Language-related episodes and pair dynamics in primary school CLIL learners: A comparison between proficiency-matched and student-selected pairs

María Basterrechea
University of the Basque Country, Vitoria-Gasteiz, Spain
https://orcid.org/0000-0002-7993-445X
maria.basterrechea@ehu.eus

Francisco Gallardo-del-Puerto
University of Cantabria, Santander, Spain
https://orcid.org/0000-0001-8578-9861
francisco.gallardo@unican.es

Abstract

A considerable body of research within the interaction framework (Long, 1996) has centred on the language-related episodes (LREs) which occur when learners topicalize a specific linguistic item while they are engaged in meaning-focused tasks. Several studies have shown that the production of LREs may be influenced by the proficiency level of the learners (Kim & McDonough, 2008; Leeser, 2004). Sociocultural theory (Lantolf & Appel, 1994) has also explored collaborative work and the effect that pairing learners with the same proficiency levels or different *patterns of interaction* (Storch, 2002) has on the production of LREs (e.g., Mozaffari, 2017; Storch & Aldosari 2013), but little research has compared the effect of the pair formation method (student-selected vs. proficiency-matched) on young learners' production of LREs and pair dynamics. This study compares young CLIL learners (aged 10-12) in student-selected and proficiency-matched pairs in task-based interaction. Results

indicate that learners produce more meaning-based than form-based LREs, regardless of their pair formation method. The percentage of meaning-based LREs which are resolved accurately is much higher in proficiency-matched dyads than in student-selected ones. As for the patterns of interaction (Storch, 2002), the dynamics of proficiency-matched dyads are of a more collaborative nature than those of self-selected pairs.

1. Introduction

Research within the interaction framework (Long, 1996) has examined the facilitative role of collaborative interaction in providing opportunities for comprehensible input, peer feedback and output, which are important mechanisms in the process of L2 learning (Loewen & Sato, 2018). A considerable body of research has centered on the language-related episodes (LREs) which occur when learners topicalize a specific linguistic item while they are engaged in meaning-based tasks. LREs are ". . . any part of the dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct" (Swain, 1998, p. 70). Several studies have shown that the production of LREs may be influenced by the proficiency level of the learners (Basterrechea & Leeser, 2019; Benson, Pavitt, & Jenkins, 2005; Kim & McDonough, 2008; Kowal & Swain, 1994; Leeser, 2004; Malmqvist, 2005; Williams, 1999, 2001), gender (Ross-Feldman, 2007), or personality traits (Storch & Aldosari, 2013; Watanabe & Swain, 2007), factors that seem to affect the nature of interactional feedback, as well as the quantity, quality and outcome of LREs.

More recently, collaborative work has also been researched from the perspective of sociocultural theory (Lantolf & Appel, 1994). A growing number of studies within this framework explore not only how learners engage in problem-solving and knowledge building (Swain, 2000) or construct new knowledge about language collaboratively (Storch, 2007), but they also focus on how learners' behavior and affective factors impact language learning opportunities that arise in collaborative dialogue (Swain, 2000). In particular, some studies within this framework have examined the effect that pairing learners with same proficiency levels or pair dynamics has on the production of LREs in terms of amount and type (e.g., García Mayo & Imaz Aguirre, 2019; Mozaffari, 2017; Storch, 2002; Storch & Aldosari, 2013; Watanabe & Swain, 2007). Results seem to suggest that these are crucial moderating factors that influence the potential that peer collaboration may

have for language learning. Hence, the impact of pair dynamics and the learners' relative proficiency on the production of LREs need to be explored experimentally. A population that has received scant attention in LRE research is that of young learners, as most research on LREs has investigated adults in English-as-a-second-language (ESL; Benson et al., 2005), immersion (Kowal & Swain, 1994; Swain 1998; Swain & Lapkin 1998), content-based instruction (Leeser, 2004), or foreign language settings (Basterrechea & García Mayo, 2013; Basterrechea & Leeser, 2019; Kim & McDonough, 2008; García Mayo, 2002a, 2002b; García Mayo & Azkarai, 2016; Malmqvist, 2005; Storch & Aldosari, 2013). Hence, more research on child interaction is needed in order to explore if: (a) young learners focus on language form while interacting, (b) they consciously reflect on their language, and (c) pairing learners according to proficiency or friendship affects pair dynamics. With the exception of García Mayo and Imaz Aguirre (2019), no studies have investigated the impact that matched proficiency or student self-selection have on the production of LREs by young EFL learners. Therefore, this study investigates how pair formation (student-selected vs. proficiency-matched) may influence the quantity and quality of LREs produced by young content-and-language-integrated-learning (CLIL) learners of English in two oral tasks, namely a narration task and a map task. In addition, drawing on Storch's (2002) model of *patterns of interaction* in pair work, this study focuses on pair behavior by student-selected and proficiency-matched pairs in order to explore how pairing method influences pair dynamics in collaborative work as ". . . simply assigning students to work in groups or pairs will not necessarily create conditions conducive to learning " (Storch, 2002, p. 122).

2. Literature review

2.1. LREs and pair formation method

One strand of research on LREs has investigated how learner-internal factors, such as target language proficiency, affect the quantity, quality and outcome of LREs (Basterrechea & Leeser, 2019; Benson et al., 2005; Kim & McDonough, 2008; Kowal & Swain, 1994; Leeser, 2004; Malmqvist, 2005; Mozaffari, 2017; Payant, 2018; Storch & Aldosari, 2013; Watanabe & Swain, 2007; Williams, 1999, 2001). Since LREs are claimed to be a window into the learners' level of proficiency (Kowal & Swain, 1994), the assumption behind these studies is that proficiency may have an impact on the type of LREs produced. Several studies have shown that high proficiency learners produce a greater number of form-based LREs than meaning-based LREs, and correctly resolve a higher amount of LREs than high-low and low-low proficiency dyads (Basterrechea & Leeser, 2019;

Leeser, 2004). However, it has been observed that the relationship formed by dyad members or personality traits influence the number and type of LREs. For instance, Malmqvist (2005) examined the production of LREs by Swedish learners of German as a foreign language. Although lower proficiency learners attended to lexical items, it was these same learners who took a leading role in group discussion. The study concluded that the relationship formed by dyad members should be considered when pairing learners.

It is still unclear how we can best form "high performance groups" (Mozaffari, 2017, p. 496), and whether dyad members' attention to form may be influenced more by their proficiency than by the relationships they form. Mozaffari (2017) compared, among other issues, the production of LREs by student-selected and teacher-assigned pairs in adult (age 20-26) English-as-a-foreign-language (EFL; Iranian L1) learners, as well as the nature of the talk of the comparison groups. All the participants represented an intermediate level according to the *Oxford Placement Test* (Allan, 2004). Teacher-selected pairs were randomly assigned on the basis of their test scores. In the case of student-selected pairs, friendship was mentioned as the sole criterion for choosing a companion. Results revealed that teacher-assigned pairs produced significantly more LREs than student-selected ones. In terms of type, teacher-assigned pairs produced more form-based episodes and meaning-based episodes than student-selected pairs. The analysis of the nature of the talk indicated that student-selected pairs talked about matters unrelated to the task more frequently than teacher-selected pairs. The results reported by García Mayo and Imaz Aguire (2019) mirror those of Mozaffari in a study that examined the effect of the pair formation method on the production of LREs with young EFL learners, a population that has received scant attention in the LRE research (but see Gallardo-del-Puerto & Basterrechea, in press). They compared the number and type of LREs produced in an oral and an oral + written task by proficiency-paired, teacher-selected (based on their personality, according to their teacher) and student-selected pairs. Results showed that proficiency-matched pairs produced more LREs in both types of tasks, followed by the teacher-assigned pairs and finally by student-selected pairs. Experimental studies on the effect of the pair formation method on the production of LREs seem to show that the relationship formed by the dyad members may be a more relevant factor than proficiency. Thus, the relationship among LREs, proficiency and pair dynamics needs to be further explored.

2.2. Collaborative dialogue and pair dynamics

More research is needed that will bring to light the factors that can create conditions conducive to language learning during collaborative tasks. One important

issue related to linguistic interactions is the impact of pair behavior in collaborative tasks. Studies such as Donato (1998) showed that some patterns of interaction are more beneficial than others, as it was found that collaborative scaffolding (i.e., those patterns facilitating actions that a more competent peer provides to the learner [Wood, Bruner, & Ross, 1976]) did not occur in all groups. In fact, collaborative dialogue research has shown that pair dynamics plays a crucial role in the degree to which pairs collaborate (e.g., Kim & McDonough, 2008; Kowal & Swain, 1994; 1997; Nelson & Murphy, 1993). Storch (2002) examined the patterns of interaction and the quantity and quality of LREs produced by 10 pairs of adult ESL students. Based on the data, she put forward four different patterns of dyadic interaction that described the role relationships in pair work, which depicted: (a) the degree of control or authority over the task (equality), and (b) the level of engagement with the contributions of the peer (mutuality). Based on these two axes, four interactional styles or patterns of interaction were defined:

- *collaborative* – a pattern where there is high equality and mutuality; learners work together and contribute jointly to the task;
- *dominant/dominant* – a pattern where there is high equality but low mutuality; both participants contribute to the task but do not engage with each other's contribution, and hence, do not reach a consensus easily; this pattern includes a cooperative type of interaction, subsequently labelled as passive/parallel (Butler & Zeng, 2015), where both members participate but there is "division of labor," as they do not engage with each other's contribution;
- *dominant/passive* – both equality and mutuality are low; one of the members dominates the discussion and little negotiation occurs;
- *expert/novice* – a pattern where there is low equality but high mutuality; the expert or more capable peer takes control over the task but encourages the novice or less capable peer to contribute to the task.

Storch (2002) showed that the patterns that are more conducive to language learning are collaborative and expert/novice, since it was attested that the pairs that exhibited these two types of pair dynamics transferred the knowledge gained in the LREs more frequently to subsequent individual tasks.

Following Storch's (2002) model of dyadic interaction, Kim and McDonough (2008) analyzed the collaborative dialogue of intermediate and advanced learners of Korean as a second language (SL) in terms of LREs and patterns of interaction when they were paired with peers having different proficiency levels. It was found that learners who were *collaborative* with an intermediate interlocutor, were *passive* or *novice* when paired with an advanced peer. However, learners that exhibited a *dominant* role with an intermediate peer took a collaborative role with an advanced interlocutor. In addition, and unlike other studies

(Leeser, 2004; Malmqvist, 2005), learners produced more meaning-based LREs with an advanced interlocutor than with an intermediate one, whereas no differences were found in the production of form-based LREs. The researchers concluded that, as long as pair dynamics are of a collaborative type, intermediate learners can benefit from being paired with an equally intermediate or equally advanced peer.

Following the work by Leeser (2004), Storch and Aldosari (2013) investigated what effect proficiency pairing and patterns of interaction have on LRE production in an EFL classroom in Saudi Arabia with 15 pairs with similar and varying proficiency levels. Similar proficiency dyads (high-high [H-H] and low-low [L-L]) exhibited a tendency to make more collaborative contributions and produced the highest number of LREs, whereas mixed proficiency dyads (H-L) exhibited a wider range of patterns. In addition, mixed proficiency dyads produced more LREs when they formed a collaborative or expert/novice type of relationship, but when they formed non-collaborative relationships, the more proficient learner tended to take a leading role, while the less proficient learner contributed neither to the task nor to the focus on language use. In terms of production of LREs, H-H produced the greatest number of form-based and correctly solved LREs, followed by H-L and L-L pairs, as evidenced in previous studies (e.g., Leeser, 2004).

As mentioned above, García Mayo and Imaz Aguirre (2019) examined the influence of pair formation on the production of LREs by young EFL learners. To our knowledge, theirs is the only study addressing those issues with this age group. No differences were found in the different pairs (proficiency-paired, teacher-selected and self-selected groups), as all featured a collaborative type of dynamics.

In summary, the scarce research on the effect that the pairing method has on the production of LREs seems to show that pairing learners based on their relative proficiency is more beneficial in fostering attention to language form than having students choose their partners, but how pair behaviour impacts the production of LREs in the two pairing methods needs to be further explored. Thus, this study has been conducted in an attempt to shed light on this under-explored area by examining the occurrence, type and outcome of LREs in proficiency-matched and student-selected dyads of L2 English learners. The study also seeks to take a close look at the learners' interactional patterns. More specifically, it attempts to investigate what effect the pair formation method has on young CLIL learners' patterns of interaction and on their attention to formal aspects of language in a context where learners' attention to form needs further investigation. On the basis of previous research, two research questions were posed:

1. Does the pairing method have an effect on the number, type and outcome of LREs?
2. Does the pairing method exert an influence on the patterns of interaction?

3. Methodology

3.1. Participants

Twenty-seven pairs of EFL learners in Grades 5 and 6 at a primary education school in the Basque Country took part in the study. They were all Basque-Spanish bilinguals aged 10 to 12 who started learning English as a school subject in pre-primary education, and subsequently in Grade 3 (age 9) they were engaged in a CLIL program where they have been learning various content subjects in English (e.g., arts and crafts, physical education, science). On average, these beginner learners had received 832 hours of exposure to English in a formal setting. Students were divided into two different groups depending on the pair formation method, namely, a proficiency-matched group and a student-selected group. They were recruited from 5 intact classrooms: 2 of them from 5th grade, and 3 of them from 6th grade. In each grade, participants were randomly assigned to the two groups, resulting in one proficiency-matched group and one self-selected group in 5th grade, and one proficiency-matched group and two self-selected groups in 6th grade. As a result, 17 pairs were formed according to the similarity of their members' scores in the *Key English Test* (KET; Cambridge University Press, 2008), which they had taken before data collection. Ten pairs were formed on the basis of self-selection (see Mozaffari, 2017, for the same pairing method).

3.2. Instruments

Apart from a general background questionnaire and an English proficiency test (KET; Cambridge University Press, 2008), which were administered to obtain information about participants' biographical profiles and English proficiency, students completed two consecutive collaborative tasks. In the first task, the members of the dyad collaborated to put a set of pictures in order to create a story (see Mackey, 1994). Participants were instructed to find the story behind a set of pictures adapted from *Dotty's doll* activity (*Sparks 1: Teacher's Book*; House & Scott, 2009, p. 74). The pictures, which are included in Appendix A, tell the story of a girl who does not know how to put the different broken pieces of her doll together, but with the help of a friend she finally succeeds. Once the participants agreed on the order, they were asked to tell the story orally in turns. The second task (see Appendix B) was specifically designed for this study and it was made up of three different phases. In the first phase, students were shown two pictures. In the first picture, a boy named Ben finds a lost dog in a park. The dog has a photograph in its mouth, presumably of its owner, and its collar reads "I belong to J. Smith." The second picture shows some possible owners and their

professions (e.g., doctor, vet) together with the map of a town showing the different places where these people might work (e.g., hospital, vet clinic, nursing school). Students were required to agree on the owner of the dog and on where s/he works. In the second phase, students had to agree on the itinerary they had to follow around various landmarks in the map from the park to the place where the dog's owner works. In the final stage, they had to collaborate in the writing of a short note for Ben explaining who the dog's owner is and giving directions from the park to the owner's workplace so that Ben can take the dog back to its owner. In sum, both were collaborative close-ended tasks in which the learners in dyads worked towards a convergent goal, but they differed in the end product in that the first task had a final oral outcome, and the second one had a final written component, which enabled us to assess the participants in both oral and written modes.

3.3. Procedure

The two tasks described above were completed by the dyads consecutively in a quiet room at school. A researcher was in the room with the students as they performed the tasks, but s/he did not intervene in their interaction unless they so required. The researcher encouraged dyad members to work jointly in the process and to ask one another for assistance when needed. It is also important to note that students were reminded of the importance of paying attention to the language they used to accomplish the last phase in each task, that is to say, the story telling in turns in the first task and the note writing in the second one. At this stage, they were also informed about the possibility of giving each other corrective feedback during the tasks. More specifically, they were told to make sure that their pronunciation (Task 1) or spelling (Task 2) was correct and there were no mistakes in their eventual oral and written productions. On average, student pairs needed about 30 minutes to complete both tasks consecutively.

Dyadic interactions were both audiotaped and videotaped. Recorded productions were transcribed and later analyzed for the production of LREs with the help of CHILDES codification protocols (MacWhinney, 2000). LREs were identified at various stages during the completion of each task, that is, both when learners interacted in order to resolve the first phases in each task (Task 1: picture ordering; Task 2: deciding on the dog's owner), and in the course of producing the final outcomes collaboratively (story-telling in Task 1; note writing in Task 2). All cases in which students discussed a language issue or self-corrected were independently identified by two researchers, who showed a high degree of agreement in their judgements. Any occasional controversy in the classification of the LREs was solved by the two researchers together on a case-by-case basis.

3.4. Analysis

Two types of analysis were conducted: an analysis of the different types of LREs and analysis of the pair dynamics. Regarding the former, LREs were initially categorized as *meaning-based*, when the addressed issue was lexical choice or meaning, or *form-based*, when phonology, spelling, morphosyntax or prepositions were involved, following García Mayo and Azkarai's (2016) taxonomy. Each of these two categories was further classified according to Leeser (2004) as *resolved* or *unresolved*, depending on whether interactants reached a final decision on the matter discussed or left it unresolved. Finally, whenever a decision was made, the linguistic outcome was assessed for accuracy and the LRE in question was further labelled as either *target-like*, when the outcome corresponded to a correct L2 element, or *non-target-like*, when the outcome deviated from the intended English word or form. Hence, six different LRE types emerged from our classification, as illustrated in the following examples taken from the students' oral productions in our database:

> (1) Target-like resolved meaning-based LRE
>
> *CHI1:   *and later they start doing it.*
> *CHI2:   *making it.*
> *CHI1:   *making it later.*

In Excerpt 1 (1), Child 1 proposes the use of *do*, but Child 2 corrects Child 1 and suggests using *make*, which is accepted by Child 1 in the following turn, reaching a correct resolution.

> (2) Non-target-like resolved meaning-based LRE
>
> *CHI1:   is *the same* (...) *the same.*
> *CHI2:   *draw?*
> *CHI1:   *yes* (.) *the same picture.*
> *CHI1:   *bueno no.* [Well no]
> *CHI2:   *is the same.*
> *CHI1:   *the same draw.*

In Excerpt 2, Child 2 suggests the word *draw* (for the intended English word *picture*) to help Child 1 finish his/her previous utterance. Child 1 provides a correction it in the next intervention and uses the right word *picture,* but he then hesitates and finally incorporates *draw* unsuccessfully in the last turn.

> (3) Unresolved meaning-based LRE
>
> *CHI2:   *¿cómo se decía coser?* [How do you say to sew?]
> *CHI1:   *eh* (.) *ni idea.* [No idea]

In Excerpt 3, Child 2 asks about the English word for the Spanish verb *coser* 'to sew,' but Child 1 does not know the word and leaves her partner's question unanswered.

(4) Target-like resolved form-based LRE

*CHI2:   *then the girl has a new idea and* (.) *it eh and he want she wants to eh.*

In Excerpt 4, Child 2 corrects his own utterance by adding the English 3rd person singular marker *–s* in *want*, while Child 1 is not involved in the correction.

(5) Non-target-like resolved form-based LRE

*CHI1:   *children happy* (.) *eh.*
*CHI2:   *children is happy.*
*CHI1:   *children is happy.*

In Excerpt 5, Child 1 omits the copula, which is incorporated as *is* by Child 2 in the following turn, an incorrect outcome because of the lack of subject-verb agreement, which is eventually accepted by Child 1 in her last utterance.

(6) Unresolved form-based LRE

*CHI1:   *puede ser* [May be] *go to the chu church.*
*CHI2:   *¿cómo?* [What?]
*CHI2:   bueno no sé cómo se pronuncia. [Well I do not know how to pronounce it]

In Excerpt 6, Child 1 attempts the pronunciation of the word *church* while Child 2 seems not to understand his partner and finally admits that he does not know how to pronounce it either, as evidenced in the last turn.

As for pair dynamics, the patterns of dialogic interaction shown by the various dyads were classified according to the taxonomy proposed by Storch (2002). The following excerpts, also taken from our database, exemplify the different dialogic dynamics:

(7) Collaborative (high equality - high mutuality)

*CHI2:   *in the pocket he's have a* (...).
*CHI1:   *eh.*
*CHI2:   *like a snake* (.) *in a.*
*CHI1:   *a eh* (.) *a glass?*
*CHI2:   *a glass.*

In Excerpt 7, both learners help one another to elaborate the content and take into account each other's utterances so as to build up ideas which follow from

each other's contributions. Child 2 needs an English word (*cup*) and Child 1 suggests a word (*glass*), which Child 2 incorporates into his final utterance.

(8) Dominant/dominant (high equality - low mutuality)

*CHI1:   *to the* (.) *así no se escribe the.* [This is not the way *the* is written]
*CHI2:   *ah eso.* [That] *(he changes the spelling)*
*CHI1:   *the se escribe te hache e.* [*The* is written tee – aitch – e]
*CHI2:   *ya lo sé.* [Yes, I know that]

In Excerpt 8, both learners exhibit a desire to control the completion of the task in addition to a lack of engagement with each other's contribution. Child 1 reproves his partner for a word which was being spelt inaccurately, but Child 2 corrects it on his own and impolitely ignores his/her partner's correction.

(9) Dominant/passive (low equality - low mutuality)

*CHI1:   *cerca?* [Near]
*INV:[1]   *CHI2 is writing?*
*CHI2:   *cerca?* [Near]
*INV: *cerca?* [Near]
*INV: *how do you say cerca* [Near] *CHI1?*
*INV: *in English.*
*CHI1:   *cerca.* [Near] *(thinking)*
%sit:[2]   CHI2 continues writing on his own
*INV: *you speak English very well!*
*INV: *near.*
*INV: *but you have to help CHI2.*

In Excerpt 9, one of the learners has totally delegated his responsibility in the task to Child 1, who has taken the lead in writing. The researcher unsuccessfully makes an indirect attempt to get Child 1 engaged in the task, but eventually explicitly tells her that she should help her partner.

(10) Expert/novice (low equality - high mutuality)

*CHI2:   *eh* (...) *eh* (..) *esto es una chica o un chico*? [Is this a girl or a boy?]
*CHI1:   *es una chica* [it is a girl] *one girl.*
*CHI2:   *ah one girl.*
*CHI1:   *this girl.*
*CHI2:   *this girl eh* (...).

---

[1] INV = investigator.
[2] %sit = situation.

In Excerpt 10, Child 1 takes a more leading role than his partner but they both contribute to the task. Using their first language (L1), Child 2 asks about the gender of the child in the picture and Child 1 answers first in Spanish but immediately afterwards he provides the English term *girl* in an attempt to encourage his partner's contribution to the task. Child 2 incorporates the given word preceded by the numeral *one,* but Child 1 makes the better suggestion of the phrase *this girl*, which is fully reproduced by Child 2 in the last utterance.

4. Results

The results described in this section are organized according to the two research questions in the study. First, we will show the results of the analyses carried out in order to examine whether the pairing method (proficiency-matched vs. student-selected) had an impact on the quantity and quality of the LREs produced. Afterwards, the results concerning the types of pair dynamics identified in the dyadic interactions of the two research groups are shown.

Regarding the first research question (*Does the pairing method have an effect on the number and type and* outcome *of LREs?*), Tables 1-5 display the total number of the various LRE types as well as the percentages of each LRE type, the mean number of LREs per subject, and the range of each of the LRE types. Additionally, these tables offer information related to the two types of inferential statistical analyses which were carried out – independent samples *t-tests* and Mann-Whitney tests for intergroup comparisons between proficiency-matched and student-selected learners (see table rows), and dependent-samples *t-tests* or Wilcoxon tests for intragroup comparisons between the two LRE types examined in each table (see table columns). Statistical significance is marked at $p \leq$ .05 (*), $p \leq$ .01 (**), and $p \leq$ .001 (***) values in Tables 1-5.

Table 1 displays the results for the meaning-based and form-based LREs produced by the two research groups (proficiency-matched vs. student-selected). As shown by the percentages and mean scores obtained, meaning-based LREs were significantly more abundant than form-based LREs in both proficiency-matched and student-selected samples. However, student-selected pairs achieved a significantly higher percentage and mean of meaning-based LREs compared to the proficiency-matched pairs (75.86% vs. 64.07%; respectively; 11.00 vs. 6.29, respectively), whereas their percentage of form-based LREs was lower than that of proficiency-matched learners (24.14% vs. 35.93%). This difference, however did not reach significance.

Table 1 Types of LREs according to pairing method

	Proficiency-matched (N = 17) (t = 3.358, p ≤ .01**)	Student-selected (N = 10) (t = 4.858, p ≤ .001***)
Meaning-based (t = -2.889, p ≤ .05*)		
Number	107	110
%	64.07%	75.86%
Mean	6.29	11.00
Range	0-16	3-17
Form-based (t = .029, p > .05)		
Number	60	35
%	35.93%	24.14%
Mean	3.53	3.50
Range	0-11	0-10

As for the outcome of LREs, Tables 2 and 3 display the results for meaning-based and form-based LREs, respectively. As for the former, the data in Table 2 indicate that most of the meaning-based LREs produced were resolved to a significantly larger extent, with resolution percentage figures of 77.57% and 83.64% in the proficiency-matched and student-selected dyads respectively. Regarding the gaps between the means for resolved and unresolved LREs in each research group, this tendency turned out to be slightly more obvious in the student-selected sample (9.20 vs. 1.80) than in the proficiency-matched (4.82 vs. 1.41) sample. As for intergroup comparisons, proficiency-matched learners' mean of resolved LREs was statistically lower than that of student-selected pairs (4.82 vs. 9.20), whereas there were no differences between the two research groups for unresolved LRE means (1.41 vs. 1.80).

Table 2 Outcome of meaning-based LRE types according to pairing method

	Proficiency-matched (N = 17) (z = 3.415, p ≤ .001***)	Student-selected (N = 10) (z = 2.807, p ≤ .01**)
Resolved (t = -3.100, p ≤ .01**)		
Number	83	92
%	77.57%	83.64%
Mean	4.82	9.20
Range	0-14	2-15
Unresolved (z = -.854, p > .05)		
Number	24	18
%	22.43%	16.36%
Mean	1.41	1.80
Range	0-5	1-5

Regarding the outcome of form-based LREs (Table 3), it was found that both groups behaved in a very similar way, since they equally resolved most of the form-based LREs (96.67%; 97.14%), the unresolved category being underrepresented (3.33%; 2.86%). Comparisons between the means also pointed to this similarity between the two groups, since the mean of resolved LREs was significantly higher than the mean of unresolved LREs in both proficiency-matched (3.41 vs. 0.12) and student-selected (3.40 vs. 0.10) pairs. Moreover, no significant differences were found between the two research groups either for resolved (3.41 vs 3.40) or unresolved (0.12 vs. 0.10) LREs.

Table 3 Outcome of form-based LRE types according to pairing method

	Proficiency-matched ($N = 17$) ($z = 3.526$, $p \le .001$***)	Student-selected ($N = 10$) ($z = 2.524$, $p \le .05$*)
Resolved ($t = .012$, $p > .05$)		
Number	58	34
%	96.67%	97.14%
Mean	3.41	3.40
Range	0-10	0-10
Unresolved ($z = -.138$, $p > .05$)		
Number	2	1
%	3.33%	2.86%
Mean	0.12	0.10
Range	0-1	0-1

Tables 4 and 5 focus on the outcome of LREs. For each kind of pairing group, they show the data related to the (in)accuracy in the LREs which had been resolved. As for meaning-based resolved LREs (Table 4), it was observed that both proficiency-matched and student-selected dyads resolved more LREs in a target-like manner than in a non-target-like manner. What is more, this was particularly true of the proficiency-matched dyads, as their percentage of correct resolutions of meaning-based LREs amounted to 73.49% whereas that of student-selected groups only reached 54.35%. The mean scores obtained also yielded greater target-likeness differences in the proficiency-matched sample than in the student-selected one. Whereas the means of right and wrong resolutions in the latter were not significantly different (5.00 and 4.20), in the former the mean of accurately resolved meaning-based LREs was significantly higher than that of inaccurately resolved ones (3.59 vs. 1.29). As for intergroup comparisons, no statistically significant differences were found for target-like resolved meaning-based LREs between proficiency-matched and student-selected learners (3.59 vs. 5.00). However,

proficiency-matched pairs produced significantly fewer non-target-like resolved meaning-based LREs than student-selected dyads (1.29 vs. 4.20).

Table 4 Accuracy in meaning-based LRE resolution according to pairing method

	Proficiency-matched (N = 17) (z = 2.704, p ≤ .01**)	Student-selected (N = 10) (z = 1.198, p > .05)
Target-like (z = -1.192, p > .05)		
Number	61	50
%	73.49%	54.35%
Mean	3.59	5.00
Range	0-11	1-9
Non-target-like (z = -2.859, p ≤ .01**)		
Number	22	42
%	26.51%	45.65%
Mean	1.29	4.20
Range	0-5	0-10

Regarding the nature of resolved form-based LREs, the percentages and means shown in Table 5 seem to indicate that both groups behaved alike, as they resolved 75.86% and 73.57% of the form-based LREs respectively, their mean scores for target-likeness (2.59 and 2.50) and for non-target-likeness (0.82 and 0.90) being quite similar.

Table 5 Accuracy in form-based LRE resolution according to pairing method

	Proficiency-matched (N = 17) (z = 2.474, p ≤ .05*)	Student-selected (N = 10) (t = 2.058, p > .05)
Target-like (t = .097, p > .05)		
Number	44	25
%	75.86%	73.53%
Mean	2.59	2.50
Range	0-9	0-8
Non-target-like (z = -.135, p > .05)		
Number	14	9
%	24.14%	26.47%
Mean	0.82	0.90
Range	0-3	0-5

As for the second research question (*Does the pairing method exert an influence on the patterns of interaction?*), the data are presented in Table 6. The

number of student pairs classified into each of the patterns of interaction, as well as the percentages which these total numbers represent is shown for the proficiency-matched sample and the student-selected sample, separately. As can be seen in Table 6, some dyads fell into a *mixed* pattern, something which happened in the data when pairs did not behave consistently. Even if in Storch's 2002) model dyads globally fell into one of the patterns of interaction, the variability found in some of the dyads in this regard led us to consider the possibility of mixed patterns. Table 6 further specifies the types of pair dynamics involved in these mixed patterns whenever this occurred.

Table 6 Pair dynamics according to pairing method

	Proficiency-matched (N = 17)	Student-selected (N = 10)
Collaborative		
Number	14	4
%	82.35%	40.00%
Dominant/dominant		
Number	-	1
%	-	10.00%
Dominant/passive		
Number	-	-
%	-	-
Expert/novice		
Number	-	-
%	-	-
Mixed		
Number	3	5
%	17.65%	50.00%
Collaborative & dominant/passive		
Number	1	3
%	5.88%	30.00%
% over mixed dynamics	33.33%	60.00%
Collaborative & expert/novice		
Number	2	1
%	11.76%	10.00%
% over mixed dynamics	66.67%	20.00%
Dominant/passive & expert/novice		
Number	-	1
%		10.00%
% over mixed dynamics		20.00%

As evidenced by the data shown in Table 6, differences were observed according to the pairing method. Proficiency-matched dyads overwhelmingly featured a *collaborative* pattern (82.35%), with none of them being labelled as fully *dominant/dominant*, *dominant/passive* or *expert/novice*. As for the *mixed* category, 100% of the pairs included the *collaborative* pattern in the mixed composition.

Moreover, 66.67% of the mixed pairs exhibited *collaborative* plus *expert/novice* patterns, both of which are considered to foster optimal collaboration according to Storch (2002). As for the student-selected pairs, data showed that only 40.00% of the dyads were labelled as *collaborative,* whereas 10.00% of them exhibited a *dominant/dominant* pattern, which was not present in proficiency-matched dyads. However, half of the student-selected pairs were of a *mixed* nature (50.00%). The majority of mixed dynamics (80.00%) included the pernicious *dominant/dominant* pattern, and only a small percentage (20.00%) was of a beneficial *collaborative* plus *expert/novice* kind.

5. Discussion

The purpose of this study was to explore the effects of the pair formation method (proficiency-matched vs. student-selected) on both the production of LREs and the pair dynamics exhibited in young CLIL learners' dialogic interaction elicited from the joint completion of two collaborative tasks. The first research question addressed the issue of how the pairing of the learners affected the production of LREs in terms of their number, type and outcome. It was found that learners produced more meaning-based than form-based LREs, regardless of the pairing method. This is consistent with previous studies which found that low proficiency learners produced more meaning-focused LREs, independently of the criterion used for pairing them (Gallardo-del-Puerto & Basterrechea, in press; García Mayo & Imaz Aguirre, 2019), or studies that examine the effect of proficiency on the type of LREs with adults (Leeser, 2004; Malmqvist, 2005). This tendency to produce more meaning-based LREs was more prominent in the case of student-selected pairs since the differences reached significance. This result must be linked to the fact that self-selection leads to a greater incidence of meaning-based LREs in students' interactions, regardless of their outcome. We might tentatively conclude that self-selection promotes dyad members' engagement in attempting to solve the lexical difficulties encountered to move the tasks forward. Therefore, it can be suggested that friendship (following Mozaffari, 2017) makes self-selected pair members feel less shy in their collaborative interaction, at least in terms of discussion of lexical gaps. The current study provides support for the importance of interpersonal relationships in task-mediated interaction (Philp, Walter, & Basturkmen, 2010).

As for the limited occurrence of form-focused LREs attested in the data, descriptive statistics indicated that proficiency-matched pairs produced a greater proportion of LREs with a focus on form than student-selected pairs, although the difference did not reach statistical significance. It seems that pairing based on proficiency may enhance attention to language form, a finding that has

also been reported for higher proficiency learners in studies that examine the impact of proficiency level on the type of LREs produced by adolescent EFL (Basterrechea & Leeser, 2019), adult EFL (Storch & Aldosari, 2013) or learners of Spanish as a FL (Leeser, 2004). This study built on and extended previous research on the type of LREs produced by young EFL learners, as no previous studies have shown an advantage for proficiency-matched pairs in the production of form-based LREs in this population. As far as outcome is concerned, no differences were found between the two groups in form-focused episodes, as both proficiency-matched and student-selected pairs resolved the majority of the LREs that involved a focus on form, and, when doing so, they reached an accurate outcome on most occasions.

However, analyses of accuracy showed that the percentage of meaning-based LREs which were resolved in an inaccurate manner was significantly lower in proficiency-matched dyads than in student-selected ones, which again, provides support for the former type of pairing method. Both for meaning- and form-based LREs the percentage of target-like solutions was significantly higher than that of incorrect ones in the proficiency-matched group, whereas student-selected pairs did not significantly distinguish between accurately and inaccurately resolved LREs. These findings lend support to the idea that matched proficiency would be a better pairing method than self-selection as it results in more accurate resolutions. The effect of the pairing method on the target-likeness of LREs is still an underexplored area in research. Thus, our study provides a first step in showing the beneficial relationship between dyad members' matched proficiency and higher target language accuracy.

The second research question focused on the impact of the pairing method on pair dynamics. The analysis of patterns of interaction put forward by Storch (2002) also yielded interesting differences according to the pairing method. Proficiency-matched dyads predominantly fell into the *collaborative* pattern with minimal representation of the *expert-novice* or *dominant-passive* categories. Student-selected pairs, however, collaborated much less in their interactions and presented a higher proportion of less *negotiated* or *consensual* dynamics (Storch, 2002) such as *dominant-passive* and *dominant-dominant* interactive behaviours. These results contradict previous research findings conducted with young EFL (García-Mayo & Imaz Aguirre, 2019) and adult EFL learners (Mozaffari, 2017), where no differences were found in pair dynamics, as all dyads exhibited a collaborative relationship.

6. Conclusion

The present study has shown that young CLIL learners produced more meaning-based than form-based LREs, regardless of the pairing method. Self-selected

pairs also produced more meaning-focused LREs and resolved a larger number of LREs than proficiency-matched pairs. However, proficiency-matched dyads produced more meaning-based LREs which were resolved in a target-like manner than self-selected dyads. As for the patterns of interaction, the dynamics of proficiency-matched pairs were of a more collaborative nature (*collaborative* or *expert-novice*) than those of self-selected pairs, whose rate of deleterious patterns (*dominant-passive* or *dominant-dominant*) was higher.

Overall, while student-selection seems to promote the occurrence and outcome of meaning-based LREs, it is proficiency-matching that makes learners engage in optimal patterns of interaction as well as achieve higher levels of accuracy in both meaning- and form-based LREs. Matched proficiency seems to boost young CLIL learners' focus on language and collaborative dynamics during task-based interaction to a greater extent than friendship. This could lead us to speculate that pair dynamics in the student self-selected pairs might not be based on friendship, as demonstrated by previous findings (Mozaffari, 2017; Russell, 2010). It should not be taken for granted that friendship is the sole reason for student self-pairing, as there may be other factors such as expediency, gender or foreign-language competence, for instance, that might account for students' choices. Our findings call for a more complex understanding of peer relationships, namely one that should consider learners' attitudes in group discusions, especially when making errors or giving and receiving feedback, as has already been suggested in the literature (Philp et al., 2010). Hence, the data presented here should be triangulated with qualitative tools, namely think-aloud protocols providing insight into learners' reported reasons guiding the selection of their peer. While we await further research to address this issue, we may conclude that the higher variability of patterns of interaction and the higher proportion of deleterious dynamics found in student-selected pairs indicate that factors other than friendship may be playing a crucial role when young CLIL learners choose a partner for pair work. A further limitation of this study is the fact that we did not control for TL proficiency in the two members of self-selected pairs, and thus the effect of this intervening variable could not be ruled out. Student-selected dyads may have been made up of learners with the same or different proficiency, and this fact may have exerted an influence on their comparison with proficiency-matched dyads.

Further studies should also explore any task effects mediating the relationship between the independent (pairing method) and the dependent (LREs and pair dynamics) variables. Task modality (speaking vs. speaking + writing) has already been found to exert an influence on language focus during learners' interaction (Adams, 2006; Adams & Ross-Feldman, 2008; Azkarai & García Mayo, 2012; García Mayo & Azkarai, 2016; García Mayo & Imaz Aguirre, 2019; Niu,

2009; Payant & Kim, 2017). Further analyses of our two tasks independently may shed new light on the interplay among types of task, LREs, pair dynamics and pairing methods. This issue is worth investigating not only for theoretical reasons, but also because of its pedagogical implications, as it would be very useful for teachers to know which kinds of task trigger optimal LREs and advantageous patterns of interaction for different types of pair formation.

Acknowledgements

References

Adams, R. (2006). L2 tasks and orientation to form: A role for modality*? ITL - International Journal of Applied Linguistics, 152,* 7-34.

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 connections* (pp. 243-267). Ann Arbor, MI: University of Michigan Press.

Allan, D. (2004). *Oxford Placement Test.* Oxford: Oxford University Press.

Azkarai, A., & García Mayo, M. P. (2012). Does gender influence task performance in EFL? Interactive tasks and language-related episodes. In E. Alcón & P. Safont (Eds.), *Discourse and language learning across L2 instructional settings* (pp. 249-278). Amsterdam: Rodopi

Basterrechea, M., & García Mayo, M. P. (2013). Language-related episodes (LREs) during ollaborative tasks: A comparison of CLIL and EFL learners. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 25-43). Amsterdam: John Benjamins.

Basterrechea, M., & Leeser, M. (2019). Language-related episodes and learner proficiency during collaborative dialogue in CLIL. *Language Awareness, 28*(2), 97-113.

Benson, C., Pavitt, J., & Jenkins, M. (2005). The use of dictogloss to encourage discussion of language use. *Edinburgh Working Papers, 14,* 1-17.

Butler, Y. G., & Zeng, W. (2015). Young learners' interactional development in task-based paired-assessment in their first and foreign languages: A case of English learners in China. *Education, 43,* 292-321.

Cambridge University Press. (2008). *Key English test.* Cambridge: Cambridge University Press.

Donato, R. (1998). *Beyond group: A psycholinguistic rationale for collective activity in second-language learning* (Unpublished doctoral dissertation). University of Delaware, Newark, USA.

Gallardo-del-Puerto, F., & Basterrechea, M. (in press). The role of proficiency and pair formation method in language-related episodes: A study of young CLIL learners' interaction. *Language Teaching Research.*

García Mayo, M. P. (2002a). The effectiveness of two form-focused tasks in advanced EFL pedagogy. *International Journal of Applied Linguistics, 12,* 156-175.

García Mayo, M. P. (2002b). Interaction in advanced EFL pedagogy: A comparison of form-focused activities. *International Journal of Educational Research, 37,* 323-341.

García Mayo, M. P., & Azkarai, A. (2016). EFL task-based interaction: Does task modality impact language-related episodes? In M. Sato & S. Ballinger

(Eds.), *Peer interaction in second language learning* (pp. 241-266). Amsterdam: John Benjamins.

García Mayo, M. P., & Imaz Aguirre, 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.

House, S., & Scott, K. (2009). *Sparks 1: Teacher's book.* London: Richmond. Retrieved from https://www.guiassantillana.com/ingl%C3%A9s/sparks/

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,* 211-234.

Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3*(2), 73-93.

Kowal, M., & Swain, M. (1997). From semantic to syntactic processing: How can we promote metalinguistic awareness in the French immersion classroom? In R. Johnson & M. Swain (Eds.), *Immersion education: International perspectives* (pp. 284-309). Cambridge: Cambridge University Press.

Lantolf, J. P., & Appel, G. (1994). *Vygotskian approaches to second language research.* Norwood, NJ: Ablex.

Leeser, M. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research, 8,* 55-81.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2. Second language acquisition* (pp. 413-468). New York: Academic Press.

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 258-329.

Mackey, A. (1994). Targeting morpho-syntax in children's ESL: An empirical study of the useof interactive goal-based tasks. *Working Papers in Educational Linguistics, 10,* 67-88.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, MJ: Lawrence Earlbaum.

Malmqvist, A. (2005). How does group discussion in reconstruction tasks affect written language output? *Language Awareness, 14*(2-3), 128-141.

Mozaffari, S. H. (2017). Comparing student-selected and teacher-assigned pairs on collaborative writing. *Language Teaching Research, 21*(4), 496-516.

Nelson, G. L., & Murphy, J. M. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly, 27,* 135-141.

Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness, 18,* 384-402.

Payant, C. (2018). Effects of L3 learner proficiency and task types on language mediation: A sociocultural perspective. In M. Ahmadian & M. P. García

Mayo (Eds.), *Recent perspectives on task-based language learning and teaching* (pp. 99-120). Boston: De Gruyter.

Payant, C., & Kim, Y. K. (2017). Impact of task modality on collaborative dialogue among plurilingual learners: A classroom-based study. *International Journal of Bilingual Education and Bilingualism, 22*, 614-627.

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.

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). Oxford: Oxford University Press.

Russell, M. (2010). The formation of effective work groups within a FE classroom. *Research in Post-compulsory Education, 20*, 286-305.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52,* 119-158.

Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research, 11*(2), 143-159.

Storch, N., & Aldosari, A. (2013). Pairing learners in pair work activity. *Language Teaching Research, 17,* 31-48.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.

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., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal, 82*, 320-337.

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*, 121-142.

Williams, J. (1999). Learner-generated attention to form. *Language Learning, 51*, 303-346.

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System, 29*, 325-340.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry 17*, 89-100.

APPENDIX A

Pictures for Dotty's doll activity
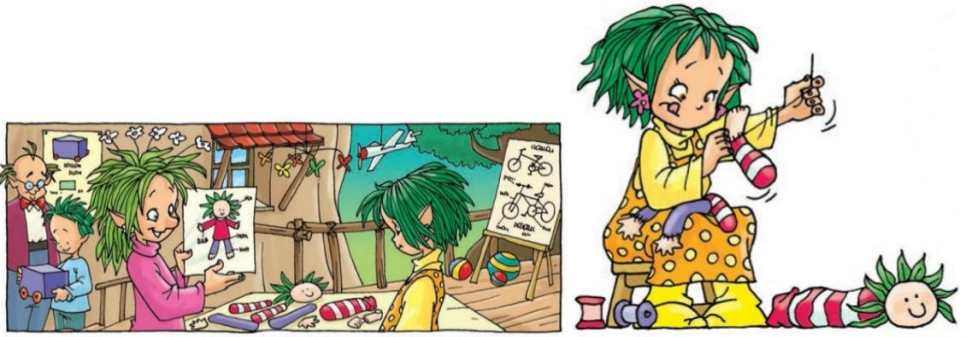
Source: *Sparks 1: Teacher's Book* (House & Scott, 2009, p. 74; reproduced with publisher's permission).

APPENDIX B

Pictures for the second task

Alberto San Emeterio Bolado© (published with author's permission)