

upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

FACULTY OF HUMANITIES, SOCIAL AND EDUCATIONAL SCIENCES
FACULTAD DE CIENCIAS HUMANAS, SOCIALES Y DE LA EDUCACIÓN
GIZA, GIZARTE ETA HEZKUNTZA ZIENTZIEN FAKULTATEA

Master's Degree in Secondary Education Teaching
Máster Universitario en Profesorado de Educación Secundaria
Unibertsitate Masterra Bigarren Hezkuntzako Irakasletzan

Final Master's Thesis
Trabajo Fin de Máster
Master Bukaerako Lana

**The Impact of CLIL Exposure on
Secondary-School Learners' Oral Proficiency
in English**

*El impacto de la exposición al AICLE en la
competencia oral en inglés de alumnos de secundaria*
*EAHIIrekiko esposizioak bigarren hezkuntzako
ikasleen ingelesezko ahozko gaitasunean duen
eragina*

Estudiante: Edurne Orce Franco

Tutor: Raúl Azpilicueta Martínez

Departamento: Ciencias Humanas y de la Educación

Especialidad: CLIL

Junio, 2023

Abstract.

This study investigates the impact of Content and Language Integrated Learning (CLIL) programmes on the oral proficiency of secondary-school learners in Spain. The study compares the performance of non-CLIL and CLIL learners, with each group consisting of 10 members and all being 13-14 years old. The non-CLIL group received 1702 hours of English instruction, and the CLIL group received 2220 hours of English instruction. The assessment of oral proficiency involved an oral task using Heaton's "Bicycle" (1966) comic strip. The results highlight significant differences between the non-CLIL and CLIL groups, particularly in overall proficiency and fluency. The CLIL group consistently achieved higher scores in these areas. Differences were also observed in vocabulary and grammar, indicating the positive influence of CLIL exposure, while disparities in coherence and pronunciation measures were less prominent. These findings underscore the beneficial effects of CLIL exposure beyond the 300-hour threshold on secondary-school learners' oral proficiency in English, emphasizing the importance of integrating content and language instruction to enhance students' language skills.

Keywords: CLIL; exposure; oral proficiency; secondary-school learners; language gains

Resumen.

Este estudio investiga el impacto de los programas de Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE) en la competencia oral de los alumnos de secundaria en España. En el estudio se compara el rendimiento de los alumnos no AICLE con el de los alumnos AICLE. Cada grupo consta de 10 miembros y todos tienen entre 13 y 14 años. El grupo no AICLE recibió 1.702 horas de enseñanza de inglés, y el grupo AICLE recibió 2.220 horas de enseñanza de inglés. La evaluación de la competencia oral consistió en una tarea oral utilizando el cómic "Bicycle" (1966) de Heaton. Los resultados ponen de manifiesto diferencias significativas entre los grupos AICLE y no AICLE, sobre todo en lo que se refiere al dominio general y la fluidez. El grupo AICLE obtuvo sistemáticamente puntuaciones más altas en estas áreas. También se observaron diferencias en vocabulario y gramática, lo que indica la influencia positiva de la exposición a AICLE, mientras que las disparidades en las medidas de coherencia y pronunciación fueron menos prominentes. Estos resultados subrayan los efectos beneficiosos de la exposición al AICLE más allá del umbral de las 300 horas en la competencia oral en inglés de los alumnos de secundaria, destacando la importancia de integrar la enseñanza de contenidos y de lenguas para mejorar las destrezas

lingüísticas de los alumnos.

Palabras clave: AICLE; exposición; competencia oral; alumnos de secundaria; progresos lingüísticos

Laburpena.

Ikerketa honek Atzerriko Edukien eta Hizkuntzen Ikaskuntza Integratuko (AEHII) programek Espainiako bigarren hezkuntzako ikasleen ahozko gaitasunean duten eragina ikertzen du. AEHII ez diren ikasleen errendimendua AEHII ikasleen errendimenduarekin alderatzen da ikerketan. Talde bakoitzak 10 kide ditu eta guztiek 13 eta 14 urte bitartean dituzte. AEHII ez den taldeak ingeleseko 1.702 ordu jaso zituen, eta AICLE taldeak ingeleseko 2.220 ordu. Ahozko gaitasunaren ebaluazioa ahozko zeregin bat izan zen, Heatonen "Bicycle" (1966) komikia erabiliz. Emaitzek agerian uzten dute alde esanguratsuak daudela AEHII taldeen eta ez AEHII taldeen artean, batez ere nagusitasun orokorrari eta arintasunari dagokienez. AEHII taldeak sistematikoki puntuazio altuagoak lortu zituen arlo horietan. Lexiko eta gramatikan ere desberdintasunak ikusi ziren, eta horrek adierazten du AEHIIrekiko esposizioak eragin positiboa izan zuela, eta koherentzia- eta ahoskera-neurrietan desberdintasunak ez zirela hain nabarmenak izan. Emaitza horiek azpimarratzen dute AEHIIri 300 orduko atalasetik gorako esposizioak eragin onuragarriak dituela bigarren hezkuntzako ikasleen ingelesezko ahozko gaitasunean, eta nabarmentzen dute garrantzitsua dela edukien eta hizkuntzen irakaskuntza integratzea ikasleen hizkuntza-trebetasunak hobetzeko.

Hitz gakoak: EAHII; esposizioa; ahozko gaitasuna; bigarren hezkuntzako ikasleak; hizkuntza-aurrerapenak

INDEX

1. INTRODUCTION	1
2. LITERATURE REVIEW	2
2.1. CLIL and overall language gains of secondary-school learners	2
2.2. CLIL and oral proficiency of secondary-school learners	4
3. METHOD	7
3.1. Participants	7
3.2. Task and procedure	9
3.3. Data collection and analysis	10
4. RESULTS	11
4.1. Global results	12
4.1.1. Proficiency scores for each subskill	12
4.1.2. Distribution by gender on global results	13
4.2. Coherence	14
4.3. Fluency	15
4.4. Grammar	15
4.5. Pronunciation	16
4.6. Vocabulary	17
4.7. Subskill results in a box plot	17
5. DISCUSSION	19
5.1. Overall oral proficiency	19
5.1.1 Coherence	20
5.1.2. Fluency	22
5.1.3. Grammar	24
5.1.4. Pronunciation	26
5.1.5. Vocabulary	28
6. CONCLUSION, LIMITATIONS AND PEDAGOGICAL IMPLICATIONS	30
6.1. Conclusion	30
6.2. Limitations	31
6.2.1. Variable limitations	31

6.2.2. Gender limitations	32
6.2.3. Voluntary participation	32
6.3. Pedagogical implications	33
7. ACKNOWLEDGEMENTS	33
REFERENCES	34
APPENDIXES	39
Task materials	39
Comic strip	39
Teacher's script	39
Holistic rubric	40

1. INTRODUCTION

Content and Language Integrated Learning (CLIL) is an approach to bilingual education that integrates the teaching of a subject through a second language, while also improving language proficiency in that language. CLIL has gained widespread acceptance in Europe, where it has been implemented in bilingual education programmes. Several studies have shown that CLIL can promote language learning, subject matter knowledge, and cognitive development (Dalton-Puffer, 2011; Goris et al, 2019). This literature review aims to discuss the benefits of CLIL on overall language gains among secondary-school students.

Prior research has investigated the critical threshold of additional exposure to the target language (TL) necessary for significant language gains to occur in CLIL groups compared to non-CLIL groups. Most of the existing evidence stems from secondary education-based studies, where a longer duration of exposure to CLIL instruction leads to better overall language gains (Muñoz, 2015). CLIL instruction has also been shown to have a positive impact on academic achievement, and language proficiency in listening, reading, writing, and speaking (Admiraal et al., 2006; Jiménez Catalán & Ruiz de Zarobe, 2009). However, the issue of exposure, CLIL, and language gains is complex, and other factors, such as students' motivation and aptitude, teacher quality, and school context, may also play a role in shaping language gains in CLIL settings (Dalton-Puffer, 2011). While CLIL instruction has shown positive results in promoting language gains, there are uncontrollable variables that may influence its effectiveness, such as socio-economic status (SES), extracurricular exposure, and "CLIL selection" (Admiraal et al., 2006).

Focusing on the impact of CLIL instruction on students' oral proficiency skills in English, various factors have to be considered, such as the amount of exposure and the type of introduction. Prior research has shown that CLIL instruction can effectively enhance students' oral proficiency skills in secondary education (Admiraal et al., 2006; Gallardo del Puerto et al., 2009; Gálvez Gómez, 2021). Studies have demonstrated a positive correlation between the amount of exposure to CLIL instruction and students' oral production outcomes, suggesting that CLIL has a positive effect on students' oral proficiency (Dalton-Puffer, 2011; Gallardo del Puerto & Gómez Lacabex, 2017; Gálvez Gómez, 2021; Goris et al. 2019; Iwashita et al., 2007). However, the question of what type of CLIL instruction is most effective in enhancing students' oral proficiency skills remains unanswered.

This research study aims to investigate the impact of CLIL on oral proficiency gains among secondary-school students and examine the effect of different types of CLIL exposure on L1-Spanish secondary-school English language learners' oral proficiency. The study will explore the critical threshold of additional exposure to the TL necessary for substantial language to occur in favour of CLIL groups compared to non-CLIL groups. The findings of this study can provide insight into the efficacy of CLIL instruction and contribute to the ongoing discussion on bilingual education.

2. LITERATURE REVIEW

2.1. CLIL and overall language gains of secondary-school learners

CLIL is a bilingual education approach that integrates the teaching of a subject through a second language, while also improving linguistic proficiency in that language. CLIL has been widely adopted in Europe, where it has been implemented in bilingual education programmes. Several studies have highlighted the effectiveness of CLIL in promoting language learning, subject matter knowledge, and cognitive development (Dalton-Puffer, 2011; Goris et al., 2019). Research indicates that CLIL can result in notable improvements in language proficiency by offering additional opportunities to use the TL without extending the total amount of time spent on the curriculum (Jiménez Catalán & Ruiz de Zarobe, 2009). However, the intensity of the CLIL programme and the initial age of instruction play a critical role in maximizing language gains (Admiraal et al., 2006; Housen, 2012). This literature review aims to discuss the benefits of CLIL on overall language gains among secondary-school students.

Researchers have investigated the amount of exposure to the TL necessary for substantial language gains to occur in CLIL groups compared to non-CLIL groups. A review of the existing literature by Muñoz (2015) established a critical threshold of 300 hours of additional exposure to the TL for meaningful language progress to be observed in CLIL groups when compared to non-CLIL groups. The additional exposure to the TL in CLIL groups over the non-CLIL ones in the studies analysed by Muñoz (2015) ranged mostly from 200 to 400 hours, with some as low as 43.2 hours and as high as 480 hours, most of the existing evidence deriving from secondary education-based studies. Furthermore, Muñoz's study (2015) on the timing of CLIL programmes showed that a longer duration of exposure to CLIL instruction leads to better overall language gains, with students who received bilingual secondary education (BSE) for a longer duration outperforming their peers who received it for a shorter duration in all language skills, including speaking, listening, reading and writing.

CLIL has been found to have a positive impact on language acquisition and academic achievements in several studies. Admiraal et al. (2006) found that students who received BSE through CLIL outperformed their peers who received traditional English as a foreign language (EFL) instruction in all language skills. Furthermore, CLIL instruction has been shown to have a positive impact on students' receptive vocabulary knowledge (Jiménez Catalán & Ruiz de Zarobe, 2009) and on language proficiency in listening, reading, writing, and speaking (Goris et al., 2019). These findings suggest that CLIL provides learners with a well-rounded language education that promotes overall language gains and is an effective approach to language learning that can benefit learners of all proficiency levels.

The issue of exposure, CLIL, and language gains is a complex one. It is challenging to attribute the positive influence in learning outcomes exclusively to the differential CLIL factor or the cumulative effect of increased exposure to the TL. Other factors, such as students' motivation and aptitude, teacher quality, and school context, can also play a role in shaping language gains in CLIL settings (Dalton-Puffer, 2011). While CLIL instruction has shown positive results in promoting linguistic acquisition, there are uncontrollable variables that may influence its effectiveness, such as SES, extracurricular exposure, and the often referred as "CLIL selection". SES has been shown to have a significant impact on language acquisition and educational achievement (Admiraal et al., 2006). Extracurricular exposure to the second language could also play a role in language gains, as students who are exposed to the language outside the classroom may have an advantage over those who are not (Gallardo del Puerto et al., 2009; Peters, 2018). Finally, the selection of students into CLIL programmes could also be a factor, as students who are selected for CLIL programmes may already have a higher level of language proficiency or a stronger motivation to learn a second language (Dalton-Puffer, 2007; de Smet et al., 2019). These uncontrollable variables need to be taken into account when evaluating the effectiveness of CLIL instruction. Moderating variables such as SES or extramural exposure to the TL, through audiovisual input or attending extracurricular foreign language lessons, can also influence the impact of CLIL time on language gains reported in previous studies. CLIL groups may have been exposed to more additional out-of-school TL time than the non-CLIL groups, leading to the interpretation of existing results with caution. Future studies should explore this issue further, controlling for moderating variables to provide a more accurate picture of the benefits of CLIL in language learning (Gallardo del Puerto & Gómez Lacabex, 2017; Goris et al., 2019).

All in all, CLIL is a promising approach for promoting language gains and content learning simultaneously, although the success of CLIL instruction is not guaranteed, as various factors could affect its effectiveness. These factors include the need for qualified teachers with the necessary language proficiency and subject matter expertise, the complexity of attributing learning outcomes to the differential CLIL factor or increased exposure to the TL, and the need to meet the needs of all students, including additional language support for students not proficient in the TL and differentiated instruction to accommodate students with different learning needs and styles. Future research should continue to investigate the effectiveness of CLIL instruction and explore ways to improve its implementation. Overall, CLIL instruction has the potential to provide students with valuable language and content knowledge, preparing them for success in an increasingly globalized world (Dalton-Puffer, 2011; Gallardo del Puerto & Gómez Lacabex, 2017; Gálvez Gómez, 2021).

2.2. CLIL and oral proficiency of secondary-school learners

CLIL provides students with authentic and meaningful contexts to practice the language they are learning. Several studies have demonstrated a positive correlation between the amount of exposure to CLIL instruction and students' oral production outcomes, suggesting that CLIL has a positive effect on students' oral proficiency. This review will discuss various research studies, including both quantitative and qualitative methods, that have investigated the relationship between CLIL and oral proficiency (Dalton-Puffer, 2011; Gallardo del Puerto & Gómez Lacabex, 2017; Gálvez Gómez, 2021; Goris et al. 2019; Iwashita et al., 2007).

Several studies have shown that CLIL instruction can effectively enhance students' oral proficiency skills in secondary education (Admiraal et al., 2006; Gálvez Gómez, 2021). However, it is worth noting that there is a relatively limited body of research specifically analysing the relationship between CLIL and oral proficiency of secondary-school learners. While the existing studies have provided valuable insight into the impact of CLIL on overall language gains, the exploration of its effects on oral proficiency seems to be less ample. This knowledge gap highlights the importance of further investigation in this area. By delving into the theme of CLIL and its influence on the oral proficiency of secondary-school learners, the aim is to contribute to the existing literature and bridge the gap in knowledge.

In a study conducted by Gallardo del Puerto and Gómez Lacabex (2017), the experimental group that received CLIL instruction outperformed the control group that

received non-CLIL instruction in some measures of oral proficiency, fluency, grammatical accuracy, and lexical complexity to be more exact. The study also found that the amount of exposure to CLIL instruction was positively associated with students' oral production outcomes, indicating that increased exposure to CLIL instruction may lead to further improvements in oral proficiency skills (Gallardo de Puerto & Gómez Lacabex, 2017). Similarly, Housen (2012) found that CLIL exposure positively impacts secondary-school students' oral proficiency skills in terms of fluency, grammatical accuracy, and lexical complexity. Higher levels of CLIL instruction led to greater improvements in oral production outcomes, underscoring the potential benefits of increased CLIL exposure for enhancing students' oral proficiency. Further research suggests that the amount of exposure to CLIL instruction is a crucial factor in promoting oral proficiency (Goris et al., 2019). That way, Goris et al. (2019) highlighted that CLIL instruction plays a crucial role in bridging the gap between language and content instruction. The research conducted involved a systematic review of longitudinal experimental studies, providing robust evidence on the effects of CLIL. The findings demonstrated that CLIL instruction offers a more integrated and effective learning experience for students. By combining language and content instruction, CLIL creates a seamless connection between the two, allowing students to develop language skills while simultaneously engaging with subject matter content. This integration promotes deeper understanding, higher engagement, and increased motivation among learners. This study is instrumental in solidifying the importance of integrating language and content instruction, ultimately enhancing the overall effectiveness of education.

One English subskill that has received considerable research attention is the impact of CLIL instruction on the development of English pronunciation. The studies conducted by Gallardo del Puerto and Gómez Lacabex (2009) found that CLIL instruction had a significant positive effect on the pronunciation of Spanish students learning English. These findings are consistent with previous research that has shown that integrating language and content instruction can lead to improvements in various aspects of language learning (Swain, 2000). Moreover, the importance of pronunciation skills in language learning cannot be overstated. Pronunciation plays a critical role in communication and can significantly affect how well a speaker is understood by their audience. As such, the positive impact of CLIL instruction on English pronunciation is particularly noteworthy, as it indicates that this approach can be an effective tool for enhancing overall oral proficiency. Gallardo del Puerto and Gómez Lacabex (2017) highlight the importance of improving pronunciation skills in language instruction, noting that it can enhance learners' confidence, increase their intelligibility, and facilitate

communication. The use of CLIL instruction in language learning programmes can help students develop the skills they need to communicate effectively in English, which can improve their academic performance and professional prospects.

In her study, Gálvez Gómez (2021) emphasized that CLIL instruction can be particularly beneficial for students who have limited exposure to English outside of the classroom, as it provides them with a more immersive language learning experience. Specifically, focusing on 4th-grade students of ESO, she observed that these CLIL learners consistently outperformed their non-CLIL counterparts across multiple dimensions, including grammatical accuracy, lexical richness, pronunciation, task fulfillment, and fluency measures. This highlights the comprehensive benefits of CLIL instruction in enhancing students' oral proficiency. By providing a more immersive language learning experience, CLIL effectively bridges the gap between language and content instruction, emphasizing the importance of integrating these components and demonstrating the long-term positive effects on students' foreign language acquisition. Moreover, in Pérez Cañado and Lancasters' longitudinal case study (2017), the effects of CLIL on secondary education students' oral proficiency are examined, by comparing a group of students in a CLIL programme with a group in a non-CLIL programme for three years. They found that CLIL instruction had a positive impact on students' oral comprehension and production skills, as the CLIL group outperformed the non-CLIL in both areas. The study suggests that the increased exposure to and use of the TL in a content-based context through the CLIL programme can improve students' confidence and motivation in using the language, contributing to their overall oral proficiency. These findings indicate that implementing CLIL in secondary education can be an effective way to enhance students' oral language skills.

Although CLIL has been recognized as a promising approach to enhance oral proficiency for secondary education students, its positive effects may depend on several factors, such as learners' age, language proficiency level, and the specific language being taught. Furthermore, uncontrollable variables such as SES and extracurricular exposure can also impact students' oral proficiency, making it difficult to isolate the effects of CLIL instruction. Therefore, appropriate support and training for teachers and scaffolding to make content accessible to all students are necessary. These findings are supported by several studies, including Dalton-Puffer (2011) and Gallardo del Puerto and Gómez Lacabex (2017). Despite these challenges, CLIL remains a promising approach to enhancing oral proficiency for secondary education students (Gallardo del Puerto et al., 2009).

Overall, CLIL is a promising approach for enhancing oral proficiency in secondary education students. Evidence shows that CLIL has a positive effect on learners' oral proficiency, cognitive and intercultural development (Gallardo de Puerto & Gómez Lacabex, 2017; Jiménez Catalán & Ruiz Ruiz de Zarobe, 2009). Implementing CLIL, however, requires careful planning and consideration of learners' specific needs and contexts. For the maximum benefits of CLIL, it is necessary to ensure that all students have access to the programme and that teachers receive adequate training to teach content in the TL (Dalton-Puffer, 2011). Scaffolding is crucial for making content accessible to learners, regardless of their language proficiency level (Gallardo del Puerto & Gómez Lacabex, 2017). Despite these challenges, CLIL has the potential to transform language learning and enhance educational outcomes for all students (Admiraal et al., 2006; Gálvez Gómez, 2021; Goris et al., 2019).

The research question that this study seeks to answer is the following:

What is the effect of CLIL exposure on secondary-school English language learners' oral proficiency?

3. METHOD

3.1. Participants

The study sample consisted of 20 secondary education students from 2nd of ESO (12-13 years old) from the same secondary school of Pamplona (Navarre, Spain), called Iparralde and which is a state school. They were divided into two equal groups of 10 members each: non-CLIL and CLIL. The students had been in their own programmes since they were 4 years old, which amounts to 11 years of education at the time of the task, as during the first year of pre-school none of the students had been introduced to the English language.

To calculate their English exposure, the pre-school and primary education weekly sessions and the secondary education weekly sessions were done separately, as the students had received a different number of sessions per week during those two academic levels. As such, the non-CLIL group received 5 sessions of EFL per week during pre-school and primary education, and 3 sessions per week during secondary education's first two academic years. On the other hand, the CLIL group had been given 5 sessions per week in EFL and 1 session per week of English as a content language (ECL) subject during pre-school and primary education. As for secondary education, they received 3 sessions per week in EFL and 3 sessions per week in ECL. Thus, the sessions in each academic level were multiplied by the weeks of instruction of each academic year (37 weeks) and the

resulting number, in turn, by the number of academic years: 8 years for pre-school and primary education and 2 years for secondary education.

- The non-CLIL group comprised 10 students, 9 boys and 1 girl. They all had attended a pre-school and primary school which offered them English as EFL, from which they had received a total amount of 1480 hours of said language, and for the first two academic years of secondary education they received 222 hours of instruction. In total, they had been exposed to 1720 hours of English, as EFL instruction.
- The CLIL group comprised 10 students, 5 boys and 5 girls. They had attended a pre-school and primary school which offered them English through a bilingual programme, which meant that they had EFL and different subjects along the academic years as ECL. During pre-school and primary school education, they had received a total of 1776 hours of English. In secondary education, they had received an equal amount of English as EFL as their non-CLIL peers (222 hours) and 222 hours of ECL. In total, they had been exposed to 2220 hours of English, both as EFL and ECL.

Table 1.

Study participants' characteristics

Group	Languages in their educational system	Participants	English exposure per week (session)		Total amount of English exposure (hours)		
			Pre-school and primary education	Secondary education	EFL	ECL	Total
non-CLIL	Basque, Spanish, English	10	5 EFL	3 EFL	1702	-	1702
CLIL		10	6 (5 EFL + 1 ECL)	6 (3 EFL + 3 ECL)	1702	518	2220

Overall, the CLIL group has been exposed to 518 more hours of English than their non-CLIL peers during their 11 years of schooling.

It must be noted that the participants in the CLIL programme were likely enrolled in the programme during pre-school by their parents. However, upon transitioning to secondary education, they had the choice to either remain in the CLIL programme or switch to the non-CLIL programme. Given that they were 11 or 12 years old at the time of enrolment in secondary education, the fact that they stayed in the CLIL programme suggests that they were willing participants who acted on their motivation and commitment to learning the

English language at a deeper level, as they use it in different content subjects. It is important to note that from the first to the second year of secondary education, students, their parents, and their teachers had the option to discontinue the students' participation in the CLIL programme. This shows the motivation, commitment and language proficiency of the participants who stayed in the CLIL programme during the second year of secondary education, when the study was conducted.

3.2. Task and procedure

For this study, Heaton's "Bicycle" (1966) was used as the task to be done by the participants. This is an integrated task that consists of a six-frame picture prompt, providing an extended stimulus for the participants to narrate a story, which aims to build participants' confidence in oral production (Iwashita et al., 2008). The participants were introduced to the task by the researcher describing the first picture (see Appendix I, section 9.1.2), after which they were asked to continue narrating the story.

The participants in this research study willingly took part in a task that involved different types of activities designed to assess their oral proficiency skills. Prior to engaging in the task, the researcher thoroughly explained the purpose of the study and the significance of having an equal representation of participants from both the non-CLIL and the CLIL programmes. The voluntary nature of participation ensured a higher level of commitment, motivation, and engagement among the participants. By adhering to these principles, we aimed to create an environment that facilitated accurate assessment of the participants' oral proficiency skills and provided valuable data for our research study. It is important to note that a letter of consent was not obtained from the participants' parents. This decision was made based on the information received from the high-school, which informed us that the parents had already provided consent by signing a letter at the beginning of the academic year. Said consent granted permission for the school and its staff to record and photograph their children. Additionally, the school explicitly stated to us that audio recordings of the students could be made for research purposes, on the condition that the recordings and the students' names would not be shared publicly.

The task employed in this study was carefully selected based on its specific characteristics. Firstly, it is a reasoning task (Prabhu, 1987). Unlike opinion or information-gap tasks, this particular one required participants to interpret and comprehend the meaning of the story. By engaging in this task, the learners were challenged to

demonstrate their ability to extract meaning and make inferences, thereby showcasing their oral proficiency in a context that demanded comprehension and reasoning skills. Secondly, it is an unfocused task (Ellis, 2009), as it did not explicitly emphasize any particular language feature or grammatical structure. Instead, its purpose was to create a communicative setting where participants could freely express their thoughts and ideas without the pressure of focusing on specific linguistic elements. By utilizing this task, we aimed to observe the participants' overall oral proficiency and their ability to communicate effectively in a spontaneous and unrestricted manner. Additionally, it is a divergent task (Pica et al., 1993), aimed to create a communicative situation where both speakers did not necessarily need to reach a common agreement or consensus. By engaging in this task, the participants were encouraged to express their individual viewpoints, fostering authentic interaction and promoting the development of oral proficiency in a context that allows for divergence and a variety of perspectives and ideas.

3.3. Data collection and analysis

Data collection for this research study took place over three lesson hours at the end of April 2023. The collected data were subsequently assessed by both the researcher and an English teacher from the participating schools. To ensure accurate evaluation, the students' oral productions were recorded using a mobile phone device and later reviewed by the raters. This approach allowed for a thorough analysis of the students' performance, enabling the researcher and the English teacher to provide comprehensive and detailed assessments of their oral proficiency. To evaluate the students' performance, a 4-point scale rubric, adapted from Azpilicueta Martínez (under review) was utilized (see Appendix I, section 9.1.3.), encompassing five key subskills: coherence, fluency, grammar, pronunciation, and vocabulary.

The qualitative data collected for this study were meticulously recorded and organized using Microsoft Excel. All pertinent information, including the participants' CLIL exposure and their scores in oral proficiency, was documented in Excel spreadsheets. By utilizing Excel's built-in functions, the mean score for each subskill was calculated based on the previously established 1 to 4 value range. However, for the purpose of assessing the participants' overall oral proficiency, the decision was made to convert these values to a more intuitive 0 to 10 point scale. This conversion was implemented as it provided a clearer and more easily understandable representation of the participants' proficiency levels. Furthermore, Excel's graphing capabilities were harnessed to create visual representations such as diagrams and

charts, which proved invaluable in illustrating the observed patterns and trends in the data. These visual aids facilitated a comprehensive and insightful analysis of the findings. Overall, Microsoft Excel served as a powerful and versatile tool for data collection, analysis, and visualization in this research study, ensuring a robust and rigorous examination of the collected data. Nonetheless, it should be mentioned that, during the process of distributing colour-coded charts for each subskill in both the non-CLIL and CLIL groups, we encountered a challenge. Several participants from each group would achieve the same result, requiring us to determine how to assign colours to differentiate the bars. To address this issue, we adopted a specific approach that involved intermixing the participants to a certain extent. For instance, in the vocabulary chart (see section 4.6.) we had 9 participants with 3 points, of which 3 were from the non-CLIL group and 6 from the CLIL group. There, we placed the 3 members of the non-CLIL group in between the other 6, resulting in 3 bars from the CLIL group at the beginning and 3 bars at the end of the 3-point range. This arrangement aimed to ensure a certain balance when analysing and discussing the top performances in each subskill. Moreover, we have incorporated a box plot for the comprehensive analysis of the subskills. This graphical representation will offer a visual depiction of the distribution and variety of scores within each subskill. By representing the results in a box plot, we can easily observe the range, median and mean scores for each subskill, thereby enhancing our understanding and interpretation of the data.

For the discussion section, we have incorporated qualitative excerpts from the students' audio recordings into the analysis. This approach has provided a comprehensive representation of the participants' English oral production, allowing for a direct comparison between the overall outcomes and conclusions reached by experts and researchers mentioned in the literature review and the actual performance of this study's participants. By integrating these excerpts, a more nuanced understanding of the results has been obtained, enabling a thorough examination of the students' language proficiency.

4. RESULTS

In this section, we aim to address our research question: the impact of CLIL exposure on oral proficiency in secondary-school learners. Firstly, we will present the overall results encompassing all five subskills. Next, we will examine the specific results for each subskill, namely, coherence, fluency, grammar, pronunciation, and vocabulary. Additional information regarding gender will also be included. Tables and charts from Excel will be utilized to visually represent the differences and similarities between groups.

4.1. Global results

Substantial differences were observed between the non-CLIL and CLIL groups on a global level, as the difference between the groups was calculated to be 1.67. The non-CLIL group exhibited the lowest proficiency score (5.87 out of 10), while the CLIL group exhibited the highest score (7.55 out of 10). Moreover, out of the top quartile comprising 10 individuals, 8 participants belong to the CLIL group, and additionally the highest 25% of performers are from this group. Consequently, the overall findings highlight a significant presence of CLIL group participants among the top scorers in our study. These results are remarkable and offer insights into the performance of both groups within the score of our research.

Table 2.

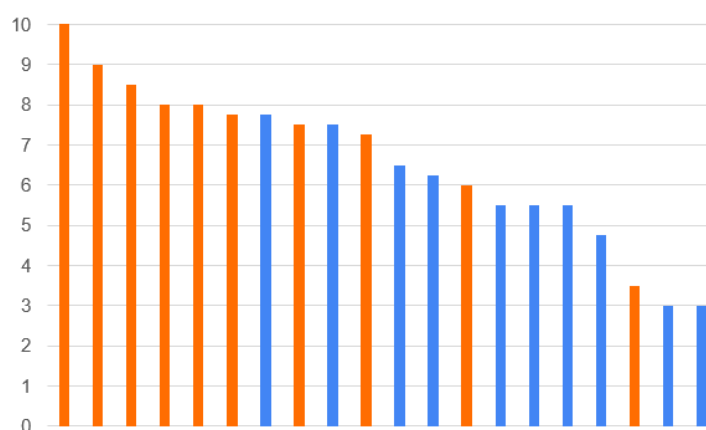
Global results

non-CLIL	5.88
CLIL	7.55

Global results for both groups

Fig. 1.

Global results



Scale representing global results of each participant for both groups (non-CLIL, blue; CLIL, orange)

4.1.1. Proficiency scores for each subskill

In terms of scores of the subskills, it is worth mentioning the separation between both groups with respect to the results obtained, with the non-CLIL group achieving scores in the lower range, while the CLIL group consistently scored higher in each subskill, occupying the upper half of the score distribution. Furthermore, focusing on the score for each subskill in both groups, it is noteworthy that pronunciation emerged with the highest result for both the non-CLIL and CLIL groups, obtaining a score of 3.4 and 2.75 respectively. However, the remaining subskills do not align between the two groups. Interestingly, the second highest score for the CLIL group is fluency (3.05), which represents the lowest score for the non-CLIL group (2.00). Similarly, the third highest score for the CLIL group is grammar (2.95), corresponding to the second lowest score for the non-CLIL group (2.05). This consistent pattern extends to the other subskills as well.

Table 3.

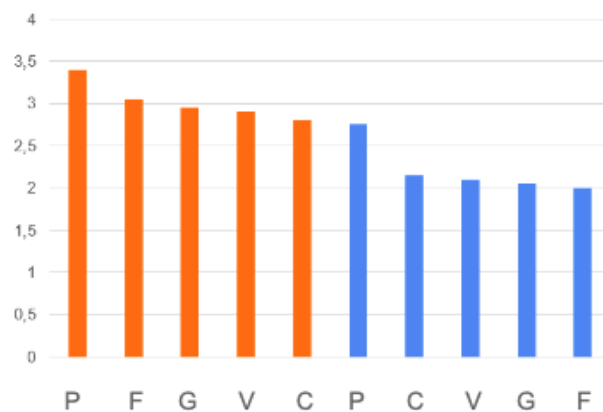
Global proficiency scores for each subskill

	Coherence	Fluency	Grammar	Pronunciation	Vocabulary
non-CLIL	2.15	2.00	2.05	2.75	2.10
CLIL	2.80	3.05	2.95	3.40	2.90

Table representing global proficiency scores for each subskill for both groups

Fig. 2.

Global proficiency scores for each subskill



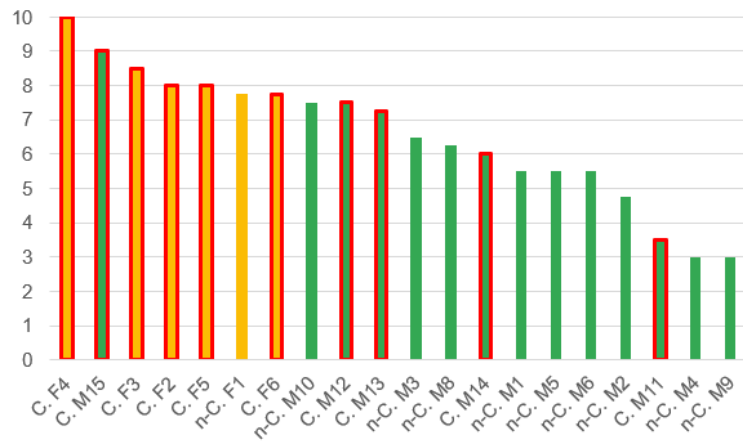
Scale representing decreasing gradient in the results for each subskill (coherence (C), fluency (F), grammar (G), pronunciation (P), vocabulary (V)) for both groups (non-CLIL, blue; CLIL, orange)

4.1.2. Distribution by gender on global results

Although not part of our research question, we wanted to point out the distribution of scores by gender on the overall scale, as the results in this respect were substantial. In the top quartile (10 individuals) 6 of the participants were female, representing moreover the total number of female participants. This gender distribution indicates an interesting representation of female participants among the highest scorers in our study. These results are noteworthy and provide information about the performance of the different genders in the context of our research.

Fig. 3.

Distribution by gender on the global results



Scale representing the distribution by gender (female, yellow; male, green) on the global results for both groups (non-CLIL, no outline; CLIL, outlined in red)

4.2. Coherence

The difference between the two groups represented a value of 0.65 in coherence subskill. The non-CLIL group achieved a score of 2.15 out of 4, while the CLIL group scored 2.80 out of 4. Nonetheless, in line with the findings of the global results, within the top 5 highest performers, and representing the highest-achieving students in the subskill of cohesion, all but one student belong to the CLIL group. Thus, the results pertaining to the coherence subskill reveal a notable presence of CLIL learners among the top performers in our study.

Table 4.

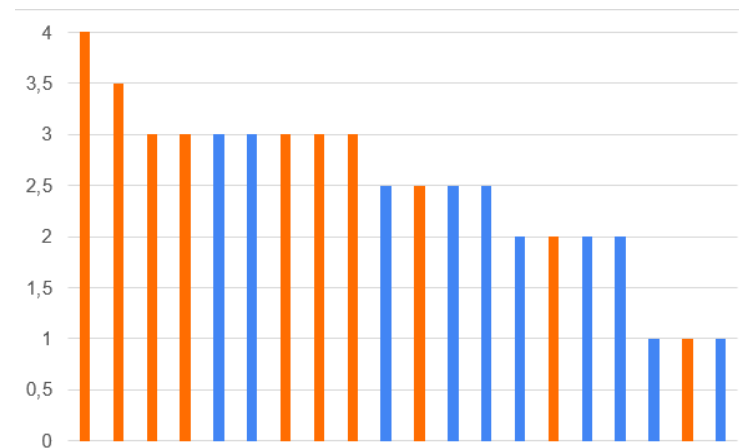
Coherence results

non-CLIL	2.15
CLIL	2.80

Coherence results for both groups

Fig. 4.

Coherence results



Scale representing coherence results of each participant for both groups (non-CLIL, blue; CLIL, orange)

4.3. Fluency

The disparity between the two groups in fluency measures amounted to a value of 1.05. The non-CLIL group attained a score of 2.00 out of 4, while the CLIL group achieved a higher score of 3.05 out of 4. However, consistent with the global results, all 5 top-performing participants, accounting for 25% of the highest-performing students, exclusively belong to the CLIL group, showcasing their exceptional performance in the fluency subskill. Hence, the findings demonstrate a solid presence of CLIL learners among the highest achievers of the fluency subskill in our study.

Table 5.

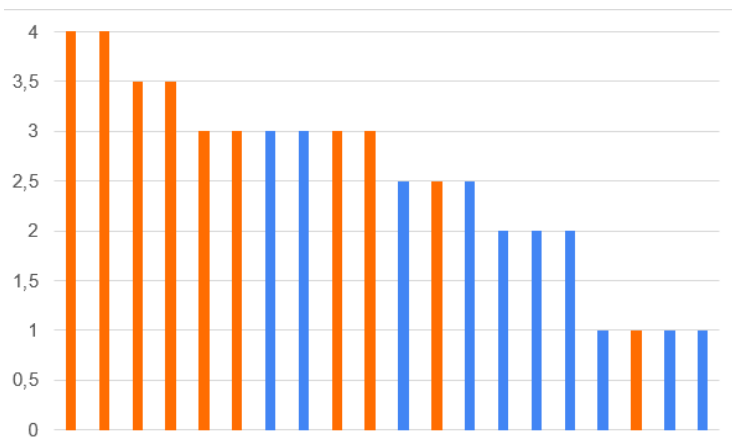
Fluency results

non-CLIL	2.00
CLIL	3.05

Fluency results for both groups

Fig. 5.

Fluency results



Scale representing fluency results of each participant for both groups

(non-CLIL, blue; CLIL, orange)

4.4. Grammar

There was a 0.90 difference observed between the two groups in the grammar subskill. The non-CLIL group obtained a score of 2.05 out of 4, while the CLIL group achieved a higher score of 2.95 out of 4. Regardless, in alignment with the overall results and the previous subskill findings, it is worth noting that the CLIL group had a remarkable representation among the top performance in the grammar subskill. Specifically, all 5 top-performing participants, constituting the 25% highest performances, belong to the CLIL group. These results highlight the exceptional performance of CLIL learners in terms of grammar. Therefore, our findings provide strong evidence of the prominent presence of CLIL learners among the highest achievers in the grammar subskill within our study.

Table 6.

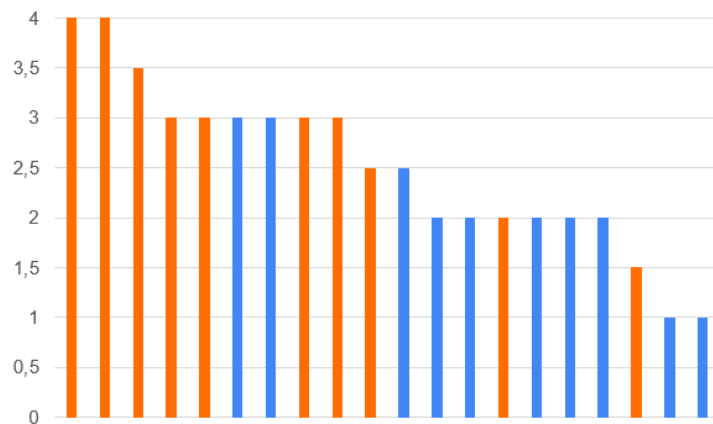
Grammar results

non-CLIL	2.00
CLIL	3.05

Grammar results for both groups

Fig. 6.

Grammar results



Scale representing grammar results of each participant for both groups (non-CLIL, blue; CLIL, orange)

4.5. Pronunciation

The contrast of the two groups in pronunciation measures yielded a value of 0.65. The non-CLIL group received a score of 2.75 out of 4, whereas the CLIL group demonstrated a higher score of 3.40 out of 4. Nonetheless, in accordance with the global results and the previous subskill findings, it is worth highlighting that all participants within the top 5 highest performers exclusively belong to the CLIL group. This fact serves as compelling evidence of their exceptional performance in the pronunciation subskill. Consequently, the results establish a substantial presence of CLIL learners among the top achievers in the pronunciation subskill in our study.

Table 7.

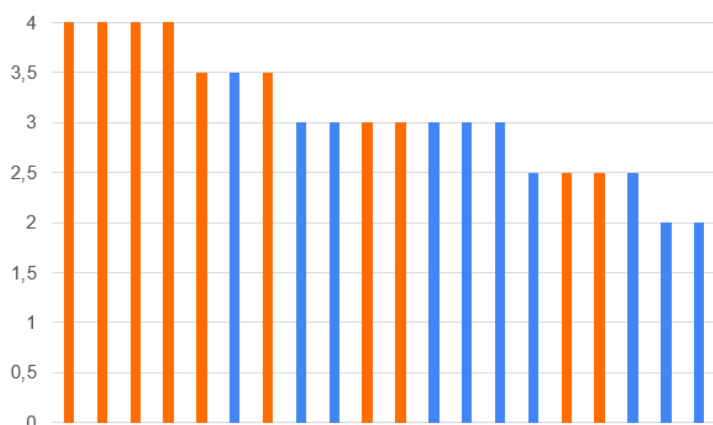
Pronunciation results

non-CLIL	2.75
CLIL	3.40

Pronunciation results for both groups

Fig. 7.

Pronunciation results



Scale representing pronunciation results of each participant for both groups
(non-CLIL, blue; CLIL, orange)

4.6. Vocabulary

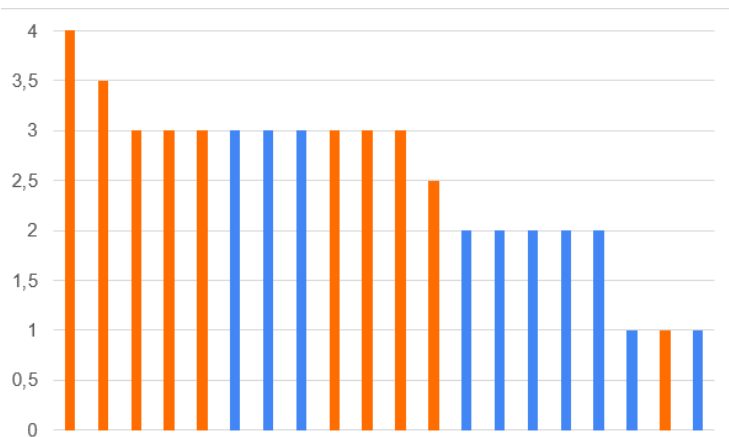
A value of 0.80 indicates the variance between the two groups in vocabulary measures. The non-CLIL group obtained a score of 2.10 out of 4, while the CLIL group demonstrated a higher score of 2.90 out of 4. Nevertheless, in line with the global results and previous subskill findings, it is notable that the participants constituting the 25% of the highest performers belong exclusively to the CLIL group. These findings strongly support their exceptional performance in the vocabulary subskill, providing compelling evidence. As a result, the scores confirm a substantial representation of CLIL learners among the top achievers in the vocabulary subskill within our study.

Table 8.
Vocabulary results

non-CLIL	2.10
CLIL	2.90

Vocabulary results for both groups

Fig. 8.
Vocabulary results



Scale representing vocabulary results of each participant for both groups
(non-CLIL, blue; CLIL, orange)

4.7. Subskill results in a box plot

Following the presentation of proficiency scores for each subskill, which provided an overview of both group performances, as well as the individual participant results for each subskill, we will now visualize the distribution of scores for each subskill using a box plot. This graphical representation will offer a comprehensive view of the results for each specific subskill.

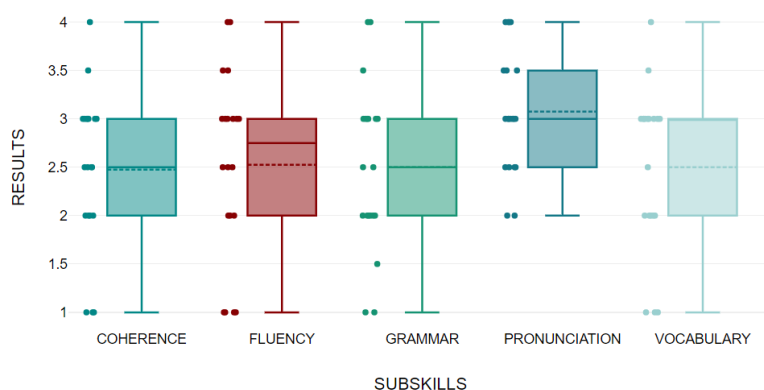
The box plot reveals interesting insights into the distribution of scores. Notably, there is a noticeable contrast in the distribution for pronunciation compared to the other four subskills (coherence, fluency, grammar, and vocabulary). 50% of participants were placed between 2.5 and 3.5 marks for the pronunciation subskill. The lower end of the bottom whisker for pronunciation starts at 2 points, which is the highest among the four boxes, indicating the lowest scores within the 50% interquartile range. However, it is important to note that the upper whisker end for all the subskills reaches 4 points, pointing at a widespread distribution in all subskills.

Furthermore, taking into account the mean score for each subskill as a whole, as in section 4.1.1. it has been mentioned the mean score for each group in each subskill. It is variable across the different subskills: coherence has a mean score of 2.47, fluency of 2.52, grammar of 2.50, pronunciation of 3.07, and vocabulary of 2.50. Overall, the mean scores fluctuate around the 2.50 points for all subskills except pronunciation, not only placing the subskills in the top quartile, but highlighting the pronunciation subskill.

Additionally, it is important to discuss the median score for each subskill. The media score represents the middle value of the distribution, indicating the point where 50% of the scores fall below and 50% fall above. The medial score is more variable in our results than the mean score. Thus, the coherence subskill has a medial score of 2.50, fluency of 2.75, grammar of 2.50, pronunciation of 3, and vocabulary of 3.

Fig. 8.

Box plot of subskills



Boxplot representing all the subskill results

5. DISCUSSION

The aim of this research study was to investigate the impact of CLIL exposure on the oral proficiency of secondary-school learners. It is crucial to comprehend the effects of CLIL exposure, or the absence thereof, on oral proficiency among secondary-school learners due to the growing prevalence and intensification of CLIL programmes, as well as their introduction at an early age, which plays a vital role in optimizing language acquisition and proficiency (Admiraal et al., 2006; Housen, 2012). Moreover, by conducting this research, we aimed to contribute to the existing body of literature by examining the specific impact of CLIL exposure on the oral proficiency of secondary-school learners. This is particularly important as CLIL programmes continue to expand and evolve, warranting a comprehensive understanding of their effects on language development and academic performance.

5.1. Overall oral proficiency

Our initial focus will be on the examination of the overall oral proficiency findings, aligning with previous research, in the context of secondary-school learners, that highlights the substantial language gains achieved through CLIL exposure exceeding 300 hours, establishing this duration as a crucial threshold for optimal outcomes (Muñoz, 2015). Additionally, various studies have studied CLIL instruction as a positive impact on language acquisition, proficiency, and academic achievement compared to learners on traditional EFL instruction (Admiraal et al., 2006; Goris et al., 2019). In line with these previous findings, our study's global results from the oral proficiency task reveal a notable advantage for the CLIL group in terms of English oral productive skills, underscoring the substantial impact of CLIL instruction on students' oral productive skills in English (see Fig. 1., section 4.1.). These differences can be attributed to the additional 518 hours of English exposure of the CLIL group. Such a substantial difference contributes to the development of greater fluency, accuracy, and complexity in the spoken English of CLIL learners (Muñoz, 2015).

Transitioning to the specific oral proficiency outcomes, our findings align with previous studies that have demonstrated the positive impact of CLIL instruction on students' oral proficiency skills in secondary education (Admiraal et al., 2006; Gálvez Gómez, 2021). Specifically, our results indicate that the CLIL group exhibited substantially greater improvement in their oral-production task compared to the non-CLIL students (see Table 2 and Fig. 1., and Table 3 and Fig. 2., sections 4.1. and 4.1.1.). These outcomes provide additional evidence supporting the effectiveness of CLIL in enhancing oral proficiency among secondary-school learners. Our study provides further evidence supporting the positive

impact of CLIL instruction on the oral proficiency of secondary-school students. The results highlight the importance of integrating CLIL programmes into secondary education curricula to enhance students' oral communication skills. By contributing to the existing literature, this research expands our understanding of the relationship between CLIL and oral proficiency, emphasizing the need for continued exploration in this area.

After reviewing the overall oral proficiency, we will now proceed to examine each subskill, providing relevant participant excerpts to substantiate our conclusions. It is important to note that, as evidenced in section 4.1.1., the subskill scores reveal intriguing patterns that differentiate the non-CLIL and CLIL groups. Pronunciation stands out as the highest-ranked subskill for both groups, indicating a shared proficiency. This raises questions about the positive impact of CLIL instruction on pronunciation skills. In contrast, the alignment of other subskills varies between the groups. Coherence ranks second highest for the non-CLIL group but lowest for the CLIL group, while fluency shows the opposite trend. These findings call for further investigation into the influence of CLIL instruction on specific subskills and the divergent trajectories observed between the groups. As for the box plot analysis of the results, seen in section 4.7, there is a contrast in score distribution, with pronunciation standing out as a subskill where participants performed relatively better compared to the other subskills (coherence, fluency, grammar, and vocabulary), both in terms of overall distribution and mean and median values. Overall, these findings suggest that pronunciation is an area of relative strength for participants, while the other subskills show varying levels of performance. The variability in median scores underscores the need to consider individual differences in language learning. It is crucial for language instruction to leverage participants' strengths in certain subskills, and provide targeted support for areas of challenge. By doing so, language educators can promote a more balanced development of oral proficiency among learners.

5.1.1 Coherence

In regards to coherence, our study did not uncover substantial disparities in the measures of this subskill between the non-CLIL and CLIL groups. Although the CLIL group achieved higher scores, the difference between the two groups was only 0.65. It is interesting to note that the coherence subskill ranked second highest for the non-CLIL group, while it ranked the lowest for the CLIL group. These findings indicate that while there is a slight advantage for the CLIL group, the difference in these measures is not substantial, as we have considered the threshold for notable difference on values exceeding 1 point.

In the broader context, it is important to acknowledge the lack of specific research focusing on coherence and revealing substantial differences between non-CLIL and CLIL participants. Although the lack of specific research limits our ability to draw definitive conclusions, our findings provide valuable insights into the coherence subskill's performance within the context of CLIL instruction. The presence of CLIL learners among the top performers suggests that CLIL instruction may have a positive influence on coherence in oral proficiency. However, further research is needed to fully understand and establish the relationship between CLIL instruction and coherence in oral production.

This is exemplified in the following excerpts, the first one narrating the whole comic strip, and the second one describing captions number 2, 3, and 4.

Example (1).

M7 (non-CLIL).

M7: (...) The driver goes with the car and pass... *nola esaten da al lado* [how do you say next to]?

Researcher (R): Next to.

M7: Next to Henry, and Henry goes to the floor. So, he tries to... *konpondu* [fix].

R: Fix.

M7: Fix the bicycle and later he passes with the bicycle from the driver car and is (...) *eh* (...) *roto* [broken].

R: Broken.

M7: Broken. So, he pass. The driver is angry and Henry is fun, happy.

F2 (CLIL):

When he, Henry, hears the honking, he get, he goes to a bush and he falls down. *Eh*, and in the next, the car continue riding, or driving, and Henry *eh*, is in the grass, or *eh*, the trees, or uhm, green things, and in the next picture, the bicycle is broken, and he has some blood in his legs.

Analysing the examples provided, it is noteworthy to consider M7's lack of cohesive elements, such as connectors or other coherence devices, to distinguish between the various captions in the narration. The students' approach to the task is characterized by a brief, superficial production, where most of the captions are brushed over. In contrast, F2 presents a clear and coherent narrative, effortlessly conveying the sequence of events. Notably, F2 utilizes narrative linkers and conjunctions, creating a continuous and intelligible narration that demonstrates cohesion.

However, it is important to acknowledge that this pattern was less recurrent compared to other subskill patterns observed in our study. Both non-CLIL and CLIL students displayed the use of narrative linkers, albeit without a substantial prevalence between the two groups. While the presence of narrative linkers indicates some level of cohesion in the students' oral production, it does not significantly differentiate both groups in terms of the coherence subskill. Furthermore, it is worth considering that many students from both groups tended to gloss over certain captions, resulting in less cohesive narrations, a technique called avoidance. This observation suggests a potential lack of descriptive skills among the students, hindering their performance in this particular subskill, as the omission or superficial treatment of captions may contribute to a less coherent overall narrative, impacting the students' ability to establish connections and maintain a cohesive structure.

In conclusion, our study highlights the importance of investigating the impact of CLIL instruction on the coherence subskill in oral proficiency. Although the observed differences between the two groups were not substantial, they contribute to the limited body of research in this area. The presence of CLIL learners among the top performers suggests the potential benefits of CLIL instruction on coherence. Future research should continue to examine the specific effects of CLL instruction on coherence in oral proficiency to deepen our understanding of its impact in language learning context.

5.1.2. Fluency

Focusing on the fluency subskill, our findings reveal substantial differences in fluency measures between the non-CLIL and CLIL groups, with the latter exhibiting higher scores. The difference results in a value of 1.05, which surpasses the previously established threshold of 1 point difference. The results from our study align with previous research conducted by various experts in the field (Gallardo del Puerto & Gómez Lacabex, 2017; Gálvez Gómez, 2021; Goris et al., 2019; Housen, 2012), reporting significant results favouring CLIL for spoken fluency, emphasizing the increased opportunity for authentic communication as a contributing factor.

In our study, we found a significant representation of CLIL learners among the highest achievers in the fluency subskill. All participants in the top 5 highest performances exclusively belonged to the CLIL group, highlighting their exceptional performance in fluency. Moreover, as expressed in section 4.3., this subskill was the second-highest for the CLIL group, representing on the other has the lowest scored subskill for the non-CLIL group.

These findings further support the notion that CLIL instruction positively impacts students' fluency in oral proficiency.

This can be further illustrated in the following excerpts, which belong to the narration of the second caption of the task.

Example (2).

M9 (non-CLIL).

R: Now, can you tell the rest of the story?

M9: Henry (...) *pues* [well] (...) *pues* [well] (...) *apartatzen da* [moves away].

R: In English?

M9: (...) He (...).

R: What happened to the car?

M9: Eh (...) *adelantarle, no sé como se dice* [overtake him, I don't know how to say it].

R: Is he driving...?

M9: Is driving very quickly (...).

R: And then, the bike...?

M9: *pues* [well] (...) *erortzen da* [it falls] (...) Henry uhm uhm (...)

R: Did he fall off of the bike?

M9: Yes, yes.

R: Ok, can you say it?

M9: Henry fall off to the bicycle.

[1 minute and 12 seconds of production]

M15 (CLIL).

[...] The man that is driving his car, *eh*, goes faster to pass Henry, and Henry goes to the, to the plants that are in his left.

[15 seconds of production]

Note how M9 did not have a fluent oral production, as he produced a single sentence for the narration of the second caption and it took him more than one minute to do it. His production was characterized by frequent pauses and the use of verbs and filler words from both Basque and Spanish. The researcher had to provide assistance to facilitate his language production. In contrast, M15 took approximately a minute less than his partner to narrate the first caption, delivering a fluent speech without any pauses or the need for intervention from the researcher. This example exemplifies the recurring pattern observed during the narration of the second caption in the task. It is plausible to attribute this to the fact that the students were not yet acquainted with the task and needed time to process and formulate their thoughts while speaking about it.

In analysing the oral production of the participants, a clear pattern emerged, showcasing notable differences between the non-CLIL and CLIL students. Specifically, the non-CLIL students demonstrated a tendency to produce incomplete sentences and rely on more frequent pauses, struggling to maintain a smooth flow of language, while the CLIL students exhibited fewer interruptions in their speech displaying a higher level of fluency. The recurrent nature of these findings highlights the impact of CLIL instruction on the fluency of the students. CLIL students consistently demonstrated a higher level of fluency and a more proficient use of the TL, whereas non-CLIL students exhibited difficulties in maintaining fluency in their oral productions. These observations shed light on the importance of CLIL programmes in promoting fluency and language skills among secondary-school students. By providing students with immersive language learning experiences, CLIL instruction facilitates the development of fluent and cohesive oral proficiency.

Overall, our study, along with the existing literature, underscores the positive influence of CLIL instruction on the fluency subskill in the oral proficiency of secondary-school students. The consistent results from various researchers strengthen the evidence base for the effectiveness of CLIL in enhancing fluency in language learning context.

5.1.3. Grammar

In terms of the grammar subskill, our study did not find significant differences in grammar measures between the non-CLIL and CLIL groups, although it must be noted that the difference in value is 0.90 in favour of CLIL, which is merely 0.10 points away from the 1 point threshold. This implies a narrow margin separating both groups, suggesting a potential advantage for the CLIL group in terms of grammar. Moreover, the presence of CLIL learners among the top performers supports the notion that CLIL instruction has a positive impact on students' grammatical accuracy in oral proficiency. This finding aligns with the results of other researchers (Gallardo del Puerto & Gómez Lacabex, 2017; Gálvez Gómez, 2021; Housein, 2012), further emphasizing the potential benefits of CLIL instruction. While our specific findings did not uncover notable distinctions between the non-CLIL and CLIL groups, the collective evidence suggests that CLIL instruction can enhance students' grammar skills.

This can be further illustrated in the following excerpts, which belong to the narration of the last two captions of the task.

Example (3).

M4 (non-CLIL).

M4: Later, the (...) the car have a accident, and Henry, *eh*, (...), *eh*, *pasatzen da* [passes].

R: He passed?

M4: He passes in front of car, *eh* (...) *por al lado* [next to].

R: Next to.

M4: Next to the car, and Henry, *eh*, *pues*, *le pita* [well, he honks at him].

R: Rang.

M4: He rings to the conductor (...) *Eh*, she is happy.

R: And why do you think Henry was happy?

M4: Because, *eh* (...) I don't know.

R: Or, the driver, he was...

M4: The driver is sad, or (...) *enfadado* [angry].

R: Or angry, yes. And why do you think the driver was sad or angry?

M4: Because he, *osea* [I mean], because the car is broke.

F3 (CLIL).

He is standing up and with the bicycle, it's good I think, so (...) It's standing up and the car, *eh*, it's broken because he, *eh*, pulls with the bike and it breaks. And the car driving, *eh*, *osea* [I mean], *eh*, the person who drives the car is angry because the bicycle, or the boy, *nola* [like], *eh*, broken the car with her bike, or (...) with the *impacto*, *osea*, *choque* [impact, I mean, crash]. And Henry goes good with the bicycle, but the man was angry because he, *eh*, he pulls with the bike.

It is important to highlight the differences in grammatical structures used by M4 and F3. Despite the researcher's use of the past simple tense, both students in this example and most of the students during the task predominantly used the present simple tense to narrate the comic strip. This preference may be due to factors such as early introduction and familiarity with the present simple tense, limited exposure to diverse language input, and a lack of awareness or understanding of verb tense usage in narrations. Regarding the specific examples provided, M4's narration exhibited simpler grammatical structures with frequent errors, making comprehension challenging. On the other hand, F3 demonstrated a relatively more grammatically complex production by incorporating relative clauses. However, it is important to note that F3's narration occasionally hindered comprehension, due to instances of hesitation and self-correction. Nonetheless, she managed to construct coherent sentences that effectively conveyed her thoughts and ideas with correct and complex grammatical structures.

In conclusion, our study emphasizes the importance of investigating the impact of CLIL instruction on the grammar subskill in oral proficiency, as the presence of CLIL learners among the top performers highlights the potential benefits of CLIL instruction on grammar. The evidence overall supports the idea that CLIL instruction provides a more immersive language learning experience, resulting in improved grammatical accuracy. Further research is necessary to uncover the specific mechanisms through which CLIL instruction facilitates the acquisition of grammar skills.

5.1.4. Pronunciation

Focusing on the pronunciation subskill, our findings revealed no substantial difference between the two groups, as the difference in favour of CLIL yielded a 0.65 value. However, it is worth highlighting that all participants within the top 5 highest performers belong to the CLIL group, representing evidence that showcases the superior intelligibility of CLIL learners in the pronunciation subskill are consistent with the findings of other researchers who have explored the impact of CLIL instruction on pronunciation skills (Gallardo del Puerto & Gómez Lacabex, 2009; Gálvez Gómez, 2021; Swain, 2000), emphasized the benefits of CLIL instructions. These findings show that CLIL learners consistently outperformed their non-CLIL counterparts in various pronunciation measurements. Furthermore, as stated in this study, the importance of pronunciation skills in language learning cannot be overstated. Pronunciation plays a critical role in effective communication and can greatly impact how well a speaker is understood by their audience. The positive impact of CLIL instruction on English pronunciation is particularly noteworthy, as it suggests that this approach can effectively enhance overall oral proficiency. Moreover, Gallardo del Puerto and Gómez Lacabex (2017) emphasize the practical implications of improving pronunciation skills in language instruction. Enhanced pronunciation can boost learners' confidence, increase their intelligibility, and facilitate communication.

In relation to the students' production, as displayed in section 4.5., our study found no significant difference between the two groups in the pronunciation subskill, as the difference between the two was 0.65. Notably, both the non-CLIL and CLIL groups exhibited the highest performance in this subskill. Therefore, we will not provide a specific comparative example. However, we would like to highlight an interesting extract that illustrates a unique occurrence. During the study, a student (F5) was presented with a new word by the researcher and mispronounced it, instead saying a different word that she might have been more familiar with. This instance not only indicates a gap in the student's lexical knowledge, but also

emphasizes the complexity of pronunciation as a subskill. The excerpt's production occurs during the narration of the fourth caption.

Example (4).

F5 (CLIL).

F5. And, *eh*, in the next one, the bicycle is broken, and he, he has some (...) how do you say *heridas* [wounds]?

/ænd, e, ɪn ðə nekst wʌn, ðə 'baɪsɪkl ɪz 'brəʊkən, ænd hi:, hi: hæz sʌm (...) haʊ du: ju: sei e'riðə [wu:ndz]?/

R: Wounds

/wu:ndz/

F5: Wounds

/wu:ndz/

R: Wounds /wu:ndz/

F5: He has some wings in his legs, but he continue riding his bicycle.

/hi: hæz sʌm wɪŋz ɪn hɪz lægz, bʌt hi: kən'tɪnju: 'raɪdɪŋ hɪz 'baɪsɪkl./

The exceptional performance of both groups in the pronunciation subskill may be attributed to various factors. Firstly, it is possible that students in both groups had access to effective pronunciation instruction within their language learning programmes, emphasizing the importance of explicit pronunciation teaching. Additionally, the use of authentic materials, exposed to English-speaking contexts, and opportunities for oral communication in both onn-CLIL and CLIL settings could have contributed to the development of pronunciation skills. Furthermore, individual motivation and effort, as well as previous exposure to English outside the classroom, might have played a role in the students' proficiency in this subskill.

In conclusion, our study adds to the growing body of evidence that supports the positive impact of CLIL instruction on the pronunciation subskill, adding another layer, as the non-CLIL group also made a great performance on this subskill. The exceptional performance of CLIL learned in this subskill, as well as the findings of previous researchers, underscores the benefits of CLIL instruction in promoting pronunciation proficiency. However, the results request further research to investigate additional factors that may contribute to the success of learners in general, and CLIL learners specifically, in this subskill.

5.1.5. Vocabulary

The findings regarding the vocabulary subskill in our study indicate no significant difference in vocabulary measures between both groups, although it is important to highlight that the difference in value, favouring the CLIL group, is 0.80, only 0.20 points shy of the 1 point threshold. This indicates a close margin between the two groups, implying a potential advantage for the CLIL group in terms of vocabulary proficiency. Moreover, there is a substantial representation of CLIL learners among the top achievers in our study, evidencing the strongly supported exceptional performance of CLIL learners in the vocabulary subskill. These findings are consistent with previous research which demonstrate that CLIL instruction positively impacted lexical measures (Gallardo del Puerto & Gómez Lacabex, 2017; Housen, 2012) and lexical richness (Gálvez Gómez, 2021) in oral proficiency.

This can be further illustrated in the following excerpts, which belong to the narration of the last three captions of the task.

Example (5).

M5 (non-CLIL).

R: What happened to the car?

M5: The car go to his house, yeah. And he has, *eh*, a lot of *heridas* [wounds].

R: Wounds.

M5: Wounds. And the bicycle is, *eh*, I don't know, *eh*, *apurtuta* [broken].

R: Broken.

M5: Broken, the bicycle is broken. And he, *eh*, *konpondu* [fix].

R: He fixes it.

M5: He fixes the bicycle and continue with the *trayect*. And look the angry, the driver of the car, *eh*, fixed his car, because its bad. And he continue and the angry driver, *eh*, I don't know, *hasertzen da gehiago* [he gets more angry].

R: He got more angry.

M5: He gets more angry, yes.

R: And how was Henry in the last picture?

M5: *Pues*, *eh*, the, *eh*, boy is very happy and the angry driver no.

R: And why do you think he was happy?

M5: Because he is good and the driver is, *eh*, very angry, because its car is, *eh*, *apurtuta* [broken].

R: Broken.

M5: Yes, the car is broken.

F4 (CLIL).

Then, Henry is left behind, with a broken bike, and in the picture it shows that he is quite hurt.

And the plant also got a little bit hurt. Then, after a little bit more walking (...) well, bicycling, he finds the angry man with a broken car. Like, his engine broke down or something like that, and he honks, so the man gets angry. It is like a payback for what he did. So, I think that now the man, he is a little bit angry at him too, because he shouldn't have done that, and he shouldn't've just honk once and see what happen.

It is crucial to note the contrasting use of vocabulary between M5 and F4 in their narrations. M5 relies heavily on basic and simple lexical items, as evident from the incorrect use of adjectives and adverbs, as it could be the use of *bad* instead of *broken*, and *no* instead of *not*, and limited lexical range, as it could be seen in repetition of the researchers TL word for broken, which the participant forgets twice. On the other hand, F4 exhibited lexical richness and complexity in her narration. She incorporated narrative linkers, adverbs and verbal expressions such as *like*, which enhanced the overall completeness and coherence of her production, as well as a wide range of correct verb tenses. Furthermore, F4 utilized a word, *payback*, which is typically associated with higher levels of English proficiency, demonstrating advanced lexical repertoire and her ability to convey meaning effectively.

The discrepancies in vocabulary use between M5 and F4 can be attributed to several factors. Firstly individual differences in vocabulary acquisition and retention might have played a role, as it is possible that F4, along with other CLIL group participants, have a stronger foundation in vocabulary knowledge or have more exposure to English language through the CLIL programme, or resources outside the classroom. In contrast, M5, along with other non-CLIL group participants, may have faced challenges in expanding their lexical repertoire, due to the reduced number of hours on English exposure, resulting in a more limited use of vocabulary.

To further support these observations, future research should investigate the influence of vocabulary instruction strategies, individual differences, and language exposure on the development of lexical knowledge in CLIL contexts. Overall, the evidence supports the notion that CLIL instruction provided a more immersive language learning experience, leading to improved lexical complexity and richness. Further research is necessary to explore the specific mechanisms through which CLIL instruction facilitates vocabulary acquisition and to gain a deeper understanding of its impact on students' oral proficiency in secondary-school settings.

6. CONCLUSION, LIMITATIONS AND PEDAGOGICAL IMPLICATIONS

6.1. Conclusion

The present study contributes to our understanding of the impact of CLIL exposure on the oral proficiency of secondary-school learners. Our findings demonstrate that CLIL programmes, particularly for learners who have previously experienced CLIL during their primary education, result in notable improvements in oral language skills. Specifically, our results indicate that the additional 518 hours of CLIL instruction received by the CLIL group led to substantial gains in oral production across all analysed measures when compared to the non-CLIL group. Moreover, certain subskills exhibited substantial differentiation between the two groups, aligning with Muñoz's assertion (2015) that surpassing the threshold of 300 hours of CLIL exposure may be sufficient for secondary-school learners to demonstrate significant language gains. These findings highlight the importance of extended exposure to CLIL instruction in secondary-school settings and emphasize its positive impact on learners' oral proficiency.

Our research study aligns with the findings of Goris et al. (2019) in the context of Spain, and more particularly in a trilingual (Basque, Spanish, English) secondary-school in Navarre. Like that study, we observed substantially positive results in our findings, as the participants were from Spain, which contrasts with the null effects of CLIL in some skills reported in other European countries (e.g. Housen, 2012). This disparity has been attributed by Goris et al. (2019) to the historical development of CLIL in Spain, where it was introduced as a top-down initiative by educational authorities with the aim of providing better EFL learning opportunities for all students, in response to the minimal quality of EFL teaching and limited access to EFL training experienced by Spanish children. In contrast, high EFL-proficiency countries with selective CLIL programmes, such as the Netherlands and Germany, showed limited gains in terms of oral proficiency. Our study reinforces the notion that the success of CLIL partly depends on the context in which it is implemented, with Spain being an example of a country where CLIL was introduced in fertile soil, addressing the need for improved EFL teaching in an increasingly internationalized market. The experience of teaching content through two languages in Spanish bilingual regions, such as the Basque Country or Navarre, might further enhance the potential for L3 acquisition, particularly in the case of CLIL and English language learning. Therefore, our research study, along with the findings of Goris et al. (2019), supports the efficacy of CLIL in enhancing oral proficiency, particularly in contexts with low EFL proficiency and a strong foundation in bilingual educational practices.

6.2. Limitations

In our study on the impact of CLIL on oral proficiency, it is important to acknowledge the limitations inherent in conducting research within a specific educational context and with a particular sample of students. Generalizing our findings to broader contexts and diverse populations may be challenging due to the unique characteristics of each educational environment, such as school policies, curricular constraints, and teacher expertise. Therefore, caution is needed when applying our results to other educational contexts. To enhance the validity and significance of the observed positive effects of CLIL on oral proficiency, future research should expand the scope of investigation to include a wider range of educational settings and diverse student populations. Exploring CLIL implementation in different contexts, such as primary schools, higher education institutions, or language academies, can provide insights into variations in the effectiveness of CLIL across different age groups and educational levels. By including a more diverse participant pool, future studies can offer a richer perspective on the role of CLIL in enhancing oral proficiency and contribute to the knowledge base in language education.

Moreover, while our study provides valuable insights into the positive effects of CLIL on oral proficiency, it is crucial to recognize the limitations imposed by our specific educational context and sample. Our findings should be interpreted within the parameters of our study, and caution should be exercised when generalizing the results. Future research should aim to investigate a broader range of contexts, diverse participants, and explore the interplay between content and language instruction. By doing so, we can further enhance our understanding of the potential benefits and implications of CLIL in promoting oral proficiency in language learning contexts.

6.2.1. *Variable limitations*

Limitations of our study should be acknowledged as they contribute to the observed differences in oral proficiency between non-CLIL and CLIL groups. The relationship between CLIL exposure and language gains is complex, making it difficult to attribute the influence solely to CLIL instruction. Factors like SES, extramural exposure, and “CLIL selection” may have impacted our results. Previous research shows the significance of SES, extramural exposure, and the potential bias introduced by the CLIL selection process (Admirall et al., 2006; Dalton-Puffer, 2007; Gallardo del Puerto et al., 2009; Peters, 2008; de Smet et al., 2019). The decision to enroll in the CLIL programme is influenced by students, parents, and teachers, potentially favouring students with higher language abilities and motivation. These limitations should be considered when interpreting our findings. Future research should

address these factors and incorporate appropriate controls to provide a comprehensive understanding of CLIL's effects on language acquisition and proficiency.

6.2.2. Gender limitations

Although not our primary focus, we analysed oral proficiency ratings by gender (see Results, section 4.1.2.). The prominence of female participants as the highest performers suggests underlying factors contributing to their academic success. Further exploration of these factors is needed to gain valuable insights. However, our analysis only captured the global distribution of scores and did not consider individual performances or other influential variables, such as SES, extramural exposure, and motivation. Additional analysis and factors should be considered for a comprehensive understanding. Moreover, the gender composition varied between the non-CLIL and CLIL groups, the latter consisting of an equal proportion of male and female participants, while the former had only one female participant. This discrepancy introduced a potential bias, as non-CLIL participants were more likely to achieve lower results. The uneven gender distribution between the groups may have influenced the observed differences in oral proficiency between the non-CLIL and the CLIL groups. In conclusion, our study revealed a significant representation of females among the highest scorers, emphasizing the need for further investigation into gender dynamics in education to create inclusive learning environments.

6.2.3. Voluntary participation

Another limitation to be acknowledged was the voluntary nature of participation. While voluntary participation has its advantages, such as ensuring participants with necessary language skills and generating more reliable data, it also presents certain drawbacks. The reliance on volunteers may have introduced a bias in our sample, as individuals who are less confident or shy may be less likely to participate, as noted in the section Gender limitations (6.2.2.). Nonetheless, it is important to recognize that participants who are genuinely interested in the study are more likely to complete tasks in a timely manner and provide accurate data, enhancing the reliability of our results (Bialystok, 2001). Additionally, language-based studies often require participants to perform challenging tasks, which can decrease motivation or commitment if participants are not fully engaged. Furthermore, volunteers are more likely to have the necessary language skills compared to participants recruited from a random sample (Gass & Mackey, 2000). Their motivation and investment in the study can lead to more reliable data and better results.

6.3. Pedagogical implications

Based on the findings of our study on the impact of CLIL exposure to secondary-school learners in English, it is recommended that teachers continue to prioritise the development of pronunciation skills regardless of the programme they are implementing. The study revealed that focusing on pronunciation consistently yielded positive outcomes for learners, regardless of whether they were enrolled in CLIL-based programmes or EFL instruction. Therefore, teachers should allocate dedicated time and resources to help students to improve their pronunciation, as it has a significant influence on overall oral proficiency. Furthermore, our study supports the adoption and implementation of CLIL-based programmes, as they demonstrated clear benefits in terms of enhancing learners' oral proficiency. The integration of content-based instruction with language learning not only deepens students' understanding of subject matter but also provides ample opportunities for meaningful language practice and real-life communication. Therefore, educators should consider incorporating CLIL principles into their teaching practices to maximize students' language learning outcomes.

7. ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and appreciation to Raúl Azpilicueta Martínez, my esteemed tutor, for his invaluable guidance, unwavering support, and meticulous feedback throughout the course of this final thesis. His expertise, dedication, and insightful comments have been instrumental in shaping the direction and quality of my work. I am truly grateful for his mentorship and for sharing his vast knowledge, which has greatly contributed to my academic and personal growth.

I extend my heartfelt thanks to the students who collaborated on this thesis. Due to safety reasons and privacy concerns, I refrain from mentioning their names individually. Their enthusiastic participation, commitment, and willingness to contribute their time and efforts have been pivotal to the development of this research. Their willingness and honesty in diligently fulfilling the task assigned to them have enriched my investigation and deepened my understanding of the issue at hand. I acknowledge their contribution with immense gratitude.

I would also like to extend my appreciation to Harri Larunbe, the English teacher at the high-school. Harri not only provided me with a conducive environment to carry out the tasks with the students, but also played a crucial role as the second ranker, or corrector, of

the students' production. Harri's expertise, attention to detail, and constructive feedback were invaluable in enhancing the accuracy and quality of my research findings. I am indebted to Harri for his guidance throughout this journey.

Furthermore, I am grateful to the high-school Iparralde DBHI for granting me permission to conduct this study within its premises. The school administration's assistance, encouragement, and provision of resources were indispensable in ensuring the smooth progress of my research.

Lastly, I would like to thank all those who have contributed to this research project in various ways, directly or indirectly. Their encouragement, suggestions, and support have played a significant role in shaping the ideas and refining the work. I am immensely grateful for their contributions. The successful completion of this research study would not have been possible without the support and assistance of the aforementioned individuals and institutions. I am profoundly grateful for their contributions and acknowledge their role in making this research a reality.

– Total number of words: 11899

REFERENCES

- Admiraal, W., Westhoff, G., de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and evaluation*, 12(1), 75-93. <https://doi.org/10.1080/13803610500392160>
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511605963>
- Dalton-Puffer, C. (2011). Content-and-Language Integrated Learning: From Practice to Principles? *Annual Review of Applied Linguistics*, 31, 182–204. <https://doi.org/10.1017/s0267190511000092>
- Dalton-Puffer, C. (2007). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. Werner Delany and Laurenz Volkmann, (Eds.) *Future Perspectives for English Language Teaching*. Heidelberg: Carl Winter.

- Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
- Gallardo del Puerto, F., García Lecumberri, M.L., Gómez Lacabex, E. (2009). Testing the effectiveness of content and language integrated learning in foreign language context: Assessment of English pronunciation. In Y. Ruiz de Zarobe & R.M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 63-80). Bristol: Multilingual Matters. <https://doi.org/10.21832/9781847691675-007>
- Gallardo del Puerto, F., & Gómez Lacabex, E. (2017). Oral production outcomes in CLIL: An attempt to manage the amount of exposure. *European Journal of Applied Linguistics*, 5(1), 31-54. <https://doi.org/10.1515/eujal-2015-0035>
- Gálvez Gómez, M. d. M. (2021). The effects of CLIL on FL learning: A longitudinal study. In M.L. Pérez Cañado (Ed.), *Content and Language Integrated Learning in monolingual settings* (pp. 141-165). Springer, Cahm. https://doi.org/10.1007/978-3-030-68329-0_8
- Gass, S. M., & Mackey, A. (2000). Stimulated recall methodology in second language research. New York: Routledge. <https://doi.org/10.4324/9781410606006>
- Goris, J., Denessen, E., & Verhoeven, L. (2019). Effects of content and language integrated learning in Europe A systematic review of longitudinal experimental studies. *European Educational Research Journal*, 18(6), 675–698. <https://doi.org/10.1177/1474904119872426>
- Housen, A. (2012). Time and amount of L2 contact inside and outside the school - Insights from the European Schools. In C. Muñoz (Ed.), *Intensive exposure experience in second language learning* (pp. 111-138). Bristol: Multilingual Matters. <https://doi.org/10.21832/9781847698063-009>
- Hüttner, J., & Rieder-Bünnemann, A. (2010). A cross-sectional analysis of oral narratives by children with CLIL and non-CLIL instruction. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 61-80). Amsterdam, the Netherlands: John Benjamins. <https://doi.org/10.1075/aals.7.04hut>

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2007). Assessed Levels of Second Language Speaking Proficiency: How Distinct? *Applied Linguistics*, 29(1), 24–49.
<https://doi.org/10.1093/applin/amm017>

Jiménez Catalán, R. M., & Agustín Llach, M. P. (2017). CLIL or time? Lexical profiles of CLIL and non-CLIL EFL learners. *System*, 66, 87–99.
<https://doi.org/10.1016/j.system.2017.03.016>

Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. *Content and language integrated learning: Evidence from research in Europe*, 41, 81.
<https://doi.org/10.21832/9781847691675-008>

Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. *The Open Applied Linguistics Journal*, 1(1), 30–41.
<https://doi.org/10.2174/1874913500801010030>

Lasagabaster, D., Doiz, A. (2016). CLIL students' perceptions of their language learning process: delving into self-perceived improvements and instructional preferences. *Language awareness*, 25(1-2), 110-126.
<https://doi.org/10.1080/09658416.2015.11222019>

Lázaro Ibarrola, A. (2012). Faster and further morphosyntactic development of CLIL vs. EFL Basque-Spanish bilinguals learning English in high-school. *International Journal of English Studies*, 12(1), 79-96. <https://doi.org/10.6018.ijes.12.1.125621>

Lo, Y. Y. (2015). A glimpse into the effectiveness of L2-content cross-curricular collaboration in content-based instruction programmes. *International Journal of Bilingual Education and Bilingualism*, 18(4), 443-462.

Lyster, R., & Ballinger, S. (2011). Content-based language teaching: Convergent concerns across divergent contexts. *Language Teaching Research*, 15(3), 279-288.
<https://doi.org/10.1177/1362168811401150>

- Maillat, D. (2010). The pragmatics of L2 in CLIL. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL* (pp. 39–60). Amsterdam, the Netherlands: John Benjamins. <https://doi.org/10.1075/aals.7.03mai>
- Martínez Agudo, J. de D. (2019). Which instructional programme (EFL or CLIL) results in better oral communicative competence? Updated empirical evidence from a monolingual context. *Linguistics and Education*, 51, 69–78. <https://doi.org/10.1016/j.linged.2019.04.008>
- Merino, J. A., & Lasagabaster, D. (2018). The effect of content and language integrated learning programmes' intensity on English proficiency: A longitudinal study. *International Journal of Applied Linguistics*, 28(1), 18–30. <https://doi.org/10.1111/ijal.12177>
- Moore, P. (2009). *On the emergence of L2 oracy in bilingual education: A comparative analysis of CLIL and mainstream learner talk*. Universidad Pablo de Olavide, Sevilla, Spain.
- Muñoz, C. (2014). Contrasting Effects of Starting Age and Input on the Oral Performance of Foreign Language Learners. *Applied Linguistics*, 35(4), 463–482. <https://doi.org/10.1093/applin/amu024>
- Muñoz, C. (2015). Time and Timing in CLIL: A Comparative Approach to Language Gains. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based Language Learning in Multilingual Educational Environments* (Vol. 23, pp. 87–102). Springer International Publishing. https://doi.org/10.1007/978-3-319-11496-5_6
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Pérez Cañado, M. L., & Lancaster, N. K. (2017). The effects of CLIL on oral comprehension and production: A longitudinal case study. *Language, Culture and Curriculum*, 30(3), 300–316. <https://doi.org/10.1080/07908318.2017.1338717>

Pérez-Vidal, C., & Roquet, H. (2015). CLIL in Context: Profiling Language Abilities. *Springer International Publishing EBooks*, 237–255.

https://doi.org/10.1007/978-3-319-11496-5_14

Peters, E. (2018). The Effect of Out-of-class Exposure to English Language Media on Learners' Vocabulary Knowledge. *Approaches to Learning, Testing, and Researching L2 Vocabulary ITL - International Journal of Applied Linguistics* 169(1), 142-68. [doi:10.1075/itl.00010.pet](https://doi.org/10.1075/itl.00010.pet)

Rallo Fabra, L., & Juan-Garau, M. (2011). Assessing FL pronunciation in a semi-immersion setting: The effects of CLIL instruction on Spanish-Catalan learners' perceived comprehensibility and accentedness. *Poznań Studies in Contemporary Linguistics*, 47. <https://doi.org/10.2478/psicl-2011-0008>

Rallo Fabra, L., & Jacob, K. (2015). Does CLIL Enhance Oral Skills? Fluency and Pronunciation Errors by Spanish-Catalan Learners of English. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based Language Learning in Multilingual Educational Environments* (Vol. 23, pp. 163–177). Springer International Publishing. https://doi.org/10.1007/978-3-319-11496-5_10

Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: a longitudinal study in the Basque country. *International CLIL research Journal*, 1, 60-73.

Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? An insight into applied linguistics research. In Y. Ruiz de Zarobe, J. M. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning: Contributions to multilingualism in European contexts* (pp. 129-153). Bern: Peter Lang. <https://doi.org/10.3726/978-3-0351-0171-3>

Ruiz de Zarobe, Y. (2015). The Effects of Implementing CLIL in Education. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based Language Learning in Multilingual Educational Environments* (Vol. 23, pp. 51–68). Springer International Publishing. https://doi.org/10.1007/978-3-319-11496-5_4

de Smet, A., Mettwie, L., Hilgsmann, P., Galand, B., & Van Mensel, L. (2019). Does CLIL shape language attitudes and motivation? Interactions with target languages and instructional levels. *International Journal of Bilingual Education and Bilingualism*, 1-20. <https://doi.org/10.1080/13670050.2019.1671308>

Swain, M. (1981). Time and timing in bilingual education. *Language learning*, 31(1), 1-15. <https://doi.org/10.1111/j.1467-1770.1981.tb01369.x>

Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching Practice and Critique*, 10(4), 116-126.

APPENDIXES

Task materials

Comic strip

The Bicycle



Teacher's script

The teacher tells the child the name of the story and describes the first picture:

Teacher: **“These pictures tell a story. It’s called “The Bicycle”. Just look at the pictures first. (Pause) Henry was riding his bicycle. An angry driver was right behind him, tooting his horn: “honk-honk”! You tell the story now”**

CAPTION	Proactive question	Reactive question
CAPTION 2	<i>What was the boy doing?</i>	<i>Was the boy scared?</i>
CAPTION 3	<i>What happened to the boy?</i> <i>What did the car driver do?</i>	<i>Did he fall off his bicycle?</i> <i>Did he keep driving after the accident?</i>
CAPTION 4	<i>What happened to the bicycle?</i>	<i>Was the bicycle broken?</i>
CAPTION 5	<i>What happened to the car?</i>	<i>Has the car broken down?</i>
CAPTION 6	<i>Were they happy then?</i>	<i>Was the boy happy because he could ride his bicycle at the end?</i>

The teacher points at the pictures if necessary.

The teacher provides positive feedback after the student's answer, with words like **“good”**, **“excellent”**, or **“that’s right”**

Holistic rubric

COHERENCE

They tell the story in a very clear and coherent way, and the sequence of events is effortlessly understood.	4
They tell the story in a clear and coherent way, and the sequence of events is understood without much effort.	3
They are able to narrate the story, although the teacher has to pay close attention to understand the sequence of events.	2
Their speech is disjointed, there is no sequence in the narration, or it is very difficult to follow what they are referring to.	1

FLUENCY

i) The speech is fluent and without pauses, or ii) does not need help from the teacher to continue.	4
i) There are hardly any pauses, and none of them are long, or	3

ii) does not need help from the teacher to continue.	
i) There are certain pauses but the speech has a reasonable pace, or ii) needs help from the teacher to continue on one occasion.	2
i) The pauses are long or frequent and/or the speech is very slow, or ii) needs help from the teacher to continue on more than one occasion.	1

GRAMMAR

Ambitious use of grammar and the mistakes they make do not hinder comprehension.	4
Ambitious use of grammar, although they make mistakes that sometimes affect comprehension.	3
Grammar is simple, and some errors make comprehension difficult.	2
Their grammar is excessively simple, and there are frequent errors that make comprehension difficult.	1

PRONUNCIATION

Their pronunciation is very clear, and it is easy to understand their speech.	4
Their pronunciation is very clear, although it is necessary to make an effort to understand some words.	3
Their pronunciation is sufficiently clear, although it requires effort on the part of the listener	2
Their pronunciation is not sufficiently clear, even if the listener makes an effort to understand the speech.	1

VOCABULARY

They demonstrate sufficient and accurate vocabulary to tell the story in a complete manner.	4
They demonstrate sufficient vocabulary to tell the story, but struggle to find individual words,	3
They demonstrate sufficient vocabulary to tell the story, but have difficulty with the vocabulary of several vignettes.	2
They have very little vocabulary, and have difficulty referring to very basic concepts.	1