to be accounted for (Muñoz, 2006; Muñoz & Singleton, 2013), such as intensity of FL contact, and quality of input. As Muñoz (2008) states, “an early starting produces long-term benefits when associated with greater time and massive exposure, as in immersion programs, but not when associated with limited time and exposure, as in typical FL learning classrooms” (p. 582).

According to the findings previously reported, starting to learn a FL at a young age seems to give an advantage to students when they are in an immersion context, with large amounts of exposure to the FL (Miralpeix, 2007). Nevertheless, in most cases the advancement of age of onset has not come in hand with a change in the context and, in many schools, learners are starting FL classes earlier, but still following FI in a FL context where the contact with English is of 2-3 hours per week at most. Therefore, the present research provides a potentially more beneficial context, since pre-primary level students are involved in a soft-CLIL program, which, although in a FL context, brings to the classroom a more meaningful and natural-like context for FL learning, from which younger learners can learn implicitly.

1.3 FL Vocabulary Acquisition

FL vocabulary acquisition has been largely studied in the last decades, as vocabulary knowledge has been seen as crucial to be able to communicate in the FL (Albaladejo, Coyle, & de Larios, 2018; Meara, 1996; Schmitt, 2008). It has been found that learners need about 8000 to 9000 words families for fluent reading, while 6000 to 7000 words families seem to be necessary for oral comprehension (Schmitt, 2008; Schmitt & Schmitt, 2020). Therefore, FL learners need to acquire a very large amount of vocabulary to be proficient in all linguistic skills. But learning a word is not an easy task, because, as pointed out by Nation (2020) it entails mastering many aspects, such as the spoken and written form, the word parts, its form and meaning, its concept and referents, its associations, its grammatical functions, collocations and uses.

Therefore, considering the large amount of vocabulary that is required to function in a language and the complexity of learning it, two of the main criteria that have been traditionally used to select the vocabulary to teach in the FL classroom are the usefulness and the frequency (Nation, 2020; Nation & Meara, 2020; Schmitt & Schmitt, 2020). By conducting a needs analysis, the teacher should be able to detect which vocabulary will be useful and engaging for the learners. Furthermore, after analyzing frequency lists, it has been seen that “a relatively small number of high-frequency items are extremely frequent” (Schmitt & Schmitt, 2020, p. 6). Thus, prioritizing the learning of high-frequency words, namely from the first three 1000 words bands from frequency lists, will enable FL learners to cover a wide range of texts and function in basic conversations (Nation, 2020; Nation & Meara, 2020).

Previous research related to word frequency and acquisition have supported such idea. In a study conducted by Alexiou (2015) on FL vocabulary acquisition with 8 and 9-year-old children, results showed that high frequency words were learnt faster and were recalled more frequently than lower-frequency words. The author claimed that this was possibly due to the fact that higher-frequency words usually refer to concepts that are more common in our daily lives and, therefore, children will probably already have a conceptual representation for such words. Similar results were reported by Shaban (2013), in a study where children aged 3 to 5 learning English as a FL seemed to acquire more “frequent, concrete, typical and basic level words faster” (p. 722). Miralpeix’s (2007) study on the effects of age and exposure on productive vocabulary knowledge in high school students also found a much higher usage of words belonging in the first 1000 band of frequency, than words belonging in less frequent bands, both in written and oral tasks.

Research in the vocabulary acquisition field has also focused on the analysis of receptive and productive vocabulary learning, and two main relevant findings are worth mentioning. Firstly, receptive vocabulary size in high level learners has been found to be a good indicator of FL proficiency (Miralpeix & Muñoz, 2018). Secondly, receptive and productive vocabulary knowledge is seen as within a continuum: while knowing a word receptively means being able to identify and understand the word in listening or reading, knowing it productively is the next step and entails knowing how to use the word in speech or writing (Meara & Miralpeix, 2021; Schmitt & Schmitt, 2020; Webb, 2020). Receptive vocabulary has been found to be acquired faster and earlier than productive vocabulary (Schmitt & Schmitt, 2020; Yongqui Gu, 2020). This may be due to, as pointed out by Nation (2020), the complexity of both kinds of knowledge: for productive vocabulary, learners need to have mastered more aspects of a word, while for receptive vocabulary word knowledge may be partial. Thus, it is natural for receptive vocabularies to precede (Meara & Miralpeix, 2021) and to be bigger than productive vocabularies (Miralpeix, 2020), especially in early EFL stages.

Within the framework just presented in the current section, a need for more research in FL acquisition with very young learners, namely pre-primary students, has been identified. More specifically, there is a need for more studies analyzing FL learning in students at an early age and in a context with limited exposure to the FL. That is precisely the niche that the present article aims to fill, by the means of analyzing receptive vocabulary development in pre-primary students following a soft-CLIL program in a FL context.

The present study seeks to analyze whether a soft-CLIL program enhances vocabulary learning in very young pre-primary learners of English as a FL and provides new data in the FL vocabulary learning domain. Thus, the effects of said soft-CLIL program will be examined in terms of the students’ ability to remember and identify the target vocabulary. Age effects will also be analyzed, by comparing the results of students in the last two grades of pre-primary education, namely 4- and 5-year-olds. Additionally, the effects of word frequency will be studied, in order to examine whether higher-frequency words are recalled easier than lower-frequency words.

Therefore, the following research questions, and respective hypotheses, have been set:

RQ1. Will pre-school students be able to acquire a wider range of receptive vocabulary in their L2 following a soft-CLIL program?

H1. According to the literature reviewed, pre-primary students will be able to acquire a greater range of receptive vocabulary in the L2 through a soft-CLIL methodology, as compared to their non-CLIL peers.

RQ2. Will older students (5-year-olds) be able to recall a larger amount of vocabulary than their younger peers (4-year-olds) within the same time period?

H2. Older students (5-year-olds) will be able to recall a larger amount of vocabulary, compared to their younger peers (4-year-olds) within the same time period.

RQ3. Does frequency of words affect vocabulary recalling?

H3. High-frequency words will be easier to recall than low-frequency words.

2. Method

2.1 Participants

A total of N=155 Catalan and Spanish bilingual pre-primary students (aged 4 and 5 years old) from two semi-private schools in Catalonia participated in the study. Students in both schools have a similar socioeconomic background and students' families were given a sociolinguistic background questionnaire, tapping onto the family languages and the exposure to English that children had inside and outside of the school.

Participants were 4- and 5- year-old students, enrolled the last two years of pre-primary education (Note 1). In each of the schools, students in each grade were divided into two classes of between 17 and 25 students each. Those groups were kept the same and used as control and experimental, thus having two groups (one control and one experimental) per grade and school. Nevertheless, in the final grouping distribution students from the same groups in each school were mixed and analyzed as part of the same group, therefore, having 4 groups, as can be seen in Table 1.

Table 1. Participants Distribution

Participant's age & grade	Group	Treatment	Participants
4-year-old learners (P4)	Group 1	Control group (FI / non-CLIL)	N=35
	Group 2	Experimental group (soft-CLIL)	N=34
5-year-old learners (P5)	Group 3	Control group (FI / non-CLIL)	N=42
	Group 4	Experimental group (soft-CLIL)	N=44

2.2 Pedagogical Intervention and Treatment

The present research was born from a needs analysis conducted with the pre-primary teachers in the schools. As it was observed, there is no established curriculum for English in pre-primary grades and teachers in those schools felt that the English content presented in each of the three grades of pre-primary was repetitive and very basic, not allowing students to learn as much as they potentially could. Therefore, a pedagogical intervention was designed by the researchers and the schoolteachers, to develop an age-appropriate curriculum that favored EFL vocabulary acquisition and that increased the amount of vocabulary that students were exposed to. As a result, a soft-CLIL methodology was adopted by the English teachers, with the aim of providing contextualized and meaningful learning. Considering that there are not separate curricular subjects in pre-primary, but that units and curricular content are worked on through topic-based projects, two main units that students learn about in their mother tongue (L1) and that are part of the general pre-primary curriculum were selected to be adapted as the soft-CLIL units: the first two units related to the seasons, namely Autumn and Halloween, and Winter and Christmas.

Control groups received instruction of English as a FL, following the basic vocabulary school curriculum; while experimental groups received the same basic instruction with the addition of more topic-related vocabulary, which was introduced through a soft-CLIL methodology. Students of both groups received the same number of hours of instruction and were exposed to English within the school the same amount of time, since CLIL hours were not extra hours, but embedded within the regular EFL classes. Thus, students were exposed to the same amount English hours within the school: the 20-minute morning routines done in English two or three days a week since the first grade of pre-primary (P3) and 2 hours per week of EFL lessons. In addition, another subject was taught in English starting from P3: Arts and Crafts in one of the schools, and Music in the other. Therefore, although participants were from two different schools, they had the same contact with English within the school, namely, around 3 hours per week.

During the development of the intervention and its implementation, there were regular meetings with the classroom teachers, as well as with the EFL teachers in both schools, to select the appropriate vocabulary and to develop the materials, in the shape of flashcards, worksheets and a list of activities and games as a resource for the English teachers. It is also important to note that in each of the schools, there was one single English teacher in charge of teaching all grades and classes in pre-primary. Therefore, the experimental and the control groups were taught by the same teacher, who followed the same teaching strategies in both classes, changing only the amount of vocabulary introduced in each group.

2.3 Design

The study followed a longitudinal design, with two testing times and a pedagogical intervention in between. The first testing time (T1) took place in October 2019, at the beginning of the school year, before beginning the CLIL pedagogical intervention in the EFL lessons, which consisted of two themed vocabulary units. T1 served as a diagnostics tool to ensure group comparability and analyze starting levels.

Each of the intervention units was worked on in the EFL sessions for two months. In each of the grades, the control groups were presented with the English basic vocabulary included in the school curriculum about those units, while the experimental groups were exposed to an increased amount of vocabulary, including not only the basic vocabulary, but also some more complex and less frequent words and expressions. Furthermore, to guarantee that there was a progression in terms of content between the two grades (P4 and P5), older students (P5) were presented with the same vocabulary as their younger peers (P4 experimental groups), as well as some more advanced vocabulary (see Appendix A). Around 60% of words in each unit were high-frequency words from the first three 1000 bands, while the other 40% of words were of lower-frequency. After the intervention, in March 2020 (T2), students took the receptive vocabulary post-tests, which included words from the two units worked on in the pedagogical interventions in the EFL and CLIL lessons.

2.4 Instruments

The standardized Peabody Picture Vocabulary Test (PPVT 4th Edition, Form A) was used as a pre-test for all students, to compare the starting levels of both schools and of the classes in each grade. Thus, the PPVT-4 test was a diagnostic tool to guarantee that there were no statistically significant differences neither between schools, nor between control and experimental groups within each grade.

A sociolinguistic background questionnaire (see Appendix B) was sent out to the families of the students to gather information about the students' contact and exposure to the language outside of the school. The questionnaire inquired on the family background and languages used at home, any extracurricular activities in English that the students took part in and any other contact with the language that the students may have, such as through a native speaker family member, babysitter, watching videos in English or listening to songs, amongst others.

As for the specific tests related to the vocabulary of the two units that the intervention focused on, two tests were developed by the researchers: a curriculum words receptive test (see Appendix C) including only the basic curriculum words, which was administered to all control and experimental groups; and a target words receptive test (see Appendix D), including the additional words for the experimental groups only. Different versions of each of the tests were developed, for each grade (P4 and P5), to include between a 50% and a 70% of the vocabulary that was presented in each grade. To select the vocabulary included in the tests, frequency (according to the British National Corpus (BNC) and New General Service List (NGSL) databases) was considered, as well as the unit that words belonged to. As regards to the format of both the curriculum and the target words receptive tests, the same format, and guidelines as in the PPVT-4 (Note 2) and the Expressive Vocabulary Test (EVT) (Note 3) were followed, in terms of picture characteristics and combination within the page.

2.5 Data Collection and Procedure

Data collection for both testing times (T1: PPVT-4 pre-test, and T2: basic curriculum and additional target words tests) was done within two weeks. For all tests, students were tested individually in a separate classroom. Instructions were given in Catalan or Spanish, the students' mother tongue, and the first four items of each test were used as training items, for the students to familiarize themselves with the test and for the researchers to make sure that participants understood the testing procedure. The curriculum and target words tests that were created by the researchers were piloted with a group of 6-year-old students in the same schools, prior to the testing time with the pre-primary students.

2.6 Data Analysis

During data collection, a record was kept with the students' answers for each of the items tested. Correct answers added a point and incorrect answers did not count. Raw scores were used in the first research question. Data was also normalized assigning standard scores, in order for raw scores to be compared.

For the first research question, IBM SPSS Statistics 26 was used to analyze the data. Group means were compared, and independent samples T-tests were carried out to compare the results of the control and the experimental groups in each test. P value was set at ($p=0.05$).

For the second and third research questions, the program R 3.5.2 was used. To describe the absolute frequencies and percentages, and analyze the relation between the categorical variables, X2 contrast was used. Additionally, two mixed generalized linear models of Poisson were adjusted, including the following variables: target word, frequency, class, school, student, age and test; from which prevalence ratios of the models were obtained. Significance levels were set at ($p=0.05$).

3. Results

3.1 Receptive Vocabulary Acquisition through Soft-CLIL

The first research question aimed at analyzing receptive vocabulary development in 4- and 5-year-old pre-primary EFL learners. Half of the participants, namely those in the control groups, followed the regular preschool EFL classes, while experimental groups were presented with a larger amount of vocabulary through a soft-CLIL program, keeping exposure time the same. It was expected for learners in the experimental groups to show higher results in their receptive vocabulary, due to their participation in the soft-CLIL program, when compared to their same grade non-CLIL peers.

Table 2. Group Statistics for Basic Curriculum Receptive Test (G1 - G2 - G3 - G4)

Group	N	Mean	Std. Deviation	Std. Error Mean
G1	35	7.11	2.259	.382
G2	34	7.62	1.826	.313
G3	42	11.79	2.637	.407
G4	44	12.55	2.406	.363

Table 3. Independent Samples Test (G1 - G2)

	Levene's Test for Equality of Variances	Test Sig.	t-test for Equality of Means						
			t	df.	Sig. (2-tailed) <i>P</i> -value	Mean Difference	Std. Error Difference	95% CI of the Difference	
	F							Lower	Upper
Equal variances assumed	.981	.326	1.016	67	<i>p</i> = .313	-.503	.495	-1.492	.485
Equal variances not assumed			-1.019	64.876	<i>p</i> = .312	-.503	.494	-1.490	.483

Table 4. Independent Samples Test (G3 - G4)

	Levene's Test for Equality of Variances	Sig.	t-test for Equality of Means						
			t	df.	Sig. (2-tailed) <i>P</i> -value	Mean Difference	Std. Error Difference	95% CI of the Difference	
								Lower	Upper
Equal variances assumed	.508	.478	-1.397	84	<i>p</i> = .166	-.760	.544	-1.841	.322
Equal variances not assumed			-1.394	82.425	<i>p</i> = .157	-.760	.545	-1.844	.325

Table 2, Table 3 and Table 4 show the raw results that each of the groups got when tested in the basic curriculum words receptively. In both grades slightly higher results were reported in the experimental CLIL groups (G2 and

G4), when compared to their same grade non-CLIL peers. In the case of the younger groups, namely 4-year-olds, the control group's (G1) mean score in the basic words receptive test was 7.11, while the mean score for the experimental group (G2) was 7.26, a difference that was not statistically significant ($p=.313$). In the older groups, namely 5-year-olds, the mean score reported in the control group (G3) was 11.79, while for the experimental group (G4) it was 12.55, such difference being not statistically significant either ($p=.166$).

Therefore, slightly higher results were seen in the CLIL groups in both grades in comparison with their non-CLIL same age peers, in the basic curriculum words test, although the differences were not statistically significant in either grade. It is also important to note that the CLIL groups were exposed to a higher amount of vocabulary than their non-CLIL peers, within the same time period. Therefore, it can be concluded, that increasing the amount of vocabulary introduced in the EFL curriculum through soft-CLIL does not hinder the learning, nor causes the children to be overwhelmed, since results showed positive tendencies in favor of the CLIL groups, although not statistically significant.

Table 5. Group Statistics in Raw Scores for Experimental Groups (G2 - G4) in Basic Curriculum and Complete (Basic + Extra) Receptive Tests

Group	N	Test	Mean	Std. Deviation	Std. Error Mean	Gains
G2	34	Basic Test	7.62	1.826	.313	+ 5.32
		Complete Test	12.94	2.741	.470	
G4	44	Basic Test	12.55	2.406	.363	+5.63
		Complete Test	18.18	3.178	.561	

Table 6. Group Statistics in Percentages for Experimental Groups (G2 - G4) in Basic Curriculum and Complete (Basic + Extra) Receptive Tests

Group	N	Test	Mean	Std. Deviation	Std. Error Mean
G2	34	Basic Test	69.2524%	16.60125	2.84709
		Complete Test	62.7059%	13.70364	2.35015
G4	44	Basic Test	73.7973%	14.15306	2.13365
		Complete Test	67.3398%	13.77243	2.07627

Additionally, it is worth mentioning that both CLIL groups (G2 and G4), were able to acquire not only the same amount of basic curriculum words as their non-CLIL peers, but also a similar percentage of the additional less frequent words presented though soft-CLIL, as can be seen in Table 5 and Table 6. The younger learners following the soft-CLIL intervention (G2) showed a mean raw score of 5.32 points higher when tested in the complete test, being their mean score for the basic test a 7.62, and their mean score for the complete test a 12.94. In terms of the older experimental group's results (G4), their mean raw scores were 5.63 points higher when tested in the complete test, as they obtained a mean score of 12.55 in the basic test, while in the complete test their mean score was of 18.18.

Table 6 shows the percentage of correct answers of the learners in the experimental groups both in the basic curriculum words test, and in the complete test, which included the basic curriculum target words and the less frequent ones introduced though soft-CLIL. Percentages of receptive vocabulary acquisition are very similar in both tests within each grade, although slightly higher in the basic test, which may be due to the fact that the additional words in the complete test were more difficult and less frequent.

3.2 Age Effects

The second research question inquired into the age effects, and it was expected that older learners (5-year-olds) would benefit more from the soft-CLIL program, when compared to their younger peers (4-year-olds). In order to analyze these differences, the results corresponding only to the basic words that were presented to all groups of both grades were analyzed.

Table 7. Poisson Regression for Age Effects

Receptive Test – Grade	% Correct answers	Prevalence ratio	CI 95%	P-value
P4 (4 years old / G1 & G2)	606/880 (68.86%)			
P5 (5 years old / G3 & G4)	777/1045(74.43%)	1.093	(0.98; 1.22)	$p=.101$

As reported in Table 7, when considering the number of words that learners of each grade were able to identify correctly from the basic words that were worked on in all grades and groups, older students (G3 and G4) could correctly identify 777 words out of 1045, while their younger peers (G1 and G2) could identify 606 out of 880 words correctly. When analyzing the percentage of correctly identified words, younger learners' results were 68.86%, while older learners' results were 74.43%, a difference that was not statistically significant ($p = .101$). Nevertheless, it is important to notice that there is a prevalence ratio of 1.093, indicating that there was some effect of age, although not statistically significant.

3.3 Word Frequency Effects

The last research question focused on analyzing the effects of word frequency in receptive vocabulary acquisition. It was hypothesized that high-frequency words would be recalled more easily, and, thus, a higher amount of high-frequency words would be identified correctly in the receptive test, when compared to the low-frequency ones.

Table 8. Descriptive Statistics (High- and Low-Frequency)

Test	Frequency	Words identified	Percentage	P-value
Receptive	High	1195 / 1680	71.17%	$p = .0001$
	Low	996 / 1636	60.88%	

Table 9. Poisson Regression for Frequency Effects

Receptive Test – Frequency	Prevalence ratio	CI 95%	P-value
High- vs Low-Frequency	1.235	(1.01; 1.50)	$p = .036$

As shown in Table 8, a higher amount of high-frequency words was correctly identified in the receptive vocabulary test, that is 1195 out of 1680 high-frequency words were identified correctly, while results were lower for the low-frequency words, of which 996 out of 1636 were identified correctly. Results in the same line were observed after data was normalized into percentages: 71.17% of high-frequency words were identified correctly, a higher result than the one obtained in the low-frequency words, 60.88%. Such difference is statistically significant ($p = .0001$), indicating a frequency effect. These results are in accordance with the Poisson Regression, in Table 9, which shows a 1.519 prevalence ratio in favor of the high-frequency words, being such difference statistically significant as well ($p = .036$).

In sum, it can be concluded that frequency appears to have an effect on FL receptive vocabulary learning, as chances to correctly recall and identify a word are significantly higher if that word is of high-frequency. Nevertheless, opposite results were seen in terms of age effects, in which no statistically significant differences were found when comparing the results of 4- and 5-year-olds.

4. Discussion

The present research sought to study whether a soft-CLIL program enhances vocabulary learning in very young pre-primary learners of English as a FL. With this objective in mind, receptive vocabulary acquisition of target words in very young EFL learners, age differences and word frequency effects have been analyzed.

The first research question inquired on the effects of such soft-CLIL program on the acquisition of basic curriculum target words, focusing on receptive vocabulary knowledge. It was expected that students in the groups following the soft-CLIL program, as compared to their same age non-CLIL peers, would be able to acquire a greater range of vocabulary, benefiting from a more contextualized learning (Lasagabaster & Sierra, 2009; Pérez-Vidal, 2011), with higher input quality, quantity, and more opportunities for interaction (Dalton-Puffer, 2008; Lasagabaster & Sierra, 2009).

Results in the present study showed that, when tested on the basic curriculum words taught over the same period of time, students in the CLIL groups had slightly higher results than their same grade non-CLIL peers, although differences were not statistically significant. Therefore, such positive tendencies in the CLIL groups, which could result in significant results over a longer treatment period, may be attributed to the increased quality and quantity of vocabulary and input that very young EFL learners are exposed to in EFL classes and to the nature of the soft-CLIL context. It is also important to note that the CLIL groups that showed positive tendencies in the basic vocabulary test results, when tested in the complete test, which included some additional words of lower-frequency that were introduced through soft-CLIL in the classroom, showed even higher results. Thus, learners following the soft-CLIL program were not overwhelmed by the introduction of a greater amount of vocabulary in the EFL classroom within the same time period. On the contrary, they managed to acquire not only

a similar amount of basic curriculum words as their non-CLIL peers, but also a similar number of the additional CLIL words. Such results lead to think that when young learners in a FL context are presented with a higher amount of target vocabulary through a natural-like approach such as soft-CLIL, they are able to reap greater benefits and acquire a larger amount of vocabulary, without being overwhelmed.

These results are in line with previous research conducted with older students, namely secondary (Artieda, Roquet, & Nicolás-Conesa, 2020; Lasagabaster, 2008), and primary education students (Jiménez-Catalán, Ruiz de Zarobe & Cenoz, 2006) following CLIL programs, in which higher results were reported in terms of vocabulary acquisition in the FI+CLIL groups, than in the FI only groups. Pérez Cañado (2018), and Agustín-Llach & Canga Alonso (2014) also reported an advantage both in primary and secondary school students' vocabulary learning through CLIL.

Therefore, CLIL has been found to provide a beneficial learning context that favors a greater development of many skills, such as vocabulary amongst others (Dalton-Puffer, 2008; Lasagabaster, 2008; Lorenzo, Casal & Moore, 2010; Pérez-Vidal, 2011). The higher results reported in FI+CLIL groups in many studies have been attributed to this context that allows for more quality and quantity of input, as well as more interaction opportunities (Dalton-Puffer, 2008; Lasagabaster & Sierra, 2009). Nevertheless, it has to be kept in mind that in many CLIL studies, the groups following the FI+CLIL program received a higher amount of hours of English instruction, since in most cases the CLIL hours were added to the FI hours. Thus, as pointed out by Agustín-Llach & Canga Alonso (2014), higher results in the FI+CLIL groups may not only be due to the rich environment that CLIL provides, but also to the increased exposure to the FL. That is however not the case in the present study, in which hours of FL instruction were kept the same in all groups, since soft-CLIL was embedded within the regular FL sessions.

The second research question aimed at analyzing whether older students, namely 5-year-olds, were able to benefit more from the soft-CLIL program than their younger peers, 4-year-olds, and thus, it was hypothesized that older students would be able to recall a higher number of words than the younger group. Nevertheless, such hypothesis cannot be fully confirmed, since there were no significant differences between the percentage of words that students in P4 and in P5 were able to identify correctly. However, some tendencies of higher results can be seen in the older learners' results when it comes to the basic words, which over a longer treatment period could eventually lead to more significant results. This is an aspect to be investigated in further research.

It has traditionally been believed that the earlier a FL is learnt, the better the chances are to reach native-like level, and, as a result, age of onset of FL learning in schools has been brought down. Nevertheless, it has been claimed that an earlier age of onset should be accompanied by an immersion-like context and great amounts of FL exposure (Miralpeix, 2007; Muñoz, 2008), in which younger learners can learn the FL implicitly through language use (Muñoz, 2006). In the present study a pedagogical intervention was designed through a soft-CLIL program to bring to the FL classroom more meaningful and contextualized input and interaction opportunities. As reported in the results, young learners were able to acquire a great amount of target words through the soft-CLIL approach, as compared to their FI peers. Slightly higher results, although not statistically significant, were reported in the older learners' groups, as compared to their one-year younger peers. However, the lack of significant results here may indicate that at such young ages, a one-year difference may not result in significant cognitive differences without a real immersion context or massive FL contact. Therefore, this is also an issue that should be further investigated within the FL context.

The third research question examined the effects of word frequency, expecting high-frequency words to be easier to recall than low-frequency ones. Target words belonging to the first three 1000 bands of the BNC and NGSL frequency lists were considered high-frequency, while words below the third band were considered low-frequency.

An effect of word frequency was found in the present research, since a bigger percentage of high-frequency words was recalled by students, with a statistically significant difference and a significant prevalence ratio. Such results are in line with previous research, in which pre-school students (Shaban, 2013), primary school students (Alexiou, 2015) and high school students (Miralpeix, 2007) were also able to learn faster and recall a larger amount of high-frequency words. Therefore, it can be claimed that there is a significant effect of word frequency, in which higher-frequency words are recalled easier than lower-frequency words. This may be due to the fact that high-frequency words are more concrete (Shaban, 2013), more present in our daily lives (Alexiou, 2015), and cover most of the basic interactions (Nation, 2020; Nation & Meara, 2020).

5. Conclusion

The present study has examined the effects of a soft-CLIL program in receptive vocabulary learning in pre-primary students learning English as a FL. In light of the results abovementioned, which showed positive tendencies in the CLIL groups, although no statistically significant differences, it could be stated that such a soft-CLIL program provides a richer context for FL vocabulary learning, with higher quality and quantity input, as well as more interaction opportunities, which could lead to significant results with a longer treatment period. This is the reason why, as stated in the previous section, we firmly believe that further research is needed to study the evolution of the tendencies reported in the current research, with longer soft-CLIL programs, since CLIL research focusing on very young EFL learners is still very scarce.

Additionally, no significant differences were found between 4- and 5-year-olds when analyzing vocabulary acquisition. The lack of significant results may be attributed to the small difference in age between the two grades, indicating that a one-year age gap may not be relevant at such young ages in a FL context with very limited FL exposure. Nevertheless, it is unknown whether some differences in favor of the group that started CLIL one year earlier would be found if those same students were followed over a longer period of time. Therefore, more longitudinal studies with very young EFL learners are needed to see if a one-year difference in the age of onset of CLIL programs is significant in the long term.

The current research has designed and implemented a soft-CLIL program for very young EFL learners and has found positive tendencies in pre-primary students' receptive vocabulary development in English as a FL. Nevertheless, early childhood CLIL programs are to date still very under researched, and more longitudinal research is needed to shed light on the possible benefits of CLIL at very young ages.

Acknowledgments

This research was supported by the project "Content and Language Integrated Learning in Pre-primary Education" Funded by Fundació Escoles Diocesanes I Parroquials Del Bisbat De Terrassa, 2018-2021. Marta Segura is a Universitat Internacional de Catalunya PhD grant student at the Institute for Multilingualism.

References

- Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: student's language proficiency. *English Educational Research and Evaluation*, 12(1), 75-93. <https://doi.org/10.1080/13803610500392160>
- Agustín-Llach, M. P., & Canga Alonso, A. (2014). Vocabulary growth in young CLIL and traditional EFL learners: evidence from research and implications for education. *International Journal of Applied Linguistics*, 26(2), 211-227. <https://doi.org/10.1111/ijal.12090>
- Albaladejo, S., Coyle, Y., & de Larios, J. R. (2018). Songs, stories, and vocabulary acquisition in preschool learners of English as a foreign language. *System*, 76, 116-128. <https://doi.org/10.1016/j.system.2018.05.002>
- Alexiou, T. (2015). Vocabulary uptake from Peppa Pig: a case study of preschool EFL learners in Greece. In C. Gitsaki & T. Alexiou (Eds.), *Current issues in second/foreign language teaching and teacher development: research and practice* (pp. 285-301). Cambridge Scholars Publishing.
- Artieda G., Roquet, H., & Nicolás-Conesa, F. (2020). The impact of age and exposure on EFL achievement in two learning contexts: formal instruction and formal instruction + content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 23(4), 1-19. <https://doi.org/10.1080/13670050.2017.1373059>
- Asensio Arjona, V. (2020). CLIL program designed to teach music and English in preschool education. In R. M. Vicente Alvarez, C. Guillaenders, J. Rodríguez, G. Romanelli & J. Pitt (Eds.), *Music education and didactic materials* (pp. 247-256). Andavira.
- Canga Alonso, A. (2015). Receptive vocabulary of CLIL and non-CLIL primary and secondary school learners. *Complutense Journal of English Studies*, 23, 57-77. https://doi.org/10.5209/rev_CJES.2015.v23.51301
- Dale, L., & Tanner, R. (2012). *CLIL activities. A resource for subject and language teaching*. Cambridge University Press.
- Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future perspectives for English language teaching* (pp. 139-157). Carl Winter.

- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-533. <https://doi.org/10.1017/S0272263100004022>
- García Esteban, S. (2015). Soft CLIL in infant education bilingual contexts in Spain. *International Journal of Language and Applied Linguistics*, 1, 30-36.
- Jiménez-Catalán, R. M., Ruiz de Zarobe, Y., & Cenoz, J. (2006). Vocabulary profiles of English foreign language learners in English as a subject and as a vehicular language. *Views*, 15(3), 23-26. <https://doi.org/10.4067/S0718-48832016000100003>
- Krashen, S., Long, M., & Scarcella, R. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly*, 13(4), 573-582. <https://doi.org/10.2307/3586451>
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1, 31-42. <https://doi.org/10.2174/1874913500801010030>
- Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and international EFL classes. *International CLIL Research Journal*, 1(2), 4-17.
- Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.
- Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics*, 31(3), 418-442. <https://doi.org/10.1093/applin/amp041>
- Mair, O. (2018). Four seasons of CLIL at two Italian pre-schools. In F. Costa, C. Cucchi, O. Mair & A. Murphy (Eds.), *English for young learners from pre-school to lower secondary: a CLIL teacher training project in Italian schools* (pp. 21-44). Universitas Studiorum.
- Meara, P. M. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjagr & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp.35-53). Cambridge University Press.
- Meara, P. M., & Miralpeix, I. (2021). Bayesian vocabulary tests. *Vigo International Journal of Applied Linguistics*, 18, 177-204. <https://doi.org/10.35869/vial.v0i18.3370>
- Miralpeix, I. (2006). Age and vocabulary acquisition in English as a foreign language (EFL). In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (Vol. 1) (pp. 89-106). Multilingual Matters. <https://doi.org/10.21832/9781853598937-006>
- Miralpeix, I. (2007). Lexical knowledge in instructed language learning: the effects of age and exposure. *International Journal of English Studies*, 7(2), 61-83.
- Miralpeix, I. (2020). L1 and L2 vocabulary size and growth. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 189-206). Routledge: Routledge Handbooks.
- Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *IRAL - International Review of Applied Linguistics in Language Teaching*, 56(1), 1-24. <https://doi.org/10.1515/iral-2017-0016>
- Muñoz, C. (2006). The effects of age on foreign language learning; the BAF Project. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (Vol. 1) (pp. 6-40). Multilingual Matters. <https://doi.org/10.21832/9781853598937-003>
- Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 29(4), 578-596. <https://doi.org/10.1093/applin/amm056>
- Muñoz, C., & Singleton, D. (2011). A critical review of age-related on L2 ultimate attainment. *Language Teaching*, 44(1), 1-35. <https://doi.org/10.1017/S0261444810000327>
- Muñoz, C., Tragant, E., & Torras, M. R. (2010). Los inicios de la producción oral en lengua extranjera de alumnos de primaria en Europa. *ELIA. Estudios de Lingüística Inglesa Aplicada*, 10, 11-39.
- Nation, I. S. P. (2020). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, I. S. P., & Meara, P. (2020). Vocabulary. In N. Schmitt & M. P. H. Rodgers (Eds.), *An introduction to applied linguistics* (pp. 35-54). Routledge.
- Ortega, L. (2015). Researching CLIL and TBLT interfaces. *System*, 54, 103-109. <https://doi.org/10.1016/j.system.2015.09.002>

- Pavón, V. (2014). Enhancing the quality of CLIL: making the best of the collaboration between language teachers and content teachers. *Encuentro*, 23(1), 115-127.
- Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton University Press. <https://doi.org/10.1515/9781400854677>
- Pérez Cañado, M. L. (2018). CLIL and educational levels: a longitudinal study on the impact of CLIL on language outcomes. *Porta Linguarium*, 29, 51-70. <https://doi.org/10.30827/Digibug.54022>
- Pérez-Vidal, C. (2011). Language acquisition in three different contexts of learning: Formal instruction, Study Abroad and Semi-immersion (CLIL). In Y. Ruiz de Zarobe, J. M. Sierra & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning: contributions to multilingualism in European contexts* (pp. 25-35). Peter Lang. <https://doi.org/10.3726/978-3-0351-0171-3>
- Pérez-Vidal, C. (2013). Perspectives and lessons from the challenge of CLIL experiences. In C. Abello-Contesse, P. M. Chandler, M. D. López-Jiménez & R. Chacón-Beltrán (Eds.), *Bilingual and multilingual education in the 21st century building on experience* (pp. 59-82). Multilingual Matters. <https://doi.org/10.21832/9781783090716>
- Pérez-Vidal, C., Escobar, C., & Roquet, H. (2013). Perspectives sobre l'aprenentatge de llengües i multilingüisme [Perspectives on the learning of languages and multilingualism]. *Butlletí Informatiu del Col·legi de Doctors i Llicenciats en Filosofia i Lletres i en Ciències de Catalunya*, 14, 4-5.
- Pérez-Vidal, C., Lorenzo, N., & Trenchs, M. (2015). Les llengües en l'educació: el plurilingüisme i la internacionalització [Languages in education: plurilingualism and internationalization]. In J. M. Vilalta (Ed.), *Reptes de l'educació a Catalunya. Anuari 2015* (pp. 139-197). Fundació Jaume Bofill.
- Roquet, H., & Pérez-Vidal, C. (2017). Do productive skills improve in content and language integrated learning (CLIL) contexts? The case of writing. *Applied Linguistics*, 38(4), 489-511. <https://doi.org/10.1093/applin/amv050>
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108569057>
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363. <https://doi.org/10.1177/1362168808089921>
- Shaban, M. (2013). Natural order of vocabulary acquisition. *European Academic Research*, 1(5), 721-733.
- Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225-239). Routledge: Routledge Handbooks.
- Yongqi Gu, P. (2020). Strategies for learning vocabulary. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 271-287). Routledge: Routledge Handbooks.

Notes

Note 1. In the Spanish education system pre-primary education is made up by three levels: P3 (3 years old), P4 (4 years old) and P5 (5 years old), which are equivalent to Foundation 3, 4 and 5.

Note 2. Dunn, L. M. & Dunn, M. (2007). *PPVT4 Peabody Picture Vocabulary Test, Fourth Edition. Manual*. Pearson: United States of America.

Note 3. Williams, K. T. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Service.

Appendix A**Vocabulary List per Unit and Group***Autumn and Halloween Unit*

Grade	Basic vocabulary for control and experimental groups		Experimental group added vocabulary	
P4	Orange	Cat		
	Brown	Rainy	Leaf / leaves	Acorn
	Yellow	Windy	Worm	Witch
	Green	Halloween	Snail	Pumpkin
	Red	Ghost	Owl	Night
	Tree	Moon		
	Squirrel	Bat		
P5	Leaf / leaves	Cat		
	Orange	Rainy	White	Mushroom
	Brown	Windy	Black	(Water) boots
	Yellow	Halloween	Chestnut	Raincoat
	Green	Ghost	Worm	Stormy
	Red	Moon	Owl	Witch
	Tree	Bat	Acorn	Night
	Snail	Pumpkin		
	Squirrel	Skeleton		

Winter and Christmas Unit

Grade	Basic vocabulary for control and experimental groups		Experimental group added vocabulary	
P4	Snow	Cold		
	Snowman	Frozen		
	Buttons	Snowy		
	Snowflake	Penguin	Carrot	Boots
	Jacket	Polar bear	Hot chocolate	Reindeer
	Gloves	Jesus	Snowball	Star
	Scarf	Three wise men	Coat	
	Hat	Bethlehem		
	Presents	Angel		
	Christmas tree	Shepherd		
P5	Snow			
	Snowman	Polar bear		
	Buttons	Carrot		
	Snowflake	Hot chocolate		
	Jacket	Snowball		
	Gloves	Coat	To ice-skate	Sledge
	Scarf	Boots	Footprints	Fairy lights
	Hat	Reindeer	Snow angel	Chimney
	Presents	Star	Snowball fight	Candy stick
	Christmas tree	Jesus		
	Cold	Three wise men		
	Frozen	Bethlehem		
	Snowy	Angel		
	Penguin	Shepherd		

Appendix B

Sociolinguistic Background Questionnaire

Dear parents,

We are carrying out a research project together with the Universitat Internacional de Catalunya and the objective of the study is to investigate the acquisition of English as a foreign language in pre-primary education. Considering that vocabulary acquisition is one of the most important aspects of language learning during the first stages, the goal of the present research project is to promote vocabulary learning in English during the early years of education.

We ask you to fill in the present questionnaire, so that we can learn about the students' language learning background and their exposure to English outside of school. Answering the questionnaire will take around 10 minutes. If you currently have more than one child in pre-school in this school, please answer this questionnaire once for each child.

All information given here will remain confidential. The data resulting from your child's participation may be available to other researchers in the future for research purposes not detailed within this consent form. In these cases, the data will contain no identifying information that could associate it with your child, or with your child's participation in any study.

If you decide to let your child participate in this project, this form will be used to record your permission. Your child's participation in this study is voluntary and you can always withdraw from the study at any stage.

If you have any questions regarding the research project, please contact your child's English teacher or tutor.

Thank you for your participation.

- Please write your child's name, grade and class: _____

- Do you consent to your child participating in this research project? ☐ Yes, I do. ☐ No, I don't.

Part 1. General information about the student

- Gender: ☐ Male ☐ Female

- Does your child have any hearing problems? ☐ Yes ☐ No ☐ Other: _____

- Date of birth of your child (day, month, year): _____

- Grade and class of your child: ☐ P4 – A ☐ P5 – A ☐ P4 – B ☐ P5 – B

- Where was your child born? _____

- If your child was not born in Catalonia, how long has he/she been living here for?

☐ 1 to 6 months ☐ 6 months to 1 year ☐ 1 to 2 years

☐ 2 to 3 years ☐ 3 to 4 years ☐ More than 4 years

- Has your child ever lived in an English-speaking country? ☐ Yes ☐ No

- If so, where, for how long and how old was he/she: _____

Part 2. Student's linguistic background

- What language(s) is the mother tongue of your child? ☐ Catalan ☐ Spanish ☐ Other: _____

- If in the previous question you have stated that your child has more than one mother tongue, which is the language that he/she uses more often? Which is his/her dominant language?

☐ Catalan ☐ Spanish ☐ Other: _____

- Does your child know any other foreign languages? ☐ Yes ☐ No

- If so, which ones? Write them in order of acquisition (start with the language that he/she learnt first, and end with the ones he/she is still learning): _____

- How often does your child use each of the languages he/she knows? Mark it with an (x) in the following table:

	Never	Rarely	Sometimes	Usually	Always
Catalan					
Spanish					
English					
Other	language: _____				

Part 3. Family background

About the student's siblings

- Does your child have any siblings? ☐ Yes ☐ No

Answer the following questions only if your child has any siblings. If he/she has no siblings, move on to the next section.

- If so, how old are they?

- ☐ Younger than 2 years old ☐ Between 2 and 5 years old ☐ Between 5 and 8 years old
☐ Between 8 and 12 years old ☐ Between 12 and 15 years old ☐ Older than 15 years old

- Do your child's siblings study English? ☐ Yes ☐ No

About the student's mother

If the child does not live with the mother or has no frequent contact with the mother, answer "Does not apply" to the questions in this section.

- Place of birth of the mother: ☐ Catalonia ☐ Does not apply ☐ Other: _____

- Educational level of the mother:

- ☐ Primary school ☐ Secondary school
☐ Professional training ☐ Undergraduate / Bachelors
☐ Masters / Postgraduate ☐ Doctorate / PhD
☐ Does not apply ☐ Other: _____

- Mother's occupation: _____

- What is the mother's mother tongue?

- ☐ Catalan ☐ Spanish ☐ Does not apply ☐ Other: _____

- What other foreign languages does the mother speak?

- ☐ Catalan ☐ Spanish ☐ English ☐ None ☐ Does not apply ☐ Other: _____

- What is the mother's overall English proficiency?

- ☐ Native level ☐ Advanced level ☐ High level ☐ Intermediate level
☐ Low level ☐ Level zero ☐ Does not apply

- How often does the mother use English?

- ☐ Every day ☐ Some times a week ☐ Once a week
☐ Some times a month ☐ Once a month ☐ Some times a year
☐ Never ☐ Does not apply

About the student's father

If the child does not live with the father or has no frequent contact with the father, answer "Does not apply" to the questions in this section.

- Place of birth of the father: ☐ Catalonia ☐ Does not apply ☐ Other: _____
- Educational level of the father:
- ☐ Primary school ☐ Secondary school
- ☐ Professional training ☐ Undergraduate / Bachelors
- ☐ Masters / Postgraduate ☐ Doctorate / PhD
- ☐ Does not apply ☐ Other: _____
- Father's occupation: _____
- What is the father's mother tongue?
- ☐ Catalan ☐ Spanish ☐ Does not apply ☐ Other: _____
- What other foreign languages does the father speak?
- ☐ Catalan ☐ Spanish ☐ English ☐ None ☐ Does not apply ☐ Other: _____
- What is the father's overall English proficiency?
- ☐ Native level ☐ Advanced level ☐ High level ☐ Intermediate level
- ☐ Low level ☐ Level zero ☐ Does not apply
- How often does the father use English?
- ☐ Every day ☐ Some times a week ☐ Once a week
- ☐ Some times a month ☐ Once a month ☐ Some times a year
- ☐ Never ☐ Does not apply

About the student's grandparents

If the child has no contact with the maternal or paternal grandparents, answer "Does not apply" to the questions below.

- What is the mother tongue of the child's maternal grandparents?
- ☐ Catalan ☐ Spanish ☐ Does not apply ☐ Other: _____
- What is the mother tongue of the child's paternal grandparents?
- ☐ Catalan ☐ Spanish ☐ Does not apply ☐ Other: _____

About the languages the student speaks with his/her family

- For the family members that appear in the following table, indicate (1) if they live with your child and (2) which language they use to speak with your child. Answer only the rows about people who are usually in contact with your child.

	Lives with the child?		Which language do they use to communicate?			
	Lives with the child	Does not live with the child	Speaks mostly Catalan with the child	Speaks mostly Spanish with the child	Speaks Catalan and Spanish with the child	Speaks another language with the child
Mother						
Father						
Mother's partner						
Father's partner						
Siblings						
Maternal grandmother						
Maternal grandfather						
Paternal grandmother						
Paternal grandfather						
Other						

- If in the previous questions you have selected "speaks another language with the child", write the person and the language he/she uses to communicate with the child. _____

Part 4. Student's contact with English

- When did your child start studying English in the school? ☐ P3 ☐ P4 ☐ P5

- Does your child attend English extracurricular classes? ☐ Yes ☐ No

- If your child attends English extracurricular activities, at what age did he/she start?

☐ Younger than 2 years old ☐ Between 2 and 3 years old

☐ Between 3 and 4 years old ☐ Between 4 and 5 years old

- Indicate in the following table how much time per week (outside of school) is your child in contact with English doing each of the following activities:

	Never	Less than 30 min / week	30 min to 1 hour / week	1 to 2 hours / week	2 to 3 hours / week	3 to 4 hours / week	4 to 5 hours / week	More than 5 hours / week
Watching television, movies or cartoons								
Watching Youtube or other videos								
Listening to songs								
Reading books								
Extracurricular or private lessons								
Other extracurricular activities								
Social interaction (with parents, other family members, friends, babysitters...)								
Interaction with a native speaker of English								
Playing games or video games								

- If your child participates in other extracurricular activities in English, name the activities: _____

Part 5. Opinions about learning English

Child's involvement in learning English

- Your child tells you about the English lesson. ☐ Yes ☐ No

- Your child practices English at home. ☐ Yes ☐ No

Child's attitudes towards learning English

- Your child is keen to learn English. ☐ Yes ☐ No

- Your child likes to talk in English. ☐ Yes ☐ No

Parents opinions about their child's learning English

- Early English learning is useful for children.

☐ Strongly agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly disagree

- Early English learning helps achieving a higher proficiency in later life.

☐ Strongly agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly disagree

- The earlier education in English starts, the better.

☐ Strongly agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly disagree

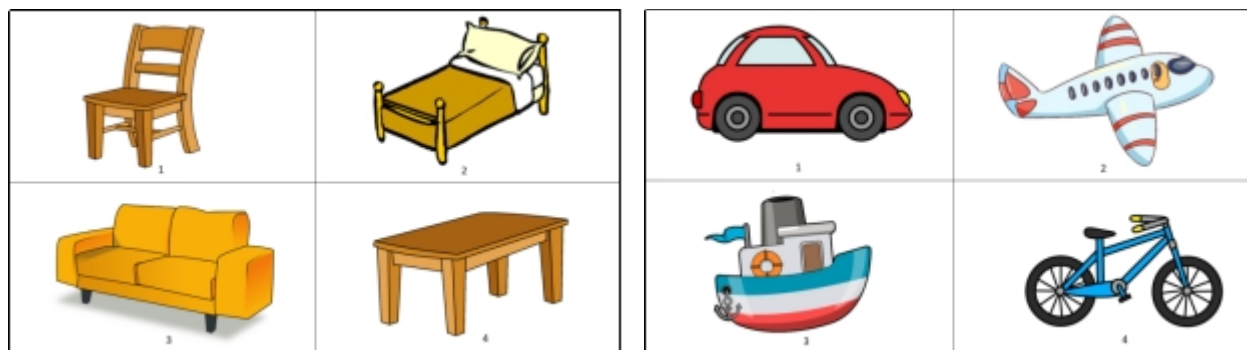
Thank you

This is the end of the questionnaire. Thank you very much for your contribution to our research project.

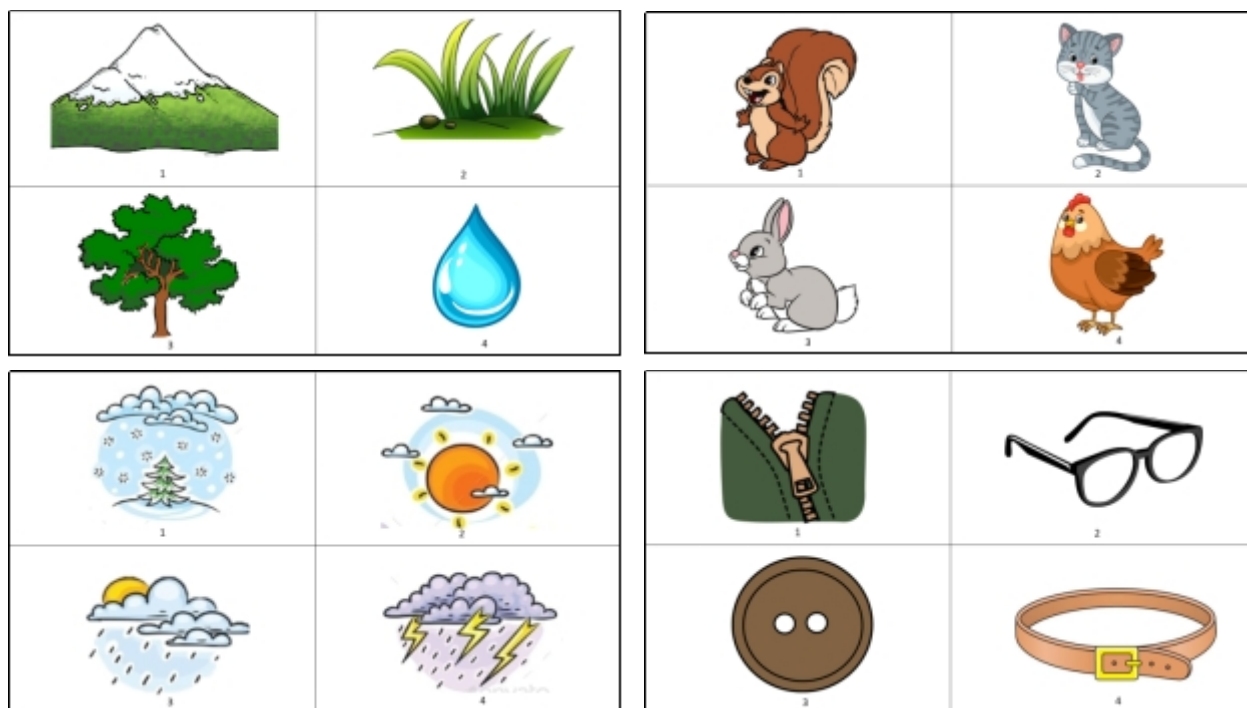
Appendix C

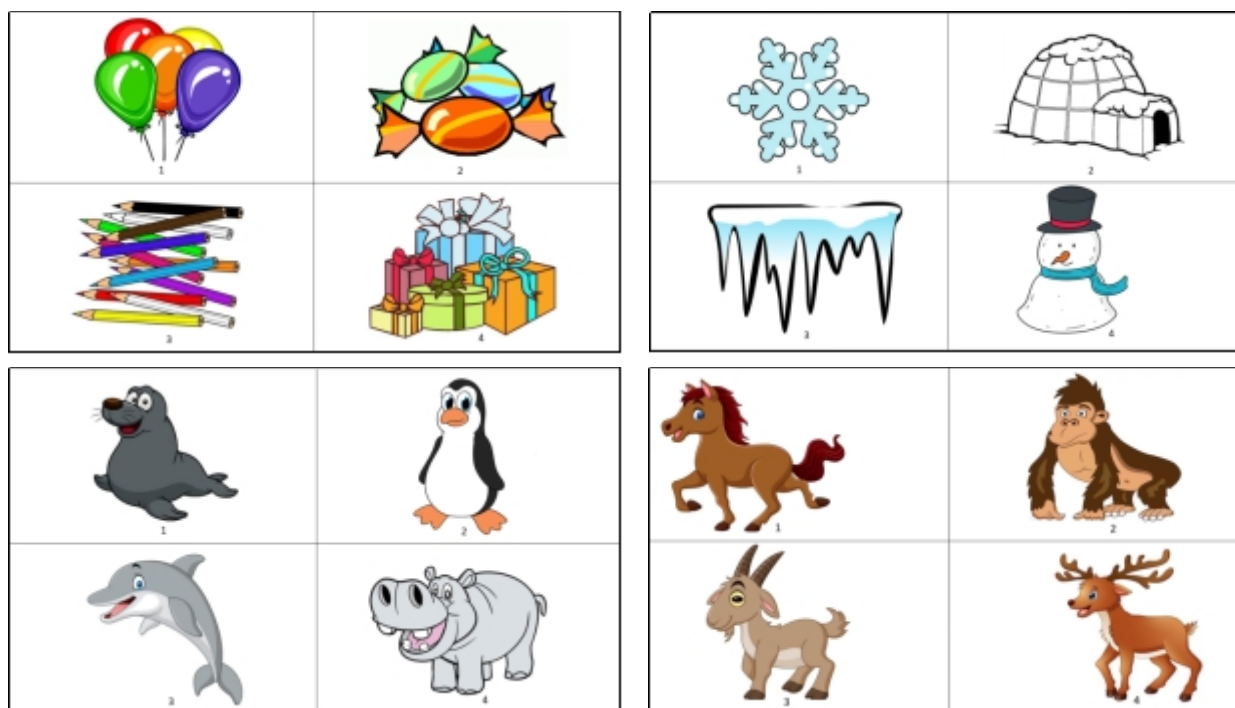
Sample of Basic Curriculum Words Receptive Test

Receptive Vocabulary Test Training Items



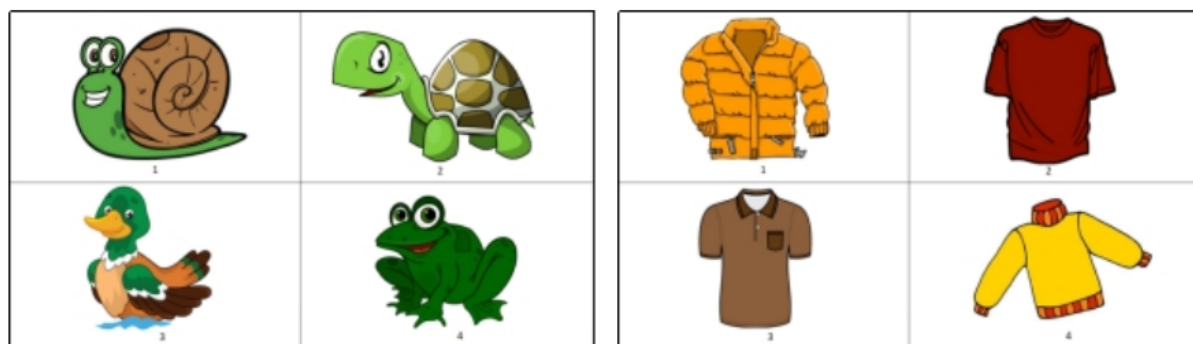
P4 Basic Curriculum Receptive Vocabulary Test





P5 Basic Curriculum Receptive Vocabulary Test (Additional to the P4 Test)





Appendix D

Sample of Additional Target Words Receptive Test

P4 Added Target Words Receptive Vocabulary Test



P5 Added Target Words Receptive Vocabulary Test (Additional to the P4 Test)**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).