Research Article

Marta Segura*, Júlia Barón, and Helena Roquet

Productive vocabulary learning in pre-primary education through soft CLIL

https://doi.org/10.1515/opli-2022-0194
received December 23, 2021; accepted May 30, 2022

Abstract: The present study aims to contribute to the field of Foreign Language (FL) acquisition by analysing English as a Foreign Language (EFL) vocabulary learning in pre-primary education learners following a Content and Language Integrated Learning (CLIL) programme. Although the present study focuses on productive vocabulary acquisition, such results are later compared to receptive vocabulary findings reported in a previous study following the same learners over the same period of time. Additionally, word frequency effects are studied. A 6-month longitudinal study was conducted with Catalan and Spanish bilingual EFL learners (N = 155) aged 4 and 5. Through the programme, two curricular units traditionally taught in the learners' mother tongue were worked on through a soft CLIL approach. Learners were administered a general vocabulary pre-test, and by the end of the intervention, receptive and productive vocabulary tests were given to the students, including the target words presented in the soft CLIL sessions. Positive trends were found in productive vocabulary learning, which may eventually turn into significant differences over a longer treatment period. A significant frequency effect was observed, as there was a higher recollection rate of higher-frequency words. When comparing receptive and productive vocabulary results, statistically significant higher scores were reported in receptive tests than in productive ones.

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

1 Introduction

Within the last few decades promoting a plurilingual education has become a priority and, thus, the European Union has promoted an early start of Foreign Language (FL) teaching in schools (Pérez-Vidal et al. 2013), as well as the introduction of teaching approaches that provide a more contextualized learning environment, such as Content and Language Integrated Learning (CLIL) (Artieda et al. 2017, Lasagabaster and Sierra 2009). Some research has already been conducted to analyse the effects of CLIL programmes of various characteristics on different personal and linguistic skills, but most of the studies conducted so far have focused on CLIL in primary and secondary levels, while research with younger pre-primary students is still very scarce (Albaladejo et al. 2018, Asensio Arjona 2020, García Esteban 2015, Mair 2018, Segura et al. 2021). Thus, the present article aims to contribute to this line of research, by analysing the effects of a soft CLIL programme on pre-primary students' learning of vocabulary.

*** Corresponding author: Marta Segura,** Department of Applied Linguistics, Institute for Multilingualism, Universitat Internacional de Catalunya, 08195, Barcelona, Catalunya, Spain, e-mail: mseguram@uic.cat
Júlia Barón: Department of Modern Languages and English Studies, Philology and Communication Faculty, Universitat de Barcelona, Catalunya, Spain
Helena Roquet: Department of Applied Linguistics, Institute for Multilingualism, Universitat Internacional de Catalunya, Barcelona, Catalunya, Spain
ORCID: Marta Segura 0000-0003-0664-0415

Open Access. © 2022 Marta Segura *et al.*, published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

1.1 Learning contexts and onset age of FL learning

One of the educational changes that has been promoted by the European Union within the last few years has been to bring down the age of onset of FL teaching in schools (Muñoz et al. 2010). Such a measure has been based on the traditional belief that the earlier a FL is learnt, the better the chances are of reaching nativelike competence. Although the Critical Period Hypothesis, first proposed by Penfield (Penfield and Roberts 1959) and later popularized by Lenneberg (Lenneberg 1967), seems to support such a statement, this cannot be generalized to all language learning contexts (Miralpeix 2006, Muñoz et al. 2010).

It has been found that younger learners who have been exposed to a naturalistic Second Language (L2) learning environment perform better than those who start later (Krashen et al. 1979), since young children learn implicitly from the big quantity of input they are surrounded with (DeKeyser 2000). However, that seems to be true only in L2 immersion contexts, with high quality and quantity of input, as well as plenty of interaction opportunities outside of the school. On the contrary, other studies have reported that in instructional FL contexts older learners have an advantage over the younger ones (Artieda et al. 2017, Cenoz 2002, Muñoz et al. 2010), since their greater cognitive development allows them to make a more effective use of explicit learning mechanisms (Muñoz 2006).

Thus, bringing down the onset age of FL teaching in schools does not seem to be a guarantee for a higher level of attainment, unless it is combined with massive contact with the FL, such as in immersion contexts (Muñoz 2008). However, immersion is not always an option, and it is within this framework that new teaching approaches have gained ground. That is the case of CLIL, which has aimed not only to increase the hours of contact with the FL within schools, but also to provide a more meaningful and contextualized FL learning environment.

1.2 CLIL

CLIL is "a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language" (Coyle et al. 2010, 1). It can take place at any educational level, although it is usually more commonly implemented in primary and secondary levels, and it can be of any intensity, namely it can involve any number of subjects (Cenoz 2015). In high-intensity CLIL programmes, often referred to as hard CLIL, around half or more of the subjects are taught through the target language, and content teachers are usually in charge of the CLIL sessions, while in lower-intensity programmes, namely soft CLIL, the language teachers do cross-curricular work by teaching both the FL and the content from other subjects within the FL sessions (Dale and Tanner 2012, García Esteban 2015). Despite the intensity of the programme, CLIL provides a contextualization for FL learning through the curricular content (Lasagabaster and Sierra 2009, Pérez-Vidal 2011), allowing students to "speak and think in an authentic, significant and relevant way in a L2" (Vallbona 2011, 152). Consequently, CLIL sessions enhance a multilingual and multicultural context, where both the language and the curricular content share the focus of attention (Dale and Tanner 2012, Ortega 2015), which in turn reduces learners' anxiety towards FL learning (Pérez-Vidal 2011).

When compared with their same-age non-CLIL peers, FL learners who enrolled in CLIL programmes have shown higher results in some linguistic skills, such as listening, reading, vocabulary, morphology (Dalton-Puffer 2008), and fluency (Pérez-Vidal 2011), as well as in some personal skills, such as creativity (Dalton-Puffer 2008, Lasagabaster 2008), and motivation (Lasagabaster and Sierra 2009). Nevertheless, some contradictory results have been reported in the areas of speaking, writing, syntax, informal language, pronunciation, and pragmatics (Dalton-Puffer 2008, Pérez-Vidal 2011).

Vocabulary acquisition has been a topic of interest within CLIL research, and some studies have reported higher results in CLIL learners' vocabulary development than in their non-CLIL peers. That is the case of Pérez Cañado (2018), who reported higher results in all linguistic skills, including vocabulary, in both primary and secondary education CLIL learners than in non-CLIL students. In a previous study,

Jiménez-Catalán et al. (2006) used vocabulary as a measuring tool to analyse reading and writing development in primary students, and higher results were found in the CLIL groups than in the Formal Instruction (FI) groups in both cases. Such findings are in line with those reported in Lasagabaster (2008), where an advantage was found in bilingual CLIL students in secondary education, when compared with their non-CLIL peers, in all linguistic skills, as well as in grammar and vocabulary.

In another study, Merikivi and Pietila (2014) analysed receptive and productive vocabulary sizes of sixth-grade (11- and 12-year-olds) CLIL and FI learners, and reported two relevant findings: the first, that CLIL students had bigger receptive and productive vocabularies than their same-age non-CLIL peers; the second, that regardless of the type of the instruction received, receptive vocabularies were larger. Two other studies (Agustín-Llach and Canga Alonso 2014, Canga Alonso 2015) focused specifically on receptive vocabulary acquisition in primary and secondary education students and concluded that CLIL did enhance receptive vocabulary learning, compared with the traditional FI settings. Nevertheless, contradictory results were found in a 4-year longitudinal study by Admiraal et al. (2006), who analysed receptive vocabulary learning in secondary education students, and no significant differences were found between the groups following FI and the bilingual education programme. In terms of productive vocabulary learning through CLIL, research is still scarce, although Canga Alonso and Arribas Garcia (2015) reported that tenth-grade (15- and 16-year-olds) CLIL learners obtained significantly better results than their same-age non-CLIL peers in the Productive Vocabulary Levels Test. In a more recent study conducted with secondary education students, Reynaert (2019) pointed out that differences between the CLIL and non-CLIL groups in productive vocabulary learning did not appear until the second year of CLIL instruction, when the experimental group showed higher productive vocabulary levels. The author attributed such results to the fact that students following CLIL may have needed more time to adapt to the new classroom dynamics and that 1 year may not be enough for positive trends to become significant differences.

CLIL programmes seem to bring to the classrooms an immersion-like setting, by increasing the quality and quantity of FL contact and by generating a more natural learning context. Previous research has attributed the positive results found in CLIL students, when compared with their non-CLIL peers, to the rich environment that CLIL provides. Nevertheless, most of the abovementioned previous research has focused on primary and secondary level students, while there is still a research gap in terms of CLIL effects on younger FL learners, namely pre-primary education children, who are still acquiring their L1, and L2 in bilingual contexts.

It is also important to note that most of the studies have not taken into account the fact that CLIL programmes usually increase the time that the children are exposed to the FL within the school (Agustín-Llach and Canga Alonso 2014, Vallbona 2011), since CLIL hours are usually additional hours, which could also be another contributing factor enhancing language acquisition, besides the context provided. Therefore, further research is needed where time in contact with the FL is accounted for or where CLIL programmes are embedded within the regular FL classroom without increasing time of contact with the language.

1.3 FL vocabulary learning: Word dimensions, vocabulary size, and frequency

When learning any language, vocabulary acquisition is essential (Meara 1996, Schmitt 2008), although complex. Full knowledge of a lexical item involves mastery of all of its dimensions –first listed by Richards (1976), later completed and popularized by Nation (1990, 2013) – namely its spoken and written forms, word parts, meaning, concepts and referents, word associations, grammatical functions, collocations, and constraints on use (Nation 2013). Due to the complexity of having full command of a lexical item, it is possible that not even native speakers have completely internalized all aspects of all words, and partial word knowledge may be “the normal state for many words” (Schmitt and Schmitt 2020, 33).

Considering that to have a fully acquired lexical item a learner should know its dimensions, it is natural to see vocabulary learning as a process or a continuum, which usually starts with the receptive knowledge

to then grow into the productive one (Meara and Miralpeix 2021). Essentially receptive vocabulary involves meaning recognition and meaning recall, while productive vocabulary goes one step further and involves form recognition and form recall (Schmitt 2010). Thus, vocabulary knowledge may be partial during the FL learning process, as a language learner may be able to understand a word when heard or read but may need more time to be able to produce its oral and written forms (Meara and Miralpeix 2021, Schmitt and Schmitt 2020, Webb 2020). Consequently, it is expected for receptive knowledge to be larger than productive knowledge (Webb 2009), since the former seems to take less time (Waring 1997) and be easier to acquire (Griffin 1992). Such learning continuum has been reported in several studies (Ellis and Beaton 1993, Laufer 2005, Shintani 2011, Waring 1997), where language learners were tested on the same vocabulary receptively and productively, and higher results were found in receptive vocabulary tests, indicating that learners had at least partial knowledge of the words at the early stages of FL learning.

According to previous research (Laufer and Ravenhorst-Kalovski 2010, Schmitt and Schmitt 2020, Van Zeeland and Schmitt 2013), language users need to know a very large amount of lexical items to be able to fully and independently communicate in a wide variety of situations in the target language. Nevertheless, previous studies (Milton 2009, Nation and Waring 1997) have also reported that learning the most frequent 2,000–3,000 word families allows learners to communicate from early FL learning stages. It has been found that there is a relatively small amount of high-frequency lexical items, namely from the first three 1,000 bands in corpus frequency lists, such as the British National Corpus (BNC) and New General Service List (NGSL), which are extremely common, and which would allow language learners to function in basic conversations (Nation 2013, Nation and Meara 2020). Some studies have found that higher-frequency words are usually learnt faster and more easily than lower-frequency ones (Alexiou 2015). Such findings have been attributed to the fact that higher-frequency items are more concrete and typical (Shaban 2013), have more common referents, and are more present in our daily lives (Alexiou 2015). In another study, Albaladejo et al. (2018) also highlighted the importance of the number of encounters that one has with a word to increase the chances to recall it and eventually produce it. Such results support the fact that vocabulary presented in the FL classes should be carefully chosen, to ensure its relevance for FL learners (Wong Kwok Shing 2006) and to maximize the opportunities of interaction. Therefore, frequency of lexical items may be a good indicator when selecting which vocabulary to start teaching in FL classes, to provide students with higher-frequency items that can allow them to communicate in basic situations, even when their vocabulary size is still very limited.

Although FL vocabulary learning has been a largely studied area, more classroom-based research is needed, to identify the strengths and shortcomings of different teaching approaches. More specifically, a need for more studies with pre-primary education level students has been identified. It is not enough for schools to bring down the onset age of FL teaching, but special attention should be paid to the teaching approaches (Muñoz 2008), to bring to the classrooms of young English as a Foreign Language (EFL) learners a more contextualized and meaningful learning (Lasagabaster and Sierra 2009, Vallbona 2011).

Thus, the current article's goal is to contribute to this field of research, focusing on the learning of vocabulary in very young EFL learners, by analysing productive vocabulary acquisition after the implementation of a soft CLIL programme, as well as word frequency effects. Additionally, the results provided will be contrasted to those reported in Segura et al. 2021.[1] In their study, Segura et al. (2021) looked into receptive vocabulary learning in the same students, who were enrolled in the same programme. The following research questions (RQ) and hypotheses (H) have been set:

RQ1. Will a soft CLIL programme lead to a greater productive vocabulary learning in pre-primary EFL learners, as compared with formal EFL instruction within the same time period?

H1. Pre-primary learners following the soft CLIL programme will show higher results in terms of productive EFL vocabulary learning when compared with their non-CLIL counterparts.

RQ2. Does frequency of words have an effect on productive vocabulary learning?

H2. High-frequency words will be easier to recall than low-frequency ones and, thus, results for high-frequency words will be higher than those for low-frequency ones.

1 Receptive vocabulary data reported in the present study are secondary data reanalysed from Segura et al. 2021, and used to contrast them with the productive vocabulary data reported here.

RQ3. Will receptive vocabulary scores be higher than productive vocabulary ones?

H3. Receptive vocabulary acquisition will precede productive acquisition and, thus, results on the receptive tests will be higher than on the productive ones.

2 Methodology

2.1 Participants

A total of $N = 187$ 4- and 5-year-old pre-primary education students in two semi-private schools in Catalonia (Spain) were tested for the purpose of the present study. Nevertheless, not all students were included in our final sample, which was of $N = 155$. Thus, $N = 32$ participants were excluded because they missed too many EFL classes or some testing sessions, or because they were exposed to English outside of the school for more than 4 h a week. The participants in our final sample were Catalan and Spanish bilingual children with no special education needs and no diagnosed speech impairments. The participants selected had a maximum of 1.5 h per week of extracurricular English classes, and a maximum of 3.5 h per week of exposure to English at home, which included reading, watching videos, listening to songs, and/or playing games in English at home. Such information was provided by the students' parents through the sociolinguistic background questionnaire (see Section 2.4 and Appendix 1).

Participants were in their last 2 years of pre-primary education[2] and distributed within each of the schools into two classes per grade. Those class divisions were kept the same and used to divide the students into a control and an experimental group per grade within each of the schools. Consequently, each of the groups, as presented in Table 1, included the students from both schools who were enrolled in the same grade and following the same treatment.

2.2 Pedagogical intervention and treatment

According to the Decree 181/2008,[3] the curriculum for the second stage of the pre-primary schooling years in Catalonia is organized around the abilities that the students are to develop, and how such abilities can be used in a transversal and interactive way in a variety of contexts and situations. Those abilities are organized around three main areas: (1) discovery of oneself and others, (2) discovery of one's surroundings, and (3) communication and languages. Within the latter, children are to experiment with the different uses and functions of the language to develop their verbal and non-verbal communication. Ultimately, they will be able to communicate ideas, as well as to participate in interactions to, for example, solve conflicts, work cooperatively, and use the language to share knowledge, as a learning tool.

Thus, in Catalonia, the curriculum for the pre-primary education grades presents the abilities that students are expected to develop organized around the aforementioned three main developmental areas. However, there is not a syllabus of the specific content within each of those areas for schools to follow. Namely, within the communication and language area, there is not a specific list of vocabulary and grammatical structures that pre-primary students are expected to learn in order to develop the abilities

2 In the Spanish educational system, pre-primary education comprises two stages: the first one, for children aged 0–3 years old; and the second one, for children aged 3–6 years old. In the present study, we focus on the second stage, which is made up of three grades: P3 (3-year-olds), P4 (4-year-olds), and P5 (5-year-olds), which would be equivalent to Foundations 3, 4, and 5.

3 Decret 181/2008, de 9 de setembre, pel qual s'estableix l'ordenació dels ensenyaments del segon cicle de l'educació infantil. Departament d'Educació, Generalitat de Catalunya. [Decree 181/2008, September 9, by which the ordinance of the teaching in the second stage of pre-primary education is established. Department of Education, Generalitat de Catalunya.]

Table 1: Participant distribution

Age (grade)	Group	Treatment	*N*
4-year-old learners (P4)	G1	Control group (FI/non-CLIL)	35
	G2	Experimental group (soft CLIL)	34
5-year-old learners (P5)	G3	Control group (FI/non-CLIL)	42
	G4	Experimental group (soft CLIL)	44

listed in the Decree. Therefore, schools have a certain degree of freedom to develop their own syllabus to promote the learners' development of such abilities.

More specifically, the two schools that participated in the present study had both pointed out that their own EFL syllabus through the last three pre-primary grades was very basic, repetitive, and that it lacked progression from one grade to the next, which was a limitation for students' English acquisition. Therefore, a needs analysis was conducted by the researchers, together with the school teachers, to develop a more complete and progressive EFL syllabus for the last two pre-primary grades.

A soft CLIL approach was chosen to be adopted by the English teachers within the EFL classes with one of the groups in each grade, to provide a more contextualized and natural learning environment. The soft CLIL programme analysed in this article was designed to bring to the EFL classroom content of projects usually taught in Catalan, namely the main language of schooling. Two units related to the four seasons (Autumn and Halloween, and Winter and Christmas) were chosen for the soft CLIL intervention. These units were selected because students were already familiarized with the content in their L1, and because they are present in the learners' daily lives. Additionally, they are worked on in all pre-primary levels, which leaves room for progression by increasing the amount of vocabulary related to the topics included in each grade.

The school's basic EFL syllabus content related to these two units was the only content presented to the control groups within the EFL classes, while the experimental groups followed a more complete syllabus, including not only that same basic EFL syllabus, but also some additional topic-related vocabulary that was added through the soft CLIL approach (see Appendix 2 for the complete list of vocabulary presented to each group and grade). Soft CLIL sessions were embedded within the regular EFL hours, which meant that all groups were in contact with English within the school for the same amount of time. That is, in total, about 3 h per week.

In terms of class dynamics, although the two units of the seasons were presented to the students in all groups, in the control groups they were not the main focus of attention, and this vocabulary was worked on with less intensity in the traditional EFL sessions only. On the other hand, experimental groups were presented with a more complete syllabus of those units within the EFL sessions, following a soft CLIL approach. The EFL soft CLIL sessions were more student centred, and promoted communication, as well as higher order thinking skills. Soft CLIL students were highly engaged in cooperative tasks that enhanced their autonomy and promoted interaction with their peers and teacher. Furthermore, for 1 h a week, students are divided into two smaller groups: half of them remain in the classroom for 30 min with an assistant English teacher playing educational computer games (ICT subject) in English, while the other half go with the English teacher to a separate room and do Arts and crafts in English. During this time, the students in the control groups attended those lessons in English, but followed the regular school's syllabus, which was not related to the units of the seasons. Students in the experimental groups worked on some crafts and played some computer games that were related to the soft CLIL units.

Each of the two units was worked on for 2 months, which adds up to a total of between 45 and 48 h of EFL instruction between the general vocabulary pre-test and the target words vocabulary post-tests (described in Sections 2.3 and 2.4). Regular meetings were held with the English teachers and classroom teachers to select the vocabulary to be taught in English in each unit, to develop the materials, and to follow up on the implementation. Each of the schools had only one English teacher for all pre-primary grades and, thus, all groups from the same school were taught by the same teacher. Both English teachers were bilingual Catalan and Spanish native speakers and had a B2 level of English.

Table 2: Frequency of target words included in the intervention

Group	Treatment	High-frequency words (%)	Low-frequency words (%)
G1 and G2	Basic words only	59	41
G3 and G4	Basic words only	50	50
G2	Additional words added to the basic ones	33	67
G4	Additional words added to the basic ones	25	75

2.3 Design

A 6-month longitudinal study was conducted, from October 2019 (pre-test at time 1, T1) until March 2020 (post-test at time 2, T2). Within the first 2 weeks, all students were administered a general English vocabulary pre-test. Following the pre-test period, the pedagogical intervention took place: about 2 months were devoted to working on the vocabulary of each of the two units selected, namely the basic school syllabus in the control groups, and that same basic vocabulary as well as the additional more complex and less frequent words in the experimental groups. Learners in both groups were exposed to the FL for the same amount of time within the school. At the end of the intervention, that is in March 2020, all students were administered a post-test of the target vocabulary of the seasons that was presented to them in the EFL classes.

In terms of the words selected, BNC and NGSL frequency lists were used to evaluate word frequency. For the purpose of this study, and according to previous publications (Nation 2013, Schmitt and Schmitt 2020), we have considered high-frequency words those lexical items belonging to the first three 1,000 bands of the abovementioned frequency lists, and low-frequency words the items belonging to the fourth and lower bands. Regarding the distribution of high- and low-frequency words that each of the groups were presented with, there was an increase in low-frequency words from the younger group to the oldest one, as well as in the additional vocabulary presented only to the experimental groups, as can be seen in Table 2.

2.4 Instruments

Before starting with the pedagogical intervention, all students were tested with the standardized Peabody Picture Vocabulary Test (PPVT 4th Edition, Form A). This test was used to guarantee that there were no statistically significant differences in terms of the starting EFL levels between the control and experimental groups within each grade.

Students' families were given a sociolinguistic background questionnaire in Catalan (its English translation can be found in Appendix 1) to collect information about the family languages and the student's contact with English outside the school. The main goal of such questionnaire was to identify any students whose mother tongue was English, who used English at home regularly, who had lived in an English-speaking country, or who had a lot of contact with English outside of the school, namely extracurricular activities, interaction with native speakers, watching videos, etc. Such information was taken into consideration for the final selection of participants, as detailed in Section 2.1.

After the pedagogical intervention, all students were tested both receptively and productively on between 50 and 70% of the target words presented in the classroom. In the receptive test, students were shown sets of four pictures and asked to point to the one matching the word spoken by the researcher. In the productive test, students were shown flashcards and asked to name in English whatever they saw in the flashcard. Some variations were made in the tests for each of the four groups, depending on the words that had been worked on in the classroom. That is, control group tests (referred to as "basic test" henceforth) were shorter, including only a percentage of the basic school curricular words, while experimental group tests (referred to as "complete test" henceforth) were longer, since they included the same basic words as in the test administered to their same-grade peers in the control group, as well as some of the more

complex and less frequent words that were added through the pedagogical intervention. Both the receptive (Appendixes 3 and 4) and the productive (Appendixes 5 and 6) vocabulary tests were developed by the researchers, following the guidelines of the PPVT-4,[4] and the Expressive Vocabulary Test (EVT),[5] respectively, in terms of format, picture selection, and picture combination within the page. Before administration, both target words tests were piloted with 6-year-old learners from the same schools. These learners were chosen for the piloting of the tests, because they were still at the beginning of first grade in primary education, namely they were only a few months older than our participants. Furthermore, they were already familiar with the target words in the tests designed. Thus, there was a lower chance that they would randomly guess the answers, which allowed us to test the design of the tests in terms of picture selection, to ensure that all the pictures were clear and elicited the appropriate target words.

2.5 Data collection procedure

Data were collected at two different testing times, within a fortnight each. The first one (T1) was prior to the pedagogical intervention and students were administered the PPVT-4 as a pre-test diagnostic tool; the second one (T2) took place after the intervention and students were administered both the productive and the receptive target words post-tests. In T2, both tests were administered consecutively, starting with the productive test, in which positive feedback was given to the students, but incorrect answers were not corrected, to avoid any transference into the receptive test, which was administered afterwards. Both in T1 and T2, students were tested individually in a separate room, instructions to the tests were given in the student's L1, and the first five items of each test were used as training items. All learners within the same group were tested on the exact same items.

2.6 Data analysis

Raw scores for each test were calculated by adding points for each correct response, while incorrect ones did not add nor deduct points. The 4-year-olds' basic test included 11 items, while the experimental group's complete test included the same 11 items and 9 additional ones, thus being 11 and 20 the highest scores possible in the younger learners' tests, respectively. The 5-year-olds' basic test included 17 items, while the experimental group's complete test included the same 17 items and 10 additional ones, thus being 17 and 27 the highest scores possible in the older learners' tests, respectively.

R 4.1.2 was used to conduct the data analysis. Regarding the general vocabulary pre-test scores obtained in the PPVT-4 test, two independent samples two-tailed *t*-tests with Welch's correction were run, one for each grade. To analyse the data in research question 1, group statistics were obtained for all groups and all tests (basic and complete), and two more independent samples two-tailed *t*-tests with Welch's correction were run to compare the results from the control and experimental groups in the basic test within each grade. As for the second research question, the same software was used to analyse word frequency effects, by describing absolute frequencies and percentages, as well as running a Chi square test and a mixed generalized linear model of Poisson, which included the variables of target word, frequency, class, school, and student; and from which prevalence ratios of the models were obtained. Regarding the third research question, paired samples *t*-tests were run to compare the mean results within each group for the receptive and productive tests. Significance levels were set at ($p = 0.05$) in all cases.

4 Dunn, Lloyd M. and Douglas M. Dunn. 2007. *PPVT4 Peabody Picture Vocabulary Test, Fourth Edition. Manual*. Bloomington: Pearson.

5 Williams, Kathleen T. 1997. *Expressive Vocabulary Test*. Circle Pines: American Guidance Service.

3 Results

3.1 Productive vocabulary learning through soft CLIL

The main focus of the present study was to analyse EFL vocabulary learning through a soft CLIL programme in pre-primary education students. Therefore, the first research question focused on whether such a programme would enhance productive vocabulary acquisition and it was expected that learners following the soft CLIL programme would show higher results when tested on the target words, as compared with their same-grade non-CLIL peers.

To ensure that the control and experimental groups from each grade were comparable, namely that students in both groups did not have significantly different starting EFL proficiency levels, independent samples two-tailed *t*-tests with Welch's corrections were run. Such test for the younger groups found no significant differences in the EFL proficiency starting levels between the control group (G1) ($M = 12.79$, $SD = 8.28$) and the experimental group (G2) ($M = 12.15$, $SD = 7.87$), $t(63.84) = 0.320$, $p = 0.749$. In the case of the older groups, no significant differences were found in the independent samples *t*-test either in the starting levels between the control group (G3) ($M = 15.78$, $SD = 9.71$) and the experimental group (G4) ($M = 15.81$, $SD = 8.92$), $t(81.01) = -0.019$, $p = 0.985$.

Regarding the post-tests focusing on the target words, two independent samples two-tailed *t*-tests with Welch's correction were run, one per each grade to analyse the results of the students in the basic test. No significant differences were found when comparing the results of the younger students in the control group (G1) ($M = 2.60$, $SD = 1.63$) with the results of their same-age peers in the soft CLIL experimental group (G2) ($M = 2.88$, $SD = 1.70$), $t(66.65) = -0.704$, $p = 0.484$. On the other hand, a bigger difference of almost one point was found between the results of the older learners in both groups, which was still not statistically significant. The independent samples *t*-test found that students in the control group (G3) had lower results in the basic test ($M = 4.19$, $SD = 2.18$) than their same-age peers in the soft CLIL experimental group (G4) ($M = 5.14$, $SD = 2.33$), $t(83.97) = -1.946$, $p = 0.054$.

The aforementioned results show slightly higher results in both groups following the soft CLIL programme in terms of productive vocabulary acquisition of the basic curricular words that were worked on in both the control and the experimental groups, although no statistically significant differences were found. Nevertheless, the results in the older learners almost represent positive trends, which could lead to significant differences over a longer period of time following the programme.

Additionally, it is worth analysing whether the soft CLIL group participants of both grades were able to not only learn the same number of basic words, but also some of the additional more complex words that were introduced through soft CLIL without increasing their exposure time to English within the school. Tables 3 and 4 show the results for the experimental groups (G2 and G4) both in the basic test and in the complete test, which was longer because it included the basic test and some additional items that were less frequent. When participants in both grades were administered the complete test, higher raw scores (Table 3) were obtained than in their basic test, which indicates that increasing the amount of vocabulary that young learners are exposed to in the EFL class does not hinder learning, but it enhances greater vocabulary acquisition. In the case of the younger group (G2), students had a mean result that was 2.059 points higher

Table 3: Group statistics in raw scores for experimental groups (G2–G4) in basic and complete (basic + extra) target words productive test

Group	*N*	Test	Mean	Std. deviation	Std. error	Gains
G2	34	Basic test (out of 11)	2.882	1.701	0.292	
		Complete test (out of 20)	4.941	2.436	0.418	+2.059
G4	44	Basic test (out of 17)	5.136	2.329	0.315	
		Complete test (out of 27)	6.523	2.808	0.423	+1.387

Table 4: Group statistics in percentages of correct answers for experimental groups (G2–G4) in basic and complete (basic + extra) target words productive test

Group	*N*	Test	Mean (%)	Std. deviation	Std. error
G2	34	Basic test	26.201	15.467	2.652
		Complete test	24.700	12.181	2.089
G4	44	Basic test	30.214	13.698	2.065
		Complete test	24.256	10.399	1.567

when tested in the complete test (M = 4.941, SD = 2.436) in comparison to the basic test (M = 2.882, SD = 1.701). And a similar trend was found in the older group (G4), which obtained a mean score 1.387 points higher in the complete test (M = 6.523, SD = 2.808) than in the basic test (M = 5.136, SD = 2.329). When transforming such scores into percentages of correct answers (Table 4), similar percentages of productive vocabulary acquisition were found between both tests within each experimental group, although percentages were slightly higher for the basic test. That could be due to the fact that the basic tests included target words that were easier and of a higher frequency in comparison to the additional words that were added to the complete tests administered to the experimental groups.

Finally, it is important to outline the fact that the younger learners in the experimental group (G2) obtained a higher mean score in the complete test (M = 4.94, SD = 2.436) than the older learners in the control group (G3) in the basic test (M = 4.19, SD = 2.18), and almost as high as the results of the older learners in the experimental group (G4) in their basic test (M = 5.14, SD = 2.33). Such results may encourage the introduction of more vocabulary in EFL pre-primary classes.

3.2 Frequency effects on productive vocabulary acquisition

The second research question focused on analysing whether there was a frequency effect that facilitated vocabulary acquisition. It was expected that high-frequency words belonging to the first three 1,000 bands of the BNC and NGSL frequency lists would be more easily learnt and, thus, that results reported for those words would be higher than those for the low-frequency ones.

Table 5 details the number of high- and low-frequency words that were correctly named by the students in all groups and all grades in the productive tests. Those raw scores show that students correctly produced a higher number of high-frequency words (481 out of 1,680, 28.63%) than low-frequency ones (314 out of 1,636, 19.19%), a difference that was statistically significant (p = 0.0001), with a prevalence ratio of 1.492 (95% CI, 1.32 to 1.69).

Additionally, an adjusted Poisson Regression test was run, as shown in Table 6, including the variables of school, student, grade, treatment (control vs experimental), and frequency. It was found that, when all variables were included in the model, there was a higher prevalence ratio of 5.092 (95% CI, 1.91 to 13.58), with a statistically significant difference (p = 0.0011). Therefore, a significant effect of word frequency was found in productive vocabulary learning, which indicates that high-frequency words are recalled more easily by young learners than lower-frequency ones.

Table 5: Descriptive statistics and p-value of Chi square test (high- and low-frequency)

Test	Frequency	Words identified	Percentage	*p*-Value	Prevalence ratio (effect size)	95% confidence interval
Productive	High	481/1,680	28.63	p = 0.0001	1.492	(1.32; 1.69)
	Low	314/1,636	19.19			

Table 6: Adjusted Poisson regression for frequency effects in productive vocabulary

Test	Frequency	*p*-Value	Prevalence ratio (effect size)	95% confidence interval
Productive	High- vs low-frequency	$p = 0.0011$	5.092	(1.91; 13.58)

3.3 Productive and receptive vocabulary acquisition

The last research question enquired on the differences between receptive and productive vocabulary acquisition. Following previous findings in articles investigating vocabulary acquisition, it was expected that receptive vocabulary scores would be higher than productive vocabulary ones, when students were tested on the same target words. To analyse such differences, students were tested on the same target words productively and receptively. As can be seen in Table 7, in the case of the control groups, only their scores on the basic test were considered, while in the experimental groups the basic test results, as well as the complete test scores, which included both the basic test and the additional words, were analysed. Participants' mean scores showed that much higher results were obtained in the receptive vocabulary tests, when compared with the productive tests. In all cases the differences were statistically significant ($p = 0.0001$).

In sum, it can be seen that results in the basic receptive and productive tests were quite similar between both groups in each grade. Consequently, the mean differences between the scores in the basic productive test and the basic receptive test were also quite similar, as shown in Table 7. Such mean differences were obtained through several paired samples *t*-tests. In the case of the younger learners' basic test, the control group's (G1) mean difference between the productive ($M = 2.600$, $SD = 1.63$) and receptive ($M = 7.114$, $SD = 2.26$) basic tests was 4.514 points, $t(34) = -16.17$, $p = 0.0001$; Cohen's $d = 2.29$ (large). For the younger learners in the experimental group (G2), the mean difference between the productive ($M = 2.882$, $SD = 1.70$) and receptive ($M = 7.618$, $SD = 1.83$) basic tests was 4.735 points, $t(33) = -15.081$, $p = 0.0001$; Cohen's $d = 2.68$ (large), which was slightly higher than the mean difference in their control group peers. On the contrary, when it comes to the results of the older participants, the mean difference of the control group's (G3) productive ($M = 4.190$, $SD = 2.18$) and receptive ($M = 11.786$, $SD = 2.64$) basic tests was 7.595 points, $t(41) = -18.699$, $p = 0.0001$; Cohen's $d = 3.14$ (large), which was slightly over the mean difference in the experimental group (G4). In the case of the older learners in the experimental group (G4), the mean difference between the productive ($M = 5.136$, $SD = 2.33$) and receptive ($M = 12.545$, $SD = 2.41$) basic tests was 7.409

Table 7: Descriptive statistics and paired samples *t*-tests for receptive and productive scores in both basic and complete tests

Group, test (*N*)	Mean	Std. deviation of mean	Mean difference (productive −receptive)	Std. error	Sig.
G1, Basic productive (35)	2.600	1.63			
G1, Basic receptive (35)	7.114	2.26	−4.514	0.367	0.0001
G2, Basic productive (34)	2.882	1.70			
G2, Basic receptive (34)	7.618	1.83	−4.735	0.372	0.0001
G2, Complete productive (34)	4.941	2.44			
G2, Complete receptive (34)	12.941	2.74	−8.000	0.463	0.0001
G3, Basic productive (42)	4.190	2.18			
G3, Basic receptive (42)	11.786	2.64	−7.595	0.335	0.0001
G4, Basic productive (44)	5.136	2.33			
G4, Basic receptive (44)	12.545	2.41	−7.409	0.327	0.0001
G4, Complete productive (44)	6.523	2.81			
G4, Complete receptive (44)	18.182	3.72	−11.659	0.407	0.0001

points, $t(43) = -21.599$, $p = 0.0001$; Cohen's $d = 3.13$ (large). It is also worth mentioning that, as expected, the mean difference between the results in both tests increased considerably when analysing the results of the experimental groups' complete test, that is when students in the experimental groups were administered the longer test with more target words. Regarding the results of the younger students (G2), the mean difference between the productive complete test ($M = 4.941$, $SD = 2.44$) and the receptive complete test ($M = 12.941$, $SD = 2.74$) was 8.000 points, $t(33) = -17.297$, $p = 0.0001$; Cohen's $d = 3.09$ (large), which almost doubles the mean difference reported for that same group in the basic test. In the case of the older learners (G4), a higher mean difference was also reported in the complete test: the mean difference between the productive complete test ($M = 6.523$, $SD = 2.81$) and the receptive complete test ($M = 18.182$, $SD = 3.72$) was 11.659 points, $t(43) = -23.054$, $p = 0.0001$; Cohen's $d = 3.54$ (large).

Therefore, it can be concluded that higher scores were obtained in all cases in the receptive vocabulary tests, all with statistically significant differences ($p = 0.0001$) and large effect sizes, as compared with the productive vocabulary tests. The differences between both tests increased when students in the experimental groups were administered the complete test, indicating that, in this case, a longer test increases the gap between receptive and productive vocabulary acquisition. Such results concurred with the fact that older participants (G3 and G4) showed a greater mean difference in all cases, when compared with their younger peers (G1 and G2), which could be due to the fact that the latter were tested in a smaller amount of target words, while the former were given a longer test to guarantee progression through the grades.

4 Discussion

The main goal of the current study was to analyse vocabulary learning in very young learners, namely 4- and 5-year-old students learning English in a FL context through a soft CLIL approach. Special attention was paid to productive vocabulary acquisition, as well as frequency effects. Furthermore, results were compared to the ones reported in Segura et al. 2021, where the same participants' receptive vocabulary acquisition was studied.

The first research question focused on productive vocabulary acquisition, and positive trends were observed, which can indicate that adding more vocabulary through soft CLIL to the regular FL sessions does not hinder language learning. On the contrary, slightly higher results were found in the groups that were presented with a wider amount of vocabulary through a soft CLIL approach, although differences were not statistically significant in either grade. Additionally, participants in the CLIL groups, when tested on the complete test including the additional more complex and less frequent vocabulary items, were also able to recall more words. Similar results were reported in Segura et al. 2021, who analysed the same learners' receptive vocabulary learning as in the current study, and similar positive trends in the CLIL groups were also observed. The lack of statistically significant differences in the basic tests may be attributed to the length of the soft CLIL programme, 6 months, which may not have given the learners enough time to fully benefit from the intervention. In a previous study, Reynaert (2019) found that significant differences in productive vocabulary acquisition through CLIL in secondary education students only appeared after 2 years of CLIL education. Therefore, the author concluded that "CLIL can be an effective way of vocabulary acquisition, but it is essential to consider certain time needed for significant productive vocabulary development" (Reynaert 2019, 158). This may have been the case in the current study, since positive trends were reported in productive vocabulary, as well as in receptive vocabulary in Segura et al. (2021), which may eventually lead to more significant results with a longer CLIL education. Therefore, further research with longer CLIL programmes of various intensities is needed to analyse the effects of such an approach in FL vocabulary learning in very young learners.

Such positive trends concur with findings reported in previous studies, where higher results were observed in FL vocabulary acquisition in the groups following CLIL compared with their non-CLIL peers (Agustín-Llach and Canga Alonso 2014, Artieda et al. 2017, Canga Alonso and Arribas Garcia 2015,

Lasagabaster 2008, Merikivi and Pietila 2014, Pérez Cañado 2018). In most cases, the advantage in the CLIL groups was attributed to the CLIL context, which provides higher quality and quantity of input, as well as a more natural and communicative context (Dalton-Puffer 2008, Lasagabaster and Sierra 2009). However, in most studies, CLIL hours were additional hours, which is yet another factor that may have contributed to the greater vocabulary learning (Agustín-Llach and Canga Alonso 2014, Vallbona 2011) in previous studies. It is important to note that in the current study exposure time to the FL was taken into account and all groups were exposed to the FL for the same amount of time, as the soft CLIL programme was embedded in the regular FL sessions. Thus, in this case, the positive trends in vocabulary learning may in fact be due to the setting that CLIL brings to the classroom. Nonetheless, it remains to be seen if pre-primary education learners would have been able to achieve higher results within the same period of time if soft CLIL hours had been additional hours to the regular FI sessions. Considering that young learners benefit more from natural learning contexts when they are exposed to massive amounts of contact with the FL (Muñoz 2008), it could be expected that higher results in FL learning would be achieved if CLIL hours were additional hours. Thus, there is a need for further research where the implementation of CLIL programmes in pre-primary provides additional hours of exposure to the FL within the school, to analyse whether very young learners can achieve higher learning outcomes within the same period of time.

The second research question aimed at analysing whether there was a frequency effect when learning lexical items. It was hypothesized that higher-frequency words belonging to the first three 1,000 bands of the BNC and NGSL frequency lists would be learnt more easily, and thus, there would be a higher percentage of production of such words in the post-test, compared with the lower-frequency ones. This hypothesis can be fully confirmed, as participants were able to recall and produce a higher percentage of higher-frequency words, with a statistically significant difference, and a prevalence ratio that was also significant. Similar results were reported in a previous study following the same students, but focusing on receptive vocabulary (Segura et al. 2021), in which learners were also able to identify a higher percentage of high-frequency words than lower-frequency ones. Nevertheless, it is important to note that in receptive vocabulary, Segura et al. (2021) reported a prevalence ratio of 1.235, while the present research showed that this prevalence ratio is even bigger in productive vocabulary, with a value of 5.092. Such results indicate that the frequency effect is bigger in productive vocabulary learning.

Previous research analysing the effects of word frequency in FL vocabulary learning reported similar findings, namely that higher-frequency words were acquired first and more easily than lower-frequency ones (Alexiou 2015, Shaban 2013). This may be due to the fact that higher-frequency lexical items have more concrete (Shaban 2013) and common referents (Alexiou 2015). Additionally, they are more present in our daily lives (Alexiou 2015), meaning that there will be more encounters with those items both inside and outside of the classroom, which may also contribute to those words being acquired faster (Albaladejo et al. 2018). Such findings have important pedagogical implications, as they may guide FL teachers when choosing the vocabulary to teach their students in the beginning stages: higher-frequency items seem to be acquired faster and they allow speakers to function in basic conversations (Nation 2013, Nation and Meara 2020).

The third research question enquired on the differences between receptive and productive vocabulary learning. Results regarding productive vocabulary were analysed in the present study and then contrasted with the receptive results reported in Segura et al. (2021). Such comparison allowed the authors to confirm the hypothesis that receptive vocabulary scores would be higher than productive ones, when the same students were tested on the same vocabulary items both receptively and productively, since higher scores were reported in receptive tests, in all groups and tests.

These results are in line with previous research on FL vocabulary learning. Considering the many dimensions of each lexical item (Nation 2013, Schmitt and Schmitt 2020), it is natural to conceive vocabulary acquisition as a complex process. Receptive vocabulary seems to precede productive vocabulary learning (Schmitt and Schmitt 2020, Yongqui Gu 2020), and, thus, one can expect receptive vocabularies to be bigger than productive vocabularies (Meara and Miralpeix 2021, Merikivi and Pietila 2014, Webb 2009). That is precisely what was reported in the present study, in which students in all grades and groups obtained higher scores in receptive tests. Such findings are relevant, not only because they confirm that

receptive vocabulary development precedes productive vocabulary development, but also because they show that young EFL learners' inability to produce a word does not necessarily mean they do not know it, but it may just mean they have partial receptive knowledge of that item (Schmitt and Schmitt 2020).

5 Conclusions

The current study has contributed to the EFL teaching field of research, by analysing a soft CLIL programme and how it fosters vocabulary learning in very young learners in a FL context. Pre-primary education levels have been the focus of attention of very few studies so far, and thus the present research provides significant data regarding vocabulary acquisition through CLIL at such early ages.

The results of this study contribute three main findings to the FL teaching field. First, some positive trends in EFL productive vocabulary learning through a 6-month soft CLIL programme in 4- and 5-year-old learners were seen. This indicates that increasing the vocabulary introduced in the EFL classroom, in this case, did not overwhelm the learners. Nevertheless, such results are not to be generalized to all CLIL contexts and more research is needed to overcome the main limitation of this study: treatment periods longer than 6 months may be necessary to see if those positive trends turn into significant advantages for the CLIL groups. Second, there is a pedagogical implication from the significant frequency effects reported. Results showed that higher-frequency lexical items are indeed easier to recall and identify than lower-frequency ones. Therefore, focusing first on higher-frequency words would facilitate vocabulary learning at the very early stages and allow for some interaction, since there is a rather small amount of vocabulary items that are very frequent and that would allow for basic communication. Finally, as expected from findings reported in previous studies, it was observed that receptive vocabulary was indeed developed before productive vocabulary, as higher results were observed in receptive vocabulary tests. However, some trends of productive vocabulary acquisition were seen after the 6-month intervention. Such results confirm the idea that vocabulary learning is a process, and that learning a lexical item starts with having partial receptive knowledge of the word to then evolve to a more complete productive knowledge. Nevertheless, there is a need for more longitudinal studies in the field of vocabulary learning to analyse whether productive vocabulary acquisition is enhanced with longer instruction time through CLIL, as compared with the FI context. Thus, it remains to be seen whether the gap between receptive and productive vocabulary can be reduced over longer treatment periods thanks to the contextualized learning environment and interaction opportunities that CLIL brings to the classroom.

Overall, the present research has aimed to further analyse vocabulary learning through a soft CLIL programme in very young EFL learners. Some positive trends were seen, but further research is needed to analyse CLIL programme effectiveness in such young learners, with longer treatment periods and a wider variety of learning contexts, such as CLIL programmes of different intensities, CLIL programmes embedded in the regular FL sessions and CLIL sessions as additional hours.

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, Department of Applied Linguistics. The authors would like to thank the reviewers for their insightful comments that have helped improve the present research article.

Funding information: Authors state no funding involved.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: Authors state no conflict of interest.

Data availability statement: All data generated or analysed during this study are included in this published article.

References

Admiraal, Wilfred, Gerard Westhoff, and Kees de Bot. 2006. "Evaluation of bilingual secondary education in the Netherlands: student's language proficiency." *English Educational Research and Evaluation* 12(1), 75–93. doi: 10.1080/13803610500392160.

Agustín-Llach, Maria Pilar and Andrés Canga Alonso. 2014. "Vocabulary growth in young CLIL and traditional EFL learners: Evidence from research and implications for education." *International Journal of Applied Linguistics* 26(2), 211–27. doi: 10.1111/ijal.12090.

Albaladejo, Sara, Yvette Coyle, and Julio Roca de Larios. 2018. "Songs, stories, and vocabulary acquisition in preschool learners of English as a foreign language." *System* 76, 116–28. doi: 10.1016/j.system.2018.05.002.

Alexiou, Thomaï. 2015. "Vocabulary uptake from Peppa Pig: a case study of preschool EFL learners in Greece." In *Current issues in second/foreign language teaching and teacher development: research and practice,* edited by Christina Gitsaki, and Thomaï Alexiou, p. 285–301. Cambridge: Cambridge Scholars Publishing.

Artieda, Gemma, Helena Roquet and Florentina Nicolás-Conesa. 2017. "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. doi: 10.1080/13670050.2017.1373059.

Asensio Arjona, Verónica. 2020. "CLIL program designed to teach music and English in preschool education." In *Music education and didactic materials*, edited by Rosa María Vicente Alvarez, Carol Guillanders, Jesús Rodríguez Rodríguez, Guilherme Romanelli, and Jessica Pitt, p. 247–56. Santiago de Compostela: Grupo STELLAE/IARTEM.

Canga Alonso, Andrés. 2015. "Receptive vocabulary of CLIL and non-CLIL primary and secondary school learners." *Complutense Journal of English Studies* 23, 57–77. doi: 10.5209/rev_CJES.2015.v23.51301.

Canga Alonso, Andrés and Mario Arribas García. 2015. "The benefits of CLIL instruction in Spanish students' productive vocabulary knowledge." *Encuentro: Revista de Investigación e Innovación en la Clase de Idiomas* 24, 15–31.

Cenoz, Jasone. 2002. "Age differences in foreign language learning." *ITL-International Journal of Applied Linguistics* 135(1), 125–42.

Cenoz, Jasone. 2015. "Content-based instruction and content and language integrated learning: The same or different?." *Language, Culture and Curriculum* 28(1), 8–24. doi: 10.1080/07908318.2014.1000922.

Coyle, Do, Phillip Hood, and David Marsh. 2010. *Content and language integrated learning*. Cambridge: Cambridge University Press.

Dale, Liz and Rosie Tanner. 2012. *CLIL activities. A resource for subject and language teaching*. Cambridge: Cambridge University Press.

Dalton-Puffer, Christiane. 2008. "Outcomes and processes in Content and Language Integrated Learning (CLIL): Current research from Europe." In *Future perspectives for English language teaching*, edited by Werner Delanoy, and Laurenz Volkmann, p. 139–57. Heidelberg: Carl winter.

DeKeyser, Robert, M. 2000. "The robustness of critical period effects in second language acquisition." *Studies in Second Language Acquisition* 22(4), 499–533. doi: 10.1017/S0272263100004022.

Ellis, Nick, and Alan Beaton. 1993. "Factors affecting foreign language vocabulary: Imagery keyword mediators and phonological short-term memory." *Quarterly Journal of Experimental Psychology* 46A(3), 533–58. doi: 10.1080/14640749308401062.

García Esteban, Soraya. 2015. "Soft CLIL in infant education bilingual contexts in Spain." *International Journal of Language and Applied Linguistics* 1, 30–6.

Griffin, Gerard Francis. 1992. *Aspects of the psychology of second language vocabulary list learning*. PhD thesis. Warwick: University of Warwick.

Jiménez-Catalán, Rosa Maria, Yolanda Ruíz de Zarobe, and Jasone Cenoz. 2006. "Vocabulary profiles of English foreign language learners in English as a subject and as a vehicular language." *Views* 15(3), 23–6. doi: 10.4067/S0718-48832016000100003.

Krashen, Stephen D., Michael A. Long, and Robin C. Scarcella. 1979. "Age, rate and eventual attainment in second language acquisition." *TESOL Quarterly* 13(4), 573–82. doi: 10.2307/3586451.

Lasagabaster, David. 2008. "Foreign language competence in content and language integrated courses." *The Open Applied Linguistics Journal* 1, 31–42. doi: 10.2174/1874913500801010030.

Lasagabaster, David, and Juan Manuel Sierra. 2009. "Language attitudes in CLIL and international EFL classes." *International CLIL Research Journal* 1(2), 4–17.

Laufer, Batia. 2005. "Focus on form in second language vocabulary learning." *EUROSLA Yearbook* 5, 223–250. doi: 10.1075/eurosla.5.11lau.

Laufer, Batia and Geke C. Ravenhorst-Kalovski. 2010. "Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension." *Reading in a Foreign Language* 22, 15–30.

Lenneberg, Eric H. 1967. *Biological foundations of language*. New York: Wiley.

Mair, Olivia. 2018. "Four seasons of CLIL at two Italian pre-schools." In *English for young learners from pre-school to lower secondary: A CLIL teacher training project in Italian schools*, edited by Francesca Costa, Constanza Cucchi, Olivia Mair, and Amanda Murphy, p. 21–44. Mantova: Universitas Studiorum.

Meara, Paul M. 1996. "The dimensions of lexical competence." In *Performance and competence in second language acquisition*, edited by Gillian Brown, Kirsten Malmkjaer, and John Williams, p. 35–53. Cambridge: Cambridge University Press.

Meara, Paul M. and Imma Miralpeix. 2021. "Bayesian vocabulary tests." *Vigo International Journal of Applied Linguistics* 18, 177–204. doi: 10.35869/vial.v0i18.3370.

Merikivi, Riika and Pïvi Pietila. 2014. "Vocabulary in CLIL and in mainstream education." *Journal of Language Teaching and Research* 5(3), 487–97. doi: 10.4304/jltr.5.3.487-497.

Milton, James. 2009. *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters. doi: 10.21832/9781847692092.

Miralpeix, Imma. 2006. "Age and vocabulary acquisition in English as a foreign language (EFL)." In *Age and the rate of foreign language learning*, edited by Carmen Muñoz, p. 89–106. Toronto: Multilingual Matters. doi: 10.21832/9781853598937-006.

Muñoz, Carmen. 2006. "The effects of age on foreign language learning: The BAF project." In *Age and the rate of foreign language learning*, edited by Carmen Muñoz, p. 6–40. Toronto: Multilingual Matters. doi: 10.21832/9781853598937-003.

Muñoz, Carmen. 2008. "Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning." *Applied Linguistics* 29(4), 578–96. doi: 10.1093/applin/amm056.

Muñoz, Carmen, Elsa Tragant, and Rosa Maria Torras. 2010. "Los inicios de la producción oral en lengua extranjera de alumnos de primaria en Europa." [The beginning stages of oral production in the foreign language in primary students in Europe] ELIA. *Estudios de Lingüística Inglesa Aplicada* 10, 11–39.

Nation, Paul. 1990. *Teaching and learning vocabulary*. Rowley, MA: Newbury House.

Nation, Paul. 2013. *Learning vocabulary in another language*. 2nd ed. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524759.

Nation, Paul, and Paul M. Meara. 2020. "Vocabulary." In *An introduction to applied linguistics*, edited by Norbert Schmitt, and Michael P. H. Rodgers, p. 35–54. London: Routledge.

Nation, Paul, and Robert Waring. 1997. "Vocabulary size, text coverage and word lists." In *Vocabulary: Description, acquisition and pedagogy*, edited by Norbert Schmitt, and Michael McCarthy, p. 6–19. Cambridge: Cambridge University Press.

Ortega, Lourdes. 2015. "Researching CLIL and TBLT interfaces." *System* 54, 103–9. doi: 10.1016/j.system.2015.09.002.

Penfield, Wilfred and Lamar Roberts. 1959. *Speech and brain mechanisms*. Princeton: Princeton University Press. doi: 10.1515/9781400854677.

Pérez Cañado, Maria Luisa. 2018. "CLIL and educational levels: A longitudinal study on the impact of CLIL on language outcomes." *Porta Linguarium* 29, 51–70. doi: 10.30827/Digibug.54022.

Pérez-Vidal, Carmen. 2011. "Language acquisition in three different contexts of learning: Formal instruction, Study Abroad and Semi-immersion (CLIL)." In *Content and foreign language integrated learning: Contributions to multilingualism in European contexts*, edited by Yolanda Ruiz de Zarobe, Juan Manuel Sierra, and Francisco Gallardo del Puerto, p. 25–35. Bern: Peter Lang. doi: 10.3726/978-3-0351-0171-3.

Pérez-Vidal, Carmen, Cristina Escobar, and Helena Roquet. 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.

Reynaert, Barbora. 2019. "CLIL and productive vocabulary acquisition in the Czech Context." *Advances in Language and Literacy Studies* 10(4), 153–60. doi: 10.7575/aiac.alls.v.10n.4p.153.

Richards, Jack C. 1976. "The role of vocabulary teaching". *TESOL Quarterly* 10(1), 77–89. doi: 10.2307/3585941.

Schmitt, Norbert. 2008. "Instructed second language vocabulary learning." *Language Teaching Research* 12(3), 329–63. doi: 10.1177/1362168808089921.

Schmitt, Norbert. 2010. *Researching vocabulary: A vocabulary research manual*. Hampshire: Palgrave Macmillan.

Schmitt, Norbert and Diane Schmitt. 2020. *Vocabulary in language teaching*. Cambridge: Cambridge University Press. doi: 10.1017/9781108569057.

Segura, Marta, Helena Roquet, and Júlia Barón. 2021. "Receptive vocabulary acquisition in pre-primary education through soft-content and language integrated learning." *English Language Teaching* 14(10), 1–22. doi: 10.5539/elt.v14n10p1.

Shaban, Muhammad. (2013). "Natural order of vocabulary acquisition." *European Academic Research* 1(5), 721–33.

Shintani, Natsuko. 2011. "A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners." *Language Teaching Research* 15(2), 137–58. doi: 10.1177/1362168810388692.

Vallbona, Anna. 2011. "Implementing CLIL in the primary classroom: Results and future challenges." In *AICLE - CLIL - EMILE Educació plurilingüe: Experiencias, research & polítiques* [*CLIL Plurilingual education: Experiences, research & policies*], edited by Cristina Escobar, Natalia Evnitskaya, Emilee Moore, and Adriana Patriño, p. 151–8. Cerdanyola del Vallès: Universitat Autònoma de Barcelona.

Van Zeeland, Hilde, and Norbert Schmitt. 2013. “Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension?.” *Applied Linguistics* 34(4), 457–79. doi: 10.1093/applin/ams074.

Waring, Robert. 1997. “A study of receptive and productive learning from word cards.” *Studies in Foreign Languages and Literature* 21(1), 94–114.

Webb, Stuart. 2009. “The effects of receptive and productive learning of words pairs on vocabulary knowledge.” *RELC Journal: A Journal of Language Teaching and Research* 40(3), 360–7. doi: 10.1177/0033688209343854.

Webb, Stuart. 2020. “Incidental vocabulary learning.” In *The Routledge handbook of vocabulary studies*, edited by Stuart Webb, p. 225–39. London: Routledge.

Wong Kwok Shing, Richard. 2006. “Enhancing English among second language learners: The pre-school years.” *Early Years* 26(3), 279–93. doi: 10.1080/09575140600898456.

Yongqi Gu, Peter. 2020. “Strategies for learning vocabulary.” In *The Routledge handbook of vocabulary studies*, edited by Stuart Webb, p. 271–87. London: Routledge.

Appendix

Appendix 1 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 min. 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

<u>About the student's siblings</u>

- 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 2and 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 with the child?			
	Lives with the child	Does not live with the child	Speaks mostly Catalan	Speaks mostly Spanish	Speaks Catalan and Spanish	Speaks another language
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:

	Time per week							
	Never	Less than 30 min	30 min to 1 hour	1 to 2 hours	2 to 3 hours	3 to 4 hours	4 to 5 hours	More than 5 hours
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 2 Vocabulary list per unit and group

Autumn and Halloween unit

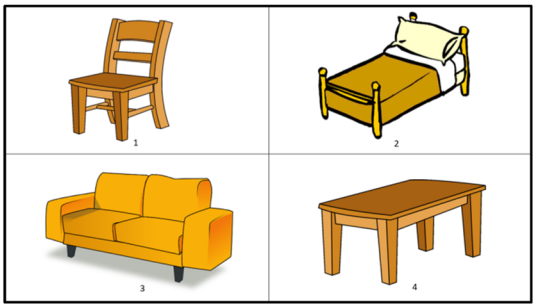
Grade	Basic vocabulary for control and experimental groups		Experimental groups added vocabulary	
P4	Orange	Cat		
	Brown	Rainy	Leaf/leaves	Acorn
	Yellow	Windy	Worm	Witch
	Green	Halloween	Snail	Pumpkin
	Red	Ghost	Owl	Night
	Tree	Moon		
	Squirrel	Bat		
P5	Leaf/leaves	Cat		
	Orange	Rainy	White	Mushroom
	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

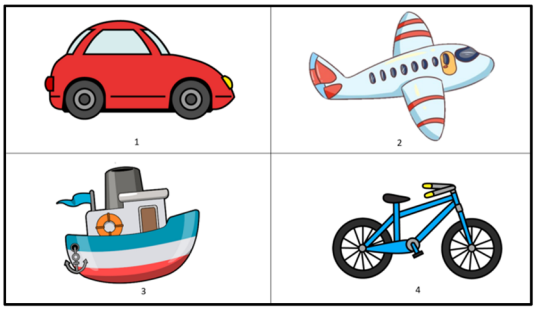
Grade	Basic vocabulary for control and experimental groups		Experimental groups added vocabulary	
P4	Snow	Cold		
	Snowman	Frozen		
	Buttons	Snowy	Carrot	Boots
	Snowflake	Penguin	Hot chocolate	Reindeer
	Jacket	Polar bear	Snowball	Star
	Gloves	Jesus	Coat	
	Scarf	Three wise men		
	Hat	Bethlehem		
	Presents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

Appendix 3 Sample of basic curriculum words receptive test

Receptive Vocabulary Test Training Items

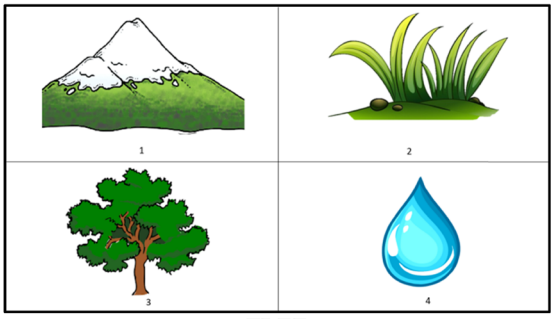


TABLE

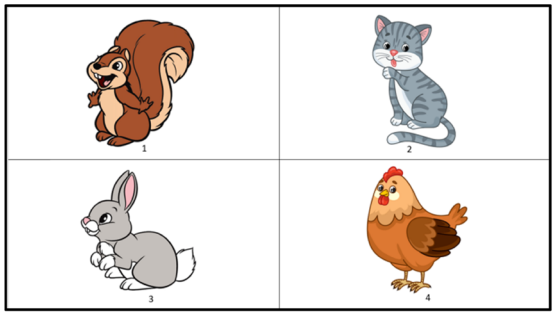


CAR

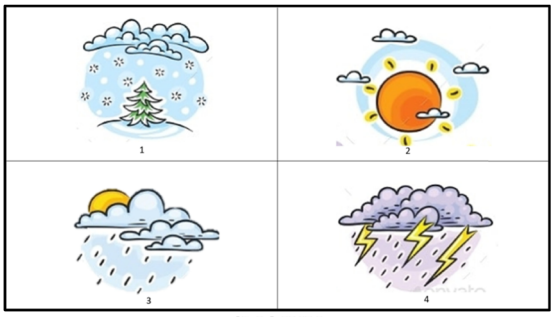
P4 Basic Curriculum Receptive Vocabulary Test



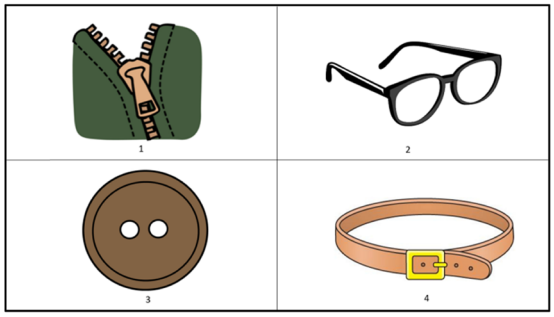
TREE



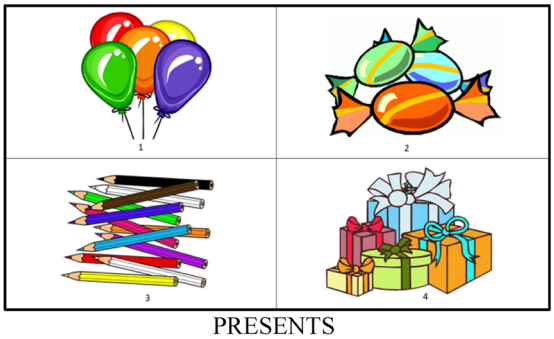
SQUIRREL



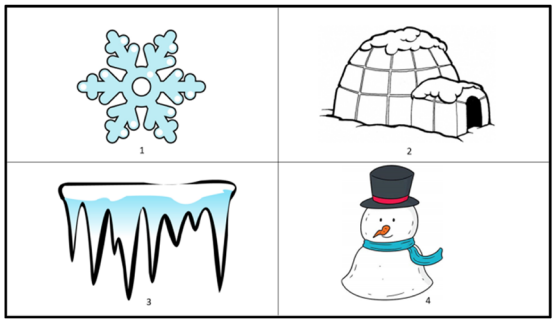
SNOWY



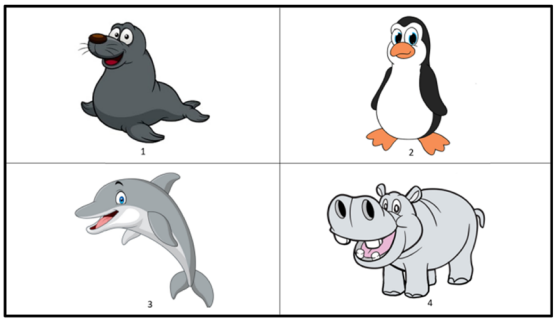
BUTTON



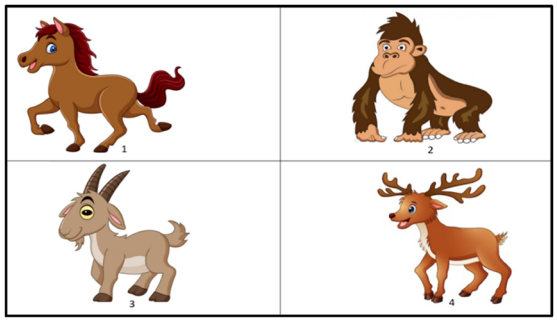
PRESENTS



SNOWMAN

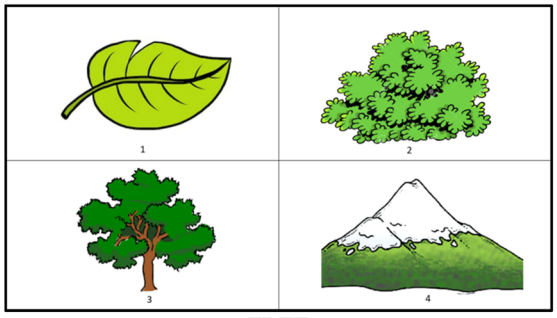


PENGUIN

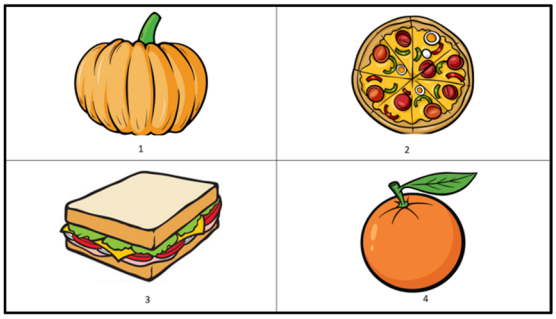


REINDEER

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

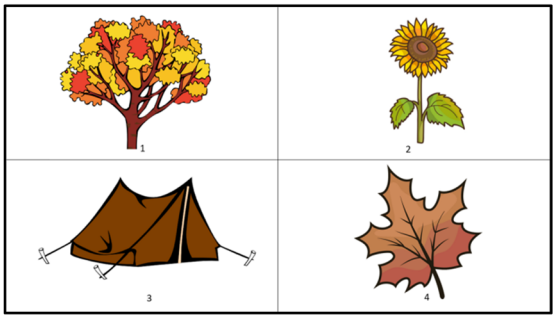


TREE

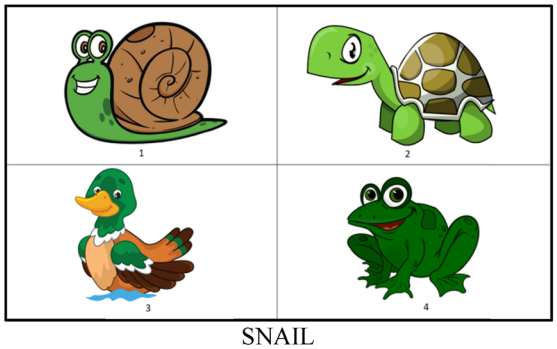


PUMPKIN

NIGHT



LEAF

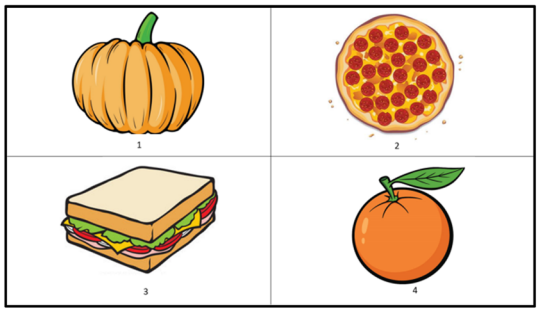


SNAIL

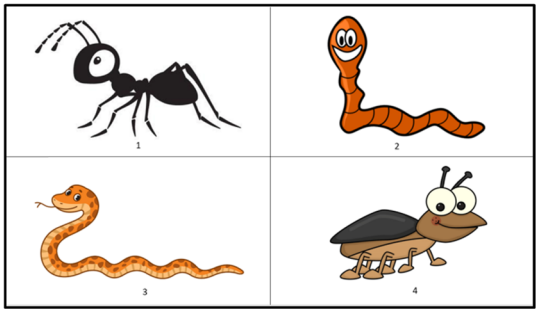
COAT

Appendix 4 Sample of additional target words receptive test

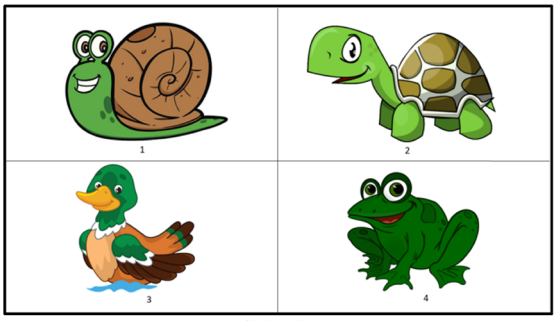
P4 Added Target Words Receptive Vocabulary Test



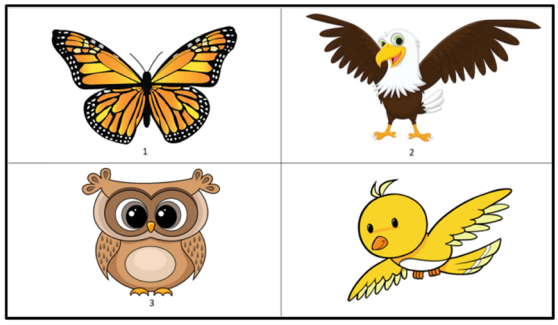
PUMPKIN



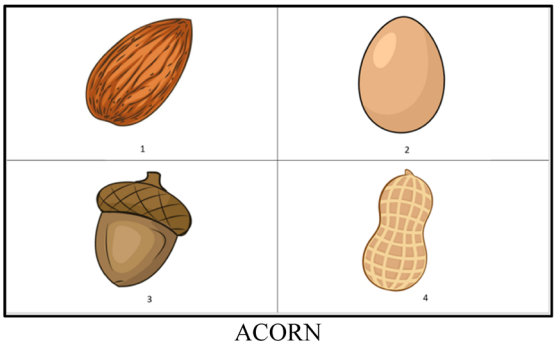
WORM



SNAIL



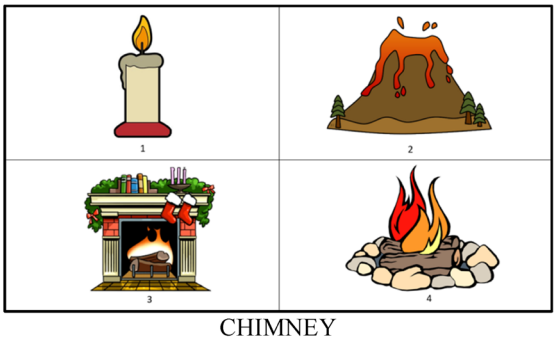
OWL



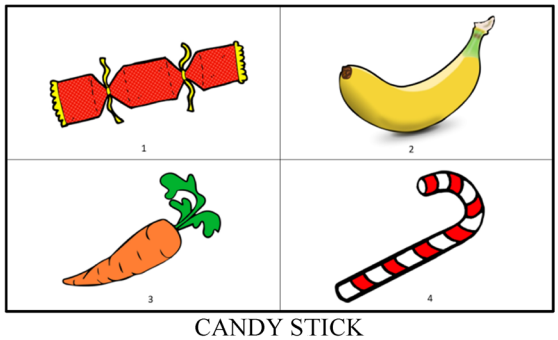
ACORN

COAT

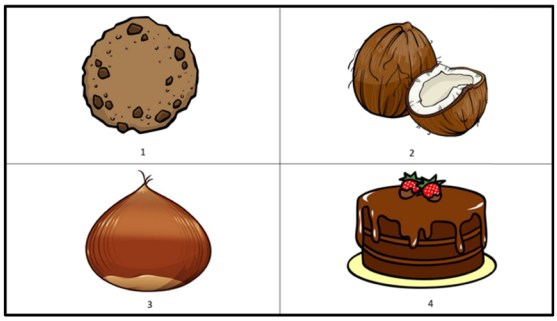
P5 Added Target Words Receptive Vocabulary Test (Additional to the P4 Test)



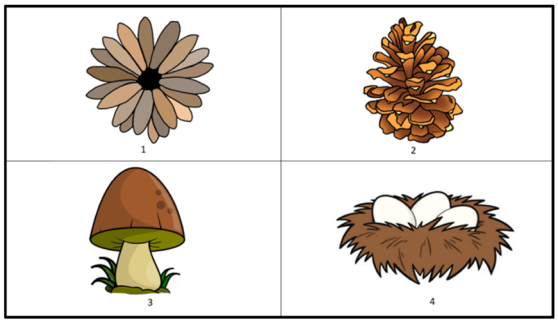
CHIMNEY



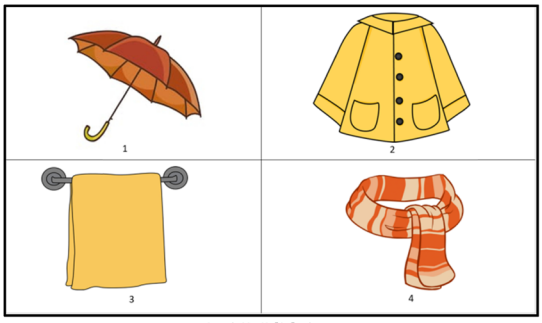
CANDY STICK



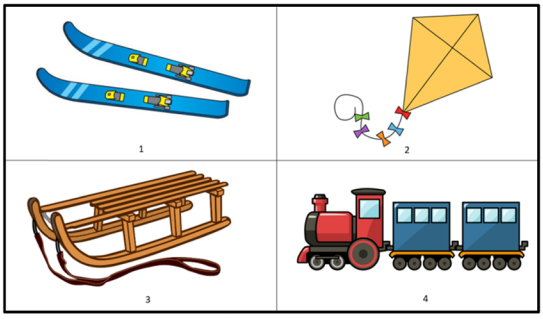
CHESTNUT



MUSHROOM



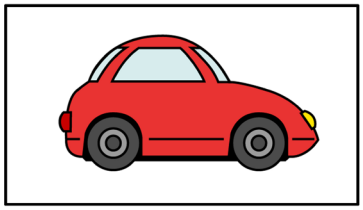
RAINCOAT

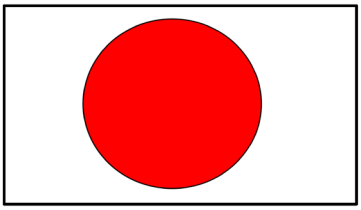


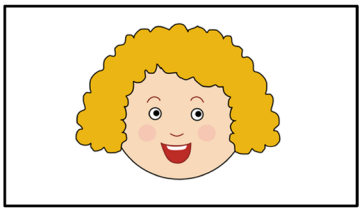
SLEDGE

Appendix 5 Sample of basic curriculum words productive test

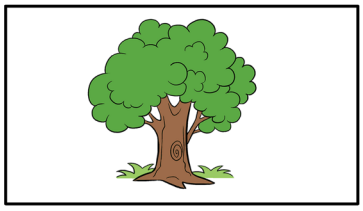
Productive Vocabulary Test Training Items

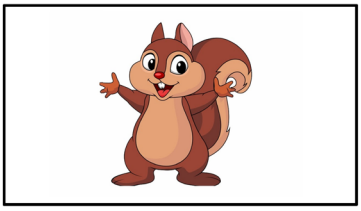

CAR

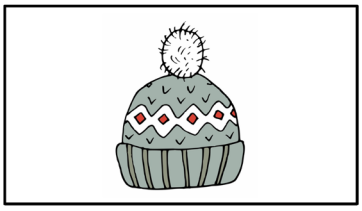

RED


HAPPY

P4 Basic Curriculum Productive Vocabulary Test


TREE


SQUIRREL

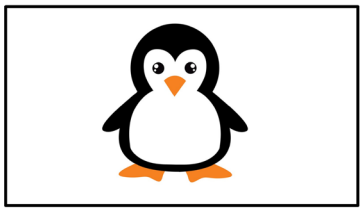

HAT

PRESENTS

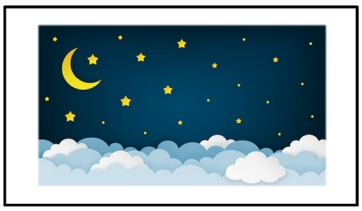
CHRISTMAS TREE

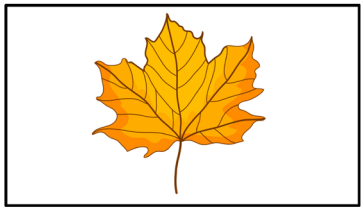
SNOWMAN

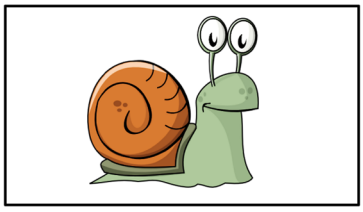
P5 Basic Curriculum Productive Vocabulary Test (Additional to the P4 Test)


PENGUIN

PUMPKIN


NIGHT

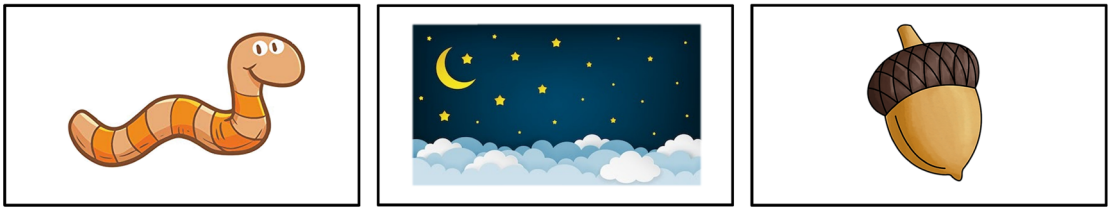

LEAF


SNAIL

COAT

Appendix 6 Sample of additional target words productive test

P4 Added Target Words Productive Vocabulary Test



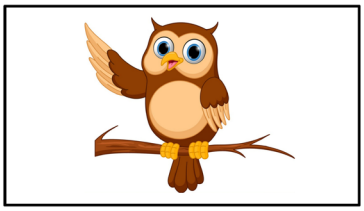
WORM

NIGHT

ACORN

COAT

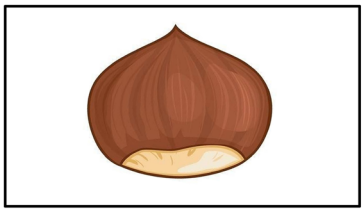
STAR



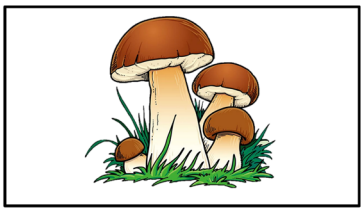
OWL

P5 Added Target Words Productive Vocabulary Test (Additional to the P4 Test)

CANDY STICK



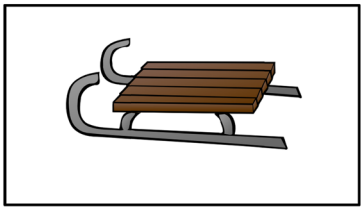
CHESTNUT



MUSHROOM

WATER BOOTS

RAINCOAT



SLEDGE