

DO GENDER-PAIRINGS AND AGE/PROFICIENCY INTERPLAY IN LRE PRODUCTION? A PRELIMINARY STUDY WITH YOUNG CLIL LEARNERS

 María Nieva-Marroquín

 María Martínez-Adrián

 María Basterrechea

University of the Basque Country, Spain

Abstract: Many studies have explored Language-Related Episodes (LREs) during collaborative dialogue. However, research on the factors affecting LRE production in children, especially regarding gender-pairings in CLIL contexts, remains limited (e.g. Martínez-Adrián & Gutiérrez-Mangado, 2022). Also, no studies have tackled the interplay between gender-pairings and age/proficiency in LRE production amongst young learners. To fill this gap, the study involved CLIL learners aged 9–11 who worked in same- and mixed-gender dyads to complete an oral picture-based narrative task. The LREs were analysed based on the amount, nature (meaning vs form) and resolution (target-like, non-target-like or unresolved). The preliminary results indicate that same-gender dyads seem to outperform mixed-gender pairs in both groups as they produce more LREs. Also, as age/proficiency increases, same-gender dyads tend to produce more target-like LREs. Thus, despite significant proficiency differences between both groups, the tendency towards working better in matched-gender pairs seems to persist throughout middle childhood.

Key words: young learners, language-related episodes, gender-pairings, proficiency, CLIL.

1. INTRODUCTION

Recently, second language acquisition (SLA) research has focused on how peer interaction during collaborative tasks contributes to second language (L2) learning (e.g. Mackey & Goo, 2012; Philp et al., 2013; Pica, 2013; Sánchez Ruiz & Lucas Pardo, 2016). The use of tasks has also been extended to Content and Language Integrated Learning (CLIL) contexts, where they serve a dual purpose: fostering the understanding of content and promoting language development in a natural, communicative way (Llinares et al., 2012). In this respect, collaborative tasks can bring about discussions about meaning or form (Basterrechea & Leeseer, 2019), operationalised by means of Language-Related Episodes (LREs). LREs are instances in which learners use language to reflect on their language use, seek explanations, ask for feedback and correct themselves or others (Swain & Lapkin, 1998).

LREs have been studied from the interactionist (Long, 1996) and the sociocultural perspective (Vygotsky, 1978). The interaction hypothesis posits that conversational interaction provides learners with comprehensible input and feedback, leading to more target-like output (Martínez-Adrián & Gallardo-del-Puerto, 2021). Also, learners make use of LREs during interaction to consciously reflect upon their language use (Leeseer, 2004). In turn, the sociocultural theory foregrounds LREs in language learning (Swain & Lapkin, 1998), stating that through peer-peer interaction, learners engage in collaborative dialogue—defined by Swain (2000:102) as “dialogue in which speakers are engaged in problem solving and knowledge building”—and use language as a tool to mediate their thinking and construct meaning (Swain & Watanabe, 2013). Thus, not only does LRE production promote L2 learning but also represents learning in progress.

Studies targeting LRE production have focused on the effect of variables such as proficiency (e.g. Basterrechea & Leeseer, 2019), pairing method based on self-selection or proficiency (e.g. Gallardo-del-Puerto & Basterrechea, 2024), pair dynamics (e.g. Kim & McDonough, 2008) or task-modality (e.g. Adams & Ross-Feldman, 2008; Gallardo-del-Puerto & Martínez-Adrián, 2022) on amount, nature and resolution of LREs. However, few have examined gender-pairings and LREs (e.g. Azkarai & García Mayo, 2012; Azkarai, 2015), especially with children (e.g. Martínez-Adrián & Gutiérrez-Mangado, 2022; Barañano-Arroyo, 2023). Given the significant role of gender in children’s socialisation and interaction (Holmes & Meyerhoff, 2003), this research addresses this gap by examining

To cite this article: Nieva-Marroquín, M., Martínez-Adrián, M., & Basterrechea, M. . (2025). “Do gender-pairings and age/proficiency interplay in LRE production? A preliminary study with young CLIL learners”. *Revista de Lingüística y Lenguas Aplicadas*, 20, 87-101. <https://doi.org/10.4995/ryla.2025.22415>

Correspondence author: maria.nieva@ehu.eus



LRE production among young learners across different gender-pairings. This line of research is key in identifying the more optimal learning conditions when performing collaborative tasks in English since the results obtained can help educators make the most of gender-pairings for collaborative tasks to enhance learning outcomes. This is particularly relevant considering teachers' concerns about the psychological and social factors that affect student engagement, such as limited participation during language learning activities (Philp et al., 2010; Sato & Csisér, 2021).

Moreover, SLA research has often overlooked childhood developmental stages, including middle childhood (7–11 years old) (Oliver & Azkarai, 2017), the age of the participants of this study. Learners within this age range possess a growing first language (L1) knowledge and show increased reflection upon language use, as they are developing greater awareness of how it works (Hakes, 1980; Pinter, 2006). By age 11–12, children exhibit a higher metalinguistic awareness (Pinter, 2006), which may greatly impact LRE production in L2 contexts. Likewise, children aged 7–11 also seem to have enhanced social abilities (Muñoz, 2007) which, combined with educational experiences, significantly influence interaction and language learning—an affectively driven (i.e. which requires motivation) and socially co-constructed process (i.e. language is learnt through social interactions with others) (The Douglas Fir Group, 2016)—. Therefore, investigations on language learning during this stage can enhance learners' opportunities for acquiring a L2 (Hidalgo, 2021).

Some studies that have considered age and proficiency jointly have revealed how older learners (aged 10–11) tend to show advantages in negotiation of meaning (García Mayo & Imaz-Agirre, 2016; Hidalgo, 2021; Pladevall-Ballester, 2021), syntactic complexity (Martínez-Adrián & Nieva-Marroquín, 2023), and vocabulary acquisition and LRE production (Iglesias-Diéguez et al., 2025) due to an accumulated language exposure and instructional hours. However, evidence on LRE production across different ages and proficiency levels remains relatively scarce.

Likewise, in CLIL contexts, most studies have focused on examining the effect of this approach on general proficiency (e.g. Lasagabaster, 2008; Merino & Lasagabaster, 2018) or specific features of language (e.g. Azpilicueta-Martínez, 2025). Despite so, variables such as the effect of gender on LRE production in CLIL have received scarce attention. Previous studies have focused on the effect of gender on motivation (e.g. Heras & Lasagabaster, 2015) or on compensatory strategies (Basterrechea et al., 2017), suggesting that CLIL seems to have a blurring effect on gender differences.

Finally, findings from research on proficiency-based pairing methods in CLIL contexts (Basterrechea & Gallardo-del-Puerto, 2020; Gallardo-del-Puerto & Basterrechea, 2024) suggest that pair formation method seemed to affect the learners' interaction since self-selected pairs produced and resolved more meaning-based LREs than proficiency-matched dyads. Building on this, and extrapolating it to the gender-based perspective, it would be helpful to investigate how different gender-pairings might similarly affect language learning opportunities, particularly if self-selected pairs tend to be same-gender.

Factoring all this in, no studies thus far have examined the interplay of gender-pairings and age/proficiency in LRE production in CLIL settings. This could shed more light on the role of gender-pairings in LREs and across learning trajectories (i.e. whether differences between same- or mixed-gender dyads emerge in the production of LREs as age/proficiency increase). This study, a follow-up of Martínez-Adrián and Nieva-Marroquín (2023), aims to bridge these gaps by examining the LRE production of matched-gender (female-female, henceforth FF; and male-male, henceforth MM) and mixed-gender (male-female, henceforth MF) in task-based dyadic interaction within two different age/proficiency groups, namely, 4th- and 6th-year CLIL students, aged 9–11. This paper will review the previous literature on gender and proficiency, present the methodology and discuss the results and pedagogical implications.

LITERATURE REVIEW

2.1 Gender and interaction

Recently, numerous investigations with native speakers have tackled the importance of gender in communication. The overall conclusion is that sex segregation is a process that starts in early childhood as children mostly interact with same-gender peers (Maltz & Borker, 1982). This shapes their communication skills and, therefore, males and females communicate differently (Tannen, 1994), all of which may affect interaction in a L2.

From an interactionist perspective, few studies have explored the impact of gender on L2 production, comprehension, and interaction, and the findings have been mixed (e.g. Gass & Varonis, 1986; Pica et al., 1989, 1991). These studies, conducted with adults, examined task-based interaction and showed that gender-pairings can influence the frequency and nature of negotiation. Gass and Varonis (1986) found more meaning negotiation in

mixed-gender dyads, while Pica et al. (1989, 1991) reported no significant differences across dyad types, although non-native males tended to modify their output more and female non-native speakers had fewer chances to interact with male native speakers.

In EFL contexts and with adult learners, gender has also received limited attention (e.g. Alcón, 1994; Alcón & Codina, 1996). Alcón (1994) observed that females negotiated more in same-gender pairs, while Alcón and Codina (1996) found no overall gender effects but noted that task-type played a role, with females negotiating more in information-gap tasks.

Child studies on gender and interaction have not received so much attention as of yet. In Australia in an ESL context, Oliver (2002) found no significant differences between same- and mixed-gender pairs during two communicative tasks in the amount of negotiation of meaning. To the authors' knowledge, the only study on gender and interaction carried out thus far in CLIL contexts is the one by Azkarai and Imaz-Agirre (2017). They discovered, through the observation of the interaction of 3rd- and 4th-year students, that the younger learners negotiated more for meaning and used more English in their discussions in same- than in mixed-gender dyads. In turn, older learners negotiated more in mixed-gender dyads, but their use of English decreased in same-gender dyads, especially in MM pairs.

Overall, mixed results have been reported with both adults and young learners. Results from the limited research conducted with young learners have shown that they seem to benefit from working in same-gender dyads (at least in the 3rd year) (Azkarai & Imaz-Agirre, 2017) while others have reported no differences between gender-pairings (Oliver, 2002).

2.2 Gender and LREs

The effect of gender on learner-initiated attention to form (i.e. LRE production) during interaction has also been explored by recent research, yet studies are still meagre. Most have looked into adult interaction in ESL (e.g. Ross-Feldman, 2005, 2007) and EFL settings (e.g. Azkarai & García Mayo, 2012; Azkarai, 2015), whilst studies addressing children and gender are scant (e.g. Martínez-Adrián & Gutiérrez-Mangado, 2022; Barañano-Arroyo, 2023).

With adults, Ross-Feldman (2005) found that female learners initiated more LREs in mixed-gender dyads, though no significant differences emerged in overall LRE frequency across genders. However, LREs initiated by males were resolved significantly more often than LREs initiated by females. Her later study (2007) showed that males initiated more LREs in mixed-gender pairs, while females resolved more LREs in same-gender dyads. Shehadeh (1994) reported similar gendered interaction patterns. In the Spanish EFL context, Azkarai and García Mayo (2012) and Azkarai (2015) found no significant gender differences in LRE frequency or nature, though mixed-gender dyads tended to show higher engagement and resolution rates. These findings suggest gender may shape interaction dynamics and opportunities for language learning, but further research with child learners is required.

Thus far, in child task-based interaction, evidence on the impact of gender-pairings on LRE production remains scarce, especially in CLIL settings, where only one study has been carried out as of now (e.g. Martínez-Adrián & Gutiérrez-Mangado, 2022). In this piece of research, through a storytelling task performed by 10–12-year-old learners, the authors found that same-gender dyads initiated significantly more LREs in total and significantly more meaning-related LREs than mixed-gender dyads. Likewise, same-gender dyads resolved significantly more LREs than mixed-gender pairs. The outcome of the resolution was not affected by gender-pairing though, as the rate of target-likeness was very low for both dyad types. Lastly, despite there not being significant differences when comparing FF and MM dyads, their findings suggest that FF dyads were mainly concerned with conveying the message and more oriented to higher accuracy when resolving meaning-focused episodes. In a similar vein yet in EFL settings, Barañano-Arroyo's (2023) study with 11–12-year-old students carrying out an oral+written task revealed that same-gender dyads initiated and resolved more LREs in a more target-like manner than mixed-gender pairs, unlike Martínez-Adrián and Gutiérrez-Mangado (2022), a result which may be explained by task-type, as the oral+written task promotes greater accuracy through the incorporation of a written component. When looking into MM and FF dyad types, the study revealed that MM resolved a larger amount of meaning-focused LREs and in a more target-like manner than FF dyads. Besides, FF outperformed MM dyads in form-related episodes, as they accurately resolved a larger amount, indicating that females seemed to take more advantage of a task that inherently promotes greater attention to formal aspects of language.

The review of the previous research targeting gender-pairings and LREs has uncovered discrepant findings between adults and children. In adults, some studies have revealed no differences in the amount of LREs between same- and mixed-gender dyads (Ross-Feldman, 2005, 2007). In resolution (i.e. resolved and unresolved), some investigations have reported differences between both dyadic types in favour of mixed-gender dyads (Azkarai, 2015). Other studies have shown that mixed-gender pairs were more profitable for males, whereas matched-gender dyads were more for females (Ross-Feldman, 2007). Others have found no gender-pairing effect at all

(Azkarai & García Mayo, 2012). In turn, children seem to benefit more from working in same- than in mixed-gender dyads (Martínez-Adrián & Gutiérrez-Mangado, 2022; Barañano-Arroyo, 2023). Further investigation with young CLIL learners is warranted as the current body of research on this population remains limited (e.g. Martínez-Adrián & Gutiérrez-Mangado, 2022).

2.3 Proficiency and interaction

A variable that has received considerable attention in L2 interaction research is proficiency. Within the interactionist perspective, a few studies have aimed at exploring the interplay between proficiency and interaction with adults (e.g. Iwashita, 2001) and with children (e.g. Oliver, 2002; Oliver & Azkarai, 2019).

Regarding research with adults, Iwashita (2001) investigated a group of university students, paired in high-high, high-low and low-low dyads, regarding a possible effect of learner proficiency on opportunities for modified input, types of interactional moves and on the actual production of modified input. No significant differences were found between the same or different proficiency groups.

With children learning English in the Australian context, mixed findings have been reported. Oliver (2002) found that the lower the proficiency of the pair, the more negotiation of meaning occurred (more clarification requests and confirmation checks) due to a greater communication breakdown. Contrary to these results, Oliver and Azkarai (2019) examined patterns of interaction—collaborative, dominant/dominant, dominant/passive and expert/novice (see Storch, 2002 for a description)—of learners with different proficiency levels performing two task-types (a one-way and a two-way task). Their results suggested that the young ESL learners engaged with the tasks and with each other no matter their proficiency level, most often in a collaborative way. However, the patterns of interaction were influenced by both proficiency and task-type, (i.e. high-proficiency learners had more opportunities to negotiate and focused more on form than low-proficiency learners in the one-way task). In a longitudinal study carried out in an EFL context, Pladevall-Ballester (2021) found that pair dynamics remained collaborative as proficiency and age increased.

In CLIL contexts, very few studies have looked into the effect of proficiency pairings (i.e. self-selected vs proficiency-matched dyads), on patterns of interaction (e.g. Basterrechea & Gallardo-del-Puerto, 2020). These authors examined peer interaction during collaborative tasks among primary-school learners, finding that proficiency-matched pairs showed more collaborative dynamics than self-selected dyads.

By and large, mixed results have been found as well in the case of the effect of proficiency on interaction with adults and young learners. With adults, some pieces of research have not found differences with matched-proficiency pairings (Iwashita, 2001). The sparse research on young learners has shown that, whereas some studies have found benefits in interaction for low-proficiency learners (Oliver, 2002), others have found advantages for high-proficiency learners (Oliver & Azkarai, 2019).

2.4 Proficiency and LREs

The role of L2 proficiency on LRE production (regarding quantity, quality and resolution) has received considerable attention in the last decades, yet most studies have focused on adults. For example, Leeser (2004) found that high proficiency L2 Spanish university student dyads produced and resolved more LREs than mixed or low proficiency pairs, with low proficiency dyads leaving more episodes unresolved. Similarly, Kim and McDonough (2008) observed that high-proficiency Korean L2 dyads resolved LREs more effectively. Watanabe and Swain (2007) also reported more LRE production among high proficiency mixed pairs, finding also that collaborative interaction positively impacted learning outcomes, i.e. when learners engaged in collaborative patterns of interaction, they were more likely to achieve better results in the post-tests.

A dearth of studies has focused on how proficiency and LRE production by young learners intertwine. In the Chilean context, Sato and Viveros (2016) compared a high- and a low-proficiency intact class with adolescents, reporting that the low-proficiency learners benefitted more as they engaged in LREs to a larger extent. In a CLIL setting with 15–16-year-olds conducting a collaborative task, Basterrechea and Leeser (2019) focused on the amount, type and resolution of LREs involving the third person singular -s morpheme by proficiency-matched pairs. They found a positive correlation between learner proficiency and the number of LREs involving the target form. They also discovered that more proficient learners resolved more LREs in a target-like way and that lower-proficiency learners predominantly prioritised meaning over formal aspects of language.

Thus far, there seems to be a lack of studies examining how proficiency affects LRE production amongst young learners in CLIL contexts since existing research (such as Basterrechea & Gallardo-del-Puerto, 2020; Gallardo-del-Puerto & Basterrechea, 2024; Martínez-Adrián & Arratibel-Irazusta, 2024) has used proficiency as a way to pair learners, i.e. proficiency-matched vs self-selected pairs, rather than explicitly investigating proficiency level as an independent variable. Despite so, Basterrechea and Gallardo-del-Puerto (2020) discovered that student-selected

pairs, who were not paired by proficiency, significantly produced and resolved more meaning-focused LREs than proficiency-matched, although the latter led to more target-like LREs. In a follow-up study, Gallardo-del-Puerto and Basterrechea (2024) confirmed that grade level (5th vs 6th) did not affect LREs, yet self-selected pairs generated and resolved more meaning-focused episodes. When looking into Previously Known Languages (PKLs) in two task-types (speaking and speaking + writing) with the same sample, Martínez-Adrián and Arratibel-Irazusta (2024) found that self-selected pairs produced significantly more overall turns in both tasks. In contrast, proficiency-matched pairs showed more effective collaboration since they used PKL strategically and stayed more on task than the self-selected pairs.

Overall, the studies on LREs and proficiency reviewed above suggest that, with adults, the higher the proficiency level of the learners, the higher the number of opportunities to negotiate (Leeser, 2004; Kim & McDonough, 2008), the higher the amount of focus on form (Leeser, 2004) and the more the opportunities for language learning (Leeser, 2004). With young learners, however, the scarce research has not yielded any conclusive results and more research is warranted with CLIL participants to shed more light on this approach.

In summary, research on proficiency effects on LRE production in young CLIL learners remains limited and the interaction between gender-pairings and LRE production across age/proficiency levels has not yet been explored. Previous studies suggest that more proficient learners show more opportunities to negotiate through LREs (at least in adults the evidence remains robust), potentially maximising the learning opportunities for matched-gender-pairings.

3. THE STUDY

Considering the gaps uncovered, the present paper will aim to fill this much-needed niche by investigating the LREs produced in an oral task by two different age/proficiency groups, namely 4th- and 6th-year CLIL learners in same- and mixed-gender pairs. The following two research questions will be addressed:

RQ1: Are there any differences in the amount, nature and resolution of the LREs produced by same- and mixed-gender dyads within the 4th-year group and within the 6th-year group?

RQ2: What are the most frequent LRE types (nature and resolution) in each gender-pairing type in the 4th- and 6th-year groups?

3.1 Participants

The participants were 28 students (15 females, 13 males) attending a state-funded primary school in a town of Castile and Leon, a monolingual autonomous community in Spain. They belonged to the 4th (CLIL I, n=16; made up of 8 boys and 8 girls) and the 6th year (CLIL II, n=12; made up of 5 boys and 7 girls) of primary education. They were aged 9 and 11, respectively. To increase the amount of exposure to the FL, the school implements CLIL lessons since the 1st year of primary education, at the age of 6. English instruction begins at age 3. In primary education, CLIL I learners receive 3h of English and 4h of CLIL per week, totalling 152h of English exposure per course. In turn, CLIL II receive 3h of English and 6h of CLIL per week, adding up to 228h of English exposure per course. Including extracurricular private English lessons (which 13 out of 16 attended in CLIL I and 9 out of 12 in CLIL II), the total number of hours of exposure to English amounted to 1,346 and 2,455, respectively. Table 1 shows a summary of the learners' characteristics.

Table 1. Participants' characteristics according to gender.

	CLIL I		CLIL II	
	Females	Males	Females	Males
N	8	8	7	5
Mean age	9.12	9	11	11
Age of onset	3	3	3	3
Years of exposure	6	6	8	8

The learners' proficiency level was measured at the outset by means of the standardised A2 Flyers test (UCLES, n.d.), which indicated that the learners' proficiency level ranged between A2.1 (CLIL I) and A2.2 (CLIL II), according to the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2020). Since the Saphiro-Wilks test yielded normally distributed results for both groups, a T-test was carried out. High statistical significance was indicated at <0.01 (*) levels. Effect size thresholds for intergroup analyses were based

on Plonsky and Oswald's (2014) guidelines for L2 acquisition: a *d* value of about 0.40 was considered 'small', 'medium' if around 0.70, and 'large' if above 1.00. The mean percents, standard deviations (SD), T-test results and the effect size are shown in Table 2.

Table 2. Mean%, SD, effect size and intergroup T-test results of the A2 Flyers test.

GROUP	Mean% (SD)	T-test		Effect size
		T	p-value	
CLIL I	29.21 (15.15)	-5.020	<0.001*	d=1.91
CLIL II	58.30 (15.20)			

The T-test indicated that the differences between the 4th and 6th year in terms of proficiency were highly significant ($t=-5.020$, $p\text{-value}<0.001$). Despite a small sample size, the Cohen's *d* value showed a very large effect size ($d=1.91$), which suggests a large difference in performance between both groups.

Due to the data being normally distributed in the case of gender according to the Shapiro-Wilks tests, subsequent T-tests were carried out to further control for possible differences in proficiency between males and females within the same school year. Table 3 shows said T-test results.

Table 3. Mean%, SD, effect size and intergroup T-test results between females and males in each school year.

GROUP	Mean% (SD)	T-test		Effect size
		T	p-value	
CLIL I	Females: 35.09 (14.17)	1.635	0.124	d=0.81
	Males: 23.33 (14.56)			
CLIL II	Females: 59.43 (16.61)	0.290	0.778	d=0.17
	Males: 56.73 (14.71)			

Although females showed higher average scores than males, the T-test did not show significant differences between gender in either the 4th or 6th year. Moreover, Cohen's *d* revealed medium ($d=0.81$) and small ($d=0.17$) effect sizes, respectively, which suggest low practical significance in performance between male and female learners in either group. These results imply that any differences or lack thereof with respect to LRE production between males and females in the subsequent intergroup analyses cannot be attributed to proficiency differences between both groups.

As in previous studies measuring the effect of gender-pairings (e.g. Martínez-Adrián & Gutiérrez-Mangado, 2022; Barañano-Arroyo, 2023), participants were paired in same- and mixed-gender-pairings but controlling proficiency level as a safeguard. In total, there were 14 dyads. In CLIL I, there were 8 dyads: 4 were matched-gender (MM and FF) and 4 were mixed-gender (MF). In CLIL II, there were 6 dyads: 3 were matched-gender (MM and FF); and 3 were mixed-gender (MF).

3.2 Instruments and procedure

Initially, the participants who agreed to partake completed a background questionnaire. Later, they sat the listening, reading and writing parts of the A2 Flyers test and, according to their results, they were matched in same-proficiency dyads in order to perform an oral narration task in the form of a picture placement and description task based upon the book *Room on the Broom* (Donaldson, 2012). The task had been teacher-approved and pilot-tested with similar-age participants. It was carried out in a separate room with the investigator and recorded for transcription. To do the task, each learner was first given a set of four vignettes, which they had to describe individually and then, in pairs, they had to order the eight vignettes and tell the story. They were encouraged to act naturally, help and speak to each other and use English as much as possible.

This type of task, with no written component, provides learners with a limited time to edit and plan what they want to say (Payant & Kim, 2019), thus focusing primarily on meaning. In such tasks, more meaning-focused LREs have been found with adults (García Mayo & Azkarai, 2016) and children (Pladevall-Ballester, 2021; Martínez-Adrián & Gutiérrez-Mangado, 2022) than in those tasks which also include a writing component, which elicit more form-focused LREs (Martínez-Adrián & Gallardo-del-Puerto, 2021; Martínez-Adrián & Gutiérrez-Mangado, 2022; Gutiérrez-Mangado & Basterrechea, 2025). However, and despite their meaning-focused nature, the aforementioned oral tasks have been widely used in previous research with adults (e.g. Azkarai & García Mayo, 2012) and children (e.g. García Mayo & Hidalgo, 2017; Basterrechea & Gallardo-del-Puerto, 2020; Pladevall-Ballester,

2021; Martínez-Adrián & Gutiérrez-Mangado, 2022) as they actively promote interaction between the learners to reach the task goal (García Mayo & Hidalgo, 2017), in this case, to order the vignettes in a meaningful way and to tell the story. Note also how these oral storytelling tasks have been profusely used when examining morphosyntactic and lexical aspects in young CLIL learners (see in this respect Martínez-Adrián & Nieva-Marroquín, 2023 with similar-age participants). Thus, the present study will contribute to reinforce the validity of this type of instrument.

3.3 Data analysis

The oral narrations were transcribed in CHILDES format (MacWhinney, 2000). Following Adams and Ross-Feldman's (2008) and García Mayo and Azkarai's (2016) LRE classification, they were first coded according to their nature (form- vs meaning-focused). Form-focused LREs deal with aspects such as grammar (phonology, morphology, syntax or spelling) as in (1), whilst meaning-focused LREs deal with problems related to word meaning or word choice as in (2).

- (1) Form-focused:
Learner 2: Fells down the *varity@s* [wand].
Learner 1: The *varity@s* [wand] fell down.
- (2) Meaning-focused:
Is the cat of the plants, trees, eh... black, red, purple, orange... Is a black y... *no sé cómo se dice "sombbrero"@s* [and... I don't know how you say "hat"]

In excerpt (1), the main focus is syntax. Learner 2 produces a structure with the wrong subject-verb order, which Learner 1 corrects. Learner 1 also corrects the verb form, from "fells" to "fell". In (2), the focus is on word choice. The learner does not know how to say the word "hat" so they produce an LRE in Spanish.

Then, LREs were coded according to their resolution (resolved target-like, resolved non-target-like or unresolved), depending on whether the learner or the interlocutors reached a correct solution (3), an incorrect one (4) or did not find any (5), respectively.

- (3) Meaning-focused, resolved target-like:
The witch lost the hat. Dog... *¿Cómo se decía "buscar"? El verbo "buscar"@s... ah, ¡ya sé!@s*. [How do you say "find"? The verb "to find". Oh, I know!]. Find the hat.
- (4) Meaning focused, resolved non-target-like:
Learner 1: This is a witch that are in the *escoba@s* [broom]. The *varita mágica@s* [magic wand], the magic...
Learner 2: wind!
- (5) Meaning-focused, unresolved:
Are rain and the witch eh... *"Pierde", ¿cómo se dice "pierde"?@s Pierde@s the varita@s* ["Lose", how do you say "lose"? She loses the wand]

In excerpt (3), the learner at first does not know how to say the verb "to find". They immediately self-correct and provide the verb in the following utterance. In (4), Learner 1 does not know how to say "wand", so Learner 2 interrupts and says a similar yet incorrect word, "wind". In (5), the learner is not able to remember the verb "to lose", so they produce the verb in Spanish.

Statistical analyses were run using IBM SPSS Statistics (IBM Corp., 2021). Due to the limited dataset, descriptive statistics (which include means and standard deviations) were calculated, as in some previous studies with LREs (see Basterrechea & García Mayo, 2013; Basterrechea & Gallardo-del-Puerto, 2023). To answer the first research question, which looked into whether LRE amount, nature, and resolution differ across same- and mixed-gender dyads, intergroup analyses were run. The second question, which focused on the most frequent LRE types by gender-pairing, required intragroup analyses.

4. RESULTS

This section will provide the results obtained in order to answer the research questions. Section 4.1 will deal with the intergroup analyses and section 4.2 will focus on the intragroup analyses.

4.1 Intergroup analyses

To answer the first research question, which inquired about the existence of any differences in the amount, nature and resolution of the LREs produced by same- and mixed-gender dyads within the 4th- and the 6th-year group, intergroup analyses were run. Table 4 displays the number, nature and resolution (resolved target-like, resolved non-target-like or unresolved) of the LREs produced by CLIL I.

Table 4. Number, percentages, means and standard deviations of LREs initiated by matched- and mixed-gender dyads.

	CLIL I	
	Matched-gender dyads (n=4)	Mixed-gender dyads (n=4)
Total number of LREs	n=13; mean=3.25; SD=1.5	n=8; mean=2 ; SD=2.82
Meaning-focused	n=13; Percentage: 100%; mean=3.25 ; SD=1.5	n=8; Percentage: 100%; mean=2; SD=2.82
Meaning-focused resolved target-like	n=2; Percentage: 15.39%; mean=0.5 ; SD=0.577	n=2; Percentage: 25%; mean=0.5; SD=1.00
Unresolved	n=11; Percentage: 84.61%; mean=2.75; SD=1.89	n=6; Percentage: 75%; mean=1.5; SD=1.91

The qualitative inspection of the means presented in Table 4 reveals that, when looking at matched- vs mixed-gender dyads in CLIL I (4th grade), matched-gender dyads produced a higher number of LREs and more meaning-focused LREs than mixed-gender pairs (3.25 vs 2). All of them were meaning-focused LREs. They both produced the same number of resolved target-like LREs (0.5). Matched-gender pairs also produced a higher number of unresolved LREs than mixed-gender dyads (2.75 vs 1.5). No form-focused or non-target-like LREs were produced by either mixed- or same-gender pairs.

Table 5 shows the number, nature and resolution of the LREs produced by CLIL II.

Table 5. Number, percentages, means and standard deviations of LREs initiated by matched- and mixed-gender dyads

	CLIL II	
	Matched-gender dyads (n=3)	Mixed-gender dyads (n=3)
Total number of LREs	n=9; mean=3; SD=3	n=7; mean=2.33 ; SD=1.52
Meaning-focused	n=9; Percentage: 100%; mean=3 ; SD=3	n=5; Percentage: 71.42%; mean=1.67; SD=2.08
Meaning-focused resolved target-like	n=4; Percentage: 44.44%; mean=1.33; SD=1.52	n=1 ; Percentage: 20%; mean=0.33; SD=0.57
Meaning-focused resolved non-target-like	n=1; Percentage: 11.11%; mean=0.33; SD=0.57	n=0; Percentage: 0%; mean=0.00; SD=0.00
Meaning-focused unresolved	n=4; Percentage: 44.44%; mean=1.33; SD=1.15	n=4; Percentage: 80%; mean=1.33; SD=1.52
Form-focused	n=0; Percentage: 0%; mean=0.00; SD=0.00	n=2; Percentage: 28.57%; mean=0.67; SD=1.15
Form-focused resolved target-like	n=0; Percentage: 0%; mean=0.00; SD=0.00	n=2; Percentage: 100%; mean=0.67; SD=1.15

When looking at the descriptive means of matched- vs mixed-gender pairs in the 6th year, matched-gender pairs produced a higher total amount of LREs than mixed-gender pairs (3 vs 2.33). Also, matched-gender dyads produced more meaning-focused LREs than mixed-gender dyads (3 vs 1.67). Same-gender dyads also produced more resolved target-like episodes than mixed-gender pairs (1.33 vs 0.33). With respect to unresolved LREs, both

dyad types produced the same number of episodes (1.33 vs 1.33). Lastly, regarding form-focused LREs (total and resolved), the mixed gender dyads produced more LREs of this type than the same-gender dyads, who did not produce any (0.67 vs 0.00).

4.2 Intragroup analyses

To answer the second research question, which aimed to analyse the most frequent LRE types (nature and resolution) in each gender-pairing in CLIL I and CLIL II, qualitative analyses were carried out.

It can be observed that, in CLIL I, for matched-gender dyads, the most frequent LRE type in terms of nature was meaning-focused (3.25). Regarding resolution, they produced a higher number of unresolved than resolved target-like LREs (2.75 vs 0.5). Likewise, more target-like LREs were resolved compared to non-target-like (0.5 vs 0.00). For mixed-gender dyads, the most frequent LRE regarding nature was also meaning-focused (2.82). In resolution, the most frequent were unresolved LREs compared to resolved target-like (1.5 vs 0.5) and, lastly, resolved target-like LREs were the most frequent ones compared to non-target-like LREs (0.5 vs 0.00).

In CLIL II, the most frequent LRE type in terms of nature for matched-gender dyads was meaning-focused (3 vs 0.00). Regarding resolution, resolved target-like LREs were as frequent as unresolved (1.33) and also, resolved target-like were more frequent than resolved non-target-like episodes (1.33 vs 0.33). For mixed-gender dyads, the most frequent LRE type was also meaning-focused compared to form-focused (1.67 vs 0.67). Within meaning-based episodes, regarding resolution, unresolved LREs were more frequent than resolved target-like and resolved non-target-like (1.33 vs 0.33 vs 0.00). In the case of form-focused LREs, all of them were resolved in a target-like-way.

5. DISCUSSION

This study set out to investigate the interplay between gender-pairings and age/proficiency in LRE production by young CLIL learners during collaborative dialogue. The first research question addressed whether differences exist in the amount, nature and resolution of the LREs produced by same- and mixed-gender dyads within the 4th- (CLIL I) and the 6th-year (CLIL II) group. Firstly, regarding the amount of LREs, in both CLIL I and CLIL II, same-gender dyads produced more LREs than mixed-gender dyads, similar to previous findings on gender and LREs with children (Martínez-Adrián & Gutiérrez-Mangado, 2022; Barañano-Arroyo, 2023). The stronger collaboration observed in same-gender pairs aligns with the results reported for self-selected dyads in Gallardo-del-Puerto & Basterrechea (2024), who produced and resolved more LREs. In this line, learners often feel more comfortable and open when working with same-gender peers (McCloskey & Coleman, 1992), and this can lead to better interaction and better teamwork—in line with what is often seen in self-selected pairs, which can be made based on friendship (see Mozaffari, 2017)—. At the same time, our findings are consistent with Basterrechea and Gallardo-del-Puerto (2020) and Martínez-Adrián and Arratibel-Irazusta (2024), who found that proficiency-matched dyads, as in the present paper, showed a better collaborative relationship. This suggests that the more effective collaboration observed in our same-gender dyads may be linked to the combined effect of gender pairing and shared proficiency. Moreover, this finding could be interesting for CLIL settings: since learners need to address both content and language, pairing them up in same-gender dyads could bring about more discussions on language whilst also focusing on content.

However, this finding contrasts with what has been found with adults in ESL (Ross-Feldman, 2005, 2007) and EFL settings (Azkarai & García Mayo, 2012; Azkarai, 2015), where no differences in the amount and nature of LREs between mixed- and matched-pairs were reported. Such difference between adult and child learners likely stems from the cognitive, metalinguistic and socio-affective differences between both groups (Muñoz, 2007; Oliver & Azkarai, 2017): while adults often seem to work better in mixed-gender dyads (Gass & Varonis, 1986) due to more developed social skills or more talkativeness (Leaper & Ayres, 2007), children may feel uncomfortable or show limited engagement with the opposite gender, which could limit LRE production.

Despite same-gender dyads producing more LREs in total, the overall low frequency of LREs in both groups is also worth mentioning. Apart from the influence of the oral modality of the task (see García Mayo & Azkarai, 2016), which will be further explained in the second research question, it is suggested that the learners' low proficiency (ranging from A2.1 to A2.2) may have limited their metalinguistic resources (see also Leeson, 2004 and Kim & McDonough, 2008 with adults and García Mayo & Imaz Agirre, 2019; Martínez-Adrián & Arratibel-Irazusta, 2020 with children). Likewise, this limited LRE production may also align with the CLIL context that the learners belong to: learners in CLIL classrooms are often focused on delivering content (Llinares & Morton, 2017) rather than talking about form, and younger learners in particular may lack the linguistic and cognitive scaffolding to elaborate LREs.

Regarding nature (meaning- vs form-based LREs), in CLIL I, all LREs were meaning-focused, regardless of the gender pairing. This is expected at age 9, when learners tend to focus on completing tasks rather than thinking about accuracy, a pattern also observed by Pladevall-Ballester (2021) with 4th-grade learners at time 1. In CLIL II, same-gender dyads still focused only on meaning, but mixed-gender dyads produced some form-focused LREs as well. This shift could be linked to the learners' growing awareness of language as they get older and their proficiency increases. This aligns with what was found by Basterrechea and Leaser (2019) when comparing high- and low-proficiency pairs with teenage learners. Also, Gallardo-del-Puerto and Martínez-Adrián (2022) found a low number of form-focused LREs with learners aged 10–11 in both oral and written tasks, reporting that more grammatical concerns appear as children grow older (see Muñoz, 2014).

Lastly, in terms of resolution of LREs (resolved target-like, resolved non-target-like and unresolved), in CLIL I, both same- and mixed-gender pairs produced the same amount of resolved target-like LREs (0.5), which was low overall. This may suggest that gender pairing did not influence the learners' ability to resolve language issues at this age and developmental stage. However, in CLIL II, same-gender dyads produced more resolved target-like episodes than the mixed-gender pairs (1.33 vs 0.33), which indicates that same-gender dyads may have achieved a more collaborative relationship than their mixed-gender counterparts at this developmental stage during communication. With regards to unresolved episodes, in CLIL I same-gender dyads left a larger amount of LREs unresolved compared to mixed-gender dyads (2.75 vs 1.5). In CLIL II, however, both matched- and mixed-gender dyads left the same amount of LREs without resolution (1.33).

The second research question purported to investigate the most frequent LRE type (nature and resolution) in each gender-pairing. The results seemingly show that, regarding nature, the most frequent LRE was meaning-based for both dyad types in both groups. This is likely due to task-type as oral tasks, unlike oral + written tasks, mainly prompt meaning-based discussions, as shown in studies with children (Pladevall-Ballester, 2021; Martínez-Adrián & Gutiérrez-Mangado, 2022) and adults (García Mayo & Azkarai, 2016). The administered task, an oral narration with picture description and picture placement, is communication-oriented, so learners need the key vocabulary to complete it (García Mayo & Azkarai, 2016). Indeed, it is important to highlight that the few instances of form-focused episodes were produced by a mixed-gender pair, in contrast with the findings by Martínez-Adrián and Gutiérrez-Mangado (2022), who found that only matched-gender pairs, MM specifically, produced form-related episodes, and by Barañano-Arroyo (2023), who found that FF dyads outscored MM dyads in amount, resolution, target-likeness and incorporation of form-focused LREs. It could be posited that, since only one dyad produced form-focused episodes, this tendency cannot be ascribed to a specific behaviour of mixed-gender pairs.

The other most frequent type of LRE for matched- and mixed-gender dyads in CLIL I were unresolved episodes, suggesting a developmental stage where linguistic gaps are not perceived as needing repair. Consequently, when learners encountered unfamiliar vocabulary, they instinctively reverted to their L1 as a compensatory strategy to move the story forward. This high occurrence of unresolved episodes in an oral task was also reported by García Mayo and Azkarai (2016) with adult learners. However, the underlying reasons likely differ: while the adults in their study actively engaged in LREs despite not reaching a solution, the participants in the present study may have left many episodes unresolved due to a more limited linguistic repertoire. As Kaspruwicz and Marsden (2018) note, young learners are less equipped with metalinguistic knowledge and young CLIL learners are not an exception in this regard. Besides, the lack of assistance between child learners, possibly due to their egoistic nature at this age (Oliver, 2009) could have contributed to the abundance of unresolved episodes. In other pieces of research with children, in contrast, it has been found that resolved episodes were more common than unresolved in matched-gender pairs (Martínez-Adrián & Gutiérrez-Mangado, 2022; Barañano-Arroyo, 2023). Nonetheless, in these two studies, the children were 11–12 and 10–12 years old, respectively, slightly older than the ones in the present paper. Thus, the larger number of unresolved episodes in the present paper may be put down to their younger age, as those in CLIL I were aged 9.

Lastly, in CLIL II, the second most frequent type of LREs were resolved target-like and unresolved for matched-gender pairs. Interestingly, the amount of resolved target-like episodes might suggest that the benefits from working in same-gender pairs seem to continue as age/proficiency increases, at least in these participants' two-year age gap, since matched-gender dyads also produced more episodes than mixed-gender pairs in CLIL II. This is consistent with the results from Pladevall-Ballester (2021), who also noted that the amount of solved episodes was higher at time 2 due to the learners' age and proficiency difference (see also Storch & Aldosari, 2013 with adults and Basterrechea & Leaser, 2019 with teenagers). Likewise, this finding may support L1 acquisition research suggesting that same-gender preferences intensify as children grow older (Hanish & Fabes, 2014). In the context of young learners, developmental factors significantly shape interaction patterns and, in this case, children aged 9 to 11 are in a phase where peer relationships and social dynamics are evolving. According to the Piagetian framework (Piaget, 1972), children in this age range are transitioning from concrete operational to formal operational stages, impacting their social interactions and collaborative capabilities. In same-gender dyads, shared social norms and communication styles may facilitate more effective collaboration and resolution of LREs than in mixed dyads, who only resolved one episode in a target-like manner.

6. CONCLUSION AND PEDAGOGICAL IMPLICATIONS

This paper has aimed to contribute to the limited body of research examining how gender-pairings influence task-based collaborative dialogue among young CLIL learners. Until now, few studies have focused on LRE production in different gender-pairings in young learners and no investigations have yet aimed at exploring the interplay between gender and age/proficiency in middle childhood, more specifically, between 9- and 11-year-olds. Albeit preliminarily, the tendencies herein reported have revealed that same-gender dyads seemingly work better than mixed-gender pairs in both groups, producing more LREs in total. Although students in this study were not allowed to choose their partners as in previous research, same-gender dyads show patterns similar to those seen in self-selected pairs: they are more engaged, produce more LREs and try to solve more of them, likely due to the greater comfort and ease of communication often found with same-gender peers at this age, especially if also paired by proficiency. Likewise, given that CLIL tasks demand simultaneous attention to content and language (Llinares et al., 2012), finding the best collaborative conditions becomes pivotal for educators. It has also been shown that as age/proficiency increases, same-gender dyads even resolve more LREs in a more accurate way. This suggests that proficiency might maximise the learning opportunities of matched-gender pairings attested in the literature.

At this point, the ecological validity of the study should be stressed. Although exploratory in nature, this paper has uncovered outcomes that could be identifiable in typical intact classes. Consequently, some pedagogical implications applicable in similar classroom environments arise from the preliminary results. Firstly, teachers should consider that working in matched-gender pairs may benefit interaction and lead to more linguistic discussions, as supported by previous research. Secondly, despite their great potential, oral tasks such as the one in this study do not seem to promote many discussions about formal linguistic aspects, noticing and therefore the acquisition of problematic features (Doughty & Williams, 1998). Consequently, teachers should combine them with other types of tasks including a written component, with focused tasks (e.g. dictogloss, see Calzada & García Mayo, 2021) or with metalinguistic explanations, which may favour a focus on more formal language aspects and greater elaboration of LREs (see Martínez-Adrián & Iglesias-Diéguez, under review). This is particularly interesting for CLIL settings, where the main emphasis is oftentimes placed on meaning and not on accuracy (Basterrechea & García Mayo, 2013). Also, oral tasks could also be accompanied by an editing phase, which may lead to more correct resolutions (see Martínez-Adrián & Gallardo-del-Puerto, 2021) and benefit the 4th-year learners even more.

The results should be interpreted with caution and some limitations should be borne in mind. Firstly, this study has dealt with a limited dataset and a reduced sample of students from only one class per school year from the same school. Future research should investigate larger groups to improve the generalisability of the results. Likewise, future studies with a higher number of MM and FF dyads will shed more in terms of gender differences and the potential blurring effect of CLIL in this regard, as attested in previous studies in primary and secondary education (e.g. Heras & Lasagabaster, 2015; Basterrechea et al., 2017). Also, this study did not explicitly control for variables like learner confidence, language anxiety or familiarity with their partner, and such factors may have influenced interactional dynamics, particularly when paired in mixed-gender settings (where learners at this young age may feel less comfortable). Lastly, whilst this study contributes to the understanding of LRE production with young CLIL learners, it is important to highlight the need to contextualise LRE production within particular developmental contexts inasmuch as children's L2 learning cannot be compared to that of adults.

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Science and Innovation (PID2020-113630-GB-I00), the Basque Government (IT1426-22) and the University of the Basque Country (PIF2022/128). Special thanks are extended to the school and, particularly, to the learners who participated in the study.

REFERENCES

- Adams, R., & Ross-Feldman, L. (2008). Does writing influence learner attention to form? In D. Belcher, & A. Hirvela (Eds.), *The oral-literate connection: Perspectives on L2 speaking, writing, and other media interactions* (pp. 243-266). Ann Arbor: The University of Michigan Press.
- Alcón, E. (1994). "The role of participation and gender in non-native speakers' classroom interaction", *Working Papers on Language, Gender, and Sexism*, 4/1, 51-68.
- Alcón, E., & Codina, V. (1996). "The impact of gender on negotiation and vocabulary learning in a situation of interaction", *International Journal of Psycholinguistics*, 12, 21-35.
- Azkarai, A. (2015). "Males and females in EFL task-based interaction: Does gender have an impact on LREs?", *Vigo International Journal of Applied Linguistics*, 12, 9-35.

- Azkarai, A., & García Mayo, M. del P. (2012). Does gender influence task performance in EFL? Interactive tasks and language related episodes. In E. Alcón Soler & M. P. Safont Jordá (Eds.), *Language learners' discourse across L2 instructional settings* (pp. 249-278). Amsterdam: Rodopi. https://doi.org/10.1163/9789401208598_013
- Azkarai, A., & Imaz-Agirre, A. (2017). Gender and age in child interaction in an EFL CLIL context: An exploratory study. In M. del P. García Mayo (Ed.), *Learning foreign languages in primary school: research insights* (pp. 103-123). Clevedon: Multilingual Matters. <https://doi.org/10.2307/jj.22730589.10>
- Azpilicuenta-Martínez, R. (2025). "Exposure or age? The effect of additional CLIL instruction on young learners' grammatical complexity while performing an oral task", *Language Teaching Research*, 0(0). <https://doi.org/10.1177/13621688241303250>
- Barañano-Arroyo, M. (2023). *Gender effects and attention to form in tasks with a written component* [Unpublished undergraduate thesis]. University of the Basque Country.
- Basterrechea, M., & Gallardo-del-Puerto, F. (2020). "Language-related episodes and pair dynamics in primary school CLIL learners: A comparison between proficiency matched and student-selected pairs", *Studies in Second Language Learning and Teaching*, 10/3, 423-447. <https://doi.org/10.14746/ssllt.2020.10.3.2>
- Basterrechea, M., & Gallardo-del-Puerto, F. (2023). "Collaborative writing and patterns of interaction in young learners: The interplay between pair dynamics and pairing method in LRE production", *Vigo International Journal of Applied Linguistics*, 20, 49-76. <https://doi.org/10.35869/vial.v0i20.4354>
- Basterrechea, M. & García Mayo, M. P. (2013). Language-related episodes (LREs) during collaborative tasks: A comparison of CLIL and EFL learners. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 24-43). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.34.05ch2>
- Basterrechea, M., & Leiser, M. J. (2019). "Language-related episodes and learner proficiency during collaborative dialogue in CLIL", *Language Awareness*, 28/2, 97-113. <https://doi.org/10.1080/09658416.2019.1606229>
- Basterrechea, M., Martínez-Adrián, M., & Gallardo-del-Puerto, F. (2017). "Gender effects on strategic competence: A survey study on compensatory strategies in a CLIL context", *ELIA: Estudios de Lingüística Inglesa Aplicada*, 17, 47-70. <https://doi.org/10.12795/elia.2017.17.03>
- Calzada, A., & García Mayo, M. P. (2021). "Child learners' reflections about EFL grammar in a collaborative writing task: When form is not at odds with communication", *Language Awareness*, 30/1, 1-16. <https://doi.org/10.1080/09658416.2020.1751178>
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing.
- Donaldson, J. (2012). *Room on the broom*. Penguin Books.
- Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press.
- The Douglas Fir Group. (2016). "A transdisciplinary framework for SLA in a multilingual world", *The Modern Language Journal*, 100, 19-47. <https://doi.org/10.1111/modl.12301>
- Gallardo-del-Puerto, F., & Basterrechea, M. (2024). "The role of proficiency and pair formation method in language-related episodes: A study of young CLIL learners' interaction", *Language Teaching Research* 28/3, 1087-1105. <https://doi.org/10.1177/13621688211024363>
- Gallardo-del-Puerto, F., & Martínez-Adrián, M. (2022). "Task-modality effects on the production and elaboration of language-related episodes: A study on schoolchildren's interactions in a foreign language", *Languages* 7/4, 306. <https://doi.org/10.3390/languages7040306>
- García Mayo, M. del P., & Azkarai, A. (2016). EFL task-based interaction: Does task-modality impact on language-related episodes? In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Research agenda and pedagogical potential* (pp. 241-266). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.45.10gar>
- García Mayo, M. del P., & Hidalgo, M. de los Á. (2017). "L1 use among young EFL mainstream and CLIL learners in task-supported interaction", *System*, 67, 132-145. <https://doi.org/10.1016/j.system.2017.05.004>
- García Mayo, M. del P., & Imaz-Agirre, A. (2016). "Task repetition and its impact on EFL children's negotiation of meaning strategies and pair dynamics: An exploratory study", *The Language Learning Journal*, 44/4, 451-466. <https://doi.org/10.1080/09571736.2016.1185799>
- García Mayo, M. del P., & Imaz-Agirre, A. (2019). "Task-modality and pair formation method: Their impact on patterns of interaction and LREs among EFL primary school children", *System* 80, 165-175. <https://doi.org/10.1016/j.system.2018.11.011>
- Gass, S. M., & Varonis, E. (1986). Sex differences in NNS/NNS interactions. In R. Day (Ed.), *Talking to learn: Conversation in second language interaction* (pp. 327-351). Cambridge, MA: Newbury House.
- Gutiérrez-Mangado, M., & Basterrechea, M. (2025). "Investigating the effect of task modality on the written and oral production of young EFL learners", *International Review of Applied Linguistics in Language Teaching*. <https://doi.org/10.1515/iral-2024-0206>

- Hakes, D. T. (1980). A study of children's metalinguistic abilities: Results and discussion. In D. T. Hakes (Ed.), *The development of metalinguistic abilities in children* (pp. 61-96). Berlin: Springer. https://doi.org/10.1007/978-3-642-67761-8_4
- Hanish, L. D., & Fabes, R. A. (2014). Peer socialization of gender in young boys and girls. *Encyclopaedia on Early Childhood Development*. Montreal, Quebec: *Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development*, 1-4. <https://www.child-encyclopedia.com/pdf/expert/gender-early-socialization/according-experts/peer-socialization-gender-young-boys-and-girls>
- Heras, A., & Lasagabaster, D. (2015). "The impact of CLIL on affective factors and vocabulary learning", *Language Teaching Research*, 19/1, 70-88. <https://doi.org/10.1177/1362168814541736>
- Hidalgo, M. de los Á. (2021). "Differences in the task-supported negotiations of younger and older EFL children: From repair into prevention", *International Review of Applied Linguistics in Language Teaching*, 59/3, 395-419. <https://doi.org/10.1515/iral-2018-0206>
- Holmes, J., & Meyerhoff, M. (2003). Different voices, different views: An introduction to current research in language and gender. In J. Holmes & M. Meyerhoff (Eds.), *Handbook of language and gender* (pp. 1-17). Oxford: Blackwell. <https://doi.org/10.1002/9780470756942.ch>
- IBM Corp. (2021). IBM SPSS Statistics for Windows (Version 28.0) [Computer software]. IBM Corp.
- Iglesias-Diéguez, K., Martínez-Adrián, M., & Gallardo-del-Puerto, F. (2025). "Peer-to-peer interaction in task-based learning and the development of young EFL learners' productive vocabulary: an exploratory study", *Language Awareness*, 1-26. <https://doi.org/10.1080/09658416.2025.2456754>
- Iwashita, N. (2001). "The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language", *System*, 29/2, 267-287. [https://doi.org/10.1016/S0346-251X\(01\)00015-X](https://doi.org/10.1016/S0346-251X(01)00015-X)
- Kasproicz, R. E., & Marsden, E. (2018). "Towards ecological validity in research into input-based practice: Form spotting can be as beneficial as form-meaning practice", *Applied Linguistics*, 39, 886-911. <https://doi.org/10.1093/applin/amw051>
- Kim, Y., & McDonough, K. (2008). "The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners", *Language Teaching Research* 12/2, 211-234. <https://doi.org/10.1177/1362168807086288>
- Lasagabaster, D. (2008). "Foreign language competence in Content and Language Integrated Courses", *The Open Applied Linguistics Journal*, 1, 30-41. <https://doi.org/10.2174/1874913500801010030>
- Leaper, C., & Ayres, M. M. (2007). "A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech", *Personality and Social Psychology Review*, 11/4, 328-363. <https://doi.org/10.1177/1088868307302221>
- Leeser, M. J. (2004). "Learner proficiency and focus on form during collaborative dialogue", *Language Teaching Research*, 8/1, 55-81. <https://doi.org/10.1191/1362168804lr134oa>
- Llinares, A. & Morton, T. (2017). *Applied linguistics perspectives on CLIL*. Amsterdam: John Benjamins. <https://doi.org/10.1075/illt.47>
- Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge, UK: Cambridge University Press.
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). New York: Academic Press. <https://doi.org/10.1016/B978-012589042-7/50015-3>
- Mackey, A. & Goo, J. (2012). Interaction approach in second language acquisition. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2748-2758). Oxford: Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0551>
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd edition). Mahwah, NJ: Lawrence Erlbaum Associates.
- Maltz, D., & Borker, R. (1982). A cultural approach to male-female miscommunication. In J. J. Gumperz (Ed.), *Language and social identity* (pp. 196-2016). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620836.013>
- Martínez-Adrián, M., & Arratibel-Irazusta, I. (2020). "The interface between task-modality and the use of previously known languages in young CLIL English learners", *Studies in Second Language Learning and Teaching*, 10/3, 473-500. <https://doi.org/10.14746/ssllt.2020.10.3.4>
- Martínez-Adrián, M., & Arratibel-Irazusta, I. (2024). "Pairing method and young learners' use of previously known languages during task-based interaction", *TASK*, 4/1, 78-110. <https://doi.org/10.1075/task.22011.mar>
- Martínez-Adrián, M., & Gallardo-del-Puerto, F. (2021). "Task-modality and language related episodes in young learners: An attempt to manage accuracy and editing", *Language Teaching Research*, 0(0). <https://doi.org/10.1177/13621688211052808>

- Martínez-Adrián, M., & Gutiérrez-Mangado, M. J. (2022). "Gender-pairings in EFL child task-based interaction", *Vigo International Journal of Applied Linguistics*, 19, 103-128. <https://doi.org/10.35869/vial.v0i19.3761>
- Martínez-Adrián, M., & Iglesias-Diéguez, K. (under review). "What role do metalinguistic explanations play on LRE production in a dictogloss task targeting the possessive determiners *his/her*?"
- Martínez-Adrián, M. & Nieva-Marroquín, M. (2023). Exploring the pseudo-longitudinal development of specific morphosyntactic features and syntactic complexity in CLIL young learners. *International Review of Applied Linguistics in Language Teaching*, 61/3, 1197-1223. <https://doi.org/10.1515/iral-2021-0137>
- McCloskey, L. A., & Coleman, L. M. (1992). "Difference without dominance: Children's talk in mixed- and same-sex dyads", *Sex Roles* 27, 241-257. <https://doi.org/10.1007/BF00289927>
- Merino, J. A., & Lasagabaster, D. (2018). "CLIL as a way to multilingualism", *International Journal of Bilingual Education and Bilingualism*, 21/1, 79-92. <https://doi.org/10.1080/13670050.2015.1128386>
- Mozaffari, S. H. (2017). "Comparing student-selected and teacher-assigned pairs on collaborative writing", *Language Teaching Research*, 21/4, 496-516. <https://doi.org/10.1177/1362168816641703>
- Muñoz, C. (2007). Age-related differences and second language learning practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 229-255). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667275.014>
- Muñoz, C. (2014). The development of language awareness at the transition from primary to secondary School. In M. del P. García Mayo (Ed.), *Learning foreign languages in primary school: Research insights* (pp. 49-68). Bristol: Multilingual Matters. <https://doi.org/10.21832/9781783098118-005>
- Oliver, R. (2002). "The patterns of negotiation for meaning in child interactions", *The Modern Language Journal*, 86/1, 97-111. <https://doi.org/10.1111/1540-4781.00138>
- Oliver, R. (2009). How young is too young? Investigating negotiation of meaning and feedback in children aged five to seven years. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction. Second language research in honor of Susan M. Gass* (pp. 135-156). New York: Routledge.
- Oliver, R., & Azkarai, A. (2017). "Review of child second language acquisition (SLA): Examining theories and research", *Annual Review of Applied Linguistics*, 37, 62-76. <https://doi.org/10.1017/S0267190517000058>
- Oliver, R., & Azkarai, A. (2019). "Patterns of interaction and young ESL learners: What is the impact of proficiency and task-type?", *Language Teaching for Young Learners*, 1/1, 82-102. <https://doi.org/10.1075/ltyl.00006.oli>
- Payant, C., & Kim, Y. (2019). "Impact of task-modality on collaborative dialogue among plurilingual learners: A classroom-based study", *International Journal of Bilingual Education and Bilingualism*, 22/5, 614-627. <https://doi.org/10.1080/13670050.2017.1292999>
- Philp, J., Adams, R., & Iwashita, N. (2013). *Peer interaction and second language learning* (1st ed.). New York: Routledge. <https://doi.org/10.4324/9780203551349>
- Philp, J., Walter, S., & Basturkmen, H. (2010). "Peer interaction in the foreign language classroom: what factors foster a focus on form?", *Language Awareness*, 19, 261-279. <https://doi.org/10.1080/09658416.2010.516831>
- Piaget, J. (1972). *The psychology of the child*. New York: Basic Books.
- Pica, T. (2013). From input, output and comprehension to negotiation, evidence, and attention: An overview of theory and research on learner interaction and SLA. In M. del P. García Mayo, J. Gutiérrez-Mangado & M. Martínez-Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 49-69). Amsterdam: John Benjamins. <https://doi.org/10.1075/aals.9.06ch3>
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). "Comprehensible output as an outcome of linguistic demands on the learner", *Studies in Second Language Acquisition*, 11/1, 63-90. <https://doi.org/10.1017/S027226310000783X>
- Pica, T., Holliday, L., Lewis, N., Berducci, D., & Newman, J. (1991). Language learning through interaction: What role does gender play? *Studies in Second Language Acquisition*, 13/3, 343-376. <https://doi.org/10.1017/S0272263100010020>
- Pinter, A. (2006). "Verbal evidence of task-related strategies: Child versus adult interactions", *System*, 34/2, 615-630. <https://doi.org/10.1016/j.system.2006.09.005>
- Pladevall-Ballester, E. (2021). "Pair dynamics and language-related episodes in child EFL task-based peer interaction", *Language Teaching for Young Learners*, 3/2, 189- 213. <https://doi.org/10.1075/ltyl.20007.pla>
- Plonsky, L., & Oswald, F. L. (2014). "How big is "big"? Interpreting effect sizes in L2 research", *Language Learning*, 64/4, 878-912. <https://doi.org/10.1111/lang.12079>
- Ross-Feldman, L. (2005). *Task-based interactions between second language learners: Exploring the role of gender*. [Unpublished doctoral dissertation]. Georgetown University.

- Ross-Feldman, L. (2007). Interaction in the L2 classroom: Does gender influence learning opportunities? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 52-77). New York: Oxford University Press.
- Sánchez Ruiz, R., & Lucas Pardo, S. (2016). "Improving English oral skills through pair work and peer-peer dialogue", *Revista de Lenguas para Fines Específicos*, 22/1, 170-189. <http://dx.doi.org/10.20420/rife.2016.309>
- Sato, M., & Csizér, K. (2021). "Introduction: Combining learner psychology and ISLA research: Intersections in the classroom", *Language Teaching Research*, 25/6, 839-855. <https://doi.org/10.1177/13621688211044237>
- Sato, M., & Viveros, P. (2016). Interaction or collaboration? Group dynamics in the foreign language classroom. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 91-112). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.45.04sat>
- Shehadeh, A. (1994). "Gender differences and second language acquisition", *Research Journal of Aleppo University (Arts and Humanities Series)* 26, 73-98.
- Storch, N. (2002). "Patterns of interaction in ESL pairwork", *Language Learning*, 52/1, 119-158. <https://doi.org/10.1111/1467-9922.00179>
- Storch, N., & Aldosari, A. (2013). "Pairing learners in pair work activity", *Language Teaching Research*, 17/1, 31-48. <https://doi.org/10.1177/1362168812457530>
- Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82/3, 320-337. <https://doi.org/10.1111/j.1540-4781.1998.tb01209.x>
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.
- Swain, M., & Watanabe, Y. (2013). Languaging: Collaborative dialogue as a source of second language learning. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). Oxford: Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0664.pub2>
- Tannen, D. (1994). *Gender and discourse*. New York: Oxford University Press.
- UCLES (n.d.). A2 Flyers. Available at: <https://www.cambridgeenglish.org/exams-and-tests/flyers/>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Watanabe, Y., & Swain, M. (2007). "Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners", *Language Teaching Research*, 11/2, 121-142. <https://doi.org/10.1177/136216880607074599>