



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

“Affordances and Challenges of Collaborative Writing in
Lower Secondary EFL Classrooms in Kosovo”

verfasst von / submitted by

Nuhi Bllaca

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Philosophie (Dr.phil.)

Wien, 2023/ Vienna, 2023

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student record sheet:

UA 792 343

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

English and American Studies

Betreut von / Supervisor:

Univ.- Prof. Mag. Dr. Christiane Dalton-Puffer

Acknowledgments

Before embarking on my PhD project, I had heard people saying that doing a PhD is both challenging and rewarding. It is exactly the same that I have experienced throughout my PhD journey. In fact, this journey was long, challenging, but rewarding. Anyway, one thing is for sure that I would have never been able to come to this point without the unwavering support and encouragement of many people who accompanied me along this journey.

First and for most, I would like to express my greatest gratitude to my thesis supervisor Prof. Christiane Dalton-Puffer who was an excellent, extremely supportive, patient, and encouraging supervisor and mentor from the beginning to the end of this endeavor. Her professional guidance, contingent and dynamic feedback, and support made it possible for me not only to find my path along the way but also to gain the most out of this wonderful project. I am and will always be grateful to you Christiane for helping me to achieve my dream.

I would also like to thank Ass. Prof. Igor Yanovich for his advice concerning the statistical analysis of my data. It was his contribution that facilitated my statistical work.

Next, I am very thankful to the Ministry of Education, Science, Technology and Innovation (MESTI) which funded my PhD program and made it possible for me to study in Vienna and grow professionally in the field of my heart - linguistics.

I humbly extend my appreciation to the school principal, Ms. Sofie Bunjaku, for granting me the permission to conduct the study as well as for her support and encouragement throughout all this journey. I am also deeply grateful to my former students who participated in this study and to their parents/guardians who allowed their children to be part of the study.

Last but not least, I would like to thank all my friends living in Vienna including Bashkim, Egzon, Bledart, Besian, Andi, Genti, Soan, and Geri with whom I spent a lot of time, cooking, eating, drinking coffee, and walking all around Vienna and beyond. I am very thankful to all of these, but especially to Bashkim Hyseni who helped me with all I needed during my stay in Vienna.

Finally, this thesis would have not been possible without the unwavering love and support of my family. Therefore, I would like to thank my father and mother for their

endless support, love and encouragement, my brother and my sisters for their help and support, and finally my lovely wife and my two wonderful sons, whose unconditional love and support made it possible for me to succeed in this, challenging but rewarding, journey. It is their love and support that pushed me strongly forward.

Abstract

Research on collaborative writing has proliferated in the recent decades. The majority of studies with adult EFL/ESL participants have demonstrated that collaborative writing benefits foreign and second language learning by promoting opportunities for L2 learning (e.g. Wigglesworth & Storch, 2009), L2 vocabulary acquisition (e.g. Kim, 2008), and the development of L2 writing proficiency (e.g. Chen, 2019). Research has also revealed that collaborative writing benefits learners' written products in terms of quantitative and qualitative measures. However, such benefits have mostly been drawn from studies with adult EFL/ESL participants (e.g. Fernández Dobao, 2012; Storch & Wigglesworth, 2007), whereas research with young and adolescent EFL learners is still in its infancy. Furthermore, and more importantly, studies that examine whether learners' L2 proficiency affects the written text outcomes and the quantity and quality of languaging (Swain, 2006) are rather scarce. Therefore, the present study sought to bridge these gaps by comparing the writing performance of 144 adolescent EFL learners (aged 13-14) of three proficiency levels (i.e. low-intermediate, intermediate, high-intermediate) as they built a story based on a visual prompt. Thus, the study analyzed 96 written texts composed individually (n=48) and in pairs (n=48) for quantitative measures of fluency, complexity and accuracy, and qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics. Additionally, the study analyzed the oral interactions of learners writing in pairs for Language-Related Episodes (LREs), tallying the number, type and outcomes of LREs produced by 24 pairs across three proficiency levels to establish whether the learners' L2 proficiency impacts on the quantity and quality of LREs produced. Following the data analysis, four major results emerged: (a) pairs compared to individuals wrote better texts in terms of fluency, complexity and accuracy, however, significant differences were detected only for accuracy and complexity (dependent clauses to clauses); (b) pairs compared to individuals obtained significantly higher scores in task achievement, coherence and cohesion, grammar and vocabulary, but not in mechanics; (c) the level of learners' L2 proficiency affected positively both the written texts produced by pairs and individuals – measured both quantitatively and qualitatively, and (d) lexical LREs were significantly more frequent than grammatical and mechanical LREs; pairs with higher L2 proficiency produced and resolved correctly a higher amount of LREs compared to pairs of lower L2 proficiency

Theoretically, findings in this study are in line with the claim of sociocultural theory (Vygotsky, 1978) that learning is a socially-bound process, that is, it emerges in interaction with others, as well as in line with the interactionist approaches that through interactional adjustments learners create favorable contexts for L2 development (Long, 1996; Mackey, 2012). Pedagogically, this study helps EFL teachers/instructors to make more informed decisions as to when and how to implement collaborative writing in their L2 classes so that it benefits their learners.

Table of Contents

Acknowledgments	iii
Abstract.....	v
Table of Contents.....	vii
List of Tables	xii
List of Figures	xiii
List of Appendices.....	xiv
List of Abbreviations	xv
Chapter 1: Introduction	1
1.1 Statement of the problem	1
1.2 The purpose of the study	6
1.3 Research questions and hypotheses	7
1.4 Significance of the study.....	8
1.5 Definitions of terms.....	9
1.6 Structure of the thesis.....	10
Chapter 2: Theoretical Orientation	11
2.1 Interaction Hypothesis and Output Hypothesis and the Sociocultural Theory of Mind.....	11
2.1.1 Interaction hypothesis and output hypothesis	11
2.1.2 Feedback in L2 settings.....	17
2.1.3 Vygotsky's Sociocultural theory of mind and the notion of languaging	23
2.1.4 Languaging: Collaborative dialogue and private speech	27
2.2 Writing approaches	31
2.2.1 Product-based approach	32
2.2.2 Process-based approach	34
2.2.3 Collaborative writing	37
2.3 Collaborative writing and peer interaction.....	39
2.3.1 Analysis of learner interaction in collaborative writing.....	41
2.3.2 Analysis of peer talk focusing on LREs.....	42
2.4 The role of L1 and L2 use in collaborative writing.....	44
2.5 Collaborative writing and opportunities for target language use	48
Chapter 3: Prior research on L2 collaborative writing	52
3.1 Factors affecting the number, type and resolution of LREs.....	52
3.1.1 Learners' L2 proficiency and LREs.....	53
3.1.2 Learner relationships and LREs	55

3.1.3	Task types and LREs.....	60
3.2	The effect of collaboration on the quality of written texts and learning gains.....	64
3.2.1	Comparing the outcomes of collaboratively and individually written texts	65
3.2.2	Collaborative writing and L2 learning gains.....	70
3.2.3	Longitudinal studies on collaborative writing.....	73
3.3	Writing processes in CMCW	77
3.4	Learners' perceptions and attitudes towards collaborative writing.....	78
3.5	Summary of literature review.....	81
Chapter 4:	Methodology	85
4.1	Introduction.....	85
4.2	Context of the present study.....	85
4.2.1	English as a foreign language in Kosovo	86
4.2.2	The English syllabus for grade eight learners	88
4.2.3	Teaching and learning material(s).....	90
4.3	Participants.....	92
4.4	Pilot study	94
4.5	The writing prompt	96
4.6	Instruments.....	97
4.6.1	Compositions	97
4.6.2	Pair talk	97
4.7	Researcher's role.....	98
4.8	Data collection procedures	99
4.9	Data analysis	101
4.9.1	Analysis of written compositions.....	101
4.9.1.1	Quantitative analysis	102
4.9.1.2	Qualitative analysis	105
4.9.2	Analysis of pair talk	106
4.9.2.1	Lexis-focused LREs	107
4.9.2.2	Form-focused LREs	108
4.9.2.3	Mechanics-focused LREs.....	109
4.9.2.4	Resolution of LREs.....	110
4.10	The rating process and interrater reliability	111
4.11	Statistical analysis of the results	113
4.12	Ethical issues.....	116
4.13	Summary of methodology.....	117

Chapter 5: Results	118
5.1 Introduction	118
5.2 Results of the quantitative analysis	118
5.2.1 Fluency	119
5.2.2 Complexity	120
5.2.3 Accuracy	121
5.3 Results of the quantitative analysis across three proficiency levels.....	122
5.3.1 Pairs vs. individuals of low-intermediate proficiency.....	122
5.3.1.1 Fluency	123
5.3.1.2 Complexity	123
5.3.1.3 Accuracy	123
5.3.2 Pairs vs. individuals of intermediate proficiency	124
5.3.2.1 Fluency	124
5.3.2.2 Complexity	125
5.3.2.3 Accuracy	125
5.3.3 Pairs vs. individuals of high-intermediate proficiency.....	126
5.3.3.1 Fluency	126
5.3.3.2 Complexity	127
5.3.3.3 Accuracy	127
5.4 Pairs and individuals within their writing conditions.....	128
5.4.1 Fluency of collaborative texts	128
5.4.2 Complexity of collaborative texts	130
5.4.3 Accuracy of collaborative texts.....	131
5.4.4 Fluency of individual texts.....	132
5.4.5 Complexity of individual texts.....	133
5.4.6 Accuracy of individual texts	134
5.5 Results of the qualitative rubrics.....	136
5.5.1 Pairs vs. individuals	137
5.6 Pairs vs. individuals across three proficiency levels	138
5.6.1 Pairs vs. individuals of low-intermediate proficiency.....	138
5.6.2 Pairs vs. individuals of intermediate proficiency	139
5.6.3 Pairs vs. individuals of high-intermediate proficiency.....	140
5.7 Pairs and individuals within their writing conditions.....	141
5.7.1 Pairs across three proficiency levels	142
5.7.2 Individual writers across three proficiency levels.....	144
5.8 Results for Language Related Episodes (LREs)	146

5.8.1	Number of LREs	147
5.8.2	Type of LREs	148
5.8.3	Resolution of LREs	149
5.9	LREs across three proficiency levels	150
5.9.1	Number and type of LREs.....	150
5.9.1.1	Time, turns and LREs	152
5.9.2	Resolution of LREs	155
5.10	Some observations on the nature of pair talk during the CW process	156
5.11	A summary of results	163
Chapter 6: Discussion.....		167
6.1	Introduction.....	167
6.2	The performance of pairs vs. individuals on CAF measures	167
6.2.1	Fluency.....	168
6.2.2	Complexity.....	169
6.2.3	Accuracy	170
6.3	Pairs vs. individuals across three proficiency levels	172
6.3.1	Pairs vs. individuals of low-intermediate proficiency.....	172
6.3.2	Pairs vs. individuals of intermediate proficiency	173
6.3.3	Pairs vs. individuals of high-intermediate proficiency.....	174
6.4	Pairs and individuals within their conditions	176
6.5	The performance of pairs vs. individuals on global measures	180
6.6	Pairs vs. individuals across three proficiency levels	182
6.6.1	Pairs vs. individuals of low-intermediate proficiency.....	182
6.6.2	Pairs vs. individuals of intermediate proficiency	183
6.6.3	Pairs vs. individuals of high-intermediate proficiency.....	185
6.7	Pairs and individuals within their conditions	187
6.8	The nature and outcome of LREs.....	189
6.9	LREs across three proficiency levels	192
6.10	Some insights on the nature of pair talk during the CW process	196
6.11	Summary	202
Chapter 7: Conclusion.....		207
7.1	Synthesis of findings	207
7.2	Implications of the Study	209
7.3	Limitations of the study and future directions	213

References.....	215
Deutsche Zusammenfassung.....	258
Appendices.....	260

List of Tables

Table 2.1 Coding scheme of LREs	43
Table 4.1 Participants.....	93
Table 4.2 Data collection schedule	101
Table 4.3 Coding T-units and clauses	104
Table 4.4 Coding LREs for their types and outcomes	106
Table 4.5 Research questions, data sources and analysis.....	115
Table 5.1. Measures of fluency across the whole dataset	119
Table 5.2 Measures of complexity across the whole dataset	120
Table 5.3 Measures of accuracy across the whole dataset	121
Table 5.4 Fluency, complexity and accuracy for low-intermediate learners	122
Table 5.5 Fluency, complexity and accuracy for intermediate learners.....	124
Table 5.6 Fluency, complexity and accuracy for high-intermediate learners	126
Table 5.7 Measures of fluency for pairs across three proficiency levels	128
Table 5.8 Measures of Complexity for pairs across three proficiency levels	130
Table 5.9 Accuracy measures for pairs across three proficiency levels.....	131
Table 5.10 Measures of fluency for individuals across three proficiency levels	132
Table 5.11 Measures of complexity for individuals across three proficiency levels	133
Table 5.12 Measures of accuracy for individuals across three proficiency levels	134
Table 5.13 Measures of holistic rubrics across the whole dataset.....	137
Table 5.14 Measures of global rubrics for low-intermediate learners	138
Table 5.15 Measures of global rubrics for intermediate learners.....	139
Table 5.16 Measures of global rubrics for high-intermediate learners	140
Table 5.17 Results of global rubrics for pairs across three proficiency levels.....	142
Table 5.18 Measures of global rubrics for individuals across three proficiency levels	144
Table 5.19 Number of LREs produced by pairs across three proficiency levels	147
Table 5.20 Type of LREs produced by 24 pairs	148
Table 5.21 The outcome of LREs	149
Table 5.22 Number and type of LREs across three proficiency levels	150
Table 5.23 Time and Turns of LREs.....	152
Table 5.24 Outcome of LREs for pairs across three proficiency levels.....	155

List of Figures

Figure 1. Boxplots - Distribution of quantitative scores	268
Figure 2. Boxplots - Distribution of holistic scores	270
Figure 3. Boxplots - Distribution of LRE data.....	272
Figure 4: A collaborative text written by Dua and Ber (hand-written).....	279
Figure 5: An individual text composed by Mul (hand-written)	281
Figure 6: A collaborative text composed by Alu and Gash (hand-written)	283
Figure 7: An individual text composed by Haz (hand-written).....	285
Figure 8: A collaborative text composed by Led and Sal (hand-written)	287
Figure 9: An individual text composed by Byt (hand-written)	289

List of Appendices

Appendix A: Visual writing prompt	260
Appendix B: The Global Rating Rubric – developed by Gassner et al., (2019).....	262
Appendix C: Writing Prompt for Pairs	265
Appendix D: Writing Prompt for Individuals	266
Appendix E: Distribution of quantitative scores	267
Appendix F: Distribution of holistic scores	269
Appendix G: Distribution of LRE data	271
Appendix H: Form of consent – in Albanian	273
Appendix I: A collaborative text written by Dua and Ber (#1 high-intermediate)	278
Appendix J: An individual text composed by Mul (#49 – high-intermediate).....	280
Appendix K: A collaborative text composed by Alu and Gash (#17 – intermediate).....	282
Appendix L: An individual text composed by Haz (#65 – intermediate)	284
Appendix M: A collaborative text composed by Led and Sal (#34 – low-intermediate).....	286
Appendix N: An individual text composed by Byt (#82 – low-intermediate)	288

List of Abbreviations

C/T	Clause per T-unit
CAF	Complexity, accuracy, fluency
CC	The core curriculum
CF	Corrective feedback
CF	the curriculum framework
CMC	Computer-mediated context
CMCW	Computer-mediated collaborative writing
CW	Collaborative writing
DC/C	Dependent clauses of total clauses
EFC	Error-free clause
EFC/C	Error-free clauses of clauses
EFL	English as a foreign language
EFT	Error-free T-unit
EFT/T	Error-free T-units of T-units
ESL	English as a second language
FTF	Face-to-face
HI	High-intermediate
I	Intermediate
IW	Individual writing
L1	First language/mother tongue
L2	Second or foreign language
LI	Low-intermediate

LP	Language proficiency
LRE	Language-related episodes
MEST/I	Ministry of education, science, technology and innovation
MO	Modified output
NNS	Non-native speaker
NS	Native speaker
SCT	Sociocultural theory
SLA	Second language acquisition
ZPD	Zone of proximal development

Chapter 1: Introduction

1.1 Statement of the problem

Pair and group work on oral activities, such as short-presentation, spot-the-difference, and role-play are very common in most foreign and second language classrooms. However, this is not the case with writing tasks which are usually done individually, presumably due to the perception of teachers and learners who consider writing as an individual rather than a collaborative activity (Storch, 2019). Besides considering writing as an individual activity, foreign language (FL) teachers in the Kosovar context often neglect writing activities in their teaching lessons due to a number of reasons, such as the limited time available to cover a wide range of language lessons with only two English classes per week (MEST, 2016a), lack of empirical evidence informing them about the benefits of writing for L2 learning, and finally, teachers may also feel unsure as to how to provide corrective feedback or assess their students' written texts (c.f. Sections 4.2.2 - 4.2.3) (McDonough, De Vleeschauwer, & Crawford, 2018). On the other hand, students also may feel reluctant with writing tasks, as they may consider it among the hardest skills to be learned in an L2 (Hyland, 2003). This has often been the case, in my teaching experience, with low L2 proficiency learners as there were times when they seemed unwilling to do any writing tasks due to the lack of target language knowledge. Thus working in an environment with many challenges, such as those listed above, the need to explore new teaching approaches that incorporate the use of writing tasks which, at the same time, promote peer interaction among learners is more than necessary to turn these barriers into learning opportunities for EFL learners in Kosovo. In this regard, using collaborative writing, which is simply defined as the writing of a single text by two or more learners (Storch, 2013), seems to be a rationale solution to satisfy the needs of our learners, as well as to help EFL teachers be more successful in their daily work with their students at pre-university level.

Collaborative writing entails both writing as a process and product, and more importantly, interaction among learners, which is one of the most distinguishing features of collaborative writing, facilitates the writing process and benefits the written products (Storch, 2021). Theoretically, the rationale behind using collaborative writing tasks in L2 settings derives from the cognitive and socio-cognitive theories of second language acquisition (SLA). From the interactionist perspective, peer-peer interaction is beneficial for L2 learning because it provides learners with opportunities to negotiate

for meaning (Long, 1996), to receive input and corrective feedback, as well as with opportunities to modify their output (Loewen, 2012; Loewen & Sato, 2018; Mackey, Abbuhl, & Gass, 2012) and thus contribute to L2 development (Gass & Mackey, 2015). Nevertheless, the exposure to L2 input on its own is not sufficient for successful L2 acquisition, especially to produce language of native-like accuracy. To reach this level of development, learners need opportunities to produce output, that is, written or spoken language (Swain, 1993), as the act of producing language promotes automatization, enables learners to notice gaps in their interlanguage, test their hypotheses and reflect on their own language produced (Swain, 1995). All of these important cognitive processes facilitate L2 learning.

From the perspective of sociocultural theory of mind (SCT), on the other hand, interaction has a greater role to play, and specifically interaction which goes beyond that of making the input comprehensible. Drawing on the work of Vygotsky (1978), this perspective considers learning as a socially-bound activity, and interaction as a means for all cognitive development including language learning. More importantly, the interaction that occurs between the expert and a novice learner is key for the latter, because it is via interaction that the expert is able to provide the novice with the support needed to complete a task which is beyond his/her competence (Storch, 2017). The attuned assistance provided by the expert to the novice has been referred to as Scaffolding (Wood, Bruner, & Ross, 1976). Interestingly research has found that learners with higher, lower or similar language proficiency were able to provide one another with scaffolding as they were assigned on collaborative tasks (e.g. Ohta, 2001; Storch, 2013; Swain, 2006). Therefore, both of these two leading theories of SLA support the use of pair and group work in the L2 classrooms, in particular tasks that combine a speaking and written component (Rouhshad & Storch, 2016).

Much research informed by the sociocultural perspective has indicated that collaborative writing tasks provide learners with interaction opportunities during the whole writing process. Such interaction involves mainly learners' deliberation over language use, and learners' talk centered around solving task problems such as dealing with organization of content, generation of ideas and task achievement (McDonough et al., 2016; Wigglesworth & Storch, 2009). Studies that have examined learner-learner interaction have generally employed Language-Related Episodes (LREs) as the unit of analysis and coded the pair talk for the number, type (e.g. form-focused, meaning-

focused, mechanics-focused) and resolution (e.g. correctly resolved, incorrectly resolved, unresolved) of LREs produced by assigning learners in different collaborative tasks (e.g. Fernández Dobao, 2012, 2014; García Mayo & Zeitler, 2017; Watanabe & Swain, 2007; Wigglesworth & Storch, 2009). LREs are defined as instances of learners' conversation where they "talk about the language they are producing, question their language use, or correct themselves or others" (Swain & Lapkin, 1998, p. 326). Interaction that covers learners' deliberation over language use as they attempt to make meaning in the process of text co-construction has initially been termed collaborative dialogue (Swain, 2000) and subsequently changed to languaging, which Swain (2006) defines as "the process of making meaning and shaping knowledge and experience through language" (p. 89).

In order to identify whether meaning-focused tasks or grammar-focused tasks provide learners with more opportunities for languaging operationalized as LREs, research with adult EFL/ESL participants has compared the number, type and outcomes of LREs generated by each type of task and has revealed that overall grammar-focused tasks tend to generate more form-focused than meaning-focused LREs (e.g. Alegría de la Colina & García Mayo, 2007; Storch, 2001). However, studies with younger participants from immersion programs (e.g. Swain & Lapkin, 2001), with intermediate EFL secondary learners (e.g. Basterrechea & Leaser, 2019; Villarreal & Gil-Sarratea, 2019), and with intermediate adult EFL students working in pairs, and in small groups (e.g. Fernández Dobao, 2012) could not find any significant difference between form-focused and meaning-focused LREs produced across tasks. Studies with young EFL learners also produced mixed results, while Calzada and García Mayo (2020), for example, obtained a higher number of form-focused LREs compared to meaning-focused and mechanics-focused LREs, Lázaro-Ibarrola and Hidalgo (2022) found a higher number of mechanics-focused LREs compared to meaning-focused and form-focused LREs. Given the mixed findings reported above and results from previous research suggesting that the type of tasks learners work on (e.g. Alegría de la Colina & García Mayo, 2007), and learners' L2 proficiency (e.g. Kim & McDonough, 2008; Leaser, 2004) were among the most influencing factors that affected the number and resolution of LREs produced by adult EFL/ESL participants, the present study with adolescent EFL learners compared the number, type and resolution of LREs, and

examined whether learners' L2 proficiency had any impact on the number, type, and resolution of LREs produced.

Another body of research that has compared the performance of learners working in pairs, in small groups, and individually by examining the written text outcomes for quantitative measures of complexity, accuracy and fluency (CAF) found that overall texts produced by pairs and small groups were better in terms of accuracy (e.g. Fernández Dobao, 2012; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009), and complexity (Storch, 2005), but not in terms of fluency. Surprisingly, somewhat conflicting results have been reported for both complexity and fluency. While Storch (2005) found that pairs outperformed individuals not only in terms of accuracy, but also in terms of complexity, Bueno-Alastuey and Lizarrondo Larumbe (2017), and Zabihi and Bayan (2020) found that collaborating students outperformed individual students in all measures of complexity, accuracy and fluency. In addition, studies employing longitudinal designs (Chen, 2019; Sarkhosh & Najafi, 2020a; Soleimani et al., 2017) also reported better results for pairs over individuals both in terms of accuracy and fluency, contradicting previous research concerning the measure of fluency (e.g. Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009). Thus, what these studies have found is that they obtained converging results for accuracy, but diverging results for complexity and fluency. Also studies that have compared the written texts of learners working in pairs and individually for qualitative measures using global rating rubrics (e.g. content or task-achievement, organization, vocabulary and mechanics) have shown that texts produced by pairs have generally obtained higher scores in content, organization, vocabulary, and mechanics than texts produced by individuals (e.g. Chen, 2019; Shehadeh, 2011; Villarreal & Gil-Sarratea, 2019). All in all, despite overall positive results in favor of collaboration, findings reported above derive mostly from studies with EFL/ESL participants from university contexts (e.g. Alegría de la Colina & García Mayo, 2007; Chen, 2019; Fernández Dobao, 2012; Leaser, 2004; Shehadeh, 2011; Storch, 2005; Wigglesworth & Storch, 2009). Furthermore, research with younger participants from EFL contexts has generally compared only the learners' written compositions for quantitative measures of accuracy and fluency (e.g. Sarkhosh & Najafi, 2020a), and of complexity, accuracy and fluency (e.g. Bueno-Alastuey & Lazarrondo Larumbe, 2017; Soleimani et al., 2017) excluding not only qualitative

measures of holistic ratings, but also pair or group talk which is extremely important to understand the affordances and challenges learners face in the process of text co-construction. Additionally, research comparing the written products of adolescent EFL learners (13-14 years) working in pairs and individually for quantitative and qualitative measures including an examination of learners' interaction for LREs is still limited (e.g. Villarreal & Gil-Sarratea, 2019). Therefore, the present study sought to bridge this gap, by comparing the writing performance of lower secondary EFL learners assigned to work in pairs and individually to complete a picture story task (write a story based on a set of photographs), and examining the peer interaction during the task completion process.

1.2 The purpose of the study

The present study is designed for two goals. First, the main goal of this dissertation study is to investigate the effect of collaboration on the writing performance of adolescent EFL learners. Drawing on the mixed results reported from previous studies that compared the written compositions of learners working in pairs and individually for quantitative measures of complexity, accuracy, and fluency (CAF) (e.g. Soleimani et al., 2017; Storch, 2005; Wigglesworth & Storch, 2009), the present study compared the writing performance of adolescent EFL learners assigned to work in pairs and alone to write a story based on a set of photographs. Drawing on earlier studies on collaborative writing (e.g. Calzada & García Mayo, 2021; Storch, 2005; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009), the written texts produced by two groups of learners were assessed quantitatively and qualitatively. Quantitatively texts were assessed employing quantitative measures of complexity, accuracy and complexity (CAF). Qualitatively texts were assessed employing holistic measures that rated the written products for task achievement, coherence and cohesion, grammar, vocabulary, and mechanics based on a rubric developed to assess writing skills of EFL learners of similar age in Austrian lower-secondary schools (Gassner, et al., 2019). This rubric is similar to rubrics used by other studies (e.g. Calzada & García Mayo, 2021; Shehadeh, 2011; Villarreal & Munarriz-Ibarrola, 2021) with some modifications due to the nature of the writing prompt.

Second this dissertation study is designed to examine the oral interaction of learners working in pairs, operationalizing it for Language-Related Episodes (LREs). The LREs produced as learners engaged in collaborative writing were analyzed for their focus (e.g. meaning-focused, form-focused, mechanics-focused) number and resolution (e.g. correctly resolved, incorrectly resolved, and unresolved) to establish whether there was a difference between pairs of lower and higher L2 proficiency in terms of LREs produced. Since previous research with adolescent and adult participants from ESL/EFL and CLIL contexts found that learners' L2 proficiency affects the number and resolution of LREs they produce (Basterrechea & Leaser, 2019; Kim & McDonough, 2008; Watanabe & Swain, 2007), the present study aimed at investigating whether this was the case with adolescent EFL learners as well.

Additionally, while previous research (e.g. Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009) using meaning-focused tasks with adult ESL learners found a higher proportion of meaning-focused than form-focused LREs produced, research with young and adolescent EFL learners found a similar amount of meaning-focused and form-focused LREs (e.g. Villarreal & Gil-Sarratea, 2019; Villarreal & Munarriz-Ibarrola, 2021). Therefore, drawing on this small body of research with adolescent EFL learners, the present study addressed this issue by examining the number and focus of LREs generated as adolescent EFL learners were assigned to complete a meaning-focused task.

In sum this dissertation aimed at investigating the affordances and challenges of collaborative writing in lower-secondary EFL classrooms in Kosovo, by: (a) comparing the writing performance of learners writing in pairs and individually for quantitative and qualitative measures (b) analyzing the peer interactions for the number, type and outcome of LREs.

1.3 Research questions and hypotheses

This study was guided by the following research questions and hypotheses:

1. What are the differences between the written texts produced by pairs and individuals in terms of quantitative measures of fluency, complexity and accuracy?
 - a) Are there any differences between the texts produced by pairs and individuals across three proficiency levels?
 - b) Are there any differences between the texts produced by pairs and individuals across three proficiency levels within their writing conditions?
2. What are the differences between the written texts produced by pairs and individuals in terms task achievement, coherence and cohesion, grammar, vocabulary and mechanics?
 - a) Are there any differences between the texts produced by pairs and individuals across three proficiency levels?
 - b) Are there any differences between the texts produced by pairs and individuals across three proficiency levels within their writing conditions?

3. What are the differences between LREs produced by pairs in terms of number, type and resolution?

- a) Are there any differences between LREs produced by pairs across three proficiency levels?

H1: Texts produced by pairs are better in terms of accuracy and fluency than those produced by learners writing on their own.

H2: Texts produced by pairs obtain higher scores for task achievement, coherence and cohesion, grammar and vocabulary.

1.4 Significance of the study

The dissertation study is significant because it attempts to fill gaps in the current literature on collaborative writing in L2. Theoretically, the present study is supposed to provide additional research evidence regarding the benefits and challenges of collaborative writing for second language acquisition, with a special focus on a population of adolescent EFL learners which is more rarely explored in terms of writing performance in general, and in the Kosovar context in particular. The analysis of pair talk will reveal important writing processes and the effects of collaboration through learner-learner interaction on the writing performance of adolescent EFL learners written products measured through quantitative and qualitative measures. Additionally, as attention to form and meaning is considered vital for L2 acquisition, the present study will bring more conclusive evidence as to whether meaning-focused tasks are suitable for drawing adolescent EFL learners' attention to form and meaning while in the process of text co-construction. The present study will also contribute to the existing research and hopefully bring new evidence with regards to the role of learners' L2 proficiency on the number, type and resolution of LREs produced as they engage to complete the task, as well as on the written text outcomes.

Practically, understanding the role of collaboration in the field of L2 writing is important because EFL teachers worldwide, and in particular those working in the Kosovar context need to be informed with research evidence regarding the benefits and challenges of collaborative writing, so that they can take decisions more easily whether to assign their students to write individually or collaboratively. More importantly, influenced by communicative language teaching approaches, EFL teachers often

employ activities which foster communication in the target language, therefore providing them with research evidence over the benefits of collaborative writing for genuine communication in L2, foreign language learning, and writing development can highly contribute to the process of achieving EFL teachers' goals in this respect.

On the personal level, this dissertation study will help the researcher shape his professional life by advancing his research skills and gaining the adequate knowledge to further contribute to the field of L2 writing research. In addition, as an EFL teacher, the researcher will be far more informed over the benefits and challenges of collaborative writing for adolescent EFL learners, not only in terms of L2 acquisition but also in terms of written text outcomes.

1.5 Definitions of terms

Collaborative writing (CW). In this study collaborative writing refers to an activity where two learners work together to produce a single text (Storch, 2013). During the production process, learners writing together engage in interaction to deliberate over the language use, and to discuss about finding the best words to express their ideas to produce a joint text. Learners writing together, collaborate to produce a single composition, interact during the whole writing process, and share the decision making process and responsibility over the text they are producing during all stages of writing.

Writing performance. Includes writing performance of learners writing collaboratively and individually. The writing performance of both groups is assessed quantitatively and qualitatively. Quantitative measures include linguistic complexity, accuracy, and fluency. Qualitative measures include rating scores on task-achievement, coherence and cohesion, grammar, vocabulary, and mechanics.

Pair talk. In the current study pair talk refers to the interaction occurring among two learners assigned to work together to write a single text. Such talk includes self-directed and other-directed talk which emerges as a result of learners' attempts to complete their assigned task during the collaboration process.

L1 use. In this study L1 refers to learners' use of their mother tongue to solve or discuss task or language-related problems during the collaborative writing process.

L2 use. In the current study L2 use refers to learners' use of English during the task completion process. L2 use can include learners' deliberation over language use, discussion of ideas, content or organization of their writing.

L2 learning. In this study, L2 learning refers to the learning of English both as a foreign or a second language.

1.6 Structure of the thesis

This thesis is organized into seven chapters. Chapter one introduces the topic of the thesis, statement of the problem, the purpose of the study along with its significance and research questions that guided this study, ending with the definition of terms and the structure of the study. Chapter two discusses the theoretical orientation of the study and the major constructs derived from two leading theories of second language acquisition (SLA), namely cognitive theories, such as Long's (1981, 1983) interaction hypothesis, Swain's (1985, 1993) output hypothesis, and socio-cognitive theory, that is, Vygotsky's (1978, 1981) sociocultural theory of mind (SCT). It is in this chapter that the role of interaction is discussed drawing on previous research guided under the theoretical perspectives noted above. Chapter three critically reviews prior research on L2 collaborative writing with a special focus on the studies comparing collaborative work with individual work in face-to-face settings, including a small number of studies in computer-mediated context.

Additionally, studies reporting on influencing variables that might affect the quantity and quality of learners' talk when assigned on collaborative writing will be reviewed. Chapter four outlines the methodology of the study, describing the pilot study, the participants, research site, data gathering instruments, and data collection and data analysis procedure. Chapter five presents findings of the present study in full compliance with research questions guiding the thesis. First, the findings drawn from the analysis of written compositions are presented, and then, the results from the analysis of pair talk. Chapter six discusses findings from the present study relating them to findings from prior research. Chapter seven concludes the thesis with a synthesis of the findings, presenting some theoretical and pedagogical implications, limitations of the study and suggestions for future research.

Chapter 2: Theoretical Orientation

2.1 Interaction Hypothesis and Output Hypothesis and the Sociocultural Theory of Mind

The issue of how humans acquire a second language has received much attention and treatment by different language learning theories in the long history of SLA. Today, Krashen's (1981, 1982, 1985) Comprehensible Input Hypothesis, Long's (1981, 1983) Interaction Hypothesis, Swain's (1985, 1993) Comprehensible Output Hypothesis, and Vygotsky's (1978, 1981) Sociocultural Theory of Mind are considered as the most influential language learning theories in the field of SLA. Hence, in the following section these cognitive and socio-cognitive theories will be discussed in more detail.

2.1.1 Interaction hypothesis and output hypothesis

Early claims of Evelyn Hatch (1978) that “language learning evolves out of learning how to carry on conversations, out of learning how to communicate” (p. 63), and Krashen's (1981) Input Hypothesis which posits that exposure to sufficient comprehensible input leads to second language acquisition have served as an influential factor for the development of Long's (1981) Interaction Hypothesis. According to Krashen's (1985) Input Hypothesis one possible way to acquire a second language is by receiving comprehensible input, input which is just beyond learner's current level of knowledge ($i+1$) in the target language (Krashen, 1982, 1985). Long's (1981) Interaction Hypothesis, on the other hand, holds that “participation in conversation with native speakers made possible through modification of interaction, is the necessary and sufficient condition for SLA” (p. 275). From this perspective, conversational modifications/adjustments provided via interactional feedback are perhaps even more relevant than just exposure to sufficient comprehensible input for L2 learning. During interactions learners receive important information regarding the correctness or incorrectness of their utterances. When they produce an incorrect utterance they may receive negative evidence through interactional feedback which signals learners over their problematic linguistic items seeking modification or reformulations of their speech (Gass & Mackey, 2015). Interactional feedback provided during conversations between speakers may contribute to better understand the input delivered to them. In fact, early research informed by the interaction hypothesis has shown that when native speakers engage in interaction with non-native speakers, they use several discourse moves to

make the input more comprehensible and thus avoid communication breakdown. In one early study, Long (1983), for example, compared the conversations of pairs composed of native speakers and non-native speakers (NS-NNS) with conversations of pairs composed of native speakers (NS-NS). He found that NS-NNS dyads used several conversational moves, such as clarification requests, confirmation checks and comprehension checks to keep the conversation going which was not the case with conversations of NS-NS dyads. Conversational moves such as these noted above are grouped under the umbrella term *negotiation for meaning*. These moves serve the purpose of avoiding communication breakdowns between speakers as well as the purpose of making the input more comprehensible. Negotiation for meaning, however, does not ensure linguistic accuracy. Therefore, researchers working under the interactionist framework have proposed the term *negotiation of form* which refers to learners' linguistic corrections aimed at improving linguistic accuracy (e.g. Ellis et al. 2001a; Lyster et al. 2013). According to Lyster (1998b, p. 190) *negotiation of form* often involves feedback targeted to improve problematic linguistic forms/realizations with a focus on linguistic accuracy and precision of language use without interfering the flow of conversation. Loewen and Sato (2018), for example, claim that corrective feedback is given by a proficient speaker or a teacher, in response to the learner's grammatically incorrect utterance even if the meaning of the utterance is clear. In fact, corrective feedback is a category of negative evidence (feedback) given following learners' incorrect utterances, or their "ill-formed [...] L2 production" (Loewen, 2012, p. 22). This type of feedback may be given on oral or written language production. While oral feedback may be provided following learners' immediate oral production, written feedback may be given following learners' production of a text (Loewen, 2012). Regardless of the feedback types, several meta-analyses have provided overall positive evidence for feedback provided (e.g. Li, 2010; Lyster, Saito, & Sato, 2013). All in all, despite the effectiveness of all interaction components, negotiation of meaning has received special attention in the field because of its role for L2 learning. Long (1996) in his revised hypothesis suggested that "negotiation of meaning, and especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways" (p. 451–452).

However, despite previous claims over the relevance of input, it has been seen as insufficient for L2 development on its own. This claim was based on findings of studies with French Immersion students which revealed that despite long exposure to L2 comprehensible input, learners included in such programmes were not able to speak the L2 with native like fluency. Therefore, Swain (1985, 1995) proposed her Output Hypothesis, claiming that for successful L2 learning to occur, learners need to be able to produce output (written or spoken language). She argues that the act of producing language enables learners not only to benefit in terms of language semantic comprehension, but also to produce language which is grammatically correct. Swain (1995) in her own words claimed that, “output may stimulate learners to move from the semantic, open-ended nondeterministic, strategic processing prevalent in comprehension to the complete grammatical processing needed for accurate production. Output, thus, would seem to have a potentially significant role in the development of syntax and morphology” (p. 128).

The importance of output for L2 acquisition has been confirmed by research exploring learners’ language gains from input and output conditions. Izumi (2002) and Izumi and Bigelow (2000) found that learners who were exposed to input immediately after they were required to produce written output did significantly better than their counterparts that had to only process the input for comprehension. Izumi (2002) also noted that output production may push learners to process language more deeply and thus create a more durable memory trace. Moreover, the production of output serves three important functions as identified by Swain (1995). Such functions include noticing, hypothesis testing, and the metalinguistic function. The first function of output is noticing, which refers to learners’ understanding of gaps in the target language. Thus as learners attempt to speak or write in the L2, they may notice gaps, that is, they realize what they know and what they miss to express a certain idea. The second function of output is hypothesis testing, which basically means finding a useful way of expressing an idea by testing multiple possibilities of joining words to make meaning. The last function of output is a metalinguistic function, whereby learners’ can reflect on their own language use, written or spoken. Swain (1995) noted this as a favorable context for L2 learners “enabling them to control and internalize linguistic knowledge” (p.126).

Attention, being cognitive in nature, is another construct which has a special place in the interactionist approach. Despite diverging views with regards to the nature of

cognitive constructs of attention, awareness and noticing, there is a general consensus among interactionist researchers that “the cognitive constructs of attention, awareness, and the related construct of noticing are part of the interaction–L2 learning process” (Gass & Mackey, 2015, p. 191). The role of attention and noticing for L2 development has been pointed out by a large number of researchers (e.g. Robinson, 2001; 2003; Robinson et al., 2012; Schmidt & Frota, 1986; Schmidt, 1993). According to Robinson et al. (2012) learning cannot take place without awareness and noticing. Similarly, Schmidt’s (1990; 2001) noticing hypothesis highlights the role of attention for learning, claiming that no learning can take place without attention. Even though this claim has been questioned by other scholars (e.g. Gass, 1997) it has been generally admitted that selective attention has a special role to play in the SLA process. In fact, through interactional feedback learners’ attention is drawn to some specific parts of the target language (language forms) which help learners notice the gaps between their interlanguage and the target language (Schmidt, 2001). Similarly, Mackey (2012) observes that through interactional modifications including recasts, clarification requests and the production of output learners’ attention to form and meaning can be drawn preparing the ground for learners to further advance their L2 knowledge.

Identifying and using appropriate tasks that promote interaction is another important issue that has received increased attention within the L2 research community (Mackey, 2012). Within the interactionist approach the most commonly used tasks aimed at engaging learners in conversations are closed and open tasks, focused and unfocused tasks, and one-way and two-way tasks (Mackey, 2012). A closed task is a task in which only one correct answer is possible. For instance, a spot-the-difference task is one such task where learners work together to find out the number of differences between two pictures. In this type of task, learners are usually told in advance the exact number of differences they have to find (Gass & Mackey, 2007). An open-task, on the other hand, is a task in which more than one correct answer/alternative is possible. In this type of task, learners are supposed to negotiate over a problem and can come to more than one solution. For example, learners may discuss about things they can choose from a restaurant menu. This type of task may be useful to expand learners’ vocabulary and language structures used in certain situations (Mackey, 2012). As far as benefits for second language learning are concerned, closed tasks have been emphasized as better tasks because they push learners to come to a solution over a problem and thus may

promote more negotiation of meaning as compared to open tasks in which learners may not necessarily engage in such negotiations (See for example, Ellis, 2003; Long, 1996).

Focused and unfocused tasks comprise another type of task type used in interaction. While the focused tasks seek to engage the learner in meaning-focused interaction with a slight element of form-focused interaction, the main purpose of unfocused tasks, however, is to provide learners with opportunities to practise general communication skills in L2 (Ellis, 2003; Nassaji & Fotos, 2004; Nobuyoshi & Ellis, 1993). In this respect, it is interesting to note that apart from the purpose, there is no other major distinction between the two types of tasks. According to Nassaji and Fotos (2004) focused tasks may be more useful to teach learners simplified linguistic structures, and easily taught grammar rules, while, unfocused tasks are far more effective to increase learners' awareness of complex grammatical structures that would be too hard to teach by means of formal instruction alone.

Last but not least, the one way-task, also known as a non-reciprocal task, is another task type defined as a task in which the information is conveyed from one person to another. An example of this type of task is a picture description task, where one learner describes the picture to a second learner who cannot see it but simply listens to the description provided by his/her interlocutor. Following the description, the recipient learner reproduces his/her narration of the picture on his/her own (Mackey, 1999; Mackey & Philp, 1998; Swain & Lapkin, 2000). On the other hand, two-way tasks which also include spot-the-difference tasks (Mackey, 1999; Mackey & Philp, 1998), story completion tasks (Bigelow, Delmas, Hansen, & Tarone, 2006; Mackey, Adams, Stafford, & Winke, 2010; Mackey & Oliver, 2002) and jigsaw tasks (Gass, et al., 2005) require information exchange between learners engaged in such activities.

An information gap task, for example, considered as a two-way information exchange task, requires that both learners share information with one another so as to complete the task successfully (Mackey, 2012). When comparing task types (one-way or two-way tasks) that provide learners with more opportunities for negotiation of meaning, researchers (e.g. Long, 1983a, 1983b; Pica, 1988) argue that two-way tasks, also known as reciprocal tasks, provide learners with more opportunities to engage in negotiation of meaning. Pica (1988) for instance, compared a one-way task (in which the exchange of information was not required) with a two-way task (in which the

information exchange was required) and found that the two-way task produced more modification of interaction than the one way task.

However, later investigations of the same type of tasks produced contradicting results (Iwashita, 1999; Shehadeh, 1999). In Shahadeh's (1999) study, the researcher compared the instances for modified output in a picture dictation task (one way task) and an opinion-exchange task (two way task) where learners talked about a newspaper article. The study found that the one-way task provided significantly more instances for modified output than the two way task. Similarly, Gass and Varonis (1985), for example, found that in a one-way task more instances for interaction modifications were produced compared to the two-way task. This led researchers to conclude that perhaps the familiarity of the background knowledge may have prohibited learners from engaging in more negotiation of meaning instances. Fewer instances of negotiation of meaning were also reported by later studies conducted with learners of similar level of L2 proficiency or learners of the same L1 compared to learners of mixed L2 proficiency or different L1, see for example, Iwashita (2001). Furthermore, in terms of producing more or less negotiation of meaning, classroom arrangement is believed to play a role. Fotos and Ellis (1991), for example, found that having students work in pairs rather than small groups produced more instances for negotiation of meaning.

Moreover, studies have shown that when learners engage in interaction they tend to primarily negotiate over meaning (lexis) rather than form. In a study by Pica (1992), for instance, it was found that communication breakdowns occurred as a result of problems with lexis, and thus learners engaged in negotiation of meaning. Whereas problems with morphosyntax (form) did not cause communication breakdowns and therefore no opportunities for negotiation of meaning were produced. For a similar study see, for example, Williams (1999). So, for learners to develop their interlanguage, they need to be engaged in tasks that provide them with opportunities to negotiate not only over language meaning (lexis) but also language form (morphosyntax), as this can encourage them to pay attention to the language they are producing (Kowal & Swain, 1994). Therefore, using tasks which require learners to engage not only in spoken but also in written output is an alternative suggestion put forward by a number of researchers (e.g. Storch, 2016; Williams, 2012).

According to Storch (2016) learners engaged in speaking tasks may feel a lot of pressure to keep the conversation going, and as a result, they may not be able to pay attention to language form and meaning and to notice the feedback they receive. In contrast, in writing tasks, learners have more processing time, are more relaxed from pressure, are able to reflect on their produced language, and thus can pay more attention to language form (Williams, 2012). Furthermore, as learners engaged in writing tasks do not receive immediate feedback or signals of non-comprehension, as in the case of speaking, they can devote more attention to the language structure (morphosyntax) in order to produce more accurate written output. In addition, in writing tasks, learners have more time to focus on language accuracy, deliberate over the language use and find the best words to express their intended meaning. Similarly they have more time to evaluate and review their produced output and incorporate the feedback they may receive until they produce a satisfactory draft (Storch, 2016; Williams, 1999).

In short, despite the fact that Swain's (1985, 1993) output hypothesis encompassed both spoken and written language, oral tasks seem to dominate in SLA interaction research (see for example, Loewen & Sato, 2018; Mackey, 2012; Mackey, et al., 2012). Moreover, research informed by Long's (1996) revised interaction hypothesis provides a rationale for the use of pair or small group work in L2 classrooms. Learners working in pairs or small groups may be encouraged to engage in negotiations of meaning and form while doing their tasks; such negotiations can trigger cognitive processes important for L2 development. This is in full compliance with Long's (1983, 1996) interaction hypothesis which claims that receiving comprehensible input and interactional feedback while negotiating for meaning and form facilitates second language acquisition.

2.1.2 Feedback in L2 settings

As discussed thus far, interaction among speakers or learners provides opportunities for feedback which in turn may be effective for second language acquisition. Despite the fact that feedback can be negative and positive (García Mayo & Alcón Soler, 2013), it is negative feedback which has received increased attention in the last decades (Ammar & Spada, 2006; Ellis, Loewen, & Erlam, 2006; Leeman, 2007; Loewen & Philp, 2006; Lyster & Izquierdo, 2009; Lyster & Ranta, 1997; Mackey, 1999; MacKey & Philp, 1998). Researchers working in the field of SLA have brought various definitions

concerning the notion of feedback. One that highlights its general characteristics is put forward by Leeman (2007) who defines feedback as “a mechanism which provides the learner with information regarding the success or failure of a given process” (p.112). By definition, feedback is responsive and thus can occur only after a given process. Regarding the dichotomy between positive and negative feedback, researchers have presented explanatory definitions framing these two types of feedback. Thus positive feedback, for example, has been defined as “an affirmation of the content or correctness of a learner utterance” that intends to provide the receiver with affective support (Nassaji, 2015, p. 11). On the other hand, negative feedback, also known as corrective feedback, has been defined as the “information provided to learners about the ill-formedness of their L2 production” (Loewen, 2012, p. 24). Moreover, Loewen explains that while oral negative feedback can be provided immediately following learners’ erroneous utterances, written corrective feedback can be provided following a production of a written text (ibid). Although corrective feedback can be oral or written, oral corrective feedback tends to be perceived as the central component of interaction and has received far more attention than written corrective feedback in recent years (Loewen & Sato, 2018), as shown by several reviews and meta-analyses (e.g. Brown, 2016; Li, 2010; Lyster & Saito, 2010; Lyster, Saito, & Sato, 2013; Russell & Spada, 2006).

Corrective feedback became an area of investigation following the claim for focus on form (Long, 1991, 1996) as an important strategy to bring learners’ attention to linguistic items that were associated with non-target-like features in meaning-focused interaction. As meaning-focused instruction was primarily concerned with linguistic fluency, attention to form which was believed to develop learners’ linguistic accuracy was strongly neglected, therefore, corrective feedback was seen as a useful means that can help draw learners’ attention to form and meaning and thus promote L2 acquisition (Loewen, 2012). After the conceptualization of corrective feedback as an important construct in L2, researchers have identified its occurrence in various contexts (e.g. EFL, ESL and CLIL), though with varying degrees of frequency (Ellis et al., 2001; Lyster & Ranta, 1997; Yoshida, 2008) as well as its facilitative role for L2 learning in both EFL and ESL settings (Li, 2010). The rationale behind using corrective feedback in classroom settings is to reach teaching objectives that aim for the consolidation of learners’ L2 knowledge (Lyster et al., 2013). In addition, corrective feedback as a tool

enables teachers to turn their students' errors into opportunities for L2 learning (Mackey et al., 2016). According to Lyster et al. (2013) accumulating research has shown that corrective feedback "plays a pivotal role in the kind of scaffolding that teachers need to provide to individual learners to promote continuing L2 growth" (p. 1).

In the late 1970s, oral corrective feedback was initially classified into the following general categories, such as repetition of the error with or without changes, prompts, and explanations (Chaudron, 1977). This categorization was further expanded by Lyster and Ranata (1997) who classified negative feedback into six other categories that included, recasts, elicitation, repetition, metalinguistic feedback, explicit correction, and clarification requests. More recently, Sheen and Ellis (2011) suggested that oral corrective feedback can be distinguished as input-providing and output-prompting, and as implicit and explicit. Input-providing feedback refers to the feedback that provides the learner with a correct linguistic form, whereas output-prompting feedback is a type of feedback that encourages the learner for self-correction of a non-target-like utterance. Implicit versus explicit feedback, on the other hand, differ in terms of whether the learner is asked with a clarification request immediately following his/her erroneous utterance or whether the learner is provided with a correct linguistic form or metalinguistic explanation on his/her non-target-like utterance.

Input-providing feedback consists of four feedback moves, such as conversational recasts, didactic recasts, explicit correction and explicit correction with metalinguistic explanation. Output-prompting feedback, on the other hand, entails five feedback strategies, such as repetitions, clarification requests, metalinguistic clues, elicitations and paralinguistic signals (Mackey et al., 2016). Since corrective feedback does not comprise the main goal of this thesis, only some of the categories mentioned above such as recasts, explicit feedback, clarification requests and repetition will be discussed in what follows assuming that these feedback types are more likely to occur in an EFL classroom with adolescent EFL learners assigned to work in pairs to write a story.

One frequent type of corrective feedback are recasts which belong to the category of input-providing feedback. A recast is a reformulation of all or part of a learner's non-target-like utterance into a more target-like one without interfering to the learner's intended meaning (Loewen & Philp, 2006).

The following excerpt from (Mackey, 1999, p. 561) is an example of a recast.

NNS: Your picture how many how many cat your picture?

NS: How many cats are there in my picture? ← **Recast**

NNS: Yeah how many cats?

Recasts have been further subcategorized into conversational recasts and didactic recasts. Conversational recasts are explicit in nature and described as the reformulation of learners' erroneous utterances into correct linguistic forms. Conversational recasts occur after a communication breakdown between interlocutors, and they can frequently take the form of confirmation checks as shown by the following question tag "*Oh, so you were sick, were you?*" italics in original (Sheen & Ellis, 2011, p. 594). Didactic recasts, on the other hand, are implicit in nature, and defined as a partial or complete reformulation of a learner's non-target-like utterance aimed at drawing the learner's attention to the specific point of the error that occurred. In addition, didactic recasts occur not because of a communication problem, but because the teacher or interlocutor wants to draw learners' attention to a linguistic form. As such they serve a pedagogical function (Sheen, 2011). The following examples taken from (Sheen, 2011, p. 3) illustrate the two types of recasts.

Example 1: Conversational recast

S: How much weigh?

T: What?

S: How weight are you?

T: How much do I weigh? (conversational recasts)

Example 2: Didactic recast

S: Women are kind than men.

T: Kinder. (partial recast)

Another type of input-providing feedback is explicit feedback which provides the learner with a correct linguistic form following a non-target-like utterance (Loewen, 2012). Explicit feedback can be distinguished as explicit correction or explicit correction with metalinguistic explanation. The distinction between these two rests on that whether the learner is provided only with a correct linguistic form or with an additional metalinguistic explanation (Sheen & Ellis, 2011). The following examples illustrate the two explicit feedback strategies:

An example of explicit correction

S: One upon a time

T: No we say once.

S: Once. (Oliver & Mackey, 2003, p. 524).

An example of explicit correction with metalinguistic explanation

S: Fox was clever.

T: The fox was clever. You should use the definite article 'the' because fox has been mentioned. (Sheen, 2011, p. 3).

While feedback strategies noted above were identified as input-providing, the following feedback moves, such as clarification requests, repetition, elicitation, metalinguistic clues and paralinguistic signals belong to the category of output-prompting feedback (Sheen & Ellis, 2011).

Clarification request is a conversational move which occurs when a learner's utterance is not fully comprehended by his/her teacher or interlocutor. As a result the teacher or interlocutor uses a specific question word or phrase (e.g. *what*, *excuse me*, *I'm sorry*) serving as a clarification request to urge the learner to provide further information concerning his/her utterance (Nassaji, 2015; Sheen, 2011). An example of a clarification request taken from (McDonough, 2005, p. 86) is presented below:

Learner: what happen for the boat?

NS: what?

Learner: what's wrong with the boat?

Repetition is another feedback move used by the teacher or interlocutor who can repeat all or only a part of the learner's erroneous utterance with a rising intonation. Using the repetition as a feedback move, the teacher or interlocutor can encourage the learner to reflect on his/her non-target-like utterance. Like clarification requests, repetition as a move does not provide the learner with a target-like utterance, but provides an opportunity for the learner to self-correct his/her utterance (Nassaji, 2015; Sheen, 2011; Sheen & Ellis, 2011). The following example from (Choi & Li, 2012, p. 337) illustrates a repetition as a feedback move:

S: he got caught by

T: he got... (elicitation)

S: caught.

As discussed thus far, different feedback moves have been identified in the literature. Despite the fact that each feedback move can play a facilitative role in SLA and can

occur with varying degrees in L2 learning contexts, research has found that recasts are one of the most frequent type of feedback provided in the classroom (Lyster & Ranta, 1997; Sheen, 2004), and that they benefit second language acquisition (Han, 2002; MacKey & Philp, 1998). Long (1996, 2007) argues that recasts benefit L2 learning because they provide learners with a target-like utterance immediately after committing an error without distracting their focus on communication. The immediacy of feedback provided enables them to make a distinction between correct and incorrect L2 forms, and potentially to notice the gap in their interlanguage process (Swain, 1995). As Lyster et al. (2013) claimed recasts occurring during conversational interaction can trigger the noticing of gaps between non-target-like and target-like utterances, a highlighted feature that is theoretically well-supported by Swain's (1993) output hypothesis. Similarly, Nassaji (2015) claims that recasts used with corrective purpose do not occur due to the communication problems but serve the purpose of treating learners' errors more specifically by drawing their attention to form, providing them with a correct form even when the meaning is quite clear. Finally corrective feedback can occur at least in two different situations: (i) it occurs when there is a communication breakdown between interlocutors, in this case they engage in negotiation of meaning to solve their communication problems, and (ii) it occurs when there are no communication problems, but the teacher or interlocutor wants to draw learners attention to form, a typical case in EFL settings (Sheen & Ellis, 2011). In each case, corrective feedback has been found beneficial for second language acquisition. It is the latter that is more relevant here because pair work or group work creates a useful environment for learners to provide one another with corrective feedback, in particular while writing collaboratively (Storch, 2013). In fact, many studies (e.g. Brooks & Swain, 2009; Villarreal & Gil-Sarratea, 2019) conducted, under the framework of sociocultural theory, have documented that when learners work in pairs they are able to provide corrective feedback to one another, feedback which is quite similar to teacher feedback reported above. The following is one example to illustrate how two EFL learners from a secondary school context provide one another with corrective feedback, clarification requests, and confirmation checks as they work in pairs on an essay writing task (Villarreal & Gil-Sarratea, 2019, p.11).

Example 1: two EFL learners working in pairs

EG1: in the one hand

EG2: on the one hand

EG1: on?
EG2: yeah
EG1: or in?
EG2: on ... on the one hand....

Thus, this example shows that corrective feedback, which emerged among secondary school learners, helped them to come to a correct linguistic solution for their given task. Irrespective of the person giving the feedback (teacher vs. learner), it is considered important for SLA to occur. As noted in the introductory chapter, the present study analyzes peer interaction for language-related episodes, including instances of corrective feedback provided during a collaborative writing task. In the following section peer interaction from the perspective of sociocultural theory of mind is discussed in more detail.

2.1.3 Vygotsky's Sociocultural theory of mind and the notion of languaging

As suggested by the interaction approach, interaction that occurs between NS-NNS interlocutors is thought to facilitate L2 learning in several ways, such as, by providing opportunities for comprehensible input, corrective feedback and modified output. From the sociocultural theory (SCT) point of view, however, interaction is perceived to play a much greater role not only in terms of language learning but also in terms of all cognitive development. For Vygotsky's (1981) sociocultural theory, learning is socially situated, and all cognitive development including higher order cognitive functions, such as voluntary attention, intentional memory and language learning, emerge in social context. The development of such higher order cognitive abilities occurs in interaction between members of the community i.e. between an expert (e.g. teacher, adult, or a more knowledgeable person) and a novice (e.g. a child, peer, or a less knowledgeable person). For the interaction to take place, it has to be mediated by physical tools (e.g. artefacts, computers, or signs) or by symbolic tools (e.g. gestures, language). It is due to the use of such tools (e.g. language, gesture, any form of communication) that interaction takes place, and humans are able to resolve problems they encounter and develop their cognitive abilities (Storch, 2017). It is interaction which occurs between the expert (e.g. teacher) and the novice (e.g. learner) that enables the former to provide attuned assistance to the latter. Contrary to other psychological theories of cognitive development (e.g. Piaget, 1977), SCT views social context to be the source of cognitive development, namely the development occurs from social to individual, that is, it starts

in social context via communication with people (e.g. peers or adults) and then becomes internalized within the individual (Ibid.). Internalization is defined as the process through which cultural artefacts, including language, take on a psychological function (Kozulin, 1990) as cited in (Lantolf, Thorne, & Poehner, 2015). Vygotsky (1981) described the process of internalization in the following way:

Any function in the child's cultural development appears twice, or on two planes. First it appears on the social plane, and then on the psychological plane. First it appears between people as an interpsychological category, and then within the child as an intrapsychological category. This is equally true with regard to voluntary attention, logical memory, the formation of concepts, and the development of volition (p. 163).

Internalization depicts the developmental process through which people increase their abilities to perform complex cognitive and physical-motor abilities, first entirely supported by external mediation (e.g. the novice guided under full support/assistance of the expert), relying then on internal mediation (e.g. relying on his/her own abilities) (Lantolf et al., 2015). It should also be noted, however, that SCT does not view internalisation as a process whereby the novice simply imitates the expert, but rather as a transformative process from social to individual, with other words, the knowledge co-constructed with the expert becomes an independent knowledge of the novice (Lantolf & Thorne, 2006).

As mentioned above, it is through interaction that the expert provides assistance to the novice. However, not all assistance provided by the expert is successful. For the assistance to be effective it must be finely tuned and delivered properly. An appropriate amount of assistance can positively contribute to the developmental process, on the other hand, an inappropriate amount of assistance (i.e. too much or too little) can be detrimental to the development process (Lantolf & Thorne, 2006). Vygotsky (1978) notes that for the assistance provided to be effective, one needs to take into account the novice's current state of knowledge and potential achievable level with the assistance provided. In fact, the assistance provided should not only be contingently responsive to the novice needs, but it must also be a form of assistance which encourages the novice to take an active role in the activity and enable him/her perform subsequent tasks independently (Van Lier, 2004). The distance between the novice's current level of

knowledge and the target level of achievement is referred to in the literature as the Zone of Proximal Development (ZPD) (Storch, 2017). Within the ZPD, the novice has a role to play as well. Since the nature of assistance is dynamic, the novice may contribute to the process of knowledge co-construction by signaling their needs and thus help the expert provide adequate assistance to the learner (Roth & Radford, 2010). Learners should not be seen as passive recipients of the assistance but rather as agentive and responsive in the process of co-construction (Poehner, 2008). Careful and adequate assistance offered by the expert as well as his encouragement for a greater responsibility of the novice is commonly referred to in the literature as scaffolding (Wood et al., 1976). Scaffolding portrays a temporary structure which serves for the learning to take place. In the case of language learning, it enables the learner to perform beyond his/her actual competence with the assistance of the expert/a more knowledgeable person. However, such support should be removed once the learner has developed his/her own knowledge and thus can complete the task independently (Wilson & Devereux, 2014).

In the L2 learning context, researchers have evidenced that learners can provide scaffolding to peers (e.g. Antón & Dicamilla, 1999; Donato, 1994; Guerrero & Villamil, 2000; Ohta, 2001; Storch, 2001, 2002a; Swain & Lapkin, 1998; Watanabe & Swain, 2007). Such evidence comes from the often cited study by Donato, (1994). Analyzing protocols of students of French at an American university, working on a collaborative planning task for ten weeks, Donato (1994) found that peers were able to provide scaffolding or mutual support to each other to solve linguistic problems. Evidence drawn from this study showed that learners were also able to create collective scaffolding for one another and co-construct new knowledge despite their low linguistic abilities (ibid.). Evidence of collective scaffolding was also found in Ohta's (1995, 2000) studies. In her 2000 study, she examined the interactional cues to which peers were oriented to provide developmentally appropriate assistance and also looked at the role of provided assistance in the developmental process of two learners studying Japanese in a university context. While completing an oral translation task, learners engaged in collaborative dialogue concerning their form-focused language use as well as providing and receiving form-related assistance (Ohta, 2000). The analysis of the pair talk revealed that learners provided assistance to each other based on interlocutor's indicated needs for support. Becky, considered as the learner with lower language proficiency experienced greater difficulty in completing the task, but the attuned

assistance provided by Hal, the other learner with higher language proficiency, made it possible for her to overcome the difficulties and improved drastically in using complex language constructions. Hal also showed evidence of development (*ibid.*). This study showed that despite their varied language proficiency, learners were able to provide each other with scaffolding. More importantly, as Ohta (2000) posits “the provision of developmentally appropriate assistance is not only dependent upon attention to what the peer interlocutor is able to do, but also upon sensitivity to the partner’s readiness for help, which is communicated through subtle interactional cues” (p. 53). Other studies have compared the effectiveness of collective scaffolding between learners and teachers. For instance, Guk and Kellogg's (2007) study compared the evidence of scaffolding between group work interactions and whole-class teacher-student interactions in an EFL primary class in Korea. Results of the study showed more evidence of scaffolding in group interactions than in whole-class teacher-students interactions. Interestingly, similar evidence was reported by a small-scale study with four ESL intermediate students (Brooks & Swain, 2009). In this study, researchers compared the effects of peer feedback and expert feedback to solve linguistic problems as students were assigned to write a story collaboratively. The study revealed that peer feedback was more effective than expert feedback in solving linguistic problems. As Brooks and Swain (2009) suggested one reason for this may be that peers were able to provide each other with more attuned assistance and the assistance provided was in compliance with their own needs and developmental capacity, on the other hand, the assistance provided by the expert was perhaps beyond the learners’ developmental stage.

Building on such research findings, researchers have extended further the original definition of expert-novice scaffolding to peer-scaffolding and self-scaffolding (e.g. Granott, 2005; Holton & Clarke, 2006; Mascolo, 2005). Self-scaffolding as Granott (2005) claims can occur “through the dialogic self where other people are represented in the multi-voiced self” (p. 148). When learners engage in self-scaffolding they demonstrate several forms of it, for instance, they tackle the easiest problems first, dismantling them into sub-problems, using their existing resources, to finally focusing on the more complex ones (Bickhard, 2005). Holton and Clarke (2006) claimed that self-scaffolding can be in the form of conversation directed to self, where the learner questions his/her epistemic self. Bickhard further observed that the primary goal of

education should be the development of self-scaffolding skills and approaches of learning to learn, as this may be crucial for learners and students in the future (Ibid.). Research indicates that some learners make use of different tools to self-scaffold while others may fail to do so. The study of Knouzi et al. (2010), which investigated the evidence of self-scaffolding in the L2 learning context, showed that some learners made successful use of self-explanation and self-regulation as a tool to scaffold themselves while some others did not. The researchers claimed that students' awareness should be raised in order for them to consider self-scaffolding as an effective self-regulatory tool which may create opportunities for subsequent development.

2.1.4 Languaging: Collaborative dialogue and private speech

In the L2 learning context, the term languaging is used to refer to the use of language as a tool that mediates learning. Swain (2006) defined languaging as “a process of making meaning and shaping knowledge through language” (p. 98). When learners are confronted with difficult tasks or concepts, languaging serves as a psychological tool enabling learners to think through a solution for a given problem they encounter in their L2 learning efforts (Suzuki & Storch, 2020). It is languaging that enables learners to transform their thoughts into artefacts upon which they can reflect, analyze or deliberate to finally aid their deeper understanding of complex tasks (Swain et al., 2015). Languaging can be of two forms, that is, collaborative dialogue – talk directed to another person about a problem we encounter in an attempt to come to a solution, and private speech – talk directed to oneself, which may be whispering or talking aloud in an attempt to come to a solution for a problem. Thus, regardless whether talking or writing with others, or talking or writing to oneself, the goal of languaging is to facilitate the solution of a complex cognitive problem using language to mediate the process (Swain & Watanabe, 2020).

Languaging as a collaborative dialogue emerges when learners face a difficult problem, that is, they use language as a psychological tool to mediate their thinking about a problem that needs a solution, and which may not necessarily be about language (Swain, 2010a). Thus, from this claim it is perceived that collaborative dialogue can be about anything, including difficult concepts in science, mathematics and language (Swain & Watanabe, 2013). Collaborative dialogue helps knowledge building. Swain (2000) being one of the first to explore the benefits of collaborative dialogue, highlights

its importance in terms of gaining new knowledge claiming that collaborative dialogue is a “dialogue in which speakers are engaged in problem solving and knowledge building” (p. 106). From the sociocultural perspective, languaging, in particular collaborative dialogue, helps the socially co-constructed knowledge to become individually internalized knowledge (Suzuki & Storch, 2020).

Private speech, as defined above, is speech directed to the self. In exploring private speech, at least two definitions have been identified, one concerns its self-addressed form, and the other its function. Looking at its self-addressed form, Ohta (2001) defines private speech as a dialogue which occurs within the individual, while, Bivens and Berk (1990) define private speech as "speech spoken out loud that is addressed either to the self or to no one in particular" (p. 443). As such, private speech may be assumed to be not specifically addressed or tuned to anybody except the self (Berk & Garvin, 1984). As per its function, Appel and Lantolf (1994) claim that private speech serves the functions of ‘*mediating learners cognitive abilities*’ and ‘*speaking to understand*’(p. 437). From this perspective, private speech is assumed to play a facilitative role in helping the speaker to better understand so that speech becomes a tool for thought. It is through private speech that we regulate our mental functioning. In the social context, we communicate by appropriating the meaning and deciphering the speech to then use it inwardly to mediate our mental activity (Lantolf & Beckett, 2009).

All things considered, research in L2 learning contexts has shown that irrespective of the forms of languaging (i.e. collaborative dialogue or private speech), it mediates learning (e.g. Swain et al., 2002; Swain et al., 2009; Watanabe & Swain, 2007). However, as the present study analyzes peer interaction to establish whether learners’ L2 proficiency impacts on its quantity and quality, a greater focus will be on studies reporting on the benefits of collaborative dialogue. Thus, as explained earlier, collaborative dialogue is a dialogue which emerges among learners when they confront a complex cognitive task or concept (Suzuki & Storch, 2020). In the L2 learning context, researchers working under the framework of sociocultural theory have operationalized collaborative dialogue in terms of Language-Related Episodes (LREs). LREs are defined as “any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others” (Swain & Lapkin, 1998, p. 326). It is through collaborative dialogue that learners form

and test their hypotheses about the appropriate and correct use of the target language and reflect on their language produced (Swain & Watanabe, 2020).

In an attempt to establish the benefits of collaborative dialogue, researchers have operationalized learners' oral interaction in terms of LREs. Among others, Swain and Lapkin (1998), for example, analyzed the LREs generated by a pair of students working on a collaborative task and focused on particular linguistic items learners discussed during the LREs. Using a tailor-made test, the researchers were able to find a correlation between the items discussed during the LREs and the posttest scores. Donato's (1994) study, mentioned earlier, also found that some of the forms learners constructed collaboratively were subsequently reused while performing the task. Likewise, Fernández Dobao (2014a), who investigated whether writing a task performed collaboratively in pairs versus small groups benefits learners' L2 vocabulary, found that collaborative dialogue helped learners from both conditions benefit in terms of L2 vocabulary learning, though the posttest results indicated that learners involved in small groups outperformed those in pairs. This was attributed to the more possibilities of pooling knowledge from more members in groups vs. pairs. These studies suggest that languaging (i.e. collaborative dialogue) benefits L2 learning.

Studies that examined the impact of both forms of languaging (e.g. collaborative dialogue vs. private speech – operationalized as think-aloud protocols) have found that collaborative dialogue is superior to the other forms of languaging. Kim (2008), for example, analyzed the pair talk and think-aloud protocols for LREs of Korean L2 learners working in pairs and individually on a dictogloss task. Using a pre-and posttest design, Kim found that although both groups of learners produced a similar number of LREs, pairs outperformed individuals on the vocabulary test. Similarly, Taguchi and Kim (2016), who examined the impact of collaborative dialogue and think aloud protocols in learning L2 pragmatics of requests, found that pairs produced more pragmatic-related episodes and made more successful use of the target language than learners working alone while thinking aloud. In a more recent study, Watanabe (2019), who examined the impact of collaborative dialogue and think-aloud protocols on written text outcomes, found that the relationship learners formed when working collaboratively affected the amount and quality of collaborative dialogue as well as the written text outcomes. Thus, when pairs formed collaborative patterns of interaction, they obtained higher writing scores and outperformed individuals, on the other hand,

when pairs formed non-collaborative patterns of interaction they performed poorly and therefore gained lower writing scores compared to individuals. In other words, when pairs failed to maximize the potential of collaborative dialogue, they failed to outperform individuals. What this suggests is that the engagement in collaborative dialogue per se does not necessarily benefit L2 learning, it is the quantity and quality of languaging that learners engage in that distinguishes between failure and success (ibid.). The importance of quantity and quality of languaging has been documented even in studies investigating languaging of learners working in solitary condition. Thus, Swain et al. (2009), who examined the amount and type of languaging produced by university students of French as they attempted to understand the concept of voice, found that students who engaged in more languaging - operationalized as self-explaining demonstrated more accurate and a deeper understanding of the concept of voice in French than students who did less so. The study also revealed that students with frequent engagement in self-talk were able to use the newly gained knowledge in the subsequent tasks which was not the case with students demonstrating low or middle engagement in such talk (Ibid.). Thus from these studies, in particular the latter, it becomes obvious that not only languaging, but also the quantity and quality of languaging plays a critical role in mediating learning. Additionally, even though both forms of languaging, i.e. collaborative dialogue and private speech are important, it is the former that has proved more effective than the latter – operationalized/labeled as self-directed talk, speech for the self, or think-aloud protocols.

Being aware of the relevance of collaborative dialogue for L2 learning, researchers (e.g. Alegría de la Colina & García Mayo, 2007; Leeser, 2004; Watanabe & Swain, 2007) who operationalized it for language-related episodes have identified several contextual factors that affect both the quantity and quality of peer interaction. Among others, such factors include learners' L2 proficiency, the relationships learners form when completing tasks collaboratively, and the type of tasks learners work in (c.f. Chapter 3, section 3.1 for a detailed description). In addition, even though to a lesser extent, both the amount and quality of peer interaction has been found to play a role on the quality of the learners' written texts (e.g. Watanabe, 2019).

Thus from the sociocultural theory, interaction that occurs either between individuals or within individuals is considered as a useful tool for learning. Unlike Long's (1996) negotiations for meaning emerging as a result of non-comprehensible

input, the collaborative dialogue (languageing) does not necessarily aim to make the input comprehensible, but rather it emerges when learners attempt to solve complex linguistic problems. In collaborative dialogue learners deliberate over the best way of expressing their ideas, pooling their linguistic resources and thus consolidating their existing knowledge or co-constructing new L2 knowledge, which can be further internalized for their own use (Storch, 2013).

Despite the facilitative role of interaction in language learning, one concern raised in the literature is whether oral tasks or written tasks provide more opportunities for interaction among learners. From the perspective of the interaction approach, the majority of studies in SLA have been focused on oral tasks such as story completion, spot the difference, and map tasks (Mackey et al., 2012). However, Storch (2017, 2020) argues that collaborative writing may be a better task for learners to engage in negotiations for meaning and collaborative dialogue. As such, collaborative writing provides learners with opportunities for peer interaction, whereby they can give and receive feedback, offer suggestions and counter suggestions, and deliberate about language use and ideas to be included in their written texts.

Finally, when learners write collaboratively they may engage in self-directed talk and other-directed talk (Storch, 2017). Though the engagement of learners in self-directed talk does not provide an automatic solution for the problems they encounter, learner's engagement in such talk provides more opportunities for them to better understand the problem (Lantolf, 2005). On the other hand, collaborative dialogue has shown to be beneficial for L2 learners especially in terms of vocabulary acquisition, even for learners who were not participating actively in such dialogues, but simply listening to their peers engaged in other-directed talk (Fernández Dobao, 2014, 2016).

2.2 Writing approaches

Writing has been and is still a central topic among scholars in applied linguistics. Writing as a skill represents a complex, multifaceted nature and has constantly produced debate among research communities about its adequate description and explanation as well as about clarifications how writing works and how it should best be taught (Hyland, 2009). However, despite growing interest in teaching English as a second language for many years, writing as one of the most difficult and complex skills to be taught was not given serious attention until the beginning of 1990s (Matsuda, 2003b). The teaching of

writing may be influenced by various factors including the roles of individuals in society, the nature of education, and perspectives foregrounding the link between language and mind. In principle, education classes are supposed to be based on insights provided by research and theory but quite frequently teachers hesitate to abandon their own teaching experiences to replace them with new ones (Williams, 2003). Teaching and researching of writing has been approached from different perspectives, but the most dominating perspectives that have prevailed are two broad approaches that focus on writing as a product and writing as a process (Hyland, 2016). These two approaches draw on cognitive and sociocultural theories of SLA. Additionally, collaborative writing has been seen as a fruitful approach that can be used to benefit learning of writing in L1 and L2, and most importantly as a site for L2 development (Storch, 2016). Therefore in the rest of this section, these three perspectives will be discussed.

2.2.1 Product-based approach

This perspective of learning writing in L2 has mainly been informed by studies conducted in L1 composition and by studies on English for Specific Purposes and English for Academic Purposes and has largely been influenced by composition classes and writing programmes in the US, such as Writing Across Curriculum (WAC)- (Hirvela, Hyland, & Manchón, 2016). Approaches that tend to view writing as a product rather than as a process or activity on itself dominate in the majority of methodologies to writing instruction. In general product-based approaches consider writing as an end-textual product arranged in orderly used words, clauses and sentences and composed of a coherent structure, whereby learners can express their intended meaning (Ibid.). This perspective of treating texts as autonomous objects to be analyzed and described disregarding a particular context, reader or writer, seems to originate from “[... ideas inherited from structuralism and implicit in the Transformational Grammar of Noam Chomsky....]” (Hyland, 2010, p. 8). What this perspective mostly provides is the acquisition of grammatical and syntactic rules through teaching of writing. Hyland (2016) claims that the goal of teaching writing is to enable students develop their language accuracy in producing formal aspects of writing without errors. This can be achieved by providing students with writing exercises that aim at advancing their knowledge of grammar, syntax and vocabulary. In many countries of the world and in particular in those countries where English is taught either as a foreign or a second language the goal of writing classes is not to develop the writing skills per se, but rather

an effort of enabling students to show the acquisition of grammatical and syntactic rules. This was often counted as good writing (ibid.). Informed by the behavioral habit-formation theory where learning to write is based on simple language exercises such as grammar drills, gap filling, substitution and guided writing practice, provides almost nothing more than learning to demonstrate mastery of grammatical knowledge without paying much attention to a particular audience beyond the teacher or examiner (Hirvela et al., 2016).

Moreover, writing improvement has often been measured by analyzing pieces of writing in terms of number of words, clauses, and passive forms used. White (2007), for example, developed and demonstrated the efficacy of a useful instrument for measuring written language entitled “Structural Analysis of Written Language” (SAWL) which includes the analysis of morphemes per T-units (MTU), and the word efficiency ratio (WER). The relevance of this measuring instrument is that it allows teachers to measure students’ written language seamlessly no matter their language proficiency level (Ibid.). On the other hand, consequent involvement of students in writing compositions where they keep struggling to adhere to correct language use in their written texts may impact on their spoken and written language use. Thus, in a study investigating possible changes of academic writing features including hedging and formality, impersonality, and features typically occurring in spoken language, Shaw and Ting-Kun Piu (1998) found that EAP students’ language in essays showed a kind of shift from spoken-language-like conventions to more written-language-like conventions.

Generally, the product-based approach still dominates in many schools where writing classes are taught simply to enable students to demonstrate the acquisition of syntactic rules, and grammatical accuracy (Hyland, 2016). Teachers’ role is minimized to that of passing on some language knowledge to novice writers and correcting students’ surface language errors rather than teaching them appropriate ways to convey meanings (Ferris, 2006). However, constructing syntactically and grammatically accurate sentences does not necessarily indicate progress on student’s writing skills because to date there is not much evidence suggesting that syntactic complexity and grammatical accuracy constitute the best measures of good composition. On the other hand, many students can construct complex and accurate sentences but are far from producing properly written compositions (Hyland, 2010). Finally, writing texts with

fewer language errors should not always be perceived as a sign of writing improvement, but even as the lack of students' willingness to produce written texts beyond their language competence.

2.2.2 Process-based approach

The second broad and highly influential writing approach that emerged in the early 1980s in L1 pedagogy is known as the process approach. This approach sees writing as a recursive process and writers as individual creators of original ideas (Ferris & Hedgcock, 2014). Process-oriented pedagogy was largely developed as a reaction to the product-based approach that dominated writing instruction before the 1960s. From this period onwards, the proponents of process-oriented pedagogy stressed the relevance of teaching writing as process not as product, where students would be given a chance to reveal their own voice, choose their own topic, consider their proposals and ideas, benefit from peer and teacher's feedback, and encourage for more revision and share their composition with peers (Matsuda, 2003a). Generally described as 'top down' in orientation, process-oriented pedagogies do not mainly focus on isolated language errors or parts of texts, but rather on steps and procedures of identifying and solving problems, shaping creative ideas and finding appropriate ways of putting them together, and revising and reflecting on emergent text (Elbow, 1998a; Matsuda, 2003b; Zamel, 1982, 1987).

Process-oriented approaches have often been divided into two different, but overlapping categories: expressivist and cognitivist (Polio & Williams, 2009). From the expressivist point of view, writing is not seen strictly based on narrow concepts defining writing as one's ability to use grammar or language rules correctly, but rather it is perceived as an act of self-discovery giving equal relevance both to the process and product of writing (Hyland, 2010). For Hyland, "Writing is learnt, not taught" (p. 19), and teachers' role is minimized to that of a facilitator and non-directive coach, offering learners frequent feedback, creating an attractive environment for all, encouraging learners to create, share and communicate their ideas without being served with ready-made models, or provided with suggestive responses to any of the possible topics (Hyland, 2016). Similarly, Berlin (1988) considers writing as "a creative act in which the process—the discovery of the true self—is as important as the product" (p. 484). Thus, from this perspective composing assignments should be such that promote the

self-discovery of the writer with no imposing views, but rather with a freedom of expression that empowers the novice's writing skills (Clark, 2012; Elbow, 1998a, 1998b, 2012).

Expressivist process-oriented approaches aim to develop and improve writers' global writing proficiency by encouraging non-directive production of texts with the main focus on conveying meaning rather than demonstrating grammatical accuracy (Hillocks, 2008; Hirvela & Belcher, 2001; Zamel, 1983). One of the greatest advocates of uninhibited writing was Elbow (1998b) who proclaimed the idea of free writing, offering students enough space to focus on producing as much text as possible without inhibition, direction or critical voices from others (Elbow, 1999). Additionally, less formal or private writing is promoted by encouraging novice writers to experiment with multiple ways of expressing themselves in written language including expressions of creative ideas, personal opinions and feelings (Clark, 2012).

From the expressivist approach, writers are encouraged to explore their writing potential based on their personal views and beliefs, share their ideas with others, and connect with readers. However, emphasizing a very individualistic stance on writing may inhibit writers from considering other cultures that do not place the same value on 'self-expression' (Hyland, 2016). Furthermore, defining what 'good writing' is from this perspective is still hard to be accomplished because, on the one hand, there are no clear theoretical guidelines upon which 'good writing' could be evaluated, and on the other, no available advice is provided to clarify what 'good writing' consists of (North, 1987) as cited in (Hyland, 2016, p. 13). Additionally, there is no clear distinction between the processes that skilled and unskilled writers engage in when writing. However, despite the underlying limitations, the expressivist approach has contributed to the development of creative writing courses, teaching of writing in L1 classrooms mainly in the US and has paved the way for research in support of a cognitive view of writing (Hyland, 2010).

The cognitivist perspective views writing as a problem solving activity. Drawing on strategies and theories of cognitive psychology, this approach seeks to model the practical steps and techniques 'good writers' engage in when approaching a writing assignment and the available resources to solve such writing tasks (Hyland, 2009). Process oriented instruction is often characterized by modeling good writing strategies,

so that novice writers can benefit from them. For example, Williams (2003) claims that one of the major goals of the process approach is to modify student behavior and habits to comply with that of ‘good’ writers.

Similarly, Hyland (2016) maintains that process-oriented pedagogies aim to provide students with ample opportunities to compose texts modeling practical steps of expert writers. Composition pedagogy both in the first and second language domain has been heavily influenced by research under the cognitivist approach since the beginning of 1980s, when the composing processes of expert writers of English was largely emphasized by Hayes and Flower (1980), and Flower and Hayes (1980, 1981). They claimed that successful writers engage in several ‘nonlinear and recursive mental’ strategies translated into pedagogical stages such as planning, composition and revision.

Research investigating the text composition processes of skilled and unskilled writers found several cognitive operations and strategies as essential to produce texts of good quality. Among others, successful writers have been shown to approach writing by defining rhetorical problems, setting up goals, proposing and selecting solutions, deploying and monitoring the effectiveness of writing strategies, elaborating definitions through explanation and illustration, manipulating syntax, constructing sentences of different types, and evaluating and reflecting on emerging text (Flower & Hayes, 1980, 1981; Haswell, 2008; Hayes & Flower, 1980; Hillocks, 2008; Hinkel, 2011; Kucer, 2014; Wiegler, 2014).

Last but not least, the teaching of writing as a process in first and second languages under a cognitivist approach entails the implementation of several pedagogical steps and techniques, such as brainstorming, planning and drafting, revising, collaborative writing, providing and sharing feedback with peers and receiving feedback from teachers, and finally closing the composition cycle with publishing (Atkinson, 2003b; Bruffee, 1984a; Clark, 2012). To conclude despite several differences between the cognitivist and expressivist perspectives, they both share similar views on the need to develop writer’s ability to understand writing as non-linear, recursive and generative process which encourages the self-discovery of the writer (Matsuda, 2003a, 2003b; Polio & Williams, 2009).

2.2.3 Collaborative writing

Early support for collaborative writing came from L1 composition literature. However, what L1 composition scholars (e.g. DiPardo & Freedman, 1988) generally meant by collaborative writing was peer response activities where learners provide each other with feedback on their written products and cooperative writing not necessarily collaborative writing in which learners co-construct texts together (e.g. Ede & Lunsford, 1990). Storch (2013) defines collaborative writing as “an activity where there is a shared and negotiated decision making process and a shared responsibility for the production of a single text” (p.3). From this definition, peer response activities, peer feedback or group planning are activities which occur either at the pre-writing or post-writing stages and thus are not considered collaborative writing, but rather collaboration among learners at one stage of the writing process. Most importantly, unlike in cooperative writing where peer interaction occurs only in pre-writing and post-writing stages, in collaborative writing peer interaction occurs during the whole writing process (Storch, 2016).

Arguments in support of collaborative writing stem from both L1 and L2 composition literature. On the one hand, scholars working in L1 composition studies such as Bruffee (1984) or Dale (1994b) argue that collaborative writing promotes the development of students’ writing skills, provides students with opportunities for exposure to different ideas, and equips them with a better sense of writing for an audience. On the other hand, scholars working in the field of second language (L2) writing (e.g. Alegría de la Colina & García Mayo, 2007; Calzada & García Mayo, 2020a; Calzada & García Mayo, 2020; Fernández Dobao, 2012, 2014, 2016; Kim, 2008; Storch, 2013, 2016; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009) have posited that collaborative writing provides both EFL and ESL learners with opportunities for foreign and second language learning.

To understand the role of collaborative writing from the perspective of L1 and L2 scholars, two major dimensions of L2 writing have been put forward in the literature, the learning-to-write (LW), and the writing-to-learn (writing to learn the content- WLC, or language knowledge and skills WLL) dimension. The first dimension (LW) tackles the manner in which L2 learners learn how to express themselves in L2 writing, that is, learning to write, the second dimension (LWC-WLL) writing-to-learn, either the

content knowledge or language knowledge, explores whether learners' engagement in L2 writing tasks and activities contributes to the second and foreign language learning other than developing the L2 writing proficiency on its own (Manchón, 2011). Thus based on this dichotomy, writing in L1 typically aims to develop learners' writing skills i.e., it enables learners to learn writing, whereas writing in L2 entails two major goals: (i) to develop the L2 writing proficiency in general, and (ii) to create opportunities for learning the L2 through writing activities in particular. Thus, collaborative writing tasks in the L2 context are generally perceived as activities that foster the development of the L2 acquisition (Storch, 2016).

There are three important features that distinguish collaborative writing from other forms of writing, such features include: (i) the substantive interaction that occurs during the whole writing process, (ii) a shared decision-making process among the writers over the text produced, and (iii) the production of a single written text (Ede & Lunsford, 1990). Given these features, it is not surprising that collaborative writing finds support from two leading theories of SLA, the interactionist perspective (Gass, 2003; Long, 1996; Mackey, 2012) and the sociocultural perspective (Donato, 1994; Swain, 2010). In collaborative writing, writers engage in communication to solve linguistic issues which may arise during the writing process. Such communication may include negotiations of meaning, interactional feedback, positive and negative feedback, modified output, etc. From the interactionist perspective all of these communicative features have been found to facilitate the acquisition of an L2 (Mackey, 2012). From sociocultural theoretical perspective, on the other hand, interaction is perceived as an opportunity for learners to verbalize their deliberations on language use, about how to best express their ideas, and as a means of solving complex tasks during which learners can consolidate their existing knowledge and eventually co-construct new L2 knowledge (e.g. Ohta, 2001; Storch, 2002b, 2013; Swain, 2006; Swain, 2010).

Learners assigned to collaborative writing conditions have been found to engage in substantial interaction to solve language problems that arise from tasks that require written output. Several studies have analyzed learners' talk as they were assigned to complete collaborative writing tasks (e.g. Fernández Dobao, 2012; Kim & McDonough, 2011; Leaser, 2004; McDonough et al., 2016; Storch & Aldosari, 2013; Wigglesworth & Storch, 2009) and have revealed that learners discuss not only issues related to language use but also issues dealing with organizational aspects. A notable feature of

collaborative writing tasks is that they provide learners with opportunities for focus on form (Alegria de la Colina & García Mayo, 2007; García Mayo, 2002; Kowal & Swain, 1994; Swain & Lapkin, 1998). As such, research (e.g. Niu, 2009) has indicated that tasks that require the production of written output are more useful to draw learners' attention to form than tasks that require only the production of oral output. Drawing on Long's (1991) claim that second language acquisition can be best promoted when learners are provided with opportunities to attend to form while making meaning, the production of written language may be advantageous as compared to spoken language for learners to attend to form and meaning simultaneously. Researchers (e.g. Cumming, 1989; Williams, 1999, 2012) have posited that writing compared to speaking, poses lower pressure on learners because writing does not require an immediate response as is the case with speaking tasks, consequently, they are more likely to focus on language choice and accuracy. Another important aspect of collaborative writing is that it is a useful site for L2 learners to provide feedback to one another. Feedback provided by learners is immediate, and it can be on all linguistic aspects or organization of ideas expressed in the text and it is present during the whole writing process (Storch, 2013). Most importantly, the feedback provided by peers may be of a better match than the teacher feedback for the learners' linguistic and cognitive level (Daiute & Dalton, 1993). Given all these positive features of collaborative writing, Storch (2011) rightly promotes the use of tasks that combine both speaking and writing in L2 classes because this can create favorable conditions for learners to attend to form and meaning simultaneously, and thus advance in their L2 development.

2.3 Collaborative writing and peer interaction

Collaborative writing entails much more than just working together on a piece of text. Thus as mentioned in the previous section, Storch (2016) greatly emphasizes the importance of collaborative writing for L2 learning because it is a useful site for peer interaction. That is, during collaborative writing learners can engage in interaction with their peers – such interaction revolves around solving issues that arise during the whole writing process. In contrast, in cooperative writing or team writing, interaction among peers takes place only in the pre-and post-writing stages (Storch, 2013). Additionally, researchers informed by sociocultural theory have increasingly proclaimed that collaborative writing tasks are such tasks that provide learners with opportunities to produce spoken and written output, and at the same time, such tasks allow learners to

attend to meaning and form even without teacher's intervention (Alegría de la Colina & García Mayo, 2007; Fernández Dobao, 2012; Payant & Kim, 2019; Storch, 2013, 2016; Swain, 2000). In this way, interaction among peers is important because by using the language to solve linguistic and task problems, learners create language learning opportunities (Yu & Lee, 2016).

During collaborative writing, learners verbalize their thoughts that can be about language or task-related issues. In this way, learners may reflect upon their language produced, exchange and provide alternative ideas on how to move forward with the task at hand, and engage in meta-talk (Kim, 2013). Learners' engagement in metatalk has frequently been operationalized as LREs. Swain and Lapkin (1998) define LREs as “. . . any part of a dialogue in which students talk about the language they are producing, question their language use, or correct themselves or others” (p. 326). As such, LREs are a source for L2 learning because they provide opportunities for learners to talk and work together to resolve any linguistic problems during the task completion, therefore such deliberations are considered crucial in promoting learners' L2 development (Gass & Mackey, 2015; Payant & Reagan, 2018).

Insights reported above derive mainly from studies with adult EFL/ESL participants. However, despite the fact of a larger number of studies with adult EFL/ESL learners (Fernández Dobao, 2012; Shehadeh, 2011; Storch, 2002b, 2008; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009) research with young EFL/ESL learners has recently begun to receive attention (Calzada & García Mayo, 2020a; García Mayo, 2017; García Mayo, 2019; Kopinska & Azkarai, 2020) and peer interaction has proved to be beneficial for young and adolescent learners as well. Driven by the fact that findings of studies with adult learners cannot be generalized to child language learners, researchers (Azkarai & García Mayo, 2017; Garcia Mayo, 2018; García Mayo & Lázaro Ibarrola, 2015; Hidalgo & Lázaro-Ibarrola, 2020) have now turned their focus to examine the effects of interaction among young and adolescent learners while engaged in collaborative writing tasks, in both EFL and ESL settings. Studies examining peer interaction as learners engage in collaborative writing tasks (Bagheri, Rahimi, & Riasati, 2012; García Mayo & Hidalgo, 2017; Lázaro Ibarrola & Hidalgo, 2017; Rhonda, 2009; Sample & Michel, 2015) have shown that the latter group of learners, i.e. young and adolescent learners, are able to provide and receive feedback, negotiate for meaning and form, and use the target language to resolve linguistic

problems. Most importantly, opportunities to experiment with L2 use to address linguistic problems while completing the task help them consolidate their existing knowledge and co-construct new knowledge about their target language. To get a better perspective on the role of peer interaction in collaborative writing be it with young, adolescent, or adult learners, either in EFL or ESL settings, researchers have applied different approaches and units of analysis to analyze the pair talk. Thus depending on the purpose of a study, the pair talk can be analyzed as a whole or a particular aspect of it. This issue is discussed in the next section.

2.3.1 Analysis of learner interaction in collaborative writing

Interaction among peers can be analyzed from at least two perspectives. While research (e.g. Pauline Foster & Ohta, 2005; Mackey, Oliver, & Leeman, 2003) informed by Long's (1983, 1996) Interaction Hypothesis has generally used turn as the unit of analysis, coding the pair talk for negotiation moves (e.g. confirmation checks, clarification requests, comprehension checks and recasts), research on collaborative writing informed by sociocultural theory (Vygotsky, 1878) has mainly used Language-Related Episodes (LREs) as the unit of analysis (Storch, 2013).

Two types of episodes have been distinguished in the literature on collaborative writing: (i) language-related episodes (LREs) and non-language related episodes (non-LREs). LRE as a unit of analysis has been mainly employed in studies on collaborative writing examining peer interaction, focusing on particular linguistic aspects (e.g. Alegría de la Colina & García Mayo, 2007; Leaser, 2004; Storch, 1998b; Villarreal & Gil-Sarratea, 2019; Watanabe & Swain, 2007; Wigglesworth & Storch, 2009). Non-LRE, on the other hand, is generally used in studies that investigate peer interaction dealing with non-linguistic aspects such as task management, content, organization, reading and revising, off-task behavior and the like (e.g. McDonough et al., 2016; Storch, 2005). As the present study operationalizes the peer talk in terms of LREs, in the following section, it is only this unit of analysis that will be discussed in more detail along with a scheme of categorization of LREs across several studies.

2.3.2 Analysis of peer talk focusing on LREs

The majority of research on collaborative tasks has used LREs as the unit of analysis to analyze the peer interaction that emerges while learners work together to write a joint text (e.g. Leeson, 2004; Mozaffari, 2017; Villarreal & Gil-Sarratea, 2019). Interestingly, a large number of studies have shown that there are many factors which might affect the number, type and resolution of LREs produced by collaborative tasks, among others, such variables include the type of tasks learners work on (e.g. Alegría de la Colina & García Mayo, 2007; Swain & Lapkin, 2001), learners' L2 proficiency (e.g. Kim & McDonough, 2008; Leeson, 2004; Choi and Iwashita 2016), pairing method (e.g. Mozaffari, 2017), relationships between learners working in collaborative tasks (e.g. Chen & Hapgood, 2019; Storch & Aldosari, 2013; Watanabe & Swain, 2007), and the size of the group of learners (e.g. Fernández Dobao, 2012, 2014).

However, despite an extensive use of LREs in research on collaborative tasks, when it comes to their categorization, the classification schemes differ across studies. While, for instance, some studies have only distinguished between form-focused and meaning-focused LREs (e.g. Swain and Lapkin, 2001, Kim and McDonough, 2008; Alegría de la Colina & García Mayo, 2007), other studies (e.g. Fernández Dobao, 2012; Storch & Aldosari, 2013; Villarreal & Munarriz-Ibarrola, 2021) have classified them into more categories such as meaning-focused, form-focused, and mechanics-focused. The difference between these types of classification has to do with learners' language focus and discussion involving issues to be solved while in the process of text co-construction. For instance, when there are only two categories of LREs such as meaning-focused and form-focused LREs, the learners' talk revolving around spelling, punctuation and pronunciation issues has been analyzed under either form-focused or lexis-focused categorization. On the other hand, when researchers opted for more than two categories, they analyzed such issues and tallied those instances of talk (LREs) under a separate category labeled mechanics-focused LREs. Regardless of the categories used, the majority of studies have coded LREs for their outcome, such as correctly resolved, incorrectly resolved, and left unresolved. To have a better idea of how researchers coded LREs and distinguished between types of LREs, identifying peer talk that belonged to one of the categories, one might refer to the table 2.1 below. Finally, as noted earlier, LREs are instances of learners' talk over a linguistic or task-related problem. Learners may talk using their L1 or L2 depending on their preferences

and language competence, however, their L1 or L2 use may serve different functions while on collaborative writing. This is discussed in the next section.

Table 2.1 Coding scheme of LREs

Coding scheme		LRE		Outcome		
Study	L	F	M	√	X	?
Swain & Lapkin, 2001	Discussing meaning of words, considering lexical alternatives	Spelling, morphology, syntax, discourse	-	-	-	-
Alegría de la Colina & García Mayo, 2007	Discussing meaning of words, considering lexical alternatives	Spelling, morphology, syntax, discourse	-	√	X	?
Leeser, 2004	Discussing meaning/lexical items, spelling, and preposition	Morphology & syntax	-	√	X	?
Kim & McDonough, 2008	Discussing meaning of words, pronunciation and spelling issues.	Grammatical aspects: verb endings, tense, case, particles.	-	√	X	?
Fernández Dobao, 2013	Discussing meaning of words, considering lexical alternatives	Grammatical: tense, agreement, preposition.	Spelling, pronunciation, punctuation, accentuation.	√	X	?
Storch & Aldosari, 2013	Discussing meaning of words, considering lexical alternatives, and prepositions	Tense, articles,	Spelling, and punctuation issues	√	X	?
Villarreal & GilSarrate, 2019	Discussing meaning of words, considering lexical alternatives	Morphology & syntax	Spelling, and punctuation issues	√	X	?
Villarreal & Munarriz-Ibarrola, 2021	Discussing meaning of words, considering lexical alternatives	Grammatical aspects, use of articles, tense, etc.	Spelling, and punctuation issues	√	X	?
L= lexical; F= grammatical; M= mechanical. Outcome: correct = √; incorrect =X; unresolved =?						

2.4 The role of L1 and L2 use in collaborative writing

As Storch (2016) noted, one of the main traits of collaborative writing is that it enables L1 and L2 learners to engage in substantial interaction during the whole writing process. However, despite benefits of peer interaction for second language acquisition, some foreign language teachers are still reluctant to use pair and group work activities in their L2 classes due to their concern that learners may resort to their shared L1 as they engage in communicative tasks (Brooks & Donato, 1994; Carless, 2007; Tognini & Oliver, 2012).

This concern may stem from a teacher's belief that target language use in L2 classrooms enables more exposure to L2 input. The study of Levine (2003) that surveyed 163 L2 teachers with regards to L1 or L2 use, found that a larger number of L2 teachers believed that using L2 rather than L1 in L2 classes is far more rewarding for foreign language learners. A similar finding was also reported in an earlier study by Duff and Polio (1990). Investigating the use of L1 and L2 in foreign language classrooms, Duff and Polio found that teachers' belief towards L2 use dominated those towards L1. The majority of teachers involved in the study thought that using L1 in L2 classrooms restricts learners' exposure to L2 input. These teachers' perceptions are in line with Krashen's (1982) Comprehensible Input Hypothesis, which posits that second language acquisition is facilitated via sufficient exposure to L2 comprehensible input. Teachers in both EFL and ESL contexts that make use of communicative approaches to language teaching highlight the use of target language in L2 classrooms, as a way of promoting L2 learning (Atkinson, 1987; Swain & Lapkin, 2013).

Despite controversial opinions regarding L1 use in L2 classrooms (Wagner & Toth, 2013), studies investigating learners' L1 use in L2 classrooms have shown that EFL learners make a modest use of their shared L1, and that the L1 serves functions that are related to task completion (Storch & Aldosari, 2010). Research informed by sociocultural theory has found that L1 use serves as a mediating tool for learners enabling them to complete highly cognitively demanding tasks in L2 (e.g. Antón & DiCamilla, 1999; Storch & Wigglesworth, 2003). This body of research has mainly examined the functions of L1 use by analyzing learners' interaction while completing collaborative writing tasks (DiCamilla & Antón, 2012; Scott & de la Fuente, 2008; Swain & Lapkin, 2000).

One of the earliest to investigate the L1 use and the functions it serves in L2 classrooms was the study by Antón and Dicamilla (1999). Informed by sociocultural theory, in this study the researchers analyzed the oral data of 10 adult students of Spanish as a foreign language as they completed three collaborative writing tasks. The study found that using their native language students were able to discuss about the task completion, task content, address specific task requirements (such as looking for a specific word), develop strategies to deal with the task at hand, engage in metatalk, and scaffold each other while completing the tasks. More importantly, L1 played a critical role for L2 learning to occur.

Other studies have looked whether L1 use among ESL learners differs across tasks. Storch and Wigglesworth (2003), for example, examined the L1 use among 24 intermediate ESL university students assigned to complete two different collaborative writing tasks (i.e. text reconstruction and joint composition task). In this study, half of the pairs (six) shared their L1 and were encouraged to use it for interaction while completing the tasks, whereas the other half did not share an L1. The study found that despite the shared L1 of the six pairs, the majority of participants used their L1 only to a minimum extent. The data also showed that L1 use differed across tasks. While in the joint composition task participants used their L1 for clarifications and task management purposes, in the text reconstruction task, the L1 use served other functions, mainly that of negotiation of meaning and discussion of vocabulary issues.

A limited use of L1 has been reported in immersion classes as well. Swain and Lapkin (2000), for instance, examined whether task type had any impact on L1 use among grade eight students in French immersion classes, assigned to complete a dictogloss and a jigsaw task collaboratively. The study found that despite variations of L1 use among pairs, the number of total turns showing L1 use amounted up to one quarter of the total turns, and only about 12% of the L1 turns were off-task. While the data showed that pairs perceived of lower proficiency made a greater use of L1 than those of higher proficiency. On the whole, the L1 use by both groups served three main functions: (i) moving the task along (e.g. understanding, task management), (ii) focusing attention (e.g. searching for words, focusing on form) and (iii) interpersonal relations (e.g. off-task, and disagreements). Furthermore, no statistically significant differences were found regarding L1 use across the tasks, a contradictory finding with that of Storch and Wigglesworth (2003).

Several studies that have investigated other mediating variables such as learner proficiency and task type have brought important insights regarding the L1 use and the functions it serves among English as a foreign language learners. In a subsequent study Storch and Aldosari (2010), for example, examined whether learner proficiency pairing and task type had any impact on the amount of L1 use and the functions it served among EFL learners while completing three different collaborative tasks (e.g. jigsaw, text-editing, and text composition). In this study, all fifteen learners were placed in three different proficiency groupings, that is, high-high, high-low and low-low, and completed three different tasks, as mentioned above. The study found that, regardless of the proficiency pairing, L1 use was very low; however the task type had a greater impact on L1 use among all pairs. Thus, low-low pairs made the greatest use of L1 among all other pairs. The functions that L1 served during the completion of the pair work were mainly related to task management and vocabulary deliberations. That proficiency pairing has an impact on learners L1 use and the functions it serves is evident in the study of DiCamilla and Antón (2012) as well. Comparing the amount of L1 use and the functions it served among EFL learners of low proficiency and high proficiency while completing a collaborative activity in an English as a foreign language context, DiCamilla and Antón found that lower proficiency students used their L1 far more often than higher proficiency students, and their L1 use served mainly functions dealing with task management, solving grammatical and lexical problems, and social communications. Higher proficiency students, on the other hand, made a rather limited use of L1 and mostly used it for social communications. They used their target language (L2) to deal with all task demands. It seems that L2 among higher proficiency learners played an important interaction role, which fits well with Swain and Lapkin's (2013) claim that "... as L2 proficiency increases, it may be used for a wider variety of functions [...] and [...] there is less and less need to use the L1 as a cognitive tool; the L2 can "stand alone" (p. 110).

More recent research has explored whether L1 use among learners affects the quality of the text produced collaboratively. Zhang (2018), for example, examined the effects of learners' L1 and L2 interaction on the complexity, accuracy, fluency and text quality of 70 intermediate EFL students as they completed collaborative writing tasks. Participants in this study worked in pairs (n=35) and completed two argumentative essays. In the first writing task, they were instructed to use their native language

(Chinese) as the medium of interaction, whereas in the second time they were instructed to use their target language (English). The study found that students' collaborative texts contained longer clauses in the first task condition. On the other hand, no significant differences were found between co-constructed texts in terms of quality, accuracy and fluency produced by learners interacting in L1 or L2 while writing. Given that longer clauses were found in L1 interaction mode, Zhang concluded that L1 use in collaborative writing might serve as an important mediational tool helping learners to retrieve more complex linguistic forms and thus produce texts of richer syntactic features. In a subsequent study, Zhang (2019) examined the effects of L1 and L2 interaction on the production of lexical, phrasal, clausal and intermediate features (i.e. finite and passive non-finite relative clauses) of learners' co-constructed texts and looked whether learners' L1 or L2 use affects learners' focus of discussion during collaborative writing. The study found that learners interacting in L1 condition produced more lexical and phrasal features than those interacting in L2. In addition, learners interacting in L1 were more focused on the discussion of language use and task management, whereas those interacting in L2, focused more on language repetition, reading and backchannel. More recently, Zhang and Crawford (2021) in a study with adult EFL learners examined whether interacting in L1 or L2 affects the quantity and quality of LREs produced while learners completed a composition task in pairs. The study found that learners interacting in L1 not only produced more LREs but also resolved a higher number of lexis-focused and mechanics-focused LREs compared to learners interacting in L2. In addition, learners interacting in their shared L1 were able to deliberate more about language use and discuss lexis-focused LREs in a more elaborate manner than those interacting in L2.

To sum up, research informed from the sociocultural perspective, considers the use of L1 in L2 classes as an important mediational tool that enables learners to complete highly cognitive demanding tasks. Though L1 use may sometimes be a concern for language teachers, researchers working under the framework of sociocultural theory consider it an important device. For example, DiCamilla and Antón (2012) consider the use of L1 in L2 classes as "an important and often indispensable semiotic device that mediates the learning process, even though what is being learned is an L2" (p. 161). The amount of L1 use may be influenced by task demands and learners L2 proficiency. Thus, when learners engage in task activities that demand a

higher command of the target language (e.g. task management) they resort to L1 more frequently, whereas when they engage in task activities that demand lower level of language knowledge (e.g. form-focused activities) they use the L2 (e.g. Tognini and Oliver, 2012). While learners of higher L2 proficiency make less use of L1, learners of lower L2 proficiency make a greater use of their native language when assigned in language tasks (Storch & Aldosari, 2010; Storch & Wigglesworth, 2003; Swain & Lapkin, 2000). Furthermore, the use of L1 in L2 classes has been found to serve important functions helping learners to understand the task content, task criteria and deal with task organization and vocabulary issues (Storch & Aldosari, 2010; Storch & Wigglesworth, 2003; Swain & Lapkin, 2000; Tognini & Oliver, 2012). Given its important functions, Swain and Lapkin, maintain that “to insist that no use be made of the L1 in carrying out tasks that are both linguistically and cognitively complex is to deny the use of an important cognitive tool” (p. 268-9). Therefore, L1 use in L2 classes may contribute L2 learners to make better use of target language and produce written output of better quality.

2.5 Collaborative writing and opportunities for target language use

Collaborative writing tasks provide a wide range of benefits for L2 learners. Among others, one very important feature of collaborative writing is that it provides learners with opportunities for peer interaction while co-constructing their texts. More importantly, the interaction learners engage in is usually carried out in the target language, which Storch (2010) finds it extremely useful in foreign language contexts because classroom is often the only site for learners to practise their L2. On collaborative tasks, learners not only make use of target language but also use it for a wide range of functions which are generally ‘reserved’ for the teacher or perceived as being beyond learners’ scope (Storch, 2013). Studies examining the nature of peer talk during collaborative writing with adult ESL/EFL learners (e.g. de la Colina & García Mayo, 2007; Wigglesworth & Storch, 2009) and young and adolescent EFL learners (Calzada & García Mayo, 2020b; Hidalgo & García Mayo, 2019; Villarreal & Gil-Sarratea, 2019) have shown that learners use the target language for a range of functions, including negative feedback (e.g. corrective feedback, recasts, clarification requests, etc.), positive feedback (confirmation, praise), confirmation checks, grammatical explanations, word choice, discussion of task requirements. Excerpts extracted from a number of studies mentioned above, demonstrate how learners use the L2 to deliberate

on language issues within their EFL/ESL classes while completing collaborative writing tasks.

Example 1 shows how two learners assigned to work on a collaborative writing task discuss many language issues simultaneously such as deliberating over word choice (*good* vs. *efficient* and *effective*), correcting each other on grammatical forms - indefinite article (*a* vs. *an*) and plural vs. singular marker (*students* vs. *a student*) (Wigglesworth and Storch, 2009, p. 459).

Example 1: a pair working on an essay

119 Fay: it's a good...
120 Miles: it's a efficient way... it's a effective and efficient way to motive to
121 motivate students to study.
122 Fay: that's right (15 second silence). So it's an efficient and effective way
123 to motive a student to study hard.
124 Miles: yes

Example 2 shows a pair assigned to work on a dictogloss task discussing task requirements (managing the task) and offering suggestions (Alegría de la Colina & García Mayo, 2007, p. 15).

Example 2: a pair working on a dictogloss

A: Ah, but we also have to talk about the work force before this because this was the last thing, I mean, first is time, then the work force and after that the part about 'loading and unloading')
B: wouldn't it be 'handling'?

Example 3 shows young EFL learners working in pairs to co-construct a story based on a set of pictures. They deliberate over selecting the proper tense to use in their writing, present versus past. They were writing the story in the present tense. They offer suggestions and counter suggestions regarding the use of tenses. At the end, they managed to pick up the right tense after the learner (CHI1) offers some grammatical explanations to CHI2 (Hidalgo & García Mayo, 2019, p. 8).

Example 3: a pair working on a story writing

CHI2: they took.
CHI1: take the (..).
CHI2: took because the (..).
CHI1: they take took is in past and (.).
CHI2: es verdad (it is true).
CHI1: is happening right now.

In Example 4, young EFL learners working on a collaborative writing task provide corrective feedback. Specifically, the learner (CH2) provides corrective feedback (i.e. explicit correction) to his/her partner on using the appropriate object pronoun. CH1 accepts the feedback and changes the pronoun from '*his*' to '*her*' (Calzada & García Mayo, 2020, p. 8).

Example 4: a pair working on a dictogloss task

CH1: and Lucy eat all the sweets in one night and his teeth...

CH2: 'her', 'her'... 'his' is for boys.

CH1: 'her', 'her'.

All these excerpts have one common feature, which is the use of target language by L2 learners of different age and in different contexts. When it comes to the type of tasks that promote the use of target language among peers, researchers (e.g. Storch and Aldosari, 2010) have found that the jigsaw task is superior compared to other collaborative tasks (e.g. text-editing or composition) in terms of boosting the target language use between learners. This is most likely due to the need to exchange the information over the content of pictures among learners (Alegría de la Colina & García Mayo, 2009).

Nevertheless, learners' L2 proficiency may also determine the extent of the target language in peer interaction. Swain and Lapkin (2000), for instance, found that pairs composed of high-high proficiency members made far more use of L2 compared to pairs of high-low and low-low members. Similarly, DiCamilla and Antón (2012) found the greatest amount of L2 use among pairs composed of higher-proficiency learners as opposed to pairs of lower-proficiency. They also claimed that as learners' L2 proficiency advances, it can serve as a mediational tool to enable learning to occur. As such, it seems that learners' L2 proficiency may not only determine the extent of L2 use, but it also may determine the functions the L2 may have.

Talking to a peer in target language aids L2 development. Philp, Adams, and Iwashita (2013), among others, highlight four beneficial aspects of peer interaction for L2 learning, suggesting that learners who engage in peer interaction can benefit by: (i) attempting to express what they already know, (ii) revising linguistic forms they already possess by adapting them to solve difficulties they encounter, (iii) revising language by building on each other's knowledge about language, and (iv) practising and appropriating the 'formulaic language' they have learned in classroom for real

communication (p. 20). In the same vein, Macaro (2009) and Littlewood (2014) stress the relevance of target language use in EFL contexts because it increases learners' amount of exposure to L2 input, and it contributes to the natural and effortless development of learners' L2. As such, writing collaboratively may not only benefit learners in terms of reaching solutions for the problems they encounter during the task completion process, but it also benefits them in terms of target language use (Adams & Oliver, 2019). Last but not least, talking to a peer helps learners reduce their speaking anxiety and boost their confidence (McDonough, 2004), which is not always the case when talking to a teacher or the whole class (Ball, Kelly, & Clegg, 2015). Finally, even though collaborative writing may be challenging for a small number of learners, yet, it is highly appreciated by others and beneficial for language learning.

Chapter 3: Prior research on L2 collaborative writing

Collaborative writing in L2 began to be explored earnestly in the last two decades. Research has shown that when learners work on collaborative writing tasks, they need to talk to each other to reach solutions for the potential problems arising from the task at hand. As a result, they often deliberate over the language use, offer suggestions and counter suggestions, provide and receive feedback. These features are paramount for L2 development.

In the literature on L2 collaborative writing, three distinct strands of research can be identified. The first strand of research focuses on the factors affecting the nature and volume of languaging of learners engaged in collaborative tasks. Such factors include: the type of tasks learners work on, learners' L2 proficiency, and the relationship learners form when assigned to collaborative writing tasks, including studies in computer-mediated context (CMC). The second strand of research, the main focus of this thesis, examines the products and outcomes of the written texts produced by learners working individually, in pairs and small groups both in face-to-face (FTF) and computer-mediated contexts (CMC). In addition, it explores the effect of collaborative work on language learning gains. Then, it continues with some studies reporting on learners' focus (meaning vs. form) during collaborative writing in CMC. The third strand of research reports on learners' perceptions and attitudes towards collaborative writing, including studies conducted in CMC. This chapter reports on all three strands and ends with a summary of the literature review geared in particular on prior research relevant to the present study.

3.1 Factors affecting the number, type and resolution of LREs

As indicated above, the first strand of research examines several variables that have been identified as influencing variables affecting the number, type and resolution of LREs produced by learners working on collaborative writing tasks. Such factors can include learners' L2 proficiency, patterns of interaction learners form when assigned to collaborative writing tasks, and the type of tasks learners work in. These variables can also affect the language learning opportunities learners can create when working collaboratively.

3.1.1 Learners' L2 proficiency and LREs

Research investigating the impact of learner proficiency on interaction, attention to language during collaborative writing tasks, and particularly on resolution of Language Related Episodes (LREs) is limited. Leeson's (2004) study was one of the earliest to examine the impact of learner proficiency on the quantity and quality of LREs when assigned to collaborative writing tasks. In this study 42 learners of Spanish as a second language were assigned to work collaboratively on a dictogloss task. Students were assigned to work in three different proficiency groupings. Thus, there were eight pairs of similar proficiency (high-high), nine pairs of similar proficiency (low-low), and four other pairs of mixed-proficiency (high-low). The study found that learners with similar proficiency level (high-high) produced the highest number of LREs, followed by the mixed-proficiency pairs (high-low), and subsequently by (low-low) pairs that produced the smallest number of LREs. Interestingly, the study also found that learners' focus on language as they worked together varied according to their language proficiency. Thus, pairs of similar (high-high) proficiency focused mainly on grammatical forms and resolved correctly the majority of LREs, whereas dyads of similar (low-low) proficiency focused mainly on lexis and resolved the smallest number of LREs from all other groups. When the production of grammatical LREs was compared between mixed-proficiency pairs (high-low) and similar proficiency pairs (low-low), the number of grammatical LREs tended to be higher for the mixed-proficiency pairs, but it was not statistically significant. The findings led Leeson to conclude that although the dictogloss task was useful for high-high dyads as it generated a lot of attention to form, it may not be similarly useful for the pairs of low-proficiency due to the small number of correctly resolved LREs.

Other studies have reported more positive results for pairs of mixed-proficiency. Kim and McDonough (2008), for example, conducted a study with 24 participants of Korean as a Second Language (KSL), and examined learners' attention to language as they collaborated in dictogloss tasks, as well as whether learner proficiency had any impact on the resolution of language-related problems. Initially, eight pairs of similar proficiency (intermediate-intermediate) completed a dictogloss task, then, eight intermediate learners were placed to work with eight advanced learners forming thus eight mixed-proficiency dyads who completed a similar dictogloss task. The analysis of the data revealed that pairs of mixed-proficiency (intermediate-advanced) produced the

highest number of lexical LREs and were most successful in resolving such LREs correctly than dyads of similar proficiency (intermediate-intermediate). However, unlike in Leaser's (2004) study, Kim and McDonough (2008) found no significant differences concerning the number of grammatical LREs between pairs of similar vs. different language proficiency. In a similar vein, Watanabe and Swain (2007) conducted a small-scale study to examine how learners' proficiency differences and patterns of interaction affect L2 learning. Thus, they assigned four core participants to work with eight non-core participants of different proficiency, either higher or lower than their own to complete a multistage task that included peer writing, reformulation by a native-speaker, noticing gaps (pairs were given their original written text and a reformulated version of it, so that they could detect the differences), and an individual writing task. The study found that the pairs formed of core and higher proficiency learners produced the greatest number of LREs as compared to the pairs of core and lower proficiency learners, a finding that aligned with that of Leaser (2004). This suggests that as the proficiency of learners increases so does the number and outcome of LREs. A more recent study by Storch and Aldosari (2013) brought further insights regarding the impact of proficiency and patterns of interaction in collaborative tasks. In this study, the researchers investigated ways that promote the effectiveness of pair work, choosing learners' L2 proficiency as the main variable. Thirty EFL university students were placed in different proficiency groupings to complete a short composition. Thus, learners were assigned into pairs of similar proficiency, five pairs (high-high), five pairs (low-low) and five pairs of mixed-proficiency (high-low). Similar to the findings of Leaser (2004), Storch and Aldosari (2013) found that the high-high pairs produced a higher number of LREs compared to the high-low and low-low pairs. Regarding the type of LREs produced, the study found that lexical LREs dominated over grammatical and mechanical LREs regardless of the learner proficiency. These findings indicate that learners' L2 proficiency is an influencing factor that impacts on the quantity and quality of LREs produced. However, both Watanabe and Swain (2007), and Storch and Aldosari (2013) maintain that there are other possible factors (e.g. patterns of interaction) that should be taken into account along with learners' L2 proficiency that may impact on the quantity and quality of LREs produced from collaborative work (c.f. Section 3.1.2).

All in all, results reported above come from studies with EFL/ESL participants from university contexts, whereas research with adolescent EFL learners from lower secondary school context measuring the effect of proficiency on the amount, type and outcome of LREs is scarce. Additionally, to control the variable of proficiency studies reported above used mixed-proficiency pairs instead of matched-proficiency pairs. Research has shown that in mixed-proficiency pairs, the more proficient learner may take up a leading role, while the less proficient member may feel intimidated in front of a more proficient learner and thus an unbalanced contribution to the task may emerge, in particular if learners form non-collaborative patterns of interaction (Storch & Aldosari, 2013; Watanabe & Swain, 2007). Hence, little is known how adolescent EFL learners' proficiency affects the quantity, type and quality of LREs produced from collaborative tasks. Interestingly, besides proficiency, the relationships learners form when working in pairs has been found to affect the number, type and outcome of LREs produced, both in face-to-face (FTF) and computer-mediated context (CMC). This is discussed in what follows.

3.1.2 Learner relationships and LREs

Research investigating dyadic relationships in L2 classrooms is still scarce, however, studies conducted thus far have brought important evidence regarding different patterns of interaction learners form when assigned to work in pairs or small groups in face-to-face (FTF) (Sato & Viveros, 2016; Storch, 2002a, 2002b; Storch & Aldosari, 2013; Watanabe, 2008; Watanabe & Swain, 2007) and computer mediated context (CMC) (e.g. Abrams, 2019; Li & Zhu, 2017).

One of the earliest studies to investigate the nature of pair interaction was Storch's longitudinal study (2002a). Informed by Sociocultural theory, Storch examined the nature of dyadic interaction using the data from ten pairs of adult ESL learners participating in a number of language tasks (e.g. short composition, editing task and text-reconstruction task) over a semester. The study found four different patterns of interaction: collaborative, expert/novice, dominant/dominant, and dominant/passive. Storch distinguished these patterns of interaction based on a model reflecting learners' involvement and contribution to the task or activity. Thus, the notion of equality in this model reflects the learners' control over the task or activity, whereas mutuality reflects the learners' degree of engagement with each other's contribution or ideas. Using this

model, Storch found more collaborative patterns of interaction in the data of all pairs. In addition, once collaborative patterns were formed, they were stable over time and across tasks. The study also revealed more instances of a transfer of knowledge among the dyadic patterns of interaction with a collaborative and expert/novice orientation than among the patterns of dyadic interaction with dominant/dominant and dominant/passive orientation. More instances showing lack of a transfer of knowledge were observed among the dominant/dominant patterns of interaction, and more instances of missed opportunities for a transfer of knowledge were observed in the dominant/passive patterns of dyadic interaction. These findings led Storch to conclude that not all pair and group work is conducive to learning, especially when pairs form dominant/dominant and dominant/passive patterns of interaction. In contrast, language learning is more likely to occur when learners form collaborative or expert/novice dyadic relationships.

Later studies reported somewhat similar results. Drawing on Storch's (2002a) study on patterns of interaction, Watanabe and Swain (2007), for example, found that the pairs forming collaborative and expert/novice patterns of interaction produced a greater number of LREs than the pairs forming dominant/passive and expert/passive dyadic relationships. In addition, post-test results reflected a similar trend in terms of patterns of interaction. Thus, when both pair members were engaged in collaborative patterns of interaction they reached higher post-tests scores, which was not the case when they were engaged in non-collaborative patterns of interaction such as dominant/passive or expert/passive. Similar to the findings of Storch (2002a, 2002b), Watanabe and Swain (2007) argue that learning is more likely to occur when learners engage in collaborative dyadic patterns of interaction than in non-collaborative patterns of interaction. However, unlike Storch's claim that more instances of evidence showing a transfer of knowledge were observed among collaborative and expert/novice patterns of interaction, Watanabe and Swain (2007) posit that when there is a large gap of L2 proficiency between learners forming the pair, it is only the collaborative pattern of interaction that might be conducive to L2 learning for both learners. Despite learners' proficiency differences, both a higher and a lower proficiency learner can create learning opportunities for each other, provided that they form collaborative patterns of interaction.

Subsequent studies have also reported similar results that patterns of interaction constitute a major factor to be considered when assigning learners to work collaboratively. In a study examining how adult ESL learners interact with a higher or lower proficiency peer during a problem solving task, Watanabe (2008), for example, found that regardless of learners' language proficiency (lower or higher) both peers could provide learning opportunities only when they showed collaborative orientation. In addition, interview data showed that learners preferred to work with interlocutors that were collaborative and shared ideas despite their language proficiency. Confirming evidence that patterns of interaction is far more important than proficiency pairing was brought forward by Storch and Aldosari (2013) mentioned earlier. The study revealed that proficiency among learners forming the pairs was not the only variable that affected the quantity and quality of LREs produced but patterns of interaction learners formed had a positive impact on the resolution of LREs as well. As reported in the study by Watanabe and Swain (2007) who found that the pairs with collaborative orientation produced a higher number of LREs than the pairs with non-collaborative orientation such as dominant/passive and expert/passive, Storch and Aldosari (2013) found that high-low pairs outperformed low-low pairs in terms of the production of LREs only when they showed collaborative or expert/novice dyadic patterns of interaction. When they formed dominant/passive dyadic relationships, on the other hand, they produced a lower number of LREs. Likewise, a greater number of LREs was produced when high proficiency learners were paired with either a fellow high proficiency or a low proficiency learner and formed collaborative patterns of interaction, however, the number of LREs was lower when pairs formed expert/novice relationships, and even lower when they showed dominant/passive dyadic relationships orientation. Given these findings, Storch and Alodsari (2013) suggest that the relationship learners are likely to form when working together can impact on the opportunities L2 learners provide to each other as they work on collaborative tasks. Similar to the findings of Storch and Aldosari (2013), in a more recent study with young EFL learners (aged 10-12) from a CLIL context, who were assigned to work in pairs (n=12) and triads (n=3) on a map task, Basterrechea and Gallardo-del-Puerto, (2023) found that pairs exhibiting expert-novice and collaborative relationships produced a greater number of LREs and resolved correctly a higher proportion of such LREs compared to the pairs forming non-collaborative patterns of interaction (i.e. dominant-dominant, and dominant-passive). These results suggests that regardless of the participants' age, it is the relationships

learners form that may affect the volume and quality of languaging when working on collaborative tasks. Also a study by Sato and Viveros (2016) with higher secondary EFL learners brought additional insights concerning the relationships learners form when working collaboratively. Thus using L2 proficiency as an independent variable, the study examined whether learners' L2 proficiency had an impact on patterns of interaction learners formed and on the provision of corrective feedback and production of modified output (CF, MO). Additionally, the study sought to examine the relationship between CF and MO and the collaborative patterns during group work. In this study two groups of learners (one as higher proficiency and one as lower proficiency) consisted of five students each engaged in communicative activities implemented in four different stages (e.g. presentation, awareness, guided practice and communicative practice), and aimed to promote real collaborative group work among EFL learners. The study found that learners placed in the low proficiency group outperformed the learners placed in high proficiency group in terms of provision of corrective feedback and production of modified output. Furthermore, the low proficiency learners showed a more collaborative orientation leading the researchers to speculate that it was a stronger mediating factor other than proficiency that affected the results, perhaps learners' attitudes towards tasks or their working partners. The study also revealed a link between CF and MO and patterns of collaboration. Thus, also Sato and Viveros suggest that collaborative interaction is a stronger mediating variable than proficiency regarding the provision of corrective feedback and production of modified output among learners. In this sense they claim that "the more collaborative a group is, the more CF they provide and the more MO they produce regardless of the learners' proficiency levels" (p.106). Moreover, unlike Watanabe and Swain (2007) who posited that the higher proficiency learners benefited more from peer interaction, Sato and Viveros found that it was the low proficiency learners who benefited more from peer interaction. Finally patterns of interaction learners form when working in collaborative writing tasks may affect not only the volume and quality of languaging, but also the products of collaborative writing. For example, Jang and Cheung (2020), who investigated the impact of peer interaction on the process and product of collaborative writing among young ESL learners, found that learners who formed collaborative and expert/novice patterns of dyadic interaction produced more LREs and achieved higher scores on post-writing tests compared to the pairs showing dominant/dominant and dominant/moderately passive patterns of interaction. In addition, similar to the findings of Storch (2002a) and

Watanabe and Swain (2007), the study revealed that the pairs showing collaborative and expert/novice patterns of interaction not only produced a greater number of LREs but also resolved correctly a higher number of them compared with the pairs with dominant/dominant and dominant/moderately passive orientation. The study also found that the pairs who engaged in the dominant/dominant patterns of interaction produced and resolved correctly the smallest number of LREs.

Studies that have investigated whether the patterns of interaction learners form when working in computer-mediated context affect the writing products have also found that the quality of writing may be affected by the patterns of interaction learners form in such modes. Abrams (2019), for example, found that the groups of learners who exhibited collaborative patterns of interaction produced significantly better texts in terms of fluency, propositional content, and textual coherence. Similarly, Li and Zhu (2017), found that the groups which exemplified higher mutuality, respectively showing collective and expert/passive patterns of interaction produced wiki writing products of overall better quality compared to groups characterized by lower mutuality, respectively exhibiting dominant/defensive and cooperating-in-parallel (dominant/dominant) patterns of interaction. These findings are in line with those produced by studies in face-to-face settings (e.g. Watanabe, 2019) who found that pairs with collaborative patterns of interaction produced texts of better quality and reached higher posttest scores than those exhibiting non-collaborative orientation.

In sum, studies reported above revealed that not only learners' L2 proficiency, but also the patterns of interaction they form when assigned to work in pairs or small groups affect the quantity and quality of collaborative dialogue and/or the quality of written products. As such, for more learning L2 opportunities to be created and more learning to occur, it is essential that learners form collaborative or expert/novice patterns of interaction (e.g. Storch, 2002b; Watanabe & Swain, 2007; Watanabe, 2019). Though research with adolescent EFL learners is still thin on the ground, the studies reported above suggest that patterns of interaction learners form during collaborative writing affect the nature and volume of languaging, a source of second language learning (Swain, 2006, 2010).

3.1.3 Task types and LREs

As mentioned earlier, collaborative writing enables learners to work together, co-construct language knowledge, and pay attention to meaning and form without teacher intervention (Alegria de la Colina & García Mayo, 2007; Payant & Kim, 2019; Storch, 2016; Storch & Wigglesworth, 2007; Swain, 2000). Collaborative writing tasks have been distinguished in terms of meaning-focused and grammar-focused. Thus, jigsaw, data commentary report and composition are considered as meaning-focused collaborative writing tasks as they require learners to write a text based on a prompt be it graphic or pictorial (Storch, 2016). A jigsaw task is an information exchange task which requires participants to share their part of information with their co-workers in order to complete the task (e.g. Pica et al. 1993). In oral interaction research, the jigsaw has mainly been used as a task promoting negotiation of meaning (e.g. Gass & Mackey, 2007), whereas in collaborative writing it may be used as a task that allows learners to build a story based on a set of photographs (Storch, 2013). A data commentary report requires participants to write a summary report based on a graphic or pictorial prompt.

The dictogloss, on the other hand, is a grammar-focused task (Wajnryb, 1990) where learners need to follow three steps to complete the task. At first, learners simply listen to a text in order to get the general idea of it. They then listen to it for the second time and take some notes (words, uncomplete sentences). Finally, they are required to write the text as accurately as possible, individually, in pairs or small groups and so complete the task (García Mayo 2018a). Text editing and text reconstruction are more explicitly grammar-focused tasks that draw learners' attention to pre-determined language items. In a text editing task students are required to check for grammatical errors in a certain piece of writing and correct such errors (e.g. Storch, 1998). The composition can be one either written by other students in the class or written by the teacher or researcher. In a text reconstruction task learners are required to insert appropriate words within a text using proper word inflections based on specific verb-tenses (Storch, 2009).

Research on collaborative writing has found that collaborative writing tasks be it grammar-based or meaning-based have proved effective in drawing learners' attention to form and meaning (Alegria de la Colina & García Mayo, 2007; Azkarai & García Mayo, 2012; Calzada & García Mayo, 2020a, 2020b; Kowal & Swain, 1994, 1997;

Storch, 1998; Swain & Lapkin, 2001). Investigating whether grammar or meaning-focused tasks are more useful for the generation of LREs, researchers have claimed that grammar-based collaborative tasks such as dictogloss, text editing and text reconstruction tend to generate more LREs than meaning-focused collaborative tasks such as jigsaw, data commentary report or composition (e.g. Alegría de la Colina & García Mayo, 2007; Storch, 1998). It is also interesting to note that despite other possible alternatives the dictogloss, which is a language-focused task (Wajnryb, 1990) as mentioned earlier, seems to have been the most frequently used task in research on collaborative writing both in EFL and ESL settings (García Mayo, 2018). This may be due to the attempts to provide learners with opportunities for attention to form so as to encourage language learning. Meaning-focused tasks such as jigsaw, composition and data commentary report, on the other hand, have been less frequently used in research on collaborative writing (e.g. Fernández Dobao, 2012, 2014; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009).

One of the first studies to compare the number and nature of LREs produced by two different tasks was Storch's (1998) study with 9 intermediate ESL students assigned to complete a traditional grammar exercise (i.e. cloze) and collaborative writing tasks (i.e. short-composition). The study found that the grammar exercise task produced more language-focused LREs, whereas the collaborative writing task produced more lexis-focused LREs. Later studies investigating the amount and type of LREs generated by collaborative writing tasks brought no significant difference in terms of LREs produced by grammar-focused and meaning-focused collaborative tasks. Swain and Lapkin (2001), for example, examined the number and focus of LREs produced by two different collaborative tasks (e.g. dictogloss vs. jigsaw) with grade 8 students involved in French immersion programmes. Contrary to Storch's study, Swain and Lapkin (2001) found no significant difference in terms of number and focus of LREs produced by two different tasks. The researchers attributed this finding to the mini grammar lesson (i.e. a video-recorded mini-lesson on French pronominal verbs shown to students), and the lack of prior modeling of both tasks, as this has been found to affect learners' attention to language (Kim & McDonough, 2011). These findings were contradicted by later studies, though conducted in different settings and with learners of different language proficiency.

In a study that investigated the effectiveness of collaborative writing for focus on form with 24 low-proficiency EFL university students performing three different collaborative writing tasks (e.g. a jigsaw, dictogloss and text reconstruction), Alegría de la Colina and García Mayo (2007), for example, found that collaborative writing is effective in promoting learners' focus on form, as a great number of LREs were produced. Comparing the number of LREs produced across all tasks, the study found that the reconstruction task generated the greatest number of LREs, followed by the jigsaw task, and finally the dictogloss task which generated the smallest number of LREs. As per the focus of LREs, the study found a higher proportion of form-focused LREs compared to meaning-focused LREs. Given these results the authors claimed that collaborative writing tasks proved effective in drawing low-proficiency learners' attention to form and that low-proficiency learners can discuss a great deal of language issues, which is contrary to what other studies have found (e.g. Leeson, 2004). Examining whether proficiency pairing affects the amount and quality of LREs produced, Leeson found that low-proficiency learners who worked in pairs produced the smallest number of LREs, the majority of which were lexical, and only a small proportion of them were resolved correctly. Given this finding, Leeson questioned the effectiveness of the dictogloss task with low-proficiency learners as they may struggle to understand the passage being read to them, and therefore cannot offer their contributions in discussing linguistic aspects with their peers.

Another study by Calzada and García Mayo (2020) brought further insights concerning the effectiveness of collaborative writing to drawing learners' attention to language. Analyzing the pair talk of 62 EFL low-proficiency learners assigned to work in pairs while completing a dictogloss task in which the targeted embedded form was the third person singular –s morpheme, while the main focus of the analysis was on the number and outcome of LREs. Calzada and García Mayo found that the dictogloss task proved effective in drawing learners' attention to language, though admittedly not very much on the targeted form (-s morpheme). Concerning the focus and resolution of LREs produced, the study found a higher proportion of form-focused and mechanics-focused than meaning-focused LREs. The study also revealed that learners paid more attention to other grammatical forms than to the –s morpheme targeted by the authors. The number of correctly resolved LREs was higher than the number of incorrectly resolved or unresolved LREs.

The number and resolution of LREs has shown to be different when learners engage in more meaning-focused collaborative tasks. In a study with 32 EFL learners (aged 16-17 years) where half of them were assigned to work in pairs and complete an argumentative essay as a collaborative writing task Villarreal and Gil-Sarratea (2019), for instance, found that pairs focused more on lexical and grammatical than mechanical issues. However, the difference in number of form-focused and meaning-focused LREs was not statistically significant. Similarly, in a more recent study with adolescent EFL learners (aged 12-13) completing a narrative essay in pairs and in small groups, Villarreal and Munarriz-Ibarrola (2021) failed to establish any significant differences between the meaning-focused, form-focused and mechanics-focused LREs produced from the task. Likewise, in a recent study with young EFL learners (aged 11) completing a picture description task individually and in pairs, Lázaro-Ibarrola and Hidalgo (2022) failed to obtain any significant differences between the three types of LREs (i.e. meaning-focused, form-focused, and mechanics-focused) produced from the task. These findings are similar to previous research with adult EFL learners. Fernández Dobao (2012), for example, examined the nature of oral interaction of learners working in pairs (n=30) and small groups (n=60) while completing a jigsaw collaborative task. Participants had to rearrange photos to create a story. They also had to set the story in the past using past tenses, hoping that this would trigger the generation of LREs. The analysis of the data showed that both learners working in pairs and small groups produced a similar number of meaning-focused and form-focused LREs, whereas the number of mechanical LREs was smaller for both groups of learners. In addition, the amount of unresolved LREs was similar for both pairs and small groups, but it differed significantly in favor of learners working in groups in terms of correct resolution of LREs.

Similar results were also reported by a larger-scale study with ESL students, whose main goal was to examine the written text composed by pairs and individuals. Analyzing the oral data of 12 pairs as they completed an argumentative essay collaboratively, Wigglesworth and Storch (2009), found that the largest proportion of LREs produced was meaning-focused followed by form-focused, and mechanics-focused which received the least attention, similar to the findings of Fernández Dobao, and Villarreal and Gil-Sarratea. Given the higher proportion of lexical LREs produced, the authors claimed that these findings could have been affected by the task type which

was a more meaning-focused rather than grammar-focused (e.g. dictogloss). Attempting to find out whether some meaning-focused tasks are more effective in drawing learners' attention to language than some others, in an earlier study with ESL advanced students completing a data commentary report and an argumentative essay, Storch and Wigglesworth (2007), for example, found that both tasks produced a similar number of LREs, most of which were meaning-focused, followed by form-focused and mechanics-focused which again received the least attention during peer interaction. Since this study was done with advanced ESL learners, the researchers claimed that learners' need to discuss grammatical accuracy may have been lower as compared with learners of intermediate or low language proficiency involved in other studies.

Thus, taken together, studies reported above showed somewhat mixed findings in terms of types of LREs generated in collaborative tasks. The overall findings suggest that in more grammar-focused tasks, the number of grammatical LREs tends to be higher compared to lexical and mechanical LREs. While in more meaning-focused tasks the number of lexical LREs tends to be higher than grammatical and mechanical LREs. Nevertheless, studies which used meaning-focused tasks (e.g. essay, jigsaw or picture story) failed to bring clear-cut evidence in terms of the type of LREs generated in these tasks. Finally, research with young and adolescent EFL learners is still limited especially with meaning-focused tasks (e.g. jigsaw or picture story). However, when it comes to examining the effect of collaboration on the written product outcomes, meaning-focused tasks seem to prevail over grammar-focused ones. This is discussed in what follows.

3.2 The effect of collaboration on the quality of written texts and learning gains

As mentioned earlier, the main focus of this part of the literature review will be on studies examining the outcome of the written texts composed individually, in pairs and small groups both in face-to-face (FTF) and computer-mediated context (CMC), with a special focus on the former. Next, studies on collaborative writing reporting learning gains from such an experience will follow. Then, a short overview of writing processes in CMC is presented. Last but not least, studies investigating learners' perceptions and attitudes towards collaborative writing experiences from both FTF and CMC modes will follow. Finally, this part ends with a summary of literature review.

3.2.1 Comparing the outcomes of collaboratively and individually written texts

Overall, findings of studies comparing texts produced by individuals versus pairs have shown that pairs tend to produce texts of better quality in terms of task-fulfillment, linguistic accuracy and complexity (Bueno-Alastuey & Lizarrondo Larumbe, 2017; Fernández Dobao, 2012; Shehadeh, 2011; Storch, 2005; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Villarreal & Munarriz-Ibarrola, 2021; Wigglesworth & Storch, 2009). In one of her seminal studies, Storch (2005) compared the written texts of 23 ESL students assigned to write a short composition based on a picture prompt. Students in this study had the possibility to choose between working alone or in pairs. 18 out of 23 learners chose to work in pairs and 5 others worked individually. Learners working in pairs were audio-recorded, and follow-up interviews were conducted. The study found that pairs produced shorter, but overall better texts in terms of task-fulfillment, grammatical accuracy and complexity than individual learners. However, even though pairs produced more grammatically accurate and complex texts than individuals, the difference was not statistically significant. Storch noted that the lack of significant difference may be linked to the small sample size of the study and short composition participants had to write, therefore, she suggested that future research include a larger sample size.

Two years later, using a similar design but with a larger sample, Storch and Wigglesworth (2007) compared the outcomes of the written texts produced by 72 ESL students of advanced proficiency who worked individually and in pairs. Forty-eight students completed a report and an argumentative essay collaboratively, while twenty-four others completed the tasks individually. In a subsequent study, Wigglesworth and Storch (2009) compared the performance of 48 pairs and 48 individuals working on an argumentative essay. Learners working in pairs were audio-recorded in both studies. Both these two studies found that pair learners produced significantly more accurate texts than individual learners. However, no significant differences were observed in terms of linguistic fluency and complexity between the texts produced by pairs and individual learners. The analysis of the pair talk revealed that pairs spent a considerable amount of time deliberating on lexical, grammatical and mechanical LREs and managed to resolve a majority of them correctly, which may explain the reason why texts produced by pair students were more accurate than those produced by individual learners. The studies reported above involved participants of advanced language

proficiency from ESL context, however, similar findings were reported in studies with intermediate participants from EFL settings as well.

In a more recent study, using a similar study design as Storch (2005), Villarreal and Gil-Sarratea (2019), for example, examined the effectiveness of collaborative writing for intermediate EFL secondary school students. They compared the performance of 32 intermediate EFL learners assigned to write an argumentative essay, in pairs (n=16) and individually (n=16). All the texts produced were analyzed in terms of quantitative measures of complexity, accuracy and fluency, and in terms of qualitative measures, assigning global rating scores for content, structure and organization. Learners working in pairs were audio-recorded, and the pair talk was coded for number, nature and outcomes of LREs produced. Similar to previous studies (e.g. Storch, 2005; Storch and Wigglesworth, 2007; Wigglesworth and Storch, 2009) pairs tended to produce shorter, but significantly more accurate texts than individual learners. Additionally, pairs produced texts with higher rating scores in content, structure and organization than individuals. Given these findings, the authors noted that collaboration was effective for secondary school students as it offered them opportunities to pool their linguistic resources, provide and receive feedback, scaffold each other while writing and thus produce texts of higher linguistic accuracy. Interestingly, in a more recent study with young EFL learners assigned to work individually and in pairs to complete a picture description story, Lázaro-Ibarrola and Hidalgo (2022) found that texts composed by pairs were better only in terms of linguistic accuracy but not in terms of fluency and complexity, and overall text quality measured through a holistic of task fulfillment, content and structure. According to the authors, one of the reasons why pairs failed to perform better than individuals in the other areas measured could be the very little time pairs spent revising their writing. This implies that when writing in pairs or individually learners need to pay a more balanced attention to all the writing stages, that is, planning, composing, and revising.

Since the majority of studies on collaborative writing have compared the written texts produced by learners working alone and pairs, another body of research has extended this research by including learners working in small groups. Thus building on previous research (e.g. Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009), Fernández Dobao (2012), for example, conducted a classroom-based study and compared the written texts produced by students of Spanish as a foreign

language working individually (n=21), in pairs (n=15) and in small groups (n=15), who had to write a story based on a set of photographs. Texts produced by all learners were analyzed for CAF measures. The study found that learners working in small groups produced the most accurate texts, followed by pairs and individuals. While texts produced by individuals were the longest, there was no significant difference for complexity between any groups of learners. Similar to previous research (e.g. Wigglesworth and Storch, 2009), the researcher noted that the analysis of oral interaction revealed that both learners working in pairs and small groups spent a considerable amount of time deliberating on language use and thus resolved correctly the highest proportion of LREs which may explain the reason why texts produced by pairs and small groups were the most accurate.

However, a more recent classroom based-study with lower EFL learners conducted by Bueno-Alastuey and Lizarrondo Larumbe (2017) which was partly a replication of Fernández Dobao's (2012) study reported somewhat contradicting results to what has been reported in previous research in terms of CAF measures (e.g. Fernández Dobao, 2012; Storch, 2005, Wigglesworth & Storch, 2009). This study compared EFL learners' (aged 12-13) performance of writing a story based on a visual prompt working individually (n=18), in dyads (n=10) and triads (n=7). All learners were allocated the same amount of time (25 min) to complete the task. Drawing on previous research (e.g. Fernández Dobao, 2012), all the texts produced by three groups of learners were measured for complexity, accuracy and fluency. The study found that texts produced by triads and dyads were statistically more accurate than those produced by learners working on their own. However, error-free T-units was the only measure to prove statistically significant for collaborating students. In addition, while triads outperformed individuals and pairs in lexical and mechanical accuracy, the study found no significant differences in grammatical accuracy between dyads and triads. The authors attributed this finding to the small number of participants involved in the study. Interestingly, while previous empirical research (e.g. Fernández Dobao, 2012; Storch, 2005; Storch and Wigglesworth, 2007) indicated that pairs overall produced shorter texts than individuals, Bueno-Alastuey and Lizarrondo Larumbe found that pairs and triads produced relatively longer texts than individual learners, even though the difference was not statistically significant. This finding was attributed to the number of participants working in pairs and small groups which may have contributed with more

ideas than working alone. Finally, no significant difference was observed for syntactic complexity between all texts produced by any of the groups. Last but not least, in a more recent study with adolescent EFL learners (aged 12-13), Villarreal and Munarriz-Ibarrola (2021) examined the effect of collaboration by comparing the writing performance of adolescent EFL learners assigned to work individually (49) , in pairs (12) and small groups (9) to complete a narrative writing task. Texts written individually were used as pre-test, those written in collaboration as an experimental task. Regardless of the writing condition, all the texts were rated for quantitative measures of fluency (number of words per text) and accuracy (counting all the grammatical, lexical and mechanical errors) as well as in terms of holistic rubrics of adequacy, coherence and cohesion, grammatical accuracy, lexical range and mechanics. Following the data analysis, the study found that pairs wrote the longest texts, while groups the shortest, however, no significant differences were observed between the three writing conditions. In terms of grammatical accuracy, results were quite balanced, though with a positive tendency for texts written in collaboration (groups). Interestingly, individually written texts contained the largest number of grammatical errors, and the lowest number of mechanical errors. Results from global measures revealed that texts written in pairs were better than those written individually and in small groups in all rubrics. However, only the difference for vocabulary was statistically significant favoring the texts written in groups. The slightly positive results for groups over pairs and individuals were attributed to the possibility of being able to pool more linguistic resources when working in groups than in pairs and individually (Fernández Dobao, 2012; Villarreal & Munarriz-Ibarrola, 2021). The lack of improvement in mechanics led the researchers to hypothesize that collaboration does not boost mechanics, perhaps because their participants were beginner learners of English that are possibly still developing their language awareness abilities as suggested by Muñoz (2014). Another interesting study with pre-adolescent EFL learners (aged 11-12) looked whether learners or their partner's proficiency has any impact on their written products measured quantitatively and qualitatively. Thus, Calzada and García Mayo (2021), for example, assigned 30 Spanish learners of English as foreign language to complete three dictogloss tasks two individually (T1 and T3) and one collaboratively (T2). Compositions (60) from T1 and T3 were assessed for quantitative measures of complexity and accuracy and for holistic measures of adequacy, coherence and cohesion, grammatical accuracy, lexical range and mechanics, similarly as in Villarreal and Munarriz-Ibarrola's (2021) study. In task

2 learners were assigned in mixed proficiency pairs (low-high). Following the data analysis, the study found that learners' proficiency affected their written outcomes in both testing times (T1, T3), that is, high proficiency learners obtained significantly higher scores in almost all rubrics measured. However, the same result did not hold when considering the effect of partner's proficiency (low-high). Thus, the impact of partner's proficiency (after collaborative dictogloss in task 2) seemed to impact only the measures of complexity by means of coordination and subordination. This result was attributed to the collaborative dialogue that emerged between learners in collaborative dictogloss task. Similarly, the possibility of impact from procedural task repetition was not absolutely ruled out by the researchers. Thus, the effect of learners' L2 proficiency turned out to be the most pronounced fact at both testing times (T1, T3).

While research reported thus far dealt with the effect of collaboration on the written text outcomes in face-to-face settings, another relatively small body of research has examined whether collaboration in computer-mediated context benefits the written products in terms of similar measures. One of the first studies that compared collaborative writing with individual writing in a computer-mediated context was the study by Elola and Oskoz (2010). In this small-scale study, nine learners of Spanish as a foreign language were assigned to complete two argumentative essays, one individually and one in pairs using wikis. Texts produced in the two writing conditions were assessed for CAF measures (e.g. complexity, accuracy, and fluency). The study found no significant differences between collaborative and individual writing in terms of complexity, accuracy and fluency. This result was attributed to the small number of participants included in the study. Similar findings were reported in a study by Strobl (2014) who compared texts produced individually with those written collaboratively in groups of three using Google docs. Both individual and collaborative texts produced by adult learners of German as a foreign language were assessed for complexity, accuracy and fluency, and rated for appropriate content selection and organization. The results revealed no significant differences between texts produced by groups and individuals in terms of complexity, accuracy and fluency. However, collaborative texts were significantly better in appropriate content selection and organization. The analysis of students' chats revealed that their talk was centered on organization, structure and content issues which may explain the significant results obtained for content selection and organization.

To sum up, results drawn from studies that involved adult EFL/ESL participants have shown that collaboration benefits learners' written texts in terms of accuracy, while findings for fluency and complexity are far from being conclusive, regardless of the working mode i.e. face-to-face or computer-mediated context. Similarly, the limited research with young and adolescent EFL learners has produced inconclusive findings concerning the effect of collaboration and/or learners' L2 proficiency on the written text outcomes. Therefore, the present findings stress the need for more research, in particular, with young and adolescent EFL learners to fully understand the benefits of collaboration and/or learners' L2 proficiency for written text outcomes and L2 learning in general. The issue whether collaborative writing benefits L2 learning is addressed in the next section.

3.2.2 Collaborative writing and L2 learning gains

Research on L2 collaborative writing investigating whether collaborative writing tasks benefit L2 learning is still scarce. While some research (Kim, 2008; Kuiken & Vedder, 2002; Nassaji & Tian, 2010) used a pre-and posttest design to examine the effect of collaborative over individual work on L2 learning gains, other research (e.g. Fernández Dobao, 2014; 2016; Swain and Lapkin, 1998) also examined the collaborative dialogue occurring in LREs and sought to find a link between LREs and posttest results. Interestingly, the studies reviewed report a mixture of results.

One of the earliest studies examining whether collaborative writing benefits L2 learning was the study by Swain and Lapkin (1998). This small-scale study conducted with 24 grade 8 French immersion students analyzed the oral data of only a single pair assigned to complete a jigsaw task collaboratively. One week after students completed the task, they were given a posttest which included linguistic items discussed in learners' LREs during the task completion process. The study found a positive relationship between the posttest scores and the LREs produced by pairs. Findings of this study were corroborated by Kim's (2008) study that compared the performance of 32 Korean L2 learners in L2 vocabulary learning as they completed a dictogloss task. Half of the participants completed the task individually and the other half in pairs. Learners working in pairs were audio-recorded, whereas those working alone were asked to think aloud while completing the task. To measure learners' vocabulary learning, Kim used the Vocabulary Knowledge Scale Test (VKS), a test which asked

learners to self-report their knowledge over the vocabulary items presented in the task and measured their knowledge using a five point scale. The study found that pairs performed significantly better in both vocabulary posttests. The analysis of oral data showed that pairs engaged in more LREs and resolved correctly a higher proportion of them compared to individual learners. This may explain the reason why pairs performed better in both tests, immediate and delayed. However, other studies reported different results.

Nassaji and Tian (2010), for example, failed to obtain clear-cut evidence whether collaborative work results in better vocabulary learning when compared with individual work. Nassaji and Tian conducted a classroom-based study with 26 ESL students and sought to examine whether completing a cloze task and an editing task - once individually and once collaboratively has any impact on the learning of sixteen English phrasal verbs. To compare the effect of task condition, they used a pre- and posttest design similar to Kim (2008), the Vocabulary Knowledge Scale Test (VKS). The study found that students were more successful at completing both tasks collaboratively than individually. However, the results of the pretests and posttests indicated no significant differences in terms of learning of the English phrasal verbs targeted between any groups of learners. Thus, even though collaborative work resulted in a greater knowledge of vocabulary items than individual work this difference was not statistically significant. The lack of the significant difference was attributed to the difficulty that ESL learners faced in learning English phrasal verbs.

Similarly, Kuiken and Vedder (2002) reported no effect of collaborative over individual work. Kuiken and Vedder conducted a study with 34 Dutch high school learners of English who were assigned to complete two dictogloss tasks. While twenty participants completed the tasks in small groups of three to four, fourteen others completed them alone. The dictogloss tasks contained three different types of passive structures, depending on the degree of complexity, entitled as type I, II, and III. Learners working in small groups were audio-recorded. To establish learners' prior knowledge on English passive structures, participants took a pre-test consisting of 32 sentences containing passive structures which participants had to underline. The pre-test consisted of eight passive structures from three different passive types (I, II, III). After the treatment, a post-test and a delayed posttest was administered. Both posttests were similar to the pre-test. The study found that all participants showed a better performance

in both the immediate and delayed tests compared to the pre-test results, however, no statistically significant difference was obtained in the performance of passive forms between learners working individually and those working in small groups. The oral interaction data showed that peer interaction enabled learners to demonstrate a simple and an elaborate noticing of passive forms. However, the authors concluded that interaction may lead to better noticing of passive structures but not necessarily to acquisition.

In a more recent study, employing a pre-and posttest design, and analyzing learners' collaborative dialogue, Fernández Dobao (2014) examined whether working in pairs versus in small groups benefits L2 vocabulary learning. In this study 110 learners of Spanish as a foreign language were assigned to complete a jigsaw task collaboratively. While 50 learners worked in pairs, 60 others worked in small groups. Both pairs and small groups of learners were audio-recorded, and their talk was coded for the number and resolution of lexical LREs produced. To assess students' learning, two testing instruments were designed, a vocabulary task which consisted of two similar versions of lexical items but arranged differently for the purpose of the task, and an individual writing task. The vocabulary task served both as pre-and posttest and was administered with all participants individually. The individual writing task served as an individualized posttest and was used to measure learners' ability to use those lexical items individually that they had discussed in the LREs with their peers while completing the task collaboratively. The study found that groups produced significantly a higher number of lexical LREs than pairs. Groups performed significantly better at solving lexical LREs than pairs. However, there was no significant difference between groups and pairs for vocabulary learning and retention of lexical items discussed in LREs. On the other hand, when the performance of individual learners was compared for the total number of words learned, the results were significantly different. Thus learners assigned to work in groups acquired significantly more new words than those assigned in pairs. Similarly, learners assigned in groups consolidated their prior knowledge of a higher number of words than those assigned in pairs, and thus reached a significant difference. Interestingly, the study found evidence that even learners acting as observers benefited almost equally from the LREs discussed by their peers who were actively engaged at solving them. Using a similar study design, but focusing only on silent learners assigned to work in groups, Fernández Dobao (2016) analyzed a subset of data from her previous

study (Fernández Dobao, 2014), and examined whether silent learners benefited equally from observing their peers who were actively engaged at solving LREs discussed collaboratively. The study found that silent learners or observers benefited almost equally from observing their peers who triggered or contributed to solve lexical LREs discussed in their groups. Given these findings, Fernández Dobao (2016) concluded that “Silent learners, although apparently passive, were in fact active participants in the interaction” (p. 46). Posttest results also indicated that silent learners benefited more or less equally in the acquisition of new knowledge or consolidation of existing knowledge as their peers who triggered or contributed to solve LREs discussed. The researcher noted that regardless of whether working in pairs or small groups, interaction among peers may be facilitative to learning even for silent learners. Findings reported above have been drawn from studies on collaborative writing relying on a single treatment, therefore, results from longitudinal studies might offer stronger evidence. This topic is addressed next.

3.2.3 Longitudinal studies on collaborative writing

Research employing longitudinal designs which examined potential benefits of collaborative writing in L2 learning and subsequent individual writing are rather limited. In fact, to date only four longitudinal studies (e.g. Chen, 2019; Sarkhosh & Najafi, 2020a; Shehadeh, 2011; Soleimani et al., 2017) have been found in the literature which investigated whether collaborative writing benefits L2 learning or writing. Additionally, it is interesting to note that these studies have been conducted with adult EFL participants, excluding Sarkhosh and Najafi’s study which involved adolescent EFL learners.

One of the first to investigate whether collaborative writing benefits L2 learning and writing was Shehadeh’s (2011) study. The study compared the performance of 38 EFL university students assigned to work in pairs (n=18) and individually (n=20). Participants in this study were assigned to write 12 pieces of writing, each piece about a paragraph’s length over a 16-week period. The study used a pre-and posttest design. Thus before the data collection, all participants wrote a piece of writing which served as a pre-test, and took a posttest after the treatment period. To measure the effect of collaborative writing, the study employed a global rating scale rubric that assessed the texts for content, organization, vocabulary, grammar and mechanics. It is interesting to

note that participants in the experimental group were instructed to change their working partners every two to three writing sessions. This was done in order to avoid situations where students working together for a period of 16-weeks could create dominant/dominant, or dominant/passive relationships which have been found to be not conducive to learning, as noted by Storch (2002). The study found that collaboration had a positive impact on students' L2 writing, that is, students working in pairs throughout the semester wrote significantly better texts in terms of content, organization and vocabulary than those writing individually. However, collaboration did not have any impact on grammatical accuracy and mechanics. The lack of significant differences for accuracy and mechanics was attributed to the low level of L2 proficiency of participants involved in this study.

In a more recent study, Chen (2019) examined whether collaborative writing affects students' individual writing. In this study with intermediate EFL university students, Chen compared the performance of EFL students assigned to work in pairs and individually on three essay writing tasks over a semester. Following previous research (e.g. Shehadeh, 2011; Wigglesworth and Storch, 2009) all texts produced by both groups of students were assessed for quantitative measures of complexity, accuracy and fluency, and qualitative measures using a global rating rubric that assessed students' writing for content, organization, vocabulary, grammar and mechanics. In addition, to compare students' performance, the study used a pre-and posttest design, specifically a pretest prior to the data collection, and two posttests, one immediate and one delayed after the data collection. The study found that learners working in pairs over a semester outperformed those working individually in an immediate test. Thus, learners working collaboratively over a semester did significantly better in their subsequent individual writing and produced better texts in terms of accuracy, fluency, and quality (organization, vocabulary and grammar) than those who worked individually over a semester. Delayed posttests results showed a similar picture, that is, learners working collaboratively over a semester performed significantly better in their individual writing tasks in terms of accuracy, fluency and quality (vocabulary, grammar), but not better in terms of organization than those who worked on their own. These findings contradict Shehadeh's (2011) study that found a significant effect of collaborative writing on content but not on grammar. Given these findings, Chen noted that this difference may be linked to the writing task and number of words participants

were required to write in her study. While participants in Shehadeh's study were asked to write a paragraph of about 100 words in length, participants in Chen's study were asked to write an argumentative essay of about 150 words in length. This made Chen conclude that participants in her study had more opportunities for negotiation of meaning and language practice and hence produced texts of higher grammatical accuracy. In line with previous research (e.g. Storch, 2005, Wigglesworth & Storch, 2009) this study demonstrated once again that collaborative writing benefits grammatical accuracy not only when students are assigned in one single writing task, but also in subsequent individual writing assignments.

Similarly, Soleimani et al. (2017) compared the performance of pairs and individuals in terms of quantitative measures of fluency, complexity and accuracy, and looked whether proficiency can have an impact on the written products of learners assigned to complete six writing pieces individually and collaboratively. The longitudinal study involved 36 pairs composed of advanced and intermediate learners and 36 advanced and intermediate learners working individually. Following the data analysis, the study found that pairs significantly outperformed individuals in terms of accuracy and fluency. Thus, texts written by pairs were better for accuracy and fluency than those produced by learners writing on their own. In addition, the study found that pairs composed of advanced learners produced not only better texts than pairs composed of intermediate learners, but also better than learners working individually. Thus, overall texts produced by pairs were significantly better than those produced by learners working alone for the measures of accuracy and fluency, but not for complexity. Similar to previous research (e.g. Wigglesworth & Storch, 2009), the study found no significant difference for complexity between texts produced by pairs and individuals. The lack of significant difference for complexity was attributed to the instructional procedures and materials (mainly centered on grammar and vocabulary items) used for teaching classes by the language institute where the study took place.

In a more recent study, Sarkhosh and Najafi (2020a) compared the performance of 120 lower EFL learners assigned to work in pairs ($n=30$) and individually ($n=60$) in seven successive writing compositions. To measure the potential effect of collaborative writing approach, the study used a pre-and posttest design. Before the data collection, all participants took a pretest individually and then two other posttests, one immediate and one delayed posttest. Drawing on Wigglesworth and Storch's (2009) study, the

written texts produced by both groups of learners were rated only for quantitative measures of accuracy and fluency, thus excluding complexity. Following the data analysis, the study found that pairs performed significantly better and reached higher results for both measures of accuracy and fluency. Hence, texts produced by pairs were better for accuracy and fluency than those produced by individuals. Similar to what has been reported in previous research (e.g. Fernández Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007) the study found that students writing in pairs produced texts of better accuracy than those writing individually. However, contrary to the previous studies (e.g. Fernández Dobao, 2012; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009), but in line with Soleimani et al. (2017), the current study found that texts produced by pairs were better than those produced by individuals in terms of fluency. This finding was attributed to the larger sample size of the current study compared with previous studies (e.g. Fernández Dobao, 2012; Storch, 2005; Wigglesworth & Storch, 2009), and to the number of writing compositions participants had to write in this study and in previous studies which examined the effect of collaboration only by a single treatment.

Studies examining the effect of collaboration on computer-mediated collaborative writing report somewhat similar findings. For example, Wang (2015) compared the writing of adult EFL students working in groups of four in computer-mediated and face-to-face contexts. Participants in this study were divided into two groups (the wiki group and non-wiki group) and collaborated over a period of 12-weeks to complete the writing assignments for a Business English Course aimed to enhance students' writing skills. Students assigned in the wiki group collaborated during the whole writing process using wiki pages, whereas those working in non-wiki groups had to meet weekly in face-to-face for collaboration and used a personal computer to write their assignments. Texts produced by both groups were assessed for quality in the following areas: format and layout, purpose and audience, organization, content and style, and structure and grammar. The study used a pre-and posttest design to measure students' progress. Following the data analysis, the study found that students assigned in wiki groups performed significantly better than students in non-wiki groups in all areas, except for the format and layout which showed no significant difference. Wiki revision histories showed that students provided feedback to one another and contributed with ideas on

content and organization, which may explain the reason why wiki groups performed better in the posttest.

Bikowski and Vithanage (2016) examined whether working collaboratively in several writing tasks in computer-mediated mode impacts positively on students' individual writing. In a study with 59 ESL students who were assigned to complete several writing tasks collaboratively in groups of three to four and individually using Google docs over a 14-week period. To measure the potential effect of web-based collaborative writing on individual writing the study employed a pre-and posttest design, which measured the writing quality using an analytic rubric that rated the texts for content, organization, academic style and grammar. Following the comparison of pre-and posttest scores, the study revealed that web-based collaborative writing had a positive impact on individual writing. Thus, students in the collaborative group performed significantly better on the post-test than those in the individual group. Given that the collaborative writing group scored lower in their pre-tests than the individual writing group, but higher in their posttests. The researchers concluded that computer-mediated collaborative writing may be highly beneficial to students of lower language proficiency.

In short, findings drawn from studies conducted in FTF context reported positive gains for writing in terms of quantitative measures of fluency and accuracy as well as in terms of holistic rubrics. Interestingly not only collaboration but also learners' L2 proficiency was found to affect positively learners' written products. Similarly, positive findings were reported in studies investigating collaborative writing in CMC context. However, research with young and adolescent EFL learners is rather limited (i.e. Sarkhosh & Najafi, 2020a). Thus, the need for more research with adolescent learners is obvious.

3.3 Writing processes in CMCW

This strand of research deals with investigations of the writing processes that L2 learners have undergone when engaged in computer-mediated collaborative writing. One of the first studies to examine the writing process in computer-mediated context was the study by Kessler, Bikowski, and Boggs (2012). In this study 38 adult ESL learners were assigned to work collaboratively and plan and report on a research project using Google docs. The researchers sought to examine learners' language focus as they

completed the task in groups of three to four members by tracing the history of revisions that learners made. The study found that learners made more language-related contributions than non-language-related contributions, and that they focused more on meaning than on form while completing the task. In addition, the majority of language changes learners made were accurate. A similar finding had been reported in Kessler's (2009) study. However, Kost's (2011) study with intermediate learners of German brought somewhat contrasting results. In this study with intermediate learners of German who were assigned to write one or two essays collaboratively using Wikis after having seen a movie and listened to a radio play, Kost found that students made more grammar-focused revisions than content-focused revisions. In addition, as in Kessler et al. (2012), despite learners' low-language proficiency, they were accurate in the majority of changes they made. A more recent study by Abrams (2016) brought similar results to what has been reported earlier in Kessler et al. (2012) and Kessler (2009). In a study conducted with 28 beginner learners of German as a second language assigned to write an essay collaboratively using Google docs, Abrams found that learners prioritized meaning over form during their collaborative work. In addition, participants in this study mostly focused on linguistic features related to subject-verb-agreement, word order, choice of propositions, and spelling. In spite of their low-language proficiency, they were accurate in 95% of the changes they made, aligning with previous research (Kost, 2011). What these studies have shown is that when writing in a CMC context, learners generally tend to focus on meaning rather than form, with some exceptions. Still, regardless of their focus during CMCW, learners were mostly accurate in their edits made to their shared written products which reflects a high rate of accuracy in terms of written corrective feedback provided. However, whether young and adolescent learners benefit from such revisions, it is still unknown because research with participants of this age, to the best of my knowledge, is still missing.

3.4 Learners' perceptions and attitudes towards collaborative writing

Learners' perceptions and beliefs towards collaborative writing have been explored in a number of studies (e.g. Calzada & García Mayo, 2020b, 2020c; Fernández Dobao & Blum, 2013; Lin & Maarof, 2013; Shehadeh, 2011; Storch, 2005; Veramuthu & Shah, 2020). On the whole, learners were positive towards collaborative writing experiences. One of the first to explore learners' beliefs and attitudes towards collaborative writing was Storch's (2005) study. Follow-up interviews with students working in pairs

revealed that the majority of students were positive about the collaborative writing experience, noting that collaboration offered multiple opportunities to benefit not only in terms of grammatical accuracy and vocabulary learning but also in pooling of ideas that were very useful in the co-construction process. Learners working on their own, in pairs and small groups were equally positive towards collaborative writing. Fernández Dobao and Blum (2013), for example, explored attitudes and perceptions of 55 learners of Spanish as a foreign language towards collaborative writing after completing a writing assignment in pairs and small groups. The study found that students were overall positive about collaborative writing in pairs and small groups. Students found the collaborative writing experience rewarding for L2 learning, specifically for vocabulary and grammar, noting that working in pairs and small groups enabled them to produce texts of better quality by drawing on multiple linguistic resources. Interestingly, collaborative writing seems to gain support even among child EFL learners. In a more recent study, Calzada and García Mayo (2020), for example, explored pre-adolescent EFL learners' attitudes towards a collaborative dictogloss task. In this study 32 lower-secondary EFL learners (aged 12-13) completed a dictogloss task collaboratively, in pairs and small groups, and an attitude questionnaire one week after. The study found that the majority of learners were positive about the collaborative dictogloss task. They also believed that they had benefited from the dictogloss task in terms of grammatical knowledge, which made the researchers suggest that the dictogloss task was perfectly suited with learners of this age. Similarly, Veramuthu and Shah (2020) explored secondary school learners' attitudes towards collaborative writing, and sought to establish whether collaborative writing gains support as a pedagogical approach among ESL learners of this age. The majority of participants were positive about their collaborative writing, perceiving it as a beneficial for their writing skills. Participants in Villarreal and Munarriz-Ibarrola (2021) were also positive about collaborative writing. They perceived collaboration as highly valuable for L2 learning, pooling linguistic resources and benefiting from each other's ideas and language knowledge.

Similarly studies that have explored learners' perceptions and attitudes towards computer-mediated collaborative writing (CMCW) were positive. Elola and Oskoz's (2010) study, mentioned earlier, found that students were positive towards computer-mediated collaborative writing, and believed that collaboration enabled them to improve their text' quality in terms of content, organization and structure. In the same way,

Bikowski and Vithanage (2016) found that learners working in groups were positive about their collaborative writing experience and believed that collaboration in Google docs helped them to organize their texts much more easily, focus on grammar issues, and learn from each other. Students working in groups in Kost (2011), and Li and Zhu (2013) were also positive about CMCW.

Nevertheless, a limited number of participants from studies on collaborative writing either in FTF or CMC mode expressed some reservations towards collaborative writing. Storch, for example, found that some students indicated lack of confidence in using the target language and concerns of possibly hurting others' feelings when criticizing. Lin and Maaroo (2013) also reported that a small number of participants hesitated to correct their partners' errors due to the fear of hurting their feelings. Storch noted that some students perceived writing as an individual activity rather than collaborative, which may explain the reason why a group of students were reluctant about their collaborative effort, perceiving collaborative writing as a complex activity throughout which they could not concentrate on task. Similarly, some participants were negative about collaborative writing in Blum and Fernández Dobao (2013) claiming that collaboration could not contribute to their grammatical or lexical knowledge because their group members were of the same language proficiency. Calzada and García Mayo (2020) also reported some negative comments of students concerning unequal contribution to the task completion. Some students expressed their concern over some group members who did not contribute to task completion. One student, for example, declared that he had to do all the work on his own, as his partner did not contribute to the task. This contradicts Fernández Dobao and Blum (2013) who found that working in groups benefits learners by pooling more resources to draw on than working in pairs. When participants in Calzada and García Mayo (2020) were asked about their preferences whether they would work in pairs or small groups in their next assignment, they preferred pair work over group work even though the difference was statistically not significant.

Similar concerns were expressed for collaborative writing in computer-mediated context. Thus, negative comments towards CMCW revolved around difficulties in joining ideas with other pair or group members, managing the time, and constraints for individualistic perspective (Elola & Oskoz, 2010), difficulties in using Google docs (Kessler et al., 2012), lack or unwillingness for equal contribution by all group members

(Arnold et al., 2012), lack of confidence to edit other members' contributions because of their own language ability, and the like (Lee, 2010).

To sum up, students' perceptions and attitudes towards collaborative writing in FTF and CMC were, on the whole, positive. Collaborative writing was perceived as beneficial in terms of sharing knowledge, ideas, and pooling linguistic resources to achieve better task performance. Nevertheless some students were negative towards collaborative writing, probably because they could not find ways for collaboration with their partners and thus had to complete the task on their own, failing to utilize the positive features that collaborative writing entails. All in all, the majority of studies reviewed above involved adult EFL/ESL participants from university contexts, while research with young and adolescent learners is still scarce.

3.5 Summary of literature review

Research covered in this literature review indicates that learners assigned to collaborative writing tasks engage in interaction during the whole writing process (Storch, 2013), interaction which is otherwise known as collaborative dialogue or languaging. Languaging has been shown to be a useful source for L2 learning (Swain, 2006). The most frequent unit of analysis used to analyze collaborative dialogue (languaging) during collaborative writing is Language-Related Episodes (LREs). Several contextual variables have been found to influence the quantity and quality of languaging operationalized as LREs. Among others, these variables include: the type of tasks learners work on (Alegría de la Colina & García Mayo, 2007), learners' L2 proficiency (Leeser, 2004), and the relationships learners form when co-constructing texts (Storch, 2002b; Wanatabe and Swain, 2007). In addition, studies on collaborative writing have mainly used meaning-focused composition tasks, such as picture stories, composition, data commentary report, and narrative essays with adult EFL/ESL participants (e.g. Storch, 2005; Fernández Dobao, 2012), whereas research using similar tasks with pre-adolescent or adolescent EFL learners is rather limited (e.g. Bueno-Alastuey & Lizarrondo Larumbe, 2017; Villarreal & Munarriz-Ibarrola, 2021).

Research on collaborative writing analyzing peer talk for LREs (e.g., Calzada & García Mayo, 2020a; Dobao, 2012; Lázaro-Ibarrola & Hidalgo, 2022; Villarreal & Gil-Sarratea, 2019; Villarreal & Munarriz-Ibarrola, 2021) has reported mixed results concerning learners' focus during task completion, this applies to young or adolescent

EFL learners in particular. Therefore, more research is needed to establish which type of tasks generate more form-focused, meaning-focused and mechanics-focused LREs with adolescent EFL learners, and whether learners' L2 proficiency affects the type, number and resolution of LREs produced. Furthermore, studies examining peer talk have also shown that learners engaged in collaborative tasks use either their mother tongue (L1) or their target language (L2). Research (e.g. Storch & Wigglesworth, 2003) has shown that learners use the L1 as a mediating tool enabling them to complete highly cognitively demanding tasks, such as talking about task content, organization and structure. On the other hand, learners use L2 to communicate about simple language issues that do not require a high command of the target language, such as reading, re-reading, and back channel (e.g. Zhang, 2019). Interestingly, learners' L2 proficiency has frequently proved to be a decisive factor among peers which language to use L1 or L2.

Collaborative writing as a product, on the other hand, is characterized by several other important trends including: learning gains and collaborative writing products. On the whole, collaborative writing tasks have shown to benefit L2 learning in several ways. First, learners engaged in collaborative writing benefit in terms of L2 vocabulary learning (Fernández Dobao, 2014). In fact, even silent learners observing their peers engaged in interaction to solve language-related issues during collaborative tasks benefit in terms of L2 vocabulary learning (Fernández Dobao, 2016). Second, early research employing experimental designs comparing the written texts produced by learners writing in pairs with those writing individually have shown, on the whole, that pairs tend to produce shorter but better texts in terms of quality, such as task fulfillment, and grammatical accuracy and complexity (e.g. Storch, 2005). However, later studies with young, adolescent and adult EFL/ESL learners (e.g. Fernández Dobao, 2012; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009) have produced mixed findings concerning the effect of collaboration on the written text outcomes. While studies with adult participants obtained higher scores for pairs over individuals in terms of accuracy, the same effect did not hold for fluency and complexity (e.g. Fernández Dobao, 2012; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009). In fact, when learners working individually and those working in pairs were allocated the same amount of time to complete the tasks, learners writing individually have been found to produce longer texts compared with pairs (e.g.

Fernández Dobao, 2012). Research with adolescent EFL learners, on the other hand, has produced inconclusive findings concerning both quantitative and qualitative measures (e.g. Bueno-Alastuey & Lizarrondo Larumbe, 2017; Villarreal & Munarriz-Ibarrola, 2021). While results for fluency, complexity and accuracy were better for pairs and groups over individuals, the differences were not statistically significant. The same applies for qualitative rubrics, where learners from the collaborative condition outperformed those from the individual condition, but the differences in scores obtained were statistically significant only in the area of vocabulary. Results obtained from studies examining the effect of collaboration with participants of different proficiency levels revealed that learners with higher language proficiency significantly outperformed their peers with lower proficiency in almost all rubrics of measurement (Calzada & García Mayo, 2021). However, this study did not include a control group to compare the writing performance of learners with same L2 proficiency from two writing conditions i.e. collaborative vs. individual, so to get a clear picture over the effect of collaboration and/or learners' proficiency, a limitation acknowledged by the researchers themselves. Additionally, the effect of collaboration may be better understood from studies employing longitudinal designs. Both Chen (2019), and Sarkhosh and Najafi (2020b) reported that EFL/ESL learners' writing improved in terms of fluency and accuracy, and overall text quality after a sustained and prolonged collaborative writing experience. However, results from studies with young and adolescent learners are scarce, excluding the one by Sarkhosh and Najafi which did not rate the texts employing both types of measures, that is, analytic and holistic rubrics.

To sum up, results from studies above revealed that learners' L2 proficiency affects the quantity and quality of language-related episodes. However, such findings come from studies assigning learners in mixed proficiency pairs, where the likelihood of forming non-collaborative patterns of interaction is very likely (Storch & Aldosari, 2013). As such, it is unclear as to how the performance of learners with different proficiency level (low-intermediate, intermediate, high-intermediate) assigned to work in matched proficiency pairs differs from one level of proficiency to another. That is, how does the performance of low-intermediate pairs compare with that of intermediate pairs in terms of number, type, and resolution of LREs? Similarly, findings from studies comparing the writing performance in terms of analytic and holistic rubrics are inconclusive, be it with young, adolescent or adult learners. Additionally, little is known

whether and to what extent learners' L2 proficiency affects the outcome of the written texts produced by adolescent EFL learners of similar L2 proficiency working in pairs and individually. Last but not least, to the best of my knowledge, no other study has compared the writing performance of adolescent or adult EFL learners across three levels of proficiency working under the same condition, i.e. individual or collaborative to establish the potential impact of proficiency on the written product outcomes.

Therefore building on previous research (e.g. Fernández Dobao, 2012; Villarreal & Munarriz-Ibarrola, 2021; Wigglesworth & Storch, 2009), the present study aimed to fill this gap by assigning adolescent EFL learners (aged 13-14) of three proficiency levels (low-intermediate, intermediate, high-intermediate) to work in pairs and individually to write a story based on a visual prompt (set of photographs), measuring their written products in terms of quantitative measures of fluency, complexity and accuracy, as well as in terms of holistic rubrics of task achievement, coherence and cohesion, grammar, vocabulary and mechanics. Additionally, the present study coded the peer talk for the type, number and outcome of LREs produced, and sought to establish learners' focus during task completion and whether learners' L2 proficiency affected the number, type and resolution of LREs produced.

Finally, although some studies conducted in computer-mediated context were covered in this literature review, they were only reported to provide the reader with an overview of research on collaborative writing in CMC. However, investigating the effect of collaborative writing in CMC was beyond the scope of the present study. Because knowing the lack of resources (e.g. desk computers) at school where the present study was conducted and the lack of prior experience of the researcher on collaborative writing in CMC, examining the effect of collaboration on written text outcomes in CMC (e.g. Google Docs, Wikis) would have been far more challenging than in FTF mode, both for the researcher and participants involved.

Chapter 4: Methodology

4.1 Introduction

This chapter begins with an overview of the context of the present study, it continues then with a description of English as a foreign language in the Kosovar context, zooming in on the teaching and learning of English a foreign language in pre-university education, the curricula, and the elaboration of the English syllabus for learners of grade 8 as well as a short description of the participants. Next, it provides a short description of the pilot study and the lessons learned from it. The chapter continues with a description of the writing prompt and then focuses on the data collection and data analysis procedures. Subsequently, it describes the researcher's role, the rating process and statistical analysis of the results. Finally the chapter discusses the ethical issues and ends with a short summary of the methodology used in the study.

4.2 Context of the present study

The present study was conducted in a public primary and lower-secondary school located in Pristina, the capital of Kosovo. The school is attended by about 2200 students from grade 1 to 9, and it is the second largest primary and lower-secondary school in Pristina. The average number of students per class is 32-35. Despite being considered as an “elite school” by the Kosovar community living in the city center or nearby, it is not a place with classrooms equipped with up-to-date technologies such as smartboards or interactive white boards, which would promote interactive activities and help teachers to engage students in a more collaborative work (Golonka et al., 2014). On the contrary, the school still remains more of a “traditional school” whose classrooms are furnished only with blackboards or chalkboards, where the teaching of subjects' content is mostly based on the course books, while the use of computers, overhead projectors and other technological devices for teaching purposes is rare due to the lack and/or insufficiency of such equipment at school. In addition, the school does not provide access to the internet neither for the teaching staff nor for its students. The IT lab is equipped with outdated computers, where the completion of educational projects is almost impossible, in particular projects that require the use of online tools such as Google Docs that enables students' collaboration. The school owns a library where pupils have access to hard copy books with a special focus on Albanian literature, while

books in English or other foreign languages are hardly found. Finally, the library is small and there is no reading space available for readers.

4.2.1 English as a foreign language in Kosovo

English is learnt as a foreign language (EFL) in Kosovo both at pre-university and university education level. As far as the pre-university education level is concerned, English is learned in three levels, in pre-school level, primary and lower secondary level (grades 1- 9) and higher secondary level (grades 10-12)(MEST, 2016a). In full compliance with the aims of education in the Republic of Kosovo, students completing the pre-university education level are expected to master a number of key competencies such as the competency in communication and expression, thinking, learning, life, work and environment-related areas as well as the competency in personal and citizenship level as set in the Curriculum Framework (CF) of Pre-University Education (MEST, 2016a). According to the same document, learners are expected to advance the competence of communication and expression in terms of writing and speaking, not only in the mother tongue and English but also in another foreign language (German or French) (MEST, 2016). Although English, German, and French are all learned as foreign languages in Kosovo, English has a special place compared to German or French because it is learned in all the levels of pre-university education and starts as early as the pre-school level. It has a higher number of classes per week, which is not the case for either German or French. German and French, on the other hand, are learned from grade 6 (i.e. the beginning of lower secondary school) with one class per week and they have a different status. Thus, while German is learned as a compulsory subject, French is an elective subject, and more importantly, it is not learned in all schools in Kosovo but only in some lower and higher secondary schools throughout the country, always subject to a competition with other electives excluding Language Gymnasiums where it is a compulsory subject.

Despite the fact that the development of communication and expression in L1 and L2 is given a special importance by the CF, literally this hardly translates into practice, in particular regarding the communication and expression in English due to many challenges both teachers and students face along the way of the pre-university education. Firstly, although there are many factors that impact on the quantity and quality of teaching and learning English in the pre-university education system in

Kosovo, one frequently noted challenge faced by EFL teachers in the pre-university level is the problem of large classes, especially in the majority of schools in Pristina. Sometimes this translates into fewer chances for all students to produce output in L2 which is of huge importance for the development of the target language (c.f. Chapter 2, section 2.1). Large classes also pose problems in terms of working with students, which is not so much the case with small classes. Nevertheless, assigning students to work in pairs or small groups benefits language learning even in large classes (Harmer, 2007). However, to what extent Kosovar EFL teachers use tasks that promote peer interaction in L2 classrooms, which potentially leads to the development of the communication skills, is unclear up to the present day due to the lack of empirical studies reporting on this issue. Secondly, although the majority of in-service teachers in Kosovo, in particular those employed in the last 10 years, meet the criterion of holding a master's degree in the respective field of education, as one of many formal criteria to get hired in the lower-secondary education level (MESTI, 2021), quite often the lack of adequate training for the teaching staff in general and the EFL teachers in the pre-university education in particular have been noted. Both the lack of adequate training and the overpopulated classrooms have been perceived as challenges that need to be further addressed in a report produced by Friedrich-Ebert-Stiftung (FES) analyzing the Kosovo's Education System (Aliu, 2019). Thirdly, another concern often raised by school community and non-profit organizations monitoring the development of the pre-university education in Kosovo is the lack of 'technological infrastructure', that is, the majority of schools in Kosovo are not fully equipped with functional IT, Biology, Chemistry, Physics, and language labs that would ease the work of teachers and students to achieve their educational goals in the respective level of education (Aliu, 2019). Furthermore, and more importantly almost all primary, lower and higher secondary schools in Kosovo do not have access to internet, which limits the opportunities for teachers and students to benefit from using open source materials to advance their knowledge in a certain subject. Thus, all of the above challenges in a way impede the quality of teaching and learning process in the pre-university education level in Kosovo.

Nevertheless, the Curriculum Framework (CF) foresees that teachers in the pre-university education in Kosovo should use the up-to-date teaching methodologies and continuously update themselves with new training on teaching methods and techniques so that they meet their teaching goals much more easily (MEST, 2016a). More

specifically, based on the Core Curriculum of Lower-Secondary Education (CC) teachers working with lower-secondary school students (grades 6-9) are expected to use a wide range of teaching approaches and methods such as student-centered teaching and learning approach, competency-based teaching and learning approach, integrated teaching and learning approach, and differentiated teaching and learning approach (MEST, 2016b) so that the learning goals of every student are met, ensuring that all students are given equal space and opportunities in the classroom. This implies that the teacher's role in the current classroom is more of a guide leading students to learn independently and together with their peers. For this to be done, and in particular for students learning English in lower-secondary level, language tasks and activities that promote the collaboration among students should be encouraged, therefore research that informs EFL teachers about the benefits of collaborative writing for L2 learning would greatly contribute to their efforts in meeting their teaching goals in the Kosovar context.

4.2.2 The English syllabus for grade eight learners

In principle, students from grade 3 to 9 have two English classes per week (MEST, 2016a). Thus, participants of the present study, as learners of grade eight, attended two compulsory English classes per week. The main goal of the English syllabus for grade 8 is set to offer learners an enjoyable experience while learning English for effective communication and cultural understanding. At the same time, the syllabus seeks to increase learners' linguistic knowledge and language proficiency by exposing them to various cultural and linguistic topics/materials so that they can develop their communicative ability (MEST, 2019).

There are different learning outcomes set to be reached at the end of one school year. Such outcomes mainly derive from four language concepts, such as literary and non-literary texts, figurative and non-figurative language, criticism, theory and history, and language system. Topics delivered based on these concepts are supposed to develop learners' linguistic abilities, cultural understanding, and effective communication skills either in spoken or written English. The majority of topics developed for the learning purposes are mainly concerned with enhancing learners' oral communication skills and enriching their vocabulary, so that they can express themselves freely in the target language (MEST, 2019). However, there are some learning outcomes in the syllabus for the written language as well. Learning outcomes expected to be reached for written

language include the production of simple and complex sentences, paragraphs, and written texts. Specifically, at the end of grade 8 learners should be able to:

- a) produce simple sentences either in present or past tense using correct grammatical, spelling and punctuation rules;
- b) produce a short description of a person's character using good grammatical and vocabulary structures;
- c) describe a picture, person, or situation;
- d) write simple topics related to social life, living conditions, school life and the environment where children spend most of their time;
- e) write simple stories in past simple tenses; summarize information and present it graphically before an audience;
- f) write a cause and effect paragraph with appropriate topic sentences, supporting ideas, coherence and cohesion, and appropriate conclusion
- g) write a story consisting of multiple paragraphs, with topic sentences, supporting ideas, good coherence and cohesion, and conclusion;
- h) demonstrate the ability to write, including writing as a process, brainstorming, drafting, revising and editing (MEST, 2019).

To create a friendly teaching and learning environment, opportunities for peer interaction and collaboration, the syllabus promotes the use of communicative language teaching, task-based language teaching and project-based language teaching. In order to reach the learning aims and outcomes, teachers should select teaching materials from course book(s) of pre/intermediate level. Additionally, teachers and instructors are highly encouraged to use supplementary resources and tools to create a friendly environment for students of different learning styles (MEST, 2019).

However, unable to make use of open source materials and free of charge platforms from the internet, foreign language teachers in Kosovo are restricted to use the course books provided by the Ministry of Education, Science, and Technology and use them as primary sources in their daily work with learners. As an attempt to increase the quality of the teaching and learning, the Ministry of Education strives to provide training opportunities for the professional development of the in-service teachers in the public sector. However, training targeting foreign language teachers is still limited. In particular, training aimed at providing useful guidance on how to teach writing, or on

how to use writing for creating L2 learning opportunities is scarce. Surprisingly, some research evidence also suggests that EFL teachers receive only a minimum amount of writing instruction during their formal education in Kosovo (Rusinovci, 2020). This calls for more attention to writing instruction both at pre-university and university level of education. At the same time, this calls for more research exploring forms of writing that benefit L2 learning in all levels of education in general and in the pre-university education in particular. Perhaps it is because of the lack of training and research evidence that foreign language teachers in primary, lower and higher secondary education heavily rely on the teacher's book as the main guide to meet their teaching goals more efficiently. Hence, research investigating the potential of writing for L2 learning, and more importantly tasks that promote peer interaction in the L2 classrooms could contribute to overcome the existing barriers of EFL teachers in the Kosovar context so that they become more productive and successful in reaching the learning outcomes set in the English syllabus for grade eight learners.

4.2.3 Teaching and learning material(s)

Although the English Syllabus mentioned above tends to promote the use of additional resources for teaching and learning purposes, it is the lack of access to the internet that often hinders EFL teachers to make use of open source materials in their teaching work. Additionally, the lack of desk computers or laptops in the majority of public schools in Kosovo may discourage teachers to explore the use of other options beyond the respective course book(s) in their daily work with students. Thus, operating in this 'resource limited' environment, for the participants of the present study, *Live Beat 3*, a course book developed by Kilbey et al. (2015) was used throughout the year. The content of the above course book is organized into a Welcome Unit for revision and nine Core Units, organized into lessons with an interesting pattern of organization, namely units 1, 3, 5, 7 and 9 are organized into lessons as follows:

- a) three language input lessons (a, b, and c)
- b) an Across Culture Lesson (d)
- c) a Writing skills lesson (e)
- d) a Language revision lesson (d).

On the other hand, units 2, 4, 6 and 8, follow this pattern of organization:

- a) three language input lessons (a, b, and c)
- b) a Real Life issues lesson (d)
- c) a Language revision lesson (e) and
- d) a Skills revision lesson (Freebairn, et al., 2015, p. 6–7)

The **input lessons** aim at providing learners with opportunities for learning and practising grammar, vocabulary and language use, while skills development lessons focus on providing learners with opportunities for listening, speaking, reading and writing practice and seek to enforce students' knowledge on grammar, vocabulary and language use. Interestingly, there are two types of skills development lessons, one being the **Real Life Issues**, and the other **Across Cultures** (Freebairn et al., 2015). Even though both types of skills development lessons cover all four language skills such as listening, speaking, reading and writing, there is a slight difference in terms of emphasis each lesson puts into main language skills. While in the **Real Life Issues** lesson a greater emphasis is placed on reading and tips on how to read a text more efficiently, in **Across Cultures** lesson a greater emphasis is put on listening and speaking, featuring good strategies aiming at developing students' listening skills and training students to speak with confidence. Apart from a focused intention at developing learners' oral communication skills as displayed in the majority of units and lessons, the course book dedicates a great deal of content on developing learners' writing skills by providing them with tips and writing practice so that they can produce different kinds of texts such as a short stories, multiple descriptions, postcards and emails. To be able to write these types of texts students are provided with model texts, writing tips, exercises, and clear instructions on writing assignments, along with useful support and guidance throughout the whole writing process (Freebairn et al., 2015). However, despite the availability of good course books and teaching materials, the writing skill seems to be neglected among primary and lower-secondary school teachers. This negligence may be due to many reasons, among others, (a) it may be due to the lack of information over the potential of writing for L2 learning, (b) teachers may feel overwhelmed with their heavy workload to correct many compositions at a time, (c) teachers may feel unsure as to how to assess writing or provide feedback on the written compositions (McDonough et al., 2018). Additionally, the limited research conducted in Kosovo to investigate the performance of lower-secondary EFL school students on the four language skills, i.e. listening, speaking, reading and writing in Kosovo shows that students reached the

lowest scores in writing compared to scores in listening, speaking and reading (Kasumi, 2015). Therefore, more rigorous research using writing tasks is needed with adolescent EFL learners in general and in the Kosovar context in particular so that FL teachers can be more widely informed about the potential and challenges of writing for second/foreign language acquisition.

4.3 Participants

Participants in the present study were learners of English as a foreign language from a public lower-secondary school in Pristina, the capital of Republic of Kosovo. There were 144 participants from grade 8, 78 (54%) males and 66 (46%) females aged between 13-14 years old, see table 4.1. These learners were considered to be of intermediate language proficiency, namely of B1 level as defined by Common European Framework of Reference (Council of Europe, 2001).

Their English language proficiency was not determined employing any means of standardized tests, but it was conceived as intermediate level based on the recommendations of the Ministry of Education, Science and Technology (MEST) to use teaching materials and course book(s) from the pre-intermediate level for learners of this age (MEST, 2019), which corresponds to B1 (intermediate) level according to Common European Framework of Reference mentioned above. The CEFR framework distinguishes six levels of language proficiency, arranged into three broad levels, that is, the basic, intermediate and advanced level (Council of Europe, 2001).

Thus based on this recommendation, the teaching material, including the course book was selected from B1 level. Additionally, the researcher was also the subject teacher of the participants and had taught them English from grade six to grade eight, that is, at least three school years by the time of the data collection. This helped him to further categorize participants into three different language proficiency levels such as participants of low-intermediate, intermediate and high-intermediate language level. This categorization derived from a regular assessment conducted during English classes within the school year 2019/20 by means of both formative and summative assessment. The assessment covered learners' performance on four language skills, including listening, speaking, reading and writing. This type of assessment revealed that learners included in the present study were of different language levels. Kosovo's pre-university education system uses a five point grading system (1-5), 1 being the lowest and 5 being

the highest grade. All grades assigned to the current participants approximately correspond to the intermediate language level, so that grade 3 corresponds to low-intermediate, grade 4 to intermediate, and grade 5 to high-intermediate.

All participants shared the Albanian language as their L1. In addition to English, which was their first foreign language studied at school from grade 3 (at the age of 9-10), participants also studied German and French as part of their school curriculum. They started studying German from grade 6 and French from grade 8 with one class per week each. Prior to the data collection, participants were assigned to either the collaborative (n=48) or the individual (n=48) condition. Table 4.1 displays more details about the participants of the present study.

Table 4.1 Participants

Pairs														
TN	P	M	F	LP	TN	P	M	F	LP	TN	P	M	F	LP
1	Dua Ber		2	HI	17	Alu Gash	2		I	33	Alt Kull	2		LI
2	Kel Mal		2	HI	18	Era Saf	1	1	I	34	Led Sal	2		LI
3	Ore Oru		2	HI	19	Ero Naz		2	I	35	Kan Vio		2	LI
4	Eri Kra		2	HI	20	Gla Haj	1	1	I	36	Erd Svi	2		LI
5	And Uki		2	HI	21	Son Ade		2	I	37	Sej Kqi	2		LI
6	Ble Jash		2	HI	22	And Meh	2		I	38	Ard Ari	2		LI
7	Eri Gjuk	2		HI	23	Ema Mus	1	1	I	39	Edu Kry	2		LI
8	Rio Ram		2	HI	24	Jet Çit	1	1	I	40	Aid Ism	2		LI
9	Ani Lim		2	HI	25	Bli Dak	2		I	41	Mer Hav	1	1	LI
10	Din Jus		2	HI	26	Jon Syl	2		I	42	Elt Rex	2		LI
11	Erj Has		2	HI	27	Yll Gas	2		I	43	Nar Gash	2		LI
12	Lim Byt		2	HI	28	Ald Spa		2	I	44	Mir Hox		2	LI
13	Fio Pit		2	HI	29	Arg Zeq	2		I	45	Hed Mur	2		LI
14	Ili Kla	2		HI	30	Ble Vlla	2		I	46	Hek Ram	2		LI
15	Edi Mra	2		HI	31	Sar Zek		2	I	47	Org Shab	2		LI
16	Jor Ali	2		HI	32	Hen Maz		2	I	48	Die Ibr	1	1	LI

56	Bun	1	HI	72	Beq	1	I	88	Laç	1	LI
57	Aze		1 HI	73	Mal		1 I	89	Ebll	1	LI
58	Goj		1 HI	74	Ade	1	I	90	Tru	1	LI
59	Lok		1 HI	75	Baj		1 I	91	JBll	1	LI
60	Sha	1	HI	76	Zog		1 I	92	Mal	1	LI
61	Beg		1 HI	77	Mus		1 I	93	Dem	1	LI
62	Bry		1 HI	78	Pac	1	I	94	Sel		1 LI
63	Deh		1 HI	79	Mur		1 I	95	Syl	1	LI
64	Gji	1	HI	80	Sal	1	I	96	Zab		1 LI
			1								1
16		6	0	16		7	9	16		1	5

*Note. LP = language proficiency; TN = text number; P = Pseudonyms (of learners).

4.4 Pilot study

Prior to data collection for the main study, I conducted one small-scale study to test the feasibility of data gathering instruments (the writing task and peer talk), and to determine how long each data collection procedure took. In order for the pilot study to be as representative as possible, I involved 60 participants from two categories of language proficiency (intermediate and high-intermediate). The data collection procedure for the pilot study lasted three days. The pilot study was conducted at the same school context and with participants that were later involved in the main study. From the pilot study 40 compositions and 10 pair talks were produced. As the aim of the pilot study was to gauge the feasibility and flow of data gathering instruments, data produced from the pilot study was neither analyzed nor included in the main study. Two major problems were identified from the pilot study which were better administered in the main study. One of the problems identified was related to the recording process of the pair talk. Like in the main study, participants in the pilot study were assigned to work in pairs ($n = 20$) and individually ($n = 20$) to complete a writing assignment. Participants working in pairs were instructed to use their mobile phones to record their talk. However, this did not work because some of the pairs chose a wrong approach to record their talk, and as a result their talk became useless for any further analysis. Some other pairs wrote their story first and then recorded a member of the pair reading their written text aloud. Some further pairs wrote only a portion of the story (only a set of three pictures) and then read that portion of the text aloud while recording their voice. Some other pairs recorded their talk properly as they were co-constructing the story but lost their recordings unintentionally. Apart from these misunderstandings, there were

some technical problems with pair talk as well. Among others, pairs using their mobile phones to record their talk were not able to properly handle the recording either because of positioning their mobile phones too far from themselves or because of the lack of experience recording themselves while engaged in writing tasks. It is for this reason that I bought audio recorders to use in the main study. The other problem identified during the pilot study had to do with writing collaboratively. Despite the fact that participants had previously experienced writing collaboratively, the lack of experience recording themselves while writing together might have confused them. Therefore, this issue was addressed by practising writing in pairs and recording the pair talk 2-3 times before the data collection for the main study. While some studies reported in the literature review have found that pairs took longer to complete the same writing task than individuals, I allocated the same amount of time (40 min) both to pairs and individuals to do the assignment because I wanted to test whether this works with adolescent EFL learners. From the notes taken while collecting the data, it was found that both pairs and individuals spent approximately 20-30 minutes to complete the task, therefore, in the main study both pairs and individuals were allocated a similar amount of time, i.e. 30 minutes to write their story.

4.5 The writing prompt

Participants in this study had to write a story based on a set of photographs (Appendix A). This is considered as meaning-focused task (e.g. Storch, 2013), and was expected to draw learners' attention to form and meaning while writing the story. Previous research that used similar writing prompts has shown that this type of task drew learners attention to form and meaning equally as learners were assigned to write a story individually, in pairs and small groups (e.g. Fernández Dobao, 2012). In fact, in the present study learners' main task was to write a story based on a set of pictures using the past simple tense. In total there were 18 pictures spread over two pages. A set of three pictures carried one single idea. This means that there were six content points to be mentioned and elaborated in the whole story. I got the visual prompt from the course book *Exploring English* developed by Harris and Rowe (1995). One of the main reasons why I used this visual prompt in the main study was that previous research on collaborative writing with children (e.g. Bueno-Alastuey & Lizarrondo Larumbe, 2017) and adults (e.g. Fernández Dobao, 2012) used visual prompts, therefore I wanted to find out how my study results' compare to theirs in terms of quantitative measures. Another reason for selecting the picture prompt was that along the way from grade 6-8, participants of the present study occasionally practised writing compositions based on set of pictures including the data collection in the pilot study, therefore writing a story based on a set of photographs in the main study would seem a natural writing task for participants. The story they had to write was about a woman called Anna, who worked in a bank. One day as she woke up to go to work, she experienced many bad things, she woke up late, got terrified from a strange dog, got stopped by a fire accident on her way to work, and finally she got fired from her boss for being late. As mentioned earlier, a set of three photographs carried one single idea. Consequently, students were supposed to compose the story by mentioning and elaborating each set of pictures until they were finished with all six content points and thus create a comprehensive text. Learners were also supposed to be clear in terms of language use, word choice, content organization, and coherence and cohesion. As described earlier, 96 out of 144 participants worked in pairs and 48 others individually. Both pairs and individuals used an identical visual prompt, but they were instructed to compose in different learner set up, that is, in pairs vs. individually. Regardless of the writing conditions i.e., collaborative vs. individual, participants were allocated the same amount of time (30 minutes) to complete the

assignment despite earlier claims that pairs needed more time to complete the same task than individuals (e.g. Fernández Dobao, 2012; Storch, 2005). Prior to the data collection, participants experienced writing in pairs and individually in the pilot study using a picture prompt and few other times using short writing compositions (essays, paragraphs). Although a picture prompt was used in the pilot and the main study, they were not identical because some earlier research has shown that exact task-repetition may impact on the written text outcomes and/or on the quantity, type and quality of LREs produced (see for example Amiryousefi, 2016; Hidalgo & Lázaro-Ibarrola, 2020; Payant & Reagan, 2018). Therefore, given that some of the participants included in the pilot study were also involved in the main study, I used a different writing prompt in the main study to avoid such a potential influence.

4.6 Instruments

For the present study two data gathering instruments were used: (a) 96 written compositions, and (b) 24 pair talks.

4.6.1 Compositions

There were 96 written compositions in total, 48 written collaboratively and 48 individually. These compositions were composed by learners of three different proficiency groupings, i.e. learners with low-intermediate, intermediate and high-intermediate proficiency level. The number of texts produced by each proficiency level was 32 ($3 \times 32 = 96$).

4.6.2 Pair talk

Learners' oral interaction constituted the second gathering instrument for this study. As elaborated in Section 4.5, in the main study, regardless of the working conditions (collaborative vs. individual), learners were asked to write a story based on a visual prompt. They received the same writing instruction and used an identical writing prompt. Nevertheless, pairs were instructed to work collaboratively and compose the story while individuals were asked to do the same but on their own. Additionally, learners working in pairs (48) were audio recorded whereas those working individually were not. As the analysis of learners' interaction was not the core issue under investigation, only the oral data of 24 pairs were analyzed for the present study.

4.7 Researcher's role

In this study, I played the role of a data collector, data analyzer, a rater and a rater trainer. As a data collector, I mainly did all the preparations before the data collection. This included informing participants about the planned study, delivering forms of consent, assigning participants to one of the writing conditions, simulating a writing prompt with collaborating students, practising with audio-recording devices, then collecting the data for both the pilot and the main study. As a data analyzer, I ran all the necessary data analyzing procedures, including the preparation of data, selection of the appropriate statistical tests, running the tests and obtaining and interpreting the results. As a rater and rater trainer, I rated all the written compositions employing the respective rubrics and the pair talk. Before the rating process started, I also trained two other raters that were involved in the rating of compositions and pair talks (See Section 4.10 for a full description). The raters were two fellow English teachers one with Bachelor's degree and one with Master's degree in applied linguistics. Although both raters worked in the rating process, the one with the Master's degree was involved in the rating of the compositions, while the other with the Bachelor's degree was involved in the rating of the pair talks.

Although in this study I was involved more as a researcher than a teacher, being a teacher and a researcher at the same time has some advantages. The established relationship with participants of the present study may have contributed positively in terms of the 'unbiased' behavior they exhibited on the day of data collection. Because it has been claimed that the presence of an unknown person (e.g. a researcher) in classroom-based research may influence the participants' behavior, that is, participants may behave differently when observed by a stranger (Mackey, 2012). Therefore, a good rapport between participants and the researcher (their teacher) seemed to have contributed for collaborative writing assignment and the recording of the pair talk to be well-received notwithstanding that neither of these activities were a common practice in the context at hand. Admittedly, asking participants to record themselves while writing collaboratively might have been perceived as weird, still the participants had been informed about this requirement in advance, and some of them had even practised recording themselves while writing in pairs during the pilot study. Despite this, to avoid any potential effect, participants were told that they were free to talk about anything be it task-related or not, and that only the segments of the pair talk that were relevant for

the purpose of the study would be closely analyzed. In this way they felt relieved as to what they could talk with each other while writing together. Another benefit of being a teacher and a researcher is that one knows better the relationship between learners. This helps the researcher to assign learners to work with peers of their preference so to maximize the learning opportunities for all learners, always making sure to not assign the 'noisy' learners to work together in order to reduce the potential of the off-task behavior that they might exhibit when conducting research in the classroom setting.

On the other hand, being a teacher and a researcher may have some drawbacks as well. Although in the present study I cannot speak of any disadvantages of playing both roles mentioned above, still there is a small number of 'lazy' learners who sometimes may find excuses or refuse to do a specific task or they may ask for more details even though such details are not necessary for the task completion. Hence, having a stranger asking those learners to do the task, could potentially contribute positively to such 'destructive' behavior of learners hesitating to perform the task. Similarly, being a teacher and a researcher demands careful planning, work and energy in order to play both roles successfully. Thus, in the present study I worked very hard to design a clear plan and implemented it rigorously to be able to collect all the data as desired. Despite this, the day of data collection gave me a feeling as if I were operating in an unknown territory but having devised a detailed plan on all research activities that were to be carried out, the data were collected successfully. Thus, despite potential similarities between language tasks implemented as part of regular classroom activities and research activities, still there is a distinction between them in terms of learner guidance, implementation, time allocation and rigor both types of activities demand. When acting as a researcher a careful planning of activities and strict implementation of such activities is a must in order to be able to control for all the variables under examination. This also implies designing a clear procedure of data collection to be followed, a point discussed in the coming section.

4.8 Data collection procedures

The data collection for the present study lasted about five weeks. In week 1, before the data collection process started, the researcher informed participants/learners about the research activities planned for the specific weeks coming ahead, and delivered to the participants a form of consent to be signed by their parents in order for them to be allowed to participate in the study. In week 2, participants brought back the form of

consent signed and agreed themselves to participate in the study. Following this, they were divided into pairs and individuals. Pairs were self-selected and could choose their working partner within their classrooms. Individuals were also given possibilities to work with any partner from their classrooms if they wished, but since the researcher informed them that some learners were going to work in pairs and some others individually, the number of volunteers to work alone was big enough so that an equal number of texts from both conditions could be produced. Both groups of learners, i.e. pairs and individuals were informed in detail over the plan and assignment they were going to complete. Learners who were designated to work in pairs were given clear instructions over the collaboration, and were told that they were going to be audio-recorded while completing their task. In week 3, both groups of learners completed one writing assignment based on a picture prompt as a means of practice for the present study. Students working in pairs were given audio-recorders and were instructed to use them under the guidance of their teacher/researcher before they were assigned to write the composition for this study. After the final trial, and after making sure that everything was ready for the real study, the data collection began with collaborative students. In week four, students working in pairs were given a writing prompt (Appendix C) which asked them to work together and build the story based on the set of photographs given. They were also given audio-recorders (one for each pair) to record their talk while writing and were advised to switch the voice-recorder properly so that the data is saved properly. Then in week five, students working individually were given a writing prompt (Appendix D) which was similar to the one provided to pairs, but which asked students to work on their own and write the story based on the set of photographs given (c.f. Section 4.4). In both working conditions, in each particular week, the researcher was present in each classroom and helped students resolve difficulties they faced in terms of understanding the instructions clearly and precisely.

The activities done in each week are summarized in table 4.2.

Table 4.2 Data collection schedule

Week	Activities	
Week 1	a. Informed participants b. Delivered a form of consent to be signed by their parents	
Week 2	a. Received the signed form of consent b. Assigned participants to pairs and individuals	
Week 3	Pairs	Individuals
Week 3	a. practised working collaboratively	a. practised a similar writing activity
	b. practised working with audio-recorder	
Week 4	a. collected data from pairs – pairs completed the assignment	
	b. collected oral data – pairs recorded themselves while writing	
Week 5		a. collected data from individual students – students completed the assignment

4.9 Data analysis

This section describes how the data gathered using both instruments (i.e. written compositions and pair talk) were prepared, coded and analyzed.

4.9.1 Analysis of written compositions

To facilitate the analysis procedure, students' hand-written compositions were converted into Microsoft Word documents by a research assistant. To ensure the authenticity of hand-written compositions, no changes in formatting, spelling, punctuation, capitalization or language errors were made during the conversion process. Building on previous research (e.g. Calzada & García Mayo, 2021; Chen, 2019; Fernández Dobao, 2012; McDonough et al., 2018; Storch, 2005; Villarreal & Gil-Sarratea, 2019; Villarreal & Munarriz-Ibarrola, 2021; Wigglesworth & Storch, 2009) all the written texts were measured in terms of quantitative measures of fluency, accuracy and complexity. Additionally, the written texts were also rated for holistic rubrics that assessed them in terms of task-achievement, coherence and cohesion, grammar, vocabulary and mechanics. To examine the written texts through analytic and holistic rubrics specific measures were employed. These measures are discussed next.

4.9.1.1 Quantitative analysis

Firstly, similar to previous studies (e.g. Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009) the written texts were analyzed in terms of quantitative measures of fluency, accuracy, and complexity.

Fluency was measured in terms of the average number of words, T-units, and clauses per text. Complexity was measured in terms of proportion of clauses to T-units, noted as a reliable measure which correlates well with other complexity measures (Foster & Skehan, 1996), and the percentage of dependent clauses to total clauses, which measures the degree of embeddedness in a text (Wolfe-Quintero, Inagaki, & Kim, 1998). Accuracy was measured in terms of number of error-free T-units to total T-units (EFT/T) and number of error-free clauses to total clauses (EFC/C) (Fernández Dobao, 2012; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009). The global measures of accuracy were employed instead of local measures of accuracy following Ellis and Barkhuizen's claim (2005) for a close correlation between these two measures, and following Foster and Skehan's (1999) argument that global measures of accuracy (error-free T-units and error-free clauses) represent a more realistic measure of accuracy. To measure accuracy, the written texts were assessed for grammatical, lexical, and mechanical errors. Grammatical errors included errors in syntax (e.g. word order, missing elements) and errors in morphology (e.g. word forms, use of tense, use of articles, prepositions, subject-verb agreements). Lexical errors comprised errors in word choice, and were marked as errors only when the word used obscured meaning. Mechanical errors included errors in spelling and capitalization only.

Even though previous research (e.g. Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009) ignored the inclusion of mechanical errors when measuring accuracy, Fernández Dobao (2012) recommended the inclusion of mechanical errors in the analysis of accuracy as she found instances of learners' talk centered on spelling and punctuation issues in the oral data of her participants. Later studies confirmed her findings providing further evidence of learners' LREs discussing spelling and punctuation in EFL contexts (e.g. Villarreal & Gil-Sarratea, 2019; Zhang & Crawford, 2021). Therefore drawing on this body of research, in the present study, along with grammatical and lexical errors, mechanical errors were tallied as well.

To operationalize quantitative measures of fluency, accuracy and complexity, and to ensure a proper analysis of the written texts, this is the procedure that was followed:

- a. All hand-written scripts were converted into word documents (electronic texts), and the number of words was counted using the word count function within Microsoft Word.
- b. T-units were first identified and marked for their boundary with (//) within each text.
- c. Dependent and independent clauses were identified
- d. T-units and clauses (dependent and independent) were then coded for accuracy and counted, distinguishing between error-free T-units of total T-units and error-free clauses of total clauses.

As explained by (Hunt, 1966, p. 735) a T-unit is defined as “one main clause plus whatever subordinate clauses happen to be attached to or embedded within it”. As noted above, clauses were categorized as dependent and independent. A dependent clause is defined as a clause which cannot stand on its own but “must be used with another clause to form a complete grammatical construction” (Richards & Schmidt, 2010, p. 62). On the other hand, an independent clause is a clause which does not depend on another clause to form a complete grammatical meaning, even though it may be linked to other independent or dependent clauses by means of clause linkers (Richards & Schmidt, 2010). In the present study, following Foster, Tonkyn and Wigglesworth’s (2000) definition, a dependent clause is one which consists of a finite or a non-finite verb form, plus at least one additional clause elements of the following: subject, object, compliment or adverbial.

Table 4.3 below provides a short description of coding T-units and clauses.

Table 4.3 Coding T-units and clauses

T-units and clauses	Description
This morning Ana woke up at seven o'clock.	1 T-unit, error-free. 1 Independent clause, error-free.
She was taking the money/ that she dropped from the bag//.	1 T-unit composed of two clauses: 1 independent and 1 dependent. This T-unit is error free, since both clauses composing the T-unit are error-free clauses.
A few moments later the man leaved/because he was busy. //	1 T-unit composed of two clauses: 1 independent and 1 dependent. This T-unit is not error-free, because of the verb leave. However, the dependent clause is error-free. Thus the T-unit contains two clauses, one error-free and one not error-free.
She went out of apartmen with a bag an her right hand.	An independent clause containing two errors (<i>apartmen</i> and <i>an</i>), therefore this clause is not error-free. This clause is composed of 1 T-unit, which is not error-free.
She left her house, / then she dropped her bag/because a dog scared her. //	2 T-units composed of three clauses, 2 independent and 1 dependent. Both T-units and three clauses are error-free.
She got there at 9.30, / and her boss was furius,/ saying that she was late.	2 T-units, one error-free and one not error free. Three clauses, two independent (one error-free and one not error-free, because of the misspelled word 'furius') and a dependent clause which is error-free.
When she arrived in the city bank, / the boss yelled at her/ because she was too late. //	1 T-unit, composed of two clauses 1 independent and 1 dependent. The T-unit is error free. There is another dependent clause at the initial position, which is not error-free because of the preposition 'in' used instead of at. In total there are three clauses, 2 dependent 1 independent.

4.9.1.2 Qualitative analysis

The majority of studies comparing the writing performance of learners writing in pairs and individually have used analytic rubrics (e.g. error-free T-units and error-free clauses) to measure the learners' product outcomes (e.g. Fernández Dobao, 2012; Wigglesworth & Storch, 2009). However, more recent research has questioned the ecological validity of relying exclusively on such measures (e.g. McDonough & Fuentes, 2015; Villarreal & Gil-Sarratea, 2019). Hence, employing holistic rating rubrics to assess the writing performance is seen as paramount because such rubrics are not only common in secondary school contexts, but they are also aligned with current views on education where formative assessment is an established practice (Ball et al., 2015). Thus drawing on these claims, in the present study, in addition to analytic measures of accuracy, fluency and complexity, I used a holistic rubric that was developed by Gassner et al. (2019) to measure the writing performance of Austrian secondary school EFL learners. Originally, the rubric contained four areas of measurement, namely task achievement, coherence and cohesion, grammar and vocabulary, but for the purpose of the present study, I also added mechanics (Appendix B). The rubric can be used to assess short (40-70 words) and long (120-180 words) writing prompts. It uses two different sets of criteria depending on the length of text being analyzed. In the case of the present study, I used the long prompt criteria since the average length of texts produced by my participants was around 150-169 respectively. Here is a short overview of what these areas take into account.

Task achievement assesses the written texts for the number of content points whether they have been clearly mentioned and elaborated, i.e. the degree of elaboration, poor vs. great. Coherence and cohesion considers the logical organization of ideas and arguments within a text. It assesses how well learners structure and arrange their ideas and arguments expressed by words, phrases and clauses within their writing topic. Grammar takes into account the range of grammatical structures and the accuracy of language used to express ideas and arguments in the written texts. Vocabulary assesses the written texts in terms of the range of lexical items used to express ideas adequately and meaningfully. Mechanics in the present study considered only spelling and capitalization rules. Finally, in all these areas the written texts were assessed using band scores, the highest being 7 and the lowest being 1. In cases where there was not enough

to rate, for example, texts shorter than 60 words were assigned with a band score of 0 in all areas.

4.9.2 Analysis of pair talk

As previously stated, many studies have analyzed learner-learner interaction to establish what goes on when learners are assigned to write tasks collaboratively. The majority of research analyzing pair talk has employed Language-related episodes (LRE) as the unit of analysis (e.g. Fernández Dobao, 2012, 2014; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009). In addition, the majority of studies have coded LREs for their focus, i.e. lexis-focused (L-LRE), form-focused (F-LRE) and mechanics-focused (M-LRE), as well as for their outcome such as correctly resolved, incorrectly resolved, and unresolved LREs. Drawing on this body of research, the present study analyzed the pair talk for LREs and identified and categorized LREs according to their focus such as lexis-focused (L-LRE), form-focused (F-LRE), and mechanics-focused (M-LRE), as well as distinguished between LREs that were correctly resolved, incorrectly resolved and left unresolved.

The table 4.4 below illustrates how the pair talk was coded for the types and outcomes of LREs.

Table 4.4 Coding LREs for their types and outcomes

Coding Scheme		LRE			Outcome		
Study	L	F	M	√	X	?	
Storch & Aldosari, 2013	Searching for words, choosing lexical alternatives, prepositions	Grammatical aspect, use of tense, articles,	Spelling, pronunciation, punctuation.	√	X	?	
The present study	Searching for words, phrases, choosing lexical alternatives, discussing meaning of propositions	Morphology & syntax (word forms, use of tense, articles, word order, etc.)	Spelling, and capitalizations.	√	X	?	
L= lexical; F= grammatical; M= mechanical. Outcome: √ = correct, X = incorrect, ? = unresolved.							

In a first round of coding, LREs were coded according to their types.

4.9.2.1 Lexis-focused LREs

Drawing on previous research, lexis-focused LREs in this study have been identified as instances of learners talk about word choice, including prepositions, word meanings, and lexical alternatives (Leeser, 2004; Storch & Aldosari, 2013). In the case of prepositions, like in Storch and Aldosari (2013), when deliberating over prepositions, learners in this study seemed to be focused more on meaning than on form of prepositions. Hence, any talk related to prepositions was included under lexis-focused LREs.

Example 1 illustrates two learners talking about the meaning of the word scolded in turn 57, 59, 62 and 63, thus solving their dilemma whether to use the word scolded or scared.

Example 1: A lexis-focused LRE

- 57. Lim: a man passed by and scolded, um.... he scared or scolded it looks like.
- 58. Byt: yeah
- 59. Lim: scolded....
- 60. Byt: he got rid of the dog ahahahaha
- 61. Lim: no scolded, scolded eheh
- 62. Byt: scolded
- 63. Lim: scolded, I ka bertit [yelled at the dog], scolded, he scolded.

In example 2, learners engage in a long discussion about a word choice, that is, whether to use the adjective *fortunately* or *luckily*. As they talk about this choice, they engage in meta-talk, turns 12-17, as a way of convincing one another for their proposal. They also give reasons and justifications over the word they are insisting should be used (turns 12, 17 and 20).

Example 2: A lexis-focused LRE

- 1. Lim: um a fortunately a man was there
- 2. Byt: unfortunately?
- 3. Lim: fortunately a man was there
- 4. Byt: luckily...
- 5. Lim: fortunately
- 6. Byt: luckily...
- 7. Lim: fortunately
- 8. Byt: luckily....
- 9. Lim: luckily is like
- 10. Byt: yeah a man was ...
- 11. Lim: fortunately a man was there so,

12. Byt: no if you say like that it's like a bad thing
13. Lim: fortunately, I am not saying unfortunately, I'm saying fortunately
14. Byt: but it's better luckily
15. Lim: no!
16. Byt: yes
17. Lim: no I think it's better, it's sounds better if we say fortunately
18. Byt: luckily
19. Lim: fortunately a man was passing so he kinda cool it the dog... are we gona use fortunately or luckily, I think we should use fortunately
20. Byt: I think we should use luckily
21. Lim: we use, at the first word we use your um...proposition jo...not proposition
22. Byt: proposal
23. Lim: ehhheheheh s'po di mi thane [I'm not able to say it right]
24. Byt: opinion
25. Lim: at the first word we use your opinion, I think we should use my opinion right now. Ok, fortunately.

4.9.2.2 Form-focused LREs

Drawing on previous research, form-focused LREs in the present study were identified as instances of learners' talk on morphology (e.g. word forms, use of tenses, gender agreement issues, use of articles, etc.) and syntax (length of sentences, word order) (e.g. Fernández Dobao , 2012; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019).

Example 3 and 4 below illustrate two form-focused LREs.

Learners in example 3 deliberate over word order in turns 2 and 3, and on the correct use of tense (past or past continues) in turns 3-6 as a way of organizing and expressing their ideas.

Example 3: A form-focused LRE

1. Jus: while she was packing her things a man was...
2. Din: pat his dog a qysh? [Pat his dog, or what?]
3. Jus: jo a man was yelling at his dog [no, a man was yelling at his dog]
4. Din: yes a man yelled at his dog
5. Jus: was yelling at his dog...
6. Din: was yelling at his dog.

Similarly in example 4, learners talk about word order, discussing ways of organizing and expressing their ideas using an appropriate form of the verb *leave* in turns 5-11. They also give reasons and justifications over their proposal as an effort to convince their partner on a particular issue.

Example 4: A form-focused LRE

1. Din: How are we gonna write about the bus leaving?
2. Jus: attacked her...
3. Din: then Ana realized that the bus...
4. Jus: bus....
5. Din: leaving
6. Jus: no, was about to leave
7. Din: jo se [no because] the bus was already gone, because she.....no
8. Jus: was already leaving then she ran to the bus but she couldn't stop him so she waited at the bus stop.
9. Din: jo [no], but then ana realized the the bus was leaving
10. Jus: was about to leave.
11. Din: was about to leave.

4.9.2.3 Mechanics-focused LREs

Following previous research (e.g. Calzada & García Mayo, 2020a; Fernández Dobao, 2012; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009), in the present study mechanics-focused LREs were identified as instances of learners' talk about spelling, punctuation and capitalization issues.

In example 5, learners deliberate on the spelling of the word *pounds* in turns 138-139, whereas turns 140-141 represent a clarification request, ending with a punctuation agreement (turns 142-143).

Example 5: A mechanics-focused LRE

- 136 Rio: Ana as a kind person
- 137 Ram: kind person..
- 138 Rio: gave him five pounds, qysh u shkrujke pounds? [How do we spell pounds?]
- 139 Ram: p-o-u-n-d-s
- 140 Rio: a je e sigurt? [Are you sure?]
- 141 Ram: po [yes]
- 142 Rio: pike. [Full stop.]
- 143 Ram: pike. [Full stop.]

In example 6, learners talk about a punctuation issue, whether to use a full stop or a comma (turns 40-42), while the turn 40 represents a clarification request.

Example 6: A mechanics-focused LRE

38. Ram: While Ana was walking in her way to the bus station
39. Rio: to the bus station...
40. Ram: presje a me ba a me ba pike? [Shall I put a comma or a full stop?]
41. Rio: presje, [Comma]
- 42 Ram: presje. [Comma]

4.9.2.4 Resolution of LREs

Following previous research (e.g. Fernández Dobao, 2012 ; Kim & McDonough, 2008; Leeson, 2004; Mozaffari, 2017; Villarreal & Gil-Sarratea, 2019) LREs in the present study have also been coded for their outcome, such as correctly resolved, incorrectly resolved and unresolved LREs.

Thus examples 7, 8 and 9 illustrate whether learners working together reached a correct or incorrect resolution or left unresolved the language-related concerns they raised.

In example 7, learners discuss the spelling of the word *catch* (turns 1- 6). In this mechanics-focused LRE with the help of one another learners reached a correct resolution.

Example 7: A correctly resolved LRE

1. Sal: Qysh shkruhet *catch*? [How do we spell *catch*?]
2. Led: c-a-t- ch.
3. Sal: C-A ...
4. Led: T-CH
5. Sal: Prit CA.. [wait]
6. Led: T-CH,
7. Sal: Catch. Didn't catch it.

In example 8, learners raise a tense issue, namely they try to formulate a sentence using the past simple tense. They have a dilemma whether to use the past simple or past continuous tense. They end up using the past progressive incorrectly - without the auxiliary verb which would be wrong even if used with the auxiliary verb, because the context required them to use the past simple tense. In this form-focused LRE, learners reached an incorrect resolution.

Example 8: An incorrectly resolved LRE

8. Hek: E ka prit apert autobusin. And she go again at the bus stop [She waited again for the bus]
9. Ram: Jo, jo, kadal vec pak, kadal, fol qysh ke bre edhe nje here... and she waiting again [No, no, hold on, hold on, tell me how did you say again...]
10. Hek: waiting again

In example 9, learners seek out a word which would carry the meaning of scolding the dog. But because of their low language competence, they decided to skip it and said something else instead. In this meaning-focused LRE, learners were not able to reach either a correct or incorrect resolution, therefore they left it unresolved. This is an example of an unresolved lexis-focused LRE.

Example 9: An unresolved LRE

1. Hav: The boss.... Qysh osht me qortu bre? [How can we say to scold?]
2. Mer: uuu.. jo s'di. [uuu, I don't know]
3. Hav: qysh o? um.....? a e lajme krejt, e thojme boss is angry and she's late? [How? Um..? Let's skip it and put it like this, the boss is angry and she's late?]
4. Mer: ani. [OK]

4.10 The rating process and interrater reliability

Apart from other roles, the researcher in the present study served as a rater and a rater trainer. Thus, to ensure interrater reliability, the compositions and pair talks were independently analyzed by the researcher and two other raters who were involved in the coding of different types of data, i.e. compositions and pair talks (c.f. Section 4.7). Before the rating process started, both raters were trained by the researcher to code the data according to the procedures described in Sections 4.9.1 – 4.9.2.4. More specifically, the training sessions aimed at informing, instructing and practically applying the rating of the texts and coding of the pair talk. First, the researcher trained the rater who was involved in the assessment of the written texts employing both measuring rubrics i.e. quantitative and qualitative and then the rater who was involved in the coding of the pair talk. Therefore, the analysis of the written compositions is discussed first.

As for the qualitative analysis a holistic rubric (Appendix B) which had different score descriptors was used, the job of the examiner was to assign a score in each of the areas assessed, namely in task achievement, coherence and cohesion, grammar, vocabulary and mechanics. For this assessment, there was a discussion between the researcher and the second rater which included clarifications about areas being assessed, scores to be assigned, and which texts (i.e. short texts) are excluded from the assessment or are assessed with a band score of zero. After the discussion, six samples of texts (two from each level of proficiency) were rated together with the researcher to ensure a full

understanding of the procedures and to discuss any issues emerging from the practical work of examining the compositions.

For the quantitative measures, on the other hand, the training process was somewhat similar but more extensive. That is, both the researcher and the second rater discussed the analytic measures of fluency, complexity and accuracy to examine the writing performance of learners. During this discussion, the researcher informed the second rater about the purpose of the assessment and discussed about the practical steps that the quantitative analysis would undergo such as identifying the T-units, dependent and independent clauses, identifying the linguistic errors across three categories (i.e. grammar, lexis and mechanics) and marking the boundaries of T-units and clauses. After this phase, the researcher and the second rater coded six written texts for the fluency, complexity and accuracy measures to ensure a full understanding of these measures, practical steps and procedures employed to assess the written products and discussed and resolved any issues raised during the practical work of coding.

Finally, following the training sessions, both raters were given 48 randomly selected written texts (24 from collaborative and 24 from individual conditions), which made up 50% of the whole dataset, and started the rating process using the holistic rubric first and then the analytic measures of fluency, complexity and accuracy.

After the rating process ended, the ratings of both raters (the researcher and the second rater) were compared, and the interrater reliability using Cohen's Kappa procedure, a practice employed to check interrater reliability in L2 writing, was 0.87 for task achievement, 0.84 for coherence and cohesion, 0.80 for grammar, 0.90 for vocabulary, and 0.82 for mechanics, thus reaching a high level of reliability. Similarly, using the percent of agreement method to check for the coding reliability, the interrater agreement for T-unit identification was 88%, for clause identification was 85% and for linguistic error identification was 88%. Regardless of the coding procedures, any disagreements in coding between the raters were discussed and fully resolved after inspecting each case individually. The remaining data were distributed and independently coded by the two raters, while any issues raised during the rating process were discussed, resolved and included in the dataset.

Finally, the researcher trained the second rater involved in the coding of the pair talk. The training session involved discussions about the coding procedure,

identification of language-related episodes (LREs), distinguishing between LRE types and resolutions. As a practical step, the researcher and the second rater coded three pair talks together and discussed issues raised during the coding process such as whether to include laughs, pauses, and off-task behavior and any other issues related to the LRE types and outcomes. After ensuring that the coding process was fully understood, the researcher and the second rater independently coded 6 pair talks, which made up 25% of the oral data subset. Following the rating process, the interrater reliability was checked using the percent of agreement method, yielding an agreement of 83% for the identification of LREs, 78% for the focus LREs, and 92% for the outcome LREs. All disagreements between the coders were discussed and fully resolved one-by-one. The remaining data were spread among the raters and coded independently. All issues raised during the coding process were discussed and resolved between the raters and included in the dataset.

4.11 Statistical analysis of the results

Prior to running any statistical analysis, the data were first recorded on spreadsheets created specifically to save the scores obtained from all the measuring rubrics. Thus, following the assessment of compositions employing the qualitative and quantitative measures, and the analysis of the pair talks, all the scores obtained were saved on spreadsheets under specific labels. After this procedure ended, the results were imported to SPSS for further analysis.

Before running any statistical tests all the data were first checked for their normal distribution utilizing histograms and boxplots¹ (also known as box-whisker diagrams), which are helpful to examine the shape of the data distribution (Field, 2017). The data were also numerically checked for their distribution through the Shapiro-Wilk test, which is among the best options for samples sizes under 50 (Larson-Hall, 2016), and which has more power to detect deviations from normality than Kolmogorov-Smirnov test, a similar test used to check the normality of data distribution. Despite this advantage, Field (2017) claims that both these tests can be significant with large sample sizes of data even when there is a slight deviation from normal distribution, therefore when checking the data for their distribution, he recommends that in addition to the tests

¹ I am very grateful to Ass. Prof. Igor Yanovich who provided me with his professional support in checking the normality of data distribution and in running the adequate statistical tests to analyze both normally and non-normally distributed data.

of normality, histograms and boxplots should be looked at as well. Based on this principle, I also checked the normality of data distribution with histograms and boxplots. Thus, the results of Shapiro-Wilk tests found that the results obtained through the analytic rubrics and those obtained from the data of oral interactions were approximately normally distributed, with few deviations. On the other hand, the scores obtained through the global rubrics were not normally distributed (For more detail, see the data in Appendices Appendix E, Appendix F, and Appendix G). Therefore, following a common practice in the field of statistics, for normally distributed data, I used parametric tests such as Independent Samples *t*-tests for comparisons between groups, and One-Way ANOVA tests for comparisons within groups. On the other hand, for the non-normally distributed data, I used non-parametric tests, such as Mann-Whitney *U* tests for comparisons between groups, and Kruskal-Wallis tests for comparisons within groups (Field, 2017; Larson-Hall, 2016). Moreover, to be able to establish the potential differences within groups working on the same condition (collaborative or individual), in addition to One-Way ANOVA and Kruskal-Wallis tests, post-hoc tests such as Tukey post-hoc tests and pairwise comparison tests were used, respectively.

Apart from general descriptive statistics including Means, Standard Deviations, Medians and Ranges (for non-parametric tests), and Test Statistics (e.g. statistics for tests as *t*-tests, One-Way ANOVA *F*, Mann-Whitney *U* and *Z*) the effect size values were also reported. Thus, for *t*-tests and Tukey post-hoc tests, the Cohen's *d* effect size values were used for either between and within groups using the benchmarks: small: $d = 0.20$, medium $d = 0.50$ and large $d = 0.80$. Whereas, for Mann-Whitney *U* and Kruskal-Wallis tests, the Cohen (1988)'s recommended effect size guidelines were used, namely the effect size values were considered: small $r = 0.10$, medium $r = 0.30$, and large $r = 0.50$. In fact, the effect sizes for Mann-Whitney *U* and Pairwise comparison tests, a further analysis following Kruskal-Wallis test, were calculated with the help of the open access data analysis software the Jamovi Project (2021).

It is interesting to note that when comparing the differences within groups, instead of computing the effect size for One-Way ANOVA or Kruskal-Wallis test, I computed the effect sizes for within group comparisons to detect which group performed better, as this type of effect size gives a better estimation of the effect of treatment on a particular variable (Field, 2017). It should also be noted that the effect size *r* reported

in this study is not similar to correlation coefficient, but it is a type of effect size reported for non-parametric tests such as Mann-Whitney U , Kruskal-Wallis and the like. As noted above, the r effect size was calculated with the help of the open access data analysis software the Jamovi Project (2021), but it can also be calculated using the equation formula: $r = \frac{Z}{\sqrt{N}}$ as suggested by Corder and Foreman (2009) and Field (2017). In this formula, Z is the z-score that SPSS produces, whereas N is the number of participants included in the sample, in the case of the present study there were 96 observations 48 (pairs) 48 (individuals). Thus, an approximate effect size for the measure of task-achievement would be $r = -3.211/\sqrt{96} = 0.32$ (see Section 5.5.1). A similar equation can be used to calculate the effect size of Kruskal-Wallis tests or pairwise comparison tests.

Finally, for non-parametric tests *Medians* and *Ranges* (minimum and maximum) were reported instead of *Means* because according to Field (2017) this statistic is more appropriate for non-parametric tests.

In light of providing as much details as possible, table 4.5 below presents the type of data, the analysis procedures, and the tests used to give answers to the questions raised in the present study.

Table 4.5 Research questions, data sources and analysis

Research question	Data source and steps	Analysis
RQ1. What are the differences between the written texts produced by pairs and individuals in terms of quantitative measures of fluency, complexity and accuracy?	Compare 96 written texts written in pairs (48) and individually (48) for quantitative measures of fluency, complexity and accuracy	Independent Samples t-tests
RQ1a. Are there any differences between the texts produced by pairs and individuals across three levels of proficiency?	Compare 32 written texts written in pairs and individually for each level of proficiency for quantitative measures of fluency, complexity and accuracy	Independent Samples t-tests
RQ1b. Are there any differences between the texts produced by pairs and individuals across three levels of proficiency within their writing conditions?	Compare 96 written texts composed in pairs (48) and individually (48) across three levels of proficiency in terms of fluency, complexity and accuracy.	ONE-WAY ANOVA - Tukey Post-hoc test

RQ2. What are the differences between the written texts produced by pairs and individuals in terms of task achievement, coherence and cohesion, grammar, vocabulary and mechanics?	Compare 96 written texts composed in pairs (48) and individually (48) for qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics.	Mann-Whitney U tests
RQ2a. Are there any differences between the texts produced by pairs and individuals across three levels of proficiency?	Compare 32 written texts written in pairs (16) and individually (16) for each level of proficiency for qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics.	Mann-Whitney U tests
RQ2b. Are there any differences between the texts produced by pairs and individuals across three levels of proficiency within their writing conditions?	Compare 48 written texts written in pairs and 48 written individually across three levels of proficiency in terms task achievement, coherence and cohesion, grammar, vocabulary and mechanics.	Kruskal-Wallis tests Pairwise comparison tests
RQ3. What are the differences between LREs produced by pairs in terms of number, type and resolution?	Compare 24 pair talks for LREs, in terms of their number, type and resolution.	ONE-WAY ANOVA - Tukey Post-hoc test
RQ3a. Are there any differences between LREs produced by pairs across three levels of proficiency?	Compare 3 sets of pair talks (3x8 = 24) from each proficiency level.	ONE-WAY ANOVA - Tukey Post-hoc test

4.12 Ethical issues

Before the data collection began, full approval from children's parents/guardians and participants themselves was obtained. Unlike with adult participants, when seeking informed consent from minors (e.g. preschoolers, children, and adolescents) who are not able to make decisions themselves whether to participate or not in research projects, researchers should follow two important stages. As explained by Cohen, Manion, and Morrison (2007) "first, researchers consult and seek permission from those adults responsible for the prospective subjects, and second, they approach the young people themselves" (p. 54). For the present study, I first sent an informed consent to children's parents/guardians explaining the purpose of the study, data collection procedures, audio-recording devices to be used, the writing assignment to be completed, participants rights, and risks and benefits for doing the study and asked for their permission to allow their children to participate in the study (see Appendix H). Second, I presented the

research plan to the participants (children) informing them about all research activities planned for this study, and I answered every single question participants raised concerning the study and activities to be completed. I also informed both parties that participation in the present study was completely voluntary and that children's grades in English would by no means be affected regardless of their decision to participate or not. Additionally, I informed participants on their full rights to withdraw from the study at any stage, and that their potential withdrawal would in no way be questioned! Since participants' English language level was considered as intermediate, it was necessary that they fully understand the content of the informed consent form (Mackey & Gass, 2005), therefore the informed consent form was delivered to them in their mother tongue, i.e. the Albanian language. Those students who agreed to participate were assured full confidentiality over their personal data.

Secondly, after receiving the consent forms signed by participants' parents/guardians, I also obtained a full approval for data collection at my school. This was obtained following a request for approval sent to the principal of my school to conduct my research. The request form sent contained all the necessary details concerning the purpose of the study, data collection procedures, the writing task, participants, the schedule for data collection, and a pledge to protect the participants' identity.

4.13 Summary of methodology

In this chapter, I first provided an overview of the context where the study took place, and then I discussed the teaching and learning of English as a foreign language (EFL) in the Kosovar context with a special focus on the pre-university education level. In doing so, I discussed the Core Curriculum (CC) for pre-university education in Kosovo, the English Syllabus for grade 8 learners and the participants of the study. I continued with a short overview of the pilot study, followed by a clear description of the writing prompt used and the data collection and data analysis procedures. I also described my role as a researcher in this study, the rating process and the statistical analysis of the results. Finally, I described the steps taken to make sure the data collected for this study were in full compliance with the highest ethical standards, ending the chapter with a short summary of the methodology used in the study.

Chapter 5: Results

5.1 Introduction

This chapter presents findings for each research question and sub-question of the study. It is organized into four sections. The first section reports results concerning the outcomes of the written texts derived from the quantitative measures of fluency, complexity and accuracy. The second section reports results obtained from the qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics. The third section reports results concerning the pair talk, specifically it reports on the number, type and resolution of LREs produced by learners working in pairs, altogether, and then for pairs across three proficiency levels. The fourth section provides some results concerning the structure of the pair talks observed from pairs across three proficiency levels. This chapter then ends with a summary of the results reported in this study.

5.2 Results of the quantitative analysis

As already noted, the first RQ asks whether there are any differences between the written texts produced by pairs and individuals in terms of quantitative measures of fluency, complexity and accuracy.

To answer this question, I analyzed 96 written texts produced in pairs (48) and individually (48) employing the quantitative measures of fluency, complexity and accuracy. Before running any statistical analysis, I checked the normality of data distribution by running of Shapiro-Wilk test (for more details check Appendix E), which indicated that the scores obtained on CAF measures were approximately normally distributed ($p > .05$). Hence, to check any statistical differences I ran several Independent Samples *t*-tests, where the alpha level was set at $< .05$, and reported the Cohen's *d* effect size values as small: $d = 0.20$, medium $d = 0.50$ and large $d = 0.80$.

5.2.1 Fluency

Fluency was measured through the average number of words per text, t-units and clauses per text. Table 5.1 presents the descriptive statistics for fluency.

Table 5.1. Measures of fluency across the whole dataset

	Pairs (n=48)			Individuals (n=48)					
	AVE	M	SD	AVE	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Words per text	170	169.47	40.89	150	150.06	56.58	1.92	.057	.39
T-units per text	22	22	4.98	21	20.75	8	.919	.361	.18
Clauses per text	32	32.31	7.76	30	29.64	12.97	1.22	.225	.25

Note. AVE = average

As can be seen from the table, pairs produced longer texts than individuals in all three measures of fluency. However, independent samples *t*-tests revealed no significant differences for any of the measures employed between texts produced by pairs and individuals. In terms of the average number of words per text, pairs produced longer texts ($M = 169.47$, $SD = 40.89$) than individuals ($M = 150.06$, $SD = 56.58$) but could not reach a statistically different result ($p = .057$). The same goes for the average number of T-units per text: while pairs performed better ($M = 22$, $SD = 4.98$) than individuals ($M = 20.75$, $SD = 8$) no statistically significant difference was observed ($p = .361$). Finally, in terms of the average number of clauses per text, again pairs performed better thus producing texts of a higher proportion of clauses ($M = 32.31$ $SD = 7.76$) than individuals ($M = 29.64$, $SD = 12.97$) but no significant difference was reached between pairs and individuals ($p = .225$). The effect sizes for all the fluency measures were medium for words per text ($d = .40$), small for t-units per text ($d = .18$) and clauses per text ($d = .25$). These results suggested that learners working in pairs produced longer texts than individuals, however, results were not statistically significant for either group.

5.2.2 Complexity

Complexity was measured in terms of proportion of clauses to T-units and dependent clauses to total clauses. Table 5.2 displays the descriptive statistics for this measure.

Table 5.2 Measures of complexity across the whole dataset

	Collaborative n=48)			Individual (n=48)					
	Total	M	SD	Total	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Clauses	1551	32.31	7.76	1423	29.64	12.97	1.22	.225	.25
T-units	1056	22	4.98	996	20.75	8	.919	.361	.18
Dependent clauses	501	10.43	4.27	412	8.58	5.78	1.78	.077	.36
C/T		1.47	0.18		1.43	0.35	.809	.420	.16
DC/C		0.31	0.09		0.26	0.14	2.12	0.036	.43

Note. C/T = clauses to T-units; DC/C = dependent clauses of total clauses.

As can be seen from the table, pairs produced texts that were syntactically more complex than those of individuals as shown by the ratio of clauses per text: pairs ($M = 32.31$, $SD = 7.76$) individuals ($M = 29.64$, $SD = 12.97$); proportion of T-units per text: pairs ($M = 22$, $SD = 4.98$) individuals ($M = 20.75$, $SD = 8$); number of dependent clauses per text: pairs ($M = 10.43$, $SD = 4.27$) individuals ($M = 8.58$, $SD = 5.78$); proportion of clauses to T-units: pairs ($M = 1.47$, $SD = 0.18$) individuals ($M = 1.43$, $SD = 0.35$), and proportion of dependent clauses to total clauses: pairs ($M = 0.31$, $SD = 0.09$) individuals ($M = 0.26$, $SD = 0.14$). The result of an independent samples *t*-test indicated that even though pairs produced more complex texts in terms of proportion of clauses to T-units than individuals, the differences between groups were not statistically significant ($p = .420$). However, pairs produced texts with significantly higher number of dependent clauses per total clauses than individuals ($p = .036$). Despite these results, the effect sizes for both these measures were small for proportion of clauses to T-units ($d = .16$) and medium for the proportion of dependent clauses to total clauses ($d = .43$). The findings indicated that pairs produced texts that were more complex than texts of individuals measured through the proportion of clauses to T-units and percentages of dependent clauses to total clauses, but only the latter was statistically significant favoring pairs.

5.2.3 Accuracy

Accuracy was measured in terms of proportion of error-free T-units of total T-units (EFT/T) and error-free clauses of total clauses (EFC/C). Table 5.3 features the descriptive statistics for the accuracy measures.

Table 5.3 Measures of accuracy across the whole dataset

	Collaborative (n=48)			Individual (n=48)					
	Total	M	SD	Total	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
T-units	1056	22	4.98	996	20.75	8	.919	.361	.18
EFT	611	12.72	6.76	456	9.5	6.47	2.38	.019	.48
Clauses	1551	32.31	7.76	1423	29.67	12.97	1.22	.225	.25
EFC	985	20.52	10.49	745	15.52	10.27	2.36	.020	.48
EFT/T		0.56	0.24		0.42	0.26	2.47	.015	.50
EFC/C		0.61	0.23		0.48	0.25	2.37	0.28	.48

Note. EFT/T = Error-free T-units of total T-units; EFC/C = Error-free clauses of total clauses.

As can be seen from the table, pairs produced a greater number of error-free T-units ($M = 12.72$, $SD = 6.76$) than individuals ($M = 9.5$, $SD = 6.47$), and a greater number of error-free clauses ($M = 20.52$, $SD = 10.49$) than individuals ($M = 15.52$, $SD = 10.27$). Similarly, pairs produced texts with a higher proportion of error-free T-units of total T-units ($M = 0.56$, $SD = 0.24$) than individuals ($M = 0.42$, $SD = 0.26$) and with greater proportion of error-free clauses of total clauses ($M = 0.61$, $SD = 0.23$) than individuals ($M = 0.48$, $SD = 0.25$). An independent samples *t*-test revealed that pairs performed significantly better than individuals in both measures of accuracy, that is, in the proportion of error-free T-units of total T-units ($p = .015$) and error-free clauses of total clauses ($p = .028$). The effect sizes for the accuracy measures of error-free T-units of total T-units (EFT/T) ($d = .50$) and error-free clauses of total clauses (EFC/C) ($d = .45$) were medium. These findings indicated that pairs produced texts that were significantly more accurate than those of individuals when comparing the results drawn from the whole dataset. The next section reports for pairs and individuals across the three proficiency levels.

5.3 Results of the quantitative analysis across three proficiency levels

This section addresses RQ1a: Are there any differences between the texts produced by pairs and individuals across three proficiency levels?

To address this sub-question, I analyzed the written composition data from the collaborative and individual conditions concerning the writing performance of pairs and individuals from low-intermediate (LI), intermediate (I), and high-intermediate (HI) language level measured through analytic measures of fluency, complexity and accuracy. In each case a similar number of texts produced collaboratively and individually is compared (16 per proficiency level). Again, statistical differences were examined using parametric tests (Independent Samples *t*-tests) as with the case of whole dataset, reporting the Cohen's *d* effect size values as small: $d = 0.20$, medium $d = 0.50$ and large $d = 0.80$.

5.3.1 Pairs vs. individuals of low-intermediate proficiency

Table 5.4 shows the descriptive statistics for the fluency measures featuring the outcomes of the written texts produced by pairs and individuals of low-intermediate (LI) language level (see Appendix M and Appendix N for more details on their texts).

Table 5.4 Fluency, complexity and accuracy for low-intermediate learners

Pairs (16)				Individuals (16)					
Fluency	AVE	M	SD	AVE	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Words per text	134	134	35	111	111	58.52	1.37	.178	.27
T-units per text	19	19	5.06	16	16	9.72	1.09	.282	.22
Clauses per text	25	26	6.73	20	21	14.34	1.23	.228	.08
Complexity	N	M	SD	N	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Clauses	411	26	6.73	333	21	14.34	1.23	.228	.43
T-units	301	19	5.06	253	16	9.72	1.09	.282	.39
Dependent clauses	112	7	3.68	75	4.68	5.47	1.4	.171	.50
C/T-units		1.38	0.19		1.35	0.5	.274	.786	.09
DPC/C		0.26	0.11		0.19	0.18	1.451	.157	.51
Accuracy	N	M	SD	T	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
T-units	301	19	5.06	253	16	9.72	1.09	.282	.22
Clauses	411	26	6.73	333	21	14.34	1.23	.228	.08
EFT/T		0.3	0.18		0.23	0.29	.791	.435	.28
EFC/C		0.35	0.2		0.29	0.31	.732	.470	.26

AVE = Average; N = Number. C/T = clauses per T-unit; DPC/C = dependent clauses of total clauses; EFT/T = error-free T-units of total T-units; EFC/C = error-free clauses of total clauses.

5.3.1.1 Fluency

As can be seen from the table, pairs outperformed individuals in the fluency measures of average number of words per text (Pairs, $M = 134$, $SD = 35$; Individuals, $M = 111$, $SD = 58.52$), T-units per text (Pairs, $M = 19$, $SD = 5.06$; Individuals, $M = 16$, $SD = 9.72$), and clauses per text (Pairs, $M = 26$, $SD = 6.73$; Individuals, $M = 21$, $SD = 14.34$). However, an independent samples t -test result indicated that the differences between pairs and individuals were not statistically significant for any of the measures, that is, the average number of words per text ($p = .178$; $d = .27$); T-units per text ($p = .282$; $d = .22$), and clauses per text ($p = .228$; $d = .08$). Likewise, the effect sizes for all the fluency measures were small.

5.3.1.2 Complexity

Complexity was measured via the proportion of clauses to T-units and dependent clauses to total clauses. Table 5.4 presents the descriptive statistics for the complexity measures. As can be seen from the table, pairs performed better than individuals in the complexity measures of proportion of clauses to T-units (Pairs, $M = 1.38$, $SD = 0.19$; Individuals, $M = 1.35$, $SD = 0.5$), and proportion of dependent clauses to total clauses (Pairs, $M = 0.26$, $SD = 0.11$; Individuals, $M = 0.19$, $SD = 0.18$). Nevertheless, an independent samples t -test indicated that the differences between pairs and individuals in proportion of clauses to T-units ($p = .786$, $d = .09$), and proportion of dependent clauses to total clauses ($p = .159$, $d = .51$) were not statistically significant with a small to medium effect, respectively.

5.3.1.3 Accuracy

Table 5.4 also presents the descriptive statistics for the accuracy measures of error-free T-units of total T-units and error-free clauses of total clauses. As can be seen from the table, pairs outperformed individuals in the accuracy measures of error-free T-units of total T-units (Pairs, $M = 0.3$, $SD = 0.18$; Individuals, $M = 0.23$, $SD = 0.29$), and error-free clauses of total clauses (Pairs, $M = 0.35$, $SD = 0.2$; Individuals, $M = 0.29$, $SD = 0.31$). However, as with fluency and complexity, an independent samples t -test revealed no statistically significant differences between pairs and individuals in the proportion of error-free T-units of total T-units ($p = .470$, $d = .25$) and error-free clauses of total clauses ($p = .435$, $d = .27$). The effect sizes for both these measures were small. These

findings indicate that pairs were able to produce longer, linguistically more complex and grammatically more accurate texts than individuals. Yet, none of these differences between pairs and individuals were statistically significant.

5.3.2 Pairs vs. individuals of intermediate proficiency

Table 5.5 presents the descriptive statistics for all analytic measures concerning the text produced by learners of intermediate proficiency (see Appendices Appendix K and Appendix L for more details on their compositions).

Table 5.5 Fluency, complexity and accuracy for intermediate learners

Pairs (16)				Individuals (16)					
Fluency	AVE	M	SD	AVE	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Words per text	170	170	27	155	156	40.24	1.18	.244	.24
T-units per text	22	22	3.31	21	21.18	3.2	.652	.519	.23
Clauses per text	32	32	5.36	31	31	8.66	.378	.708	.13
Complexity	N	M	SD	N	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Clauses	517	32	4.84	514	31	8.66	.378	.708	.13
T-units	351	22	3.31	342	21.18	3.2	.652	.519	.23
Dependent clauses	167	9.81	3.12	152	9.18	5.07	.419	.678	.17
C/T-units		1.48	0.16		1.49	0.32	-.137	.258	-.48
DPC/C		0.31	0.07		0.28	0.09	1.153	.892	.40
Accuracy	N	M	SD	T	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
T-units	351	22	3.31	342	21.18	3.2	.652	.519	.17
Clauses	517	32	4.84	514	31	8.66	.378	.708	.02
EFT/T		0.67	0.18		0.49	0.19	2.69	.012	.95
EFC/C		0.7	0.17		0.56	0.15	2.42	.021	.85

AVE = Average; N = Number. C/T = clauses per T-unit; DPC/C = dependent clauses of total clauses; EFT/T = error-free T-units of total T-units; EFC/C = error-free clauses of total clauses.

5.3.2.1 Fluency

As can be seen from the table, pairs outperformed individuals in the fluency measures of average number of words per texts (Pairs, $M = 170$, $SD = 27$; Individuals, $M = 155$, $SD = 40.24$); T-units per text (Pairs, $M = 22$, $SD = 3.31$; Individuals, $M = 21.18$, $SD = 3.2$), and clauses per text (Pairs, $M = 32$, $SD = 5.36$; Individuals, $M = 31$, $SD = 8.66$). However, despite a better performance of pairs over individuals an independent samples *t*-test revealed no statistically significant differences in the average number of words per text ($p = .244$, $d = .24$), T-units per text ($p = .519$, $d = .23$), and clauses per text (p

= .708, $d = .13$). The effect sizes for all the fluency measures were small. These findings show that even though pairs outperformed individuals in the fluency measures, the differences between them were statistically insignificant.

5.3.2.2 Complexity

Table 5.5 also presents the descriptive statistics for complexity. As can be seen from the table, while individuals slightly outperformed pairs in the complexity measure of proportion of clauses to T-units (Pairs, $M = 1.48$, $SD = 0.16$; Individuals, $M = 1.49$, $SD = 0.32$), pairs did the same against individuals in the complexity measure of proportion of dependent clauses to total clauses (Pairs, $M = 0.31$, $SD = 0.07$; Individuals, $M = 0.28$, $SD = 0.09$). Interestingly, the proportion of clauses to T-units, as a measure of complexity is the only measure where individuals outperformed pairs of intermediate proficiency. On the other hand, pairs outperformed individuals in the proportion of dependent clauses to total clauses. An independent samples t -test revealed no statistically significant differences between pairs and individuals in any of the complexity measures of proportion of clauses to T-units ($p = .892$, $d = -.48$), and proportion of dependent clauses to total clauses ($p = .258$, $d = .40$). The effect sizes for both the complexity measures were medium. These findings indicate that collaboration did not affect the complexity of the written texts produced by intermediate proficiency learners.

5.3.2.3 Accuracy

Table 5.5 also features the descriptive statistics for accuracy. As can be seen from the table, pairs outperformed individuals in both the accuracy measures of error-free T-units of total T-units (Pairs, $M = 0.67$, $SD = 0.18$; Individuals, $M = 0.49$, $SD = 0.19$), and error-free clauses of total clauses (Pairs, $M = 0.7$, $SD = 0.17$; Individuals, $M = 0.56$, $SD = 0.15$). An independent samples t -test revealed that the differences between pairs and individuals in both the accuracy measures were statistically significant, that is, pairs produced significantly more error-free T-units ($p = .012$, $d = .95$) and more of error-free clauses ($p = .021$, $d = .85$) than individuals. The effect sizes for accuracy measures were large. These findings show that pairs outperformed individuals in all quantitative measures of fluency and accuracy but not in complexity, where individuals slightly outperformed pairs in the complexity measure of proportion of clauses to T-units. Collaboration seems to benefit learners' written texts in terms of accuracy, as indicated

by significant results favoring pairs over individuals. Now we compare results of the same measures for high-intermediate learners.

5.3.3 Pairs vs. individuals of high-intermediate proficiency

Table 5.6 presents the descriptive statistics for all analytic measures concerning the text produced by learners of high-intermediate proficiency (see Appendices Appendix I and Appendix J for more details on their compositions).

5.3.3.1 Fluency

Table 5.6 shows the descriptive statistics for the analytic measures.

Table 5.6 Fluency, complexity and accuracy for high-intermediate learners

Pairs (16)				Individuals (16)					
Fluency	AVE	M	SD	AVE	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Words per text	205	205	26	184	184	45.53	1.56	.129	.31
T-units per text	25	25	4.35	25	25	7.03	.000	.928	0
Clauses per text	38	39	4.62	35	36	10.11	1.05	.299	.37
Complexity	N	M	SD	N	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
Clauses	623	37.68	4.3	576	34.87	9.09	1.11	.272	.40
T-units	404	25.06	4.12	401	25.06	7.03	.000	1	0
Dependent clauses	222	12.68	2.52	185	10.43	4.16	1.85	.074	.65
C/T-units		1.56	0.16		1.44	0.14	2.08	.046	.73
DPC/C		0.35	0.05		0.31	0.1	1.4	.170	.50
Accuracy	N	M	SD	T	M	SD	<i>t</i>	<i>p</i>	<i>d</i>
T-units	404	25.06	4.12	401	25.06	7.03	.000	1	.0
Clauses	623	37.68	4.3	576	34.87	9.09	1.11	.272	.40
EFF/T		0.68	0.12		0.54	0.19	2.37	.024	.83
EFC/C		0.73	0.11		0.6	0.14	2.68	.012	.95

AVE = Average; N = Number. C/T = clauses per T-unit; DPC/C = dependent clauses of total clauses; EFT/T = error-free T-units of total T-units; EFC/C = error-free clauses of total clauses.

As can be seen from the table, pairs outperformed individuals in terms of fluency measures. Hence, they performed better in the average number of words per text (Pairs, $M = 205$, $SD = 26$; Individuals, $M = 184$, $SD = 45.53$), and the average number of clauses per text (Pairs, $M = 38$, $SD = 4.62$; Individuals, $M = 35$, $SD = 10.11$). However, pairs performed quite similarly to individuals in the fluency measure of average number of

T-units per text (Pairs, $M = 25$, $SD = 4.35$; Individuals, $M = 25$, $SD = 7.03$). Despite a better performance of pairs over individuals in terms of fluency measures stated above, an independent samples t -test found no significant differences between pairs and individuals in any of the fluency measures of average number of words per text ($p = .129$, $d = .31$), T-units per text ($p = .928$, $d = .0$), and clauses per text ($p = .299$, $d = .37$). The effect size values for each measure were medium. These findings indicate that collaboration had a positive impact on the learners' written texts in terms of fluency, however such an impact was not significant.

5.3.3.2 Complexity

Table 5.6 also presents the descriptive statistics for the complexity measures. As can be seen from the table, high-intermediate pairs outperformed high-intermediate individuals in both the complexity measures of proportion of clauses to T-units (Pairs, $M = 1.56$, $SD = 0.16$; Individuals, $M = 1.44$, $SD = 0.14$), and proportion of dependent clauses to total clauses (Pairs, $M = 0.35$, $SD = 0.05$; Individuals, $M = 0.31$, $SD = 0.1$). However, despite a better performance of pairs over individuals in terms of complexity measures, an independent samples t -test revealed significant differences only for the proportion of clauses to T-units ($p = .046$, $d = .73$) but not for the ratio of dependent clauses to total clauses ($p = .170$, $d = .49$). In accordance with p -values, the effect size values were close large and moderate for both complexity measures, respectively. These findings suggest that collaboration had a positive impact on the complexity of the texts produced, indicating more clauses per T-unit in favor of pair texts. However, collaboration did not have a similar effect on the number of dependent clauses to total clauses. Thus, even though pairs produced texts with a higher proportion of dependent clauses to total clauses than individuals, these differences were not statistically significant.

5.3.3.3 Accuracy

Table 5.6 displays the descriptive statistics for the accuracy measures for the two high-intermediate conditions. As can be seen from the table, pairs outperformed individuals in both the accuracy measures of proportion of error-free T-units per total T-units (Pairs, $M = 0.68$, $SD = 0.12$; Individuals, $M = 0.54$, $SD = 0.19$) and ratio of error-free clauses per total clauses (Pairs, $M = 0.73$, $SD = 0.11$; Individuals, $M = 0.60$, $SD = 0.14$). An independent samples t -test revealed that pairs significantly outperformed individuals in both the accuracy measures of proportion of error-free T-units of total T-units ($p = .025$,

$d = .83$), and rate of error-free clauses of total clauses ($p = .012$, $d = .95$). The effect sizes for the accuracy measures were large. The results indicate that pairs produced texts that were grammatically more accurate than texts of individuals, suggesting that collaboration benefits adolescent EFL learners' written texts in terms of accuracy. The next section compares the performance of learners within their writing conditions.

5.4 Pairs and individuals within their writing conditions

This section addresses RQ1b: Are there any differences between the texts produced by pairs and individuals across three proficiency levels within their writing conditions?

In this section, I present results concerning quantitative measures of fluency, complexity, and accuracy for learners working in the same writing condition, but with different language proficiency level. Here, I report findings for the written texts produced by low-intermediate (LI), intermediate (I) and high-intermediate (HI) proficiency learners assigned to work in pairs and individually. First, I report results for pairs (3 X 16 = 48) and then for individuals (3 X 16 = 48). To check for any statistical differences within groups, several One-Way ANOVA tests were run, while the effect size values using the Cohen's d benchmarks as small: $d = 0.20$, medium $d = 0.50$ and large $d = 0.80$ for each comparison (e.g. low-intermediate vs. intermediate or high-intermediate) were reported.

5.4.1 Fluency of collaborative texts

Table 5.7 features the descriptive statistics for the fluency measures.

Table 5.7 Measures of fluency for pairs across three proficiency levels

Fluency	LP	M	SD	p	t	d
Words per text	LI	134.19	34.69	3-4 = <.004	-3.41	1.21
	I	169.69	26.93	4-5 = < .005	-3.55	1.19
	HI	204.56	25.86	3-5 = <. 001	-6.76	2.39
T-units per text	LI	18.81	5.06	3-4 = <.111	-2.05	0.72
	I	21.94	3.31	4-5 = < .086	-2.18	0.76
	HI	25.25	4.35	3-5 = <. 001	-4.23	1.49
Clauses per text	LI	25.69	6.73	3-4 = <.005	-3.32	1.17
	I	32.31	5.36	4-5 = < .005	-3.32	1.17
	HI	38.94	4.62	3-5 = <. 001	-6.64	2.35

LI = low-intermediate; I = intermediate; HI = high-intermediate.

As can be seen from the table, pairs of higher proficiency produced longer texts than pairs of lower L2 proficiency as indicated by the fluency measures of words per text LI ($M = 134, SD = 35$), I ($M = 170, SD = 27$), HI ($M = 205, SD = 26$); T-units per text ($M = 19, SD = 5.06$), I ($M = 22, SD = 3.31$), HI ($M = 25, SD = 4.35$); clauses per text LI ($M = 26, SD = 6.73$), I ($M = 32.33, SD = 5.36$), HI ($M = 39, SD = 4.62$). A One-Way ANOVA test result revealed that the differences between pairs of different proficiency level were statistically significant $F = (2, 45) = 22.84, p = .001$. Thus both HI and I pairs performed significantly better than LI pairs in the measure of average number of words per text (HI-LI $p = .001, d = 2.39$; I-LI $p = .004, d = 1.29$; HI-I = .005, $d = 1.19$) with large effect sizes in favor of higher proficiency pairs.

The situation changed a bit for the other fluency measure, that is, ratio of T-units per text. Thus, an ANOVA test result indicated that only the performance between HI and LI pairs was statistically significant $F = (2, 45) = 8.94, p = .001$. Follow-up test results showed that HI pairs performed significantly better than LI pairs (HI-LI $p = .001, d = 1.49$) but not better than I pairs (HI-I $p = .086, d = 0.76$). Similarly the performance between I and LI pairs was not statistically significant (I-LI $p = .111, d = 0.72$). The effect size values were large for the differences between HI and LI pairs, and moderate for the differences between HI and I pairs.

Finally, also the results for the average number of clauses per text were significantly different as shown by a One-Way ANOVA test result $F = (2, 45) = 22.71, p = .001$. Hence, HI and I pairs performed significantly better than LI pairs (HI-LI $p = .001, d = 2.35$; I-LI $p = .005, d = 1.17$). Similarly, HI pairs performed significantly better than I pairs (HI-I = .005, $d = 1.17$). The effect size values for this measure were large in each case. These findings indicate that pairs of higher proficiency performed significantly better than pairs of lower proficiency in all the fluency measures, excluding the rate of T-units per text (e.g. LI-I and I-HI pairs) where the impact of learner proficiency seemed to be moderate.

5.4.2 Complexity of collaborative texts

Table 5.8 shows the descriptive statistics for the complexity measures.

Table 5.8 Measures of Complexity for pairs across three proficiency levels

Complexity	LP	M	SD	<i>p</i>	<i>t</i>	<i>d</i>
Clauses to T-units	LI	24.31	6.29	3-4 = <.009	-3.09	1.09
	I	30.13	4.95	4-5 = <.039	-2.53	0.89
	HI	34.88	4.54	3-5 = <.001	-5.62	1.98
Dependent clauses to total clauses	LI	7	3.68	3-4 = <.013	-2.97	1.05
	I	10.44	3.61	4-5 = <.013	-2.97	1.05
	HI	13.88	2.33	3-5 = <.001	-5.94	2.1

LI = low-intermediate; I = intermediate; HI = high-intermediate.

As can be seen from the table, pairs of higher proficiency performed better than pairs of lower proficiency in the rate of clauses per T-unit LI ($M = 24.31$, $SD = 6.29$), I ($M = 30.13$, $SD = 4.95$), HI ($M = 34.88$, $SD = 4.54$), and proportion of dependent clauses to total clauses LI ($M = 3$, $SD = 3.68$) I ($M = 10.44$, $SD = 3.61$) HI ($M = 13.88$, $SD = 2.33$). As shown by a One-Way ANOVA test result the differences between pairs in number of clauses to T-units were statistically significant $F(2, 45) = 15.85$, $p = .001$. Further, a Tukey post-hoc test result indicated that HI and I pairs significantly outperformed LI pairs (HI-I $p = .001$, $d = 1.98$; I-LI $p = .009$, $d = 1.09$) in the ratio of clauses per T-unit. Also HI pairs performed significantly better than I pairs (HI-I $p = .039$, $d = 0.89$) in the same measure. The effect size values ranged from very large to large, respectively.

Additionally, the result for the complexity measure proportion of dependent clauses to total clauses was significantly different as shown by a One-Way ANOVA test result which revealed that pairs of higher proficiency performed significantly better than those of lower proficiency $F(2, 45) = 17.66$, $p = .001$. A Tukey post-hoc test indicated that HI and I pairs significantly outperformed LI pairs in the rate of dependent clauses (HI-LI $p = .001$, $d = 2.1$; I-LI $p = .013$, $d = 1.05$). Also HI pairs performed significantly better than I pairs (HI-I $p = .013$, $d = 1.05$). The effect size values for this measure were large in each comparison.

5.4.3 Accuracy of collaborative texts

Table 5.9 features the descriptive statistics for the accuracy measures.

Table 5.9 Accuracy measures for pairs across three proficiency levels

Accuracy	LP	M	SD	<i>p</i>	<i>t</i>	<i>d</i>
Error-free T-units	LI	6	4.5	3-4 = <.001	-5.25	1.85
	I	14.81	4.84	4-5 = < .288	-1.53	0.54
	HI	17.38	4.88	3-5 = <. 001	-6.78	2.39
Error-free clauses	LI	9.88	7.42	3-4 = <.001	-5.32	1.88
	I	23.06	7.24	4-5 = < .074	-2.25	0.79
	HI	28.63	6.29	3-5 = <. 001	-7.57	2.67

LI = low-intermediate; I = intermediate; HI = high-intermediate.

As can be seen from the table, pairs of higher proficiency outperformed those of lower proficiency in the accuracy measures of error-free T-units LI ($M = 6$, $SD = 4.5$), I ($M = 14.81$, $SD = 4.84$), HI ($M = 17.38$, $SD = 4.88$), and error-free clauses LI ($M = 9.88$, $SD = 7.42$), I ($M = 23.06$, $SD = 7.24$), HI ($M = 28.63$, $SD = 7.69$). A One-Way ANOVA test result showed that the differences between pairs in both the accuracy measures, i.e. proportion of error-free T-units $F = (2, 45) = 25.27$, $p = .001$, and proportion of error-free clauses $F = (2, 45) = 30.23$, $p = .001$ were statistically significant. Thus, HI and I pairs significantly outperformed LI pairs (HI-LI $p = .001$, $d = 2.39$; I-LI $p = .001$, $d = 1.85$) in the rate of error-free T-units, however, no significant differences were observed between HI and I pairs (HI-I $p = .288$, $d = 0.54$). A similar trend was observed in the proportion of error-free clauses. Hence, HI and I pairs significantly outperformed LI pairs (HI-LI $p = .001$, $d = 2.67$; I-LI $p = .001$, $d = 1.88$). As in the case of the proportion of error-free T-units, the differences between HI and I pairs in the rate of error-free clauses were not statistically significant (HI-I $p = .074$, $d = 0.74$). The effect sizes values were large and close to large for the proportion of error-free T-units, as well as for the proportion of error-free clause, respectively. These findings indicate that HI and I pairs performed significantly better than LI pairs in both of the accuracy measures. However, the differences between HI and I pairs were not statistically significant in any of the accuracy measures.

A somewhat similar performance was observed in the individual condition, that is, individuals of higher proficiency outperformed their peers of lower proficiency. Results of the quantitative measures for the individual condition across three proficiency levels are presented next.

5.4.4 Fluency of individual texts

Table 5.10 features the descriptive statistics for the fluency measures.

Table 5.10 Measures of fluency for individuals across three proficiency levels

Fluency	LP	M	SD	<i>p</i>	<i>t</i>	<i>d</i>
Words per text	LI	110.75	58.52	3-4 = <.034	-2.59	0.91
	I	155.31	40.24	4-5 = <.227	-1.67	0.59
	HI	184.13	45.53	3-5 = <.001	-4.26	1.5
T-units per text	LI	15.81	9.72	3-4 = <.103	-2.19	0.07
	I	21.38	3.2	4-5 = <.161	-1.45	0.51
	HI	25.05	7.03	3-5 = <.013	-3.65	1.29
Clauses per text	LI	20.81	14.34	3-4 = <.021	-2.78	0.97
	I	32.13	9.31	4-5 = <.609	-0.95	0.33
	HI	36	10.11	3-5 = <.001	-3.74	1.32

LI = low-intermediate; I = intermediate; HI = high-intermediate.

As can be seen from the table, individuals of higher proficiency outperformed their peers of lower proficiency in the fluency measures as indicated by means of the average number of words per text LI ($M = 111$, $SD = 59$), I ($M = 155$, $SD = 40$), HI ($M = 184$, $SD = 45.53$); T-units per text LI ($M = 16$, $SD = 9.72$) I ($M = 21.38$, $SD = 3.2$) HI ($M = 25$, $SD = 7.03$); clauses per text LI ($M = 21$, $SD = 14.34$) I ($M = 32$, $SD = 9.31$) HI ($M = 36$, $SD = 10$).

Result of a One-Way ANOVA showed that the differences between learners of higher and lower proficiency were statistically significant in the fluency measure of the average number of words per text $F(2, 45) = 9.21$, $p = .001$. However, this counts only for the difference in performance between HI and LI learners (HI-LI, $p = .001$, $d = 1.51$) and I-LI (I-LI, $p = .034$, $d = 0.91$), but not for HI and I learners (HI-I, $p = .227$, $d = 0.59$). The effect size values for this measure were large and moderate, respectively.

Almost an identical trend was observed for the average number of T-units per text. A One-Way ANOVA test result found a statistically significant difference between learners of higher and lower proficiency $F(2, 45) = 6.74$, $p = .003$. Yet, the significant difference counts only for HI and LI learners, implying that HI learners performed significantly better than LI learners (HI-LI, $p = .013$, $d = 1.29$). However, no significant differences were observed between HI and I learners (HI-I $p = .161$, $d = 0.51$), and I and

LI learners (I-LI $p = .103$, $d = 0.77$). The effect size values for this measure were large and moderate, respectively.

Finally, learners of higher and lower language level performed differently in the proportion of clauses per text. Thus, a One-Way ANOVA test result showed that learners of higher proficiency significantly outperformed those of lower language proficiency in the rate of clauses per text $F = (2, 45) = 7.57$, $p = .001$. This difference counts at least for the learners of two proficiency levels. Specifically, HI individuals significantly outperformed LI individuals (HI-LI, $p = .001$, $d = 1.32$) but not I individuals (HI-I, $p = .609$, $d = 0.33$). Additionally, I learners performed significantly better than LI learners (I-LI, $p = .021$, $d = 0.97$). The effect size values for this measure ranged from small to large as shown above.

5.4.5 Complexity of individual texts

Table 5.11 displays the descriptive statistics for the complexity measures.

Table 5.11 Measures of complexity for individuals across three proficiency levels

Complexity	LP	M	SD	p	t	d
Clauses to T-units	LI	20	13.33	3-4 = <.069	-2.28	0.8
	I	28.44	6.52	4-5 = <.357	-1.39	0.49
	HI	33.56	10.39	3-5 = <.002	-3.67	1.29
Dependent clauses to total clauses	LI	4.69	5.47	3-4 = <.028	-2.66	0.94
	I	9.5	4.96	4-5 = <.494	-1.14	0.4
	HI	11.56	4.87	3-5 = <.001	-3.8	1.34

LI = low-intermediate; I = intermediate; HI = high-intermediate.

As can be seen from the table, learners of higher proficiency performed better than those of lower proficiency in the proportion of clauses to T-units LI ($M = 20$, $SD = 13.33$), I ($M = 28.44$, $SD = 6.52$), HI ($M = 33.56$, $SD = 10.39$), and proportion of dependent clauses to total clauses LI ($M = 4.69$, $SD = 5.47$), I ($M = 9.50$, $SD = 4.96$), HI ($M = 11.56$, $SD = 4.87$). A One-Way ANOVA test result showed that the differences between learners of different language proficiency in the proportion of clauses to T-units were statistically significant $F = (2, 45) = 6.85$, $p = .003$. Thus, HI individuals significantly outperformed their LI peers (HI-LI $p = .002$, $d = 1.29$) in this measure but not I learners (HI-I $p = .357$, $d = 0.49$). Also the difference in performance between I and LI learners

was not statistically significant (I-LI $p = .069$, $d = 0.80$). As can be seen the effect size values for this measure were medium and large, respectively.

A somewhat similar result was observed for the complexity measure of proportion of dependent clauses to total clauses. As indicated by a One-Way ANOVA test result there was a significant difference in the performance of individual learners in the proportion of dependent clauses to total clauses $F = (2, 45) = 7.62$, $p = .001$. Hence, HI individuals performed significantly better than LI individuals (HI-LI $p = .001$, $d = 1.34$). In the same vein, I learners significantly outperformed LI learners (I-LI $p = .028$, $d = 0.94$). However, the difference in performance between HI and I learners in the same measure was not statistically significant (HI-I $p = .494$, $d = 0.40$). The effect size values for this measure were small and large, respectively.

5.4.6 Accuracy of individual texts

Table 5.12 features the descriptive statistics for the accuracy measures.

Table 5.12 Measures of accuracy for individuals across three proficiency levels

Accuracy	LP	M	SD	p	t	d
Error-free T-units	LI	4.25	5.23	3-4 = <.003	-3.47	1.22
	I	10.69	4.9	4-5 = < .279	-1.55	0.54
	HI	13.56	5.59	3-5 = <. 001	-5.01	1.77
Error-free clauses	LI	6.5	8.68	3-4 = <.001	-4.15	1.46
	I	18.31	7.56	4-5 = < .454	-1.21	0.42
	HI	21.75	7.83	3-5 = <. 001	-5.36	1.89

LI = low-intermediate; I = intermediate; HI = high-intermediate.

As can be seen from the table, individuals of higher proficiency outperformed their peers of lower proficiency in both the accuracy measures of error-free T-units LI ($M = 4.25$, $SD = 5.23$), I ($M = 10.69$, $SD = 4.9$), HI ($M = 13.56$, $SD = 5.59$), and error-free clauses LI ($M = 6.5$, $SD = 8.68$), I ($M = 18.31$, $SD = 7.56$), HI ($M = 21.75$, $SD = 7.83$). A One-Way ANOVA test result revealed that the differences between learners from individual writing condition in the accuracy measures of error-free T-units of total T-units $F = (2, 45) = 13.17$, $p = .001$ and error-free clauses of total clauses $F = (2, 45) = 15.82$, $p = .001$ were statistically significant. Hence, HI and I individuals performed significantly better than LI individuals (HI-LI $p = .001$, $d = 1.77$; I-LI $p = .003$, $d = 1.22$) in the rate of error-free T-units. On the other hand, no significant differences were

observed between HI and I individuals (HI-I $p = .279$, $d = 0.54$) in the same measure. Almost an identical result was obtained for the proportion of error-free clauses of total clauses. Thus, HI and I individuals significantly outperformed LI individuals (HI-LI $p = .001$, $d = 1.89$; I-LI $p = .001$, $d = 1.46$). As in the rate of error-free T-units, no significant differences were found between HI and I individuals in the rate of error-free clauses (HI-I $p = .454$, $d = 0.42$). The effect sizes for the accuracy measures were small and large in both measures.

To sum up, overall, the findings reported above for the collaborative and individual conditions suggest that pairs' or learners' language proficiency affected positively the outcomes of the written texts measured through analytic measures of fluency, complexity and accuracy. While in the collaborative writing condition pairs of higher language proficiency performed significantly better than those of lower proficiency in almost all the fluency measures except for the average number of T-units per text (LI-I and I-HI). Learners from the individual writing condition showed a somewhat weaker performance, see for example, the results for the measures of words per text (I-HI), T-units per text (LI-I and I-HI), and clauses per text (I-HI). As such the L2 proficiency of learners seemed to have a positive effect on their writing performance with large effect sizes in almost all measures as shown in the texts of pairs, which was not always the texts of individuals, where the effect size values varied from small to large.

As with the fluency measures, both pairs and individuals of higher L2 proficiency significantly outperformed their peers of lower proficiency. Nevertheless, results indicated that learners from the individual condition did not always perform better than their peers of lower L2 proficiency in the complexity measures (cf. low-intermediate vs. intermediate; intermediate vs. high-intermediate proficiency learners). On the whole, regardless of the writing conditions, the L2 proficiency of learners seemed to positively affect their writing performance measured through the complexity measures. This was particularly true for the pair learners where the effect of proficiency seemed to always be large favoring the higher proficiency pairs over the lower proficiency ones. Nevertheless, this was not always the case for the learners from the individual condition.

Finally, results of the accuracy measures indicated that the learners' L2 proficiency positively affected their writing performance. In fact, the study found that the higher the L2 proficiency of learners, the more accurate texts they produced regardless of their writing conditions. This was in particular the case for the results between LI and HI pairs and individuals. However, in both writing conditions, collaborative vs. individual, no statistically significant results were found between HI and I pairs and individuals in terms of the accuracy measures, suggesting that the larger the proficiency gap, the bigger the difference in scores between learners. Nevertheless, the results reported thus far have been drawn from the analysis of quantitative measures. Therefore, to be able to capture the effect of collaboration and/or the learners' L2 proficiency on the written text outcomes, I also analyzed the written texts through global rubrics, the results of which are presented next.

5.5 Results of the qualitative rubrics

In this section, I address the second research question (RQ2) that asks whether there are any differences between the written texts produced by pairs and individuals in terms of qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics. Before running any statistical test to check for the potential differences, I checked the normality of data distribution using Shapiro-Wilk test ($p < .05$) (for more details check Appendix F), which indicated that the scores obtained on the holistic rubrics were not normally distributed, hence to check any differences in terms of scores obtained between groups I ran several Mann-Whitney U tests (a type of non-parametric test), where the alpha level was set at $< .05$, and reported the effect size (r) values based on Cohen's (1988) guidelines as small $r = 0.10$, medium $r = 0.30$, and large $r = 0.50$. As noted above (c.f. Chapter 4, Section 4.11), the effect size values were calculated using the open access statistical analysis software Jamovi Project (2020).

To answer this question, I first report results for 96 written texts composed in pairs (48) and individually (48) that were analyzed using the global rating rubrics of task achievement, coherence and cohesion, grammar, vocabulary and mechanics. Second, I report results comparing the writing performance of pairs and individuals across three proficiency levels, namely low-intermediate (LI), intermediate (I) and high-intermediate (HI), where 16 texts are compared for each level of proficiency. Finally, I report results of holistic rubrics for pairs and individuals within their writing conditions, comparing the performance of pairs and individuals across three proficiency levels.

5.5.1 Pairs vs. individuals

Table 5.13 displays the descriptive statistics for the writing performance of pairs and individuals on holistic rubrics.

Table 5.13 Measures of holistic rubrics across the whole dataset

	Pairs (48)			Individuals (48)					
	Mdn	Min	Max	Mdn	Min	Max	<i>U</i>	<i>p</i>	<i>r</i>
Task-achievement	5	0	7	3	0	7	721.5	.001	0.37
Coherence & cohesion	4	0	7	3	0	6	726.5	.002	0.36
Grammar	5	0	7	4	0	7	837	.019	0.27
Vocabulary	5	0	7	3	0	7	615	.001	0.47
Mechanics	4.5	0	7	4	0	7	987.5	.220	0.14

*Note: *r* = Effect size

As can be seen from the table, pairs outperformed individuals in all qualitative measures of task achievement (Pairs, *Mdn* = 5; individuals, *Mdn* = 3), coherence and cohesion (Pairs, *Mdn* = 4; individuals, *Mdn* = 3), grammar (Pairs, *Mdn* = 5; individuals, *Mdn* = 4), vocabulary (Pairs, *Mdn* = 5; individuals, *Mdn* = 3), and mechanics (Pairs, *Mdn* = 4.5; individuals, *Mdn* = 4). A Mann-Whitney U test, however, revealed that while pairs performed significantly better than individuals in all qualitative measures, this did not hold for mechanics. Thus, pairs did significantly better in task achievement ($U = 721.5$, $z = -3.211$, $p = 0.001$, $r = 0.37$), coherence and cohesion ($U = 726.5$, $z = -3.166$, $p = 0.002$, $r = 0.36$), grammar ($U = 836.5$, $z = -2.338$, $p = 0.019$, $r = 0.27$), vocabulary ($U = 615.0$, $z = -3.975$, $p = 0.001$, $r = 0.46$), but not in mechanics ($U = 987.5$, $z = -1.224$, $p = .220$, $r = 0.14$). The effect size values for all these measures were medium, except grammar and mechanics that were small. These findings indicate that collaboration benefits adolescent EFL learners' written texts in all qualitative rubrics especially in vocabulary, but not in mechanics. Never the less these results count for the whole dataset. In the following section, I present results for pairs and individuals across three proficiency levels.

5.6 Pairs vs. individuals across three proficiency levels

RQ2a: Are there any differences between the texts produced by pairs and individuals across three levels of proficiency?

To address this second sub-question, I compared results obtained by qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics for pairs and individuals across three proficiency levels. For every proficiency level, an equal number of texts composed in pairs (16) and individually (16) was compared using the same global rubrics mentioned above.

5.6.1 Pairs vs. individuals of low-intermediate proficiency

Table 5.14 presents the descriptive statistics for qualitative rubrics for pairs and individuals of low-intermediate language proficiency.

Table 5.14 Measures of global rubrics for low-intermediate learners

	Pairs (16)			Individuals (16)					
	Mdn	Min	Max	Mdn	Min	Max	<i>U</i>	<i>p</i>	<i>r</i>
Task achievement	3	0	7	2.5	0	7	82.5	.078	0.35
Coherence & Cohesion	2	0	6	1	0	5	102	.323	0.2
Grammar	2	0	5	1.5	0	4	92	.168	0.28
Vocabulary	2.5	0	5	1	0	4	55	.005	0.57
Mechanics	3	0	6	0	0	5	67	.016	0.47

*Note: *r* = Effect size

As can be seen from the table, LI pairs outperformed LI individuals in all measures of qualitative rubrics, that is, in task achievement (Pairs, *Mdn* = 3; individuals, *Mdn* = 2.5), coherence and cohesion (Pairs, *Mdn* = 2; individuals, *Mdn* = 1), grammar (Pairs, *Mdn* = 2; individuals, *Mdn* = 1.5), vocabulary (Pairs, *Mdn* = 2.5; individuals, *Mdn* = 1), and mechanics (Pairs, *Mdn* = 3; individuals, *Mdn* = 0). However, a Mann-Whitney U test revealed that pairs performed significantly better than individuals only in vocabulary ($U = 55.00$, $Z = -8.210$, $p = 0.005$, $r = 0.57$) and mechanics ($U = 67.00$, $Z = -2.438$, $p = 0.016$, $r = 0.47$), but not in task achievement ($U = 82.50$, $Z = -1.783$, $p = 0.078$, $r = 0.35$), coherence and cohesion ($U = 102.00$, $Z = -1.009$, $p = 0.323$, $r = 0.20$), and grammar ($U = 92.00$, $Z = -1.397$, $p = 0.168$, $r = 0.28$). The effect size values were large for vocabulary ($r = 0.57$), close to large for mechanics ($r = 0.47$), medium for coherence

and cohesion ($r = 0.35$) ($r = 0.20$), and close to medium for grammar ($d = 0.28$). The findings indicate that collaboration benefits pairs' written texts in vocabulary and mechanics, but not in task achievement, coherence and cohesion, and grammar.

5.6.2 Pairs vs. individuals of intermediate proficiency

Table 5.15 features the descriptive statistics for the performance of pairs and individuals of intermediate proficiency.

Table 5.15 Measures of global rubrics for intermediate learners

	Pairs (16)			Individuals (16)					
	Mdn	Min	Max	Mdn	Min	Max	<i>U</i>	<i>p</i>	<i>r</i>
Task achievement	5	3	7	3.5	2	6	73.5	.033	0.42
Coherence & Cohesion	4	2	7	3	1	5	45	.001	0.64
Grammar	5	3	7	4	2	6	77	.050	0.39
Vocabulary	6.5	4	7	3	2	6	20.5	.001	0.84
Mechanics	6	0	7	4.5	0	7	110	.503	0.14

*Note: r = Effect size

As can be seen from the table, I pairs outperformed I individuals in all qualitative rubrics of task achievement (Pairs, $Mdn = 5$; individuals, $Mdn = 3.5$), coherence and cohesion (Pairs, $Mdn = 4$; individuals, $Mdn = 3$), grammar (Pairs, $Mdn = 5$; individuals, $Mdn = 4$), vocabulary (Pairs, $Mdn = 6.5$; individuals, $Mdn = 3$), and mechanics (Pairs, $Mdn = 6$; individuals, $Mdn = 4.5$). A Mann-Whitney U test showed that pairs significantly outperformed individuals in task achievement ($U = 73.50$, $z = -3.210$, $p = 0.033$, $r = 0.42$), coherence and cohesion ($U = 45.00$, $z = -8.210$, $p = 0.001$, $r = 0.64$), grammar ($U = 77.00$, $z = -1.979$, $p = 0.050$, $r = 0.39$), and vocabulary ($U = 20.50$, $z = -4.123$, $p = 0.001$, $r = 0.84$), but not in mechanics ($U = 110.0$, $z = -688$, $p = 0.503$, $r = 0.14$). Furthermore, the effect sizes for coherence and cohesion ($r = 0.64$) and vocabulary ($r = 0.84$) were large, medium for task achievement ($r = 0.42$) and grammar ($r = .0.39$) and small for mechanics ($r = 0.14$). These findings revealed that collaboration benefits intermediate adolescent EFL written texts measured through global rubrics of task achievement, coherence and cohesion, grammar and vocabulary, but not in mechanics.

5.6.3 Pairs vs. individuals of high-intermediate proficiency

Table 5.16 shows the descriptive statistics for the performance of pairs and individuals of high-intermediate language level.

Table 5.16 Measures of global rubrics for high-intermediate learners

	Pairs (16)			Individuals (16)			<i>U</i>	<i>p</i>	<i>r</i>
	Mdn	Min	Max	Mdn	Min	Max			
Task achievement	6	3	7	4	2	7	59.5	.009	0.53
Coherence & Cohesion	5	3	6	3	1	6	61.5	.010	0.51
Grammar	6	5	7	5	3	7	60.5	.009	0.52
Vocabulary	6	2	7	4	2	7	81	.074	0.36
Mechanics	6	0	7	6.5	0	7	116.5	.619	0.08

*Note: *r* = Effect size

As can be seen from the table, HI pairs outperformed HI individuals in almost all rubrics, except for mechanics, namely pairs performed better in task achievement (Pairs, *Mdn* = 6; individuals, *Mdn* = 4), coherence and cohesion (Pairs, *Mdn* = 5; individuals, *Mdn* = 3), grammar (Pairs, *Mdn* = 6; individuals, *Mdn* = 5), vocabulary (Pairs, *Mdn* = 6; individuals, *Mdn* = 4), but not in mechanics (Pairs, *Mdn* = 6; individuals, *Mdn* = 6.5). A Mann-Whitney U test found that the differences between HI pairs and HI individuals in task achievement ($U = 59.50, z = -2.2629, p = 0.009, r = 0.53$), coherence and cohesion ($U = 61.50, z = -2.603, p = 0.010, r = 0.51$), and grammar ($U = 60.50, z = -2.649, p = 0.009, r = 0.52$) were statistically significant favoring pairs. However, pairs failed to obtain significantly higher scores in vocabulary ($U = 81.00, z = -1.808, p = 0.074, r = 0.36$) and mechanics ($U = 116.50, z = -.451, p = 0.619, r = 0.08$). The effect size values for task achievement ($r = 0.53$), coherence and cohesion ($r = 0.51$), and grammar ($r = 0.52$) were large, but they were medium for vocabulary ($r = 0.36$), and small for mechanics ($r = 0.08$). These findings indicated that while the effect of collaboration on the written texts of high-intermediate pairs was large in task achievement, coherence and cohesion, and grammar, it was medium and small on vocabulary and mechanics, respectively.

To summarize, results reported above indicated that collaboration was least beneficial for low-intermediate learners based on the performance of pairs and individuals. It seems that for them collaboration was effective only for two rubrics, namely vocabulary

and mechanics in which pairs significantly outperformed individuals. However, this was not the case for other rubrics in which LI pairs and LI individuals performed similarly. On the other hand, collaboration proved effective for intermediate proficiency learners, as the differences between I pairs and I individuals were statistically significant for all rubrics except for mechanics, where I pairs and I individuals obtained very similar scores. In all the other measures, pairs significantly outperformed individuals. Collaboration proved effective for high-intermediate learners as well. As shown by the descriptive statistics, HI pairs significantly outperformed HI individuals in at least three holistic rubrics, namely task achievement, coherence and cohesion, and grammar. Interestingly, pairs of intermediate and high-intermediate proficiency performed better than individuals of the same language level in vocabulary, but their performance in mechanics was almost identical, suggesting that collaboration benefited mechanics the least. Next I present results concerning global measures within the same writing condition, that is, collaborative or individual.

5.7 Pairs and individuals within their writing conditions

This section addresses RQ2b: Are there any differences between the texts produced by pairs and individuals across three levels of proficiency within their writing conditions? To check for any statistical differences within groups, several Kruskal-Wallis tests (a type of non-parametric test) were run. For further comparisons, pairwise comparison tests were run to check if significant differences between any of the groups compared existed (i.e. low-intermediate vs. intermediate or high-intermediate learners). The effect size (r) values were reported using Cohen's (1988) guidelines as small $r = 0.10$, medium $r = 0.30$, and large $r = 0.50$, comparing the effect of treatment for each comparison. As mentioned earlier (c.f. Chapter 4, Section 4.11), the effect size values were calculated using the open access statistical analysis software Jamovi Project (2020).

To address this sub-question, I report results drawn from qualitative rubrics for pairs and individuals across three proficiency levels within their writing conditions. First, I report results for pairs of low-intermediate (LI), intermediate (I) and high-intermediate (HI) language level. Second, I report results for individuals of the same language proficiency within their writing condition. Finally, a summary of these findings is provided.

5.7.1 Pairs across three proficiency levels

Table 5.17 presents the descriptive statistics for the performance of pairs across three levels of proficiency, namely low-intermediate (LI), intermediate (I), and high-intermediate (HI).

Table 5.17 Results of global rubrics for pairs across three proficiency levels

Rubrics	LP	Mdn	Min	Max	Z	p	r
Task achievement	LI	3	0	7	-2.803	3-4 = .015	0.40
	I	5	3	7	-1.469	4-5 = .425	0.21
	HI	6	3	7	-4.273	3-5 = .001	0.61
Coherence & Cohesion	LI	2	0	6	-3.549	3-4 = .001	0.51
	I	4	2	7	.786	4-5 = 1.00	0.01
	HI	5	3	6	-4.335	3-5 = .001	0.62
Grammar	LI	2	0	5	-3.878	3-4 = .001	0.54
	I	5	3	7	-1.305	4-5 = .575	0.18
	HI	6	5	7	-5.183	3-5 = .001	0.74
Vocabulary	LI	2.5	0	5	-3.604	3-4 = .001	0.51
	I	6.5	4	7	-1.368	4-5 = .541	0.19
	HI	6	2	7	-4.972	3-5 = .001	0.71
Mechanics	LI	3	0	6	-2.243	3-4 = .075	0.32
	I	6	0	7	-.594	4-5 = 1.00	0.08
	HI	6	0	7	-2.837	3-5 = .014	0.40

LI= low-intermediate; I = intermediate; HI = high-intermediate; Z = scores of follow-up tests.

As can be seen from the table, pairs of higher proficiency performed better than pairs of lower proficiency in all qualitative measures of task achievement LI ($Mdn = 3$), I ($Mdn = 5$), HI ($Mdn = 6$), coherence and cohesion LI ($Mdn = 2$), I ($Mdn = 4$), HI ($Mdn = 5$), grammar LI ($Mdn = 2$), I ($Mdn = 5$), HI ($Mdn = 6$), vocabulary LI ($Mdn = 2.5$), I ($Mdn = 6.5$), HI ($Mdn = 6$) and mechanics LI ($Mdn = 3$), I ($Mdn = 6$), HI ($Mdn = 6$). A Kruskal-Wallis test result indicated that the differences between pairs in task achievement were statistically significant $H(2) = 18.84$, $p = .001$. Consequently, pairwise comparisons with adjusted p -values showed that HI and I pairs performed significantly better and obtained higher scores in task achievement than LI pairs (HI-LI $p = .001$, $r = 0.61$; I-LI $p = .015$, $r = 0.40$) with a moderate effect. However no significant differences were observed between HI and I pairs in the same measure (HI-I $p = .425$, $r = 0.21$) indicating a small effect. A significantly different result was obtained for the

measure of coherence and cohesion $H(2) = 21.33, p = .001, p = .001$. Hence, pairwise comparisons with adjusted p -values indicated that HI and I pairs reached significantly higher scores in coherence and cohesion than LI pairs (HI-LI $p = .001, r = 0.62$; I-LI $p = .001, r = 0.51$) with a large effect size. However, the differences between scores of HI and I pairs in the same measure (HI-I $p = 1.00, r = 0.01$) were not statistically significant with almost no effect.

Pairs reached statistically significant scores in grammar $H(2) = 29.06, p = .001$ as well. Thus, follow-up tests with adjusted p -values showed that HI and I pairs performed significantly better than LI pairs in grammar (HI-LI $p = .001, r = 0.74$; I-LI $p = .001, r = 0.54$) with a large effect size. Nevertheless, the performance of HI and I pairs in the same measure was not statistically significant (HI-I $p = .575, r = 0.18$) showing also a small effect.

The performance of pairs also differed in qualitative measure of vocabulary. A Kruskal-Wallis test result revealed that the scores obtained among pairs of different proficiency were statistically significant $H(2) = 26.39, p = .001$. Thus, pairwise comparisons with adjusted p -values showed that HI and I pairs performed significantly better than LI pairs in vocabulary (HI-LI $p = .001, r = 0.72$; I-LI $p = .001, r = 0.51$) with a large effect size. Yet, the difference between scores obtained by HI and I pairs in the same measure was not statistically significant (HI-I $p = .541, r = 0.19$) with a small effect size.

Finally, pairs performed differently in the rubric of mechanics as well. A Kruskal-Wallis test indicated that the performance between pairs in terms of mechanics was statistically significant $H(2) = 8.95, p = .039$. Consequently, pairwise comparisons with adjusted p -values revealed that HI and I pairs obtained higher scores than LI pairs in mechanics. However, only the difference between HI and LI pairs was statistically significant (HI-LI $p = .014, r = 0.40$). The same trend was not observed between HI and I (HI-I $p = 1.00, r = 0.08$), and I and LI (HI-LI $p = .075, r = 0.32$).

To sum up, HI and I pairs outperformed LI pairs in all qualitative measures. Thus, HI and I pairs obtained higher scores than LI pairs in task achievement, coherence and cohesion, grammar, vocabulary and mechanics. However, the differences between HI and I pairs in terms of scores obtained were not statistically significant in all measures, this was especially true for scores obtained in mechanics. On the other hand, pairs of

high-intermediate and intermediate proficiency did significantly better than pairs of low-intermediate proficiency, a trend observed in quantitative rubrics as well. In the same vein, the performance of HI and I pairs was very similar, as indicated by the lack of significant differences among these two categories. Likewise, a similar trend was detected in quantitative rubrics. Finally, all these findings suggested that pairs' proficiency affected their performance in qualitative rubrics, implying that the larger the differences in proficiency the bigger the differences in their performance.

5.7.2 Individual writers across three proficiency levels

Table 5.18 features the descriptive statistics for learners working in individually across three proficiency levels.

Table 5.18 Measures of global rubrics for individuals across three proficiency levels

	LP	Mdn	Min	Max	Z	p	Effect size <i>r</i>
Task achievement	LI	2.5	0	7	-2.472	3-4 = .013	0.35
	I	3.5	2	6	-.560	4-5 = 1.00	0.08
	HI	4	2	7	-3.032	3-5 = .002	0.43
Coherence & Cohesion	LI	1	0	5	-2.315	3-4 = .062	0.33
	I	3	1	5	-1.364	4-5 = .517	0.19
	HI	3	1	6	-3.679	3-5 = .001	0.52
Grammar	LI	1.5	0	4	-3.658	3-4 = .001	0.52
	I	4	2	6	-1.324	4-5 = .556	0.19
	HI	5	3	7	-4.983	3-5 = .001	0.63
Vocabulary	LI	1	0	4	-3.373	3-4 = .002	0.48
	I	3	2	6	-1.27	4-5 = .687	0.17
	HI	4	2	7	-4.58	3-5 = .001	0.66
Mechanics	LI	0	0	5	-2.935	3-4 = .010	0.42
	I	4.5	0	7	-1.201	4-5 = .689	0.17
	HI	6.5	0	7	-4.137	3-5 = .001	0.59

LI= low-intermediate; I = intermediate; HI = high-intermediate; Z = scores of follow-up tests.

As can be seen from the table, individuals of higher proficiency outperformed those of lower proficiency in all qualitative measures of task achievement LI (*Mdn* = 2.5), I (*Mdn* = 3.5), HI (*Mdn* = 4), coherence and cohesion LI (*Mdn* = 1), I (*Mdn* = 3), HI (*Mdn* = 3), grammar LI (*Mdn* = 0), I (*Mdn* = 2), HI (*Mdn* = 3), vocabulary LI (*Mdn* = 1), I (*Mdn* = 3), HI (*Mdn* = 4), and mechanics LI (*Mdn* = 0), I (*Mdn* = 4.5), HI (*Mdn* = 6.5). To

establish whether the differences observed between groups of learners were statistically significant, several Kruskal-Wallis tests were run. These tests indicated that there were statistically significant differences between groups of learners in task achievement $H(2) = 10.41, p = .005$, coherence and cohesion $H(2) = 13.83, p = .001$, grammar $H(2) = 26.64, p = .001$, vocabulary $H(2) = 22.54, p = .001$, and mechanics $H(2) = 18.11, p = .001$.

More specifically, pairwise comparisons with adjusted p -values showed that HI and I individuals performed significantly better than LI individuals in task achievement (HI-LI $p = .002, r = 0.43$; I-LI $p = .013, r = 0.35$) with a medium effect size. However, the difference in the scores obtained by HI and I individuals in the same measure was not statistically significant (HI-I $p = 1.00, r = 0.08$) with a small effect.

A similar result was observed in coherence and cohesion. Thus, pairwise comparisons with adjusted p -values indicated that HI and I individuals outperformed their LI peers in the qualitative measure of coherence and cohesion. In particular, HI and I individuals obtained significantly higher scores than LI individuals with large and medium effect, respectively (HI-LI $p = .001, r = 0.52$; I-LI $p = .062, r = 0.33$). However, the difference in the scores obtained by HI and I individuals in the same measure was not statistically significant (HI-I $p = .517, r = 0.19$) showing thus a small effect.

HI and I individuals performed better than LI individuals in the qualitative measure of grammar. Pairwise comparisons with adjusted p -values showed that the difference in the scores obtained by HI, I and LI individuals was statistically significant (HI-LI $p = .001, r = 0.63$; I-LI $p = .001, r = 0.52$) with large effect. Nevertheless, as with the two previous measures, the difference between scores obtained by HI and I individuals was not statistically significant (HI-I $p = .556, r = 0.19$) indicating a small effect size.

Last but not least, HI and I individuals performed significantly better than LI individuals in vocabulary. Thus, as illustrated by pairwise comparisons with adjusted p -values the difference between HI and LI individuals (HI-LI $p = .001, r = 0.66$), and I and LI individuals (I-LI $p = .002, r = 0.48$) was statistically significant with a large effect size. However, no significant difference was observed between HI and I individuals in the same measure (HI-I $p = .687, r = 0.17$) with a small effect.

Finally, HI and I individuals obtained higher scores in the qualitative rubric of mechanics than LI individuals. Hence, pairwise comparisons with adjusted p -values showed that the difference in scores obtained by HI, I and LI individuals was statistically significant (HI-LI $p = .001$, $r = 0.59$; I-LI $p = .010$, $r = 0.42$) with large and moderate effect, respectively. Yet, no significant difference in terms of scores obtained by HI and I individuals was observed for the same measure (HI-I $p = .689$, $r = 0.17$) indicating thus a small effect.

To sum up, as indicated above HI and I individuals outperformed LI individuals in all qualitative measures. Thus, HI and I individuals obtained significantly higher scores than LI individuals in task achievement, coherence and cohesion, grammar, vocabulary and mechanics. On the other hand, the performance between HI and I individuals on the same measures was not significantly different, this was particularly true for scores obtained in task achievement and mechanics. A similar trend was maintained in the collaborative writing condition both in quantitative and qualitative rubrics. Hence, the present findings revealed that the deeper the gap in L2 proficiency between learners, the bigger the difference in scores obtained, this was in particular true for scores of HI and LI individuals. All the results reported thus far had to do with the written texts, whereas results presented next treat the peer interaction analyzed through Language-Related Episodes (LREs).

5.8 Results for Language Related Episodes (LREs)

This section addresses RQ3: What are the differences between LREs produced by pairs in terms of number, type and resolution?

With regards to data distribution, a Shapiro-Wilk test indicated that data were approximately normally distributed ($p < .05$) (for more information check Appendix G). Hence, to check for any statistical differences within groups, several One-Way ANOVA tests were run, where the alpha level was set at $< .05$, while the effect size values using the Cohen's d benchmarks as small: $d = 0.20$, medium $d = 0.50$ and large $d = 0.80$ for each comparison (e.g. low-intermediate vs. intermediate or high-intermediate) were reported. Tukey post-hoc tests were run to check for further differences between groups of three proficiency levels (i.e. low-intermediate vs. intermediate or high-intermediate).

To address this question, I report results concerning the number, type and outcome of LREs produced by 24 pairs across three proficiency levels.

5.8.1 Number of LREs

Table 5.19 presents the descriptive statistics for the number of LREs produced by all pairs.

Table 5.19 Number of LREs produced by pairs across three proficiency levels

Pairs	LREs	M	SD	Min	Max	<i>p</i>	<i>d</i>
LI (8)	70	8.75	4.89	1	17	LI-HI $p=.011$	0.84
I (8)	110	13.75	6.29	7	24	LI-I $p=.229$	0.84
HI (8)	146	18.25	6.36	10	29	I-HI $p=.229$	0.76
24	326	13.58	5.85	1	29		

As can be seen from the table, 24 pairs produced 326 LREs altogether. From this amount, HI pairs generated the highest number of LREs ($M = 18.35$, $SD = 6.36$), followed by I pairs ($M = 13.75$, $SD = 6.29$), and LI pairs ($M = 8.75$, $SD = 4.89$) who produced the smallest number of LREs. The number of LREs produced by pairs was statistically significant, as shown by the One-way ANOVA test result $F(2, 21) = 5.20$, $p = .015$. Furthermore, A Tukey post-hoc test found that HI pairs produced a significantly higher number of LREs than LI pairs (HI-PI $p = .011$, $d = 1.61$) with a very large effect size. However, no significant differences were observed between LI and I pairs (LI-I $p = .229$, $d = 0.84$), and HI and I pairs (HI-I $p = .299$, $d = 0.74$) despite a large and moderate effect size, respectively. The number of LREs generated across three proficiency levels varied from 1-17 LREs for low-intermediate pairs, 7-24 LREs for intermediate pairs, and 10-29 for high-intermediate pairs.

5.8.2 Type of LREs

Table 5.20 displays the descriptive statistics for the type of LREs produced by all pairs.

Table 5.20 Type of LREs produced by 24 pairs

LREs	N	M	SD	Min	Max	<i>p</i>	<i>d</i>
Lexical	146	18.25	3.73	13	23	L-G $p=.004$	1.84
Grammatical	90	11.25	4.23	3	18	L-M $p=.004$	1.84
Mechanical	90	11.25	3.45	6	15	G-M $p=1.00$	0
	326	13.58	3.80	3	23		

Concerning the type of LREs, as can be seen from the table, the largest number of LREs was lexical ($M = 18.25$, $SD = 3.73$), followed by grammatical ($M = 11.25$, $SD = 4.23$) and mechanical ($M = 11.25$, $SD = 3.45$) the number of which was equal. A One-Way ANNOVA analysis indicated that the number of LREs produced according to their type was statistically significant $F = (2, 21) = 8.95$, $p = .002$. A follow-up Tukey post-hoc test showed that the number of lexical LREs was significantly higher than grammatical and mechanical LREs (L - G $p = .004$, $d = 1.84$; L - M $p = .004$, $d = 1.84$) with a very large effect size. On the other hand, no significant difference was observed between grammatical and mechanical LREs (G - M $p = 1.00$, $d = 00$) with no effect at all.

5.8.3 Resolution of LREs

Table 5.21 presents the descriptive statistics for the resolution of LREs produced by pairs across three proficiency levels.

Table 5.21 The outcome of LREs

Type	Outcome	LI = 8		I = 8		HI = 8		Pairs = 24	
		N	%	N	%	N	%	N	%
L-LREs	Correct	25	58%	30	81%	60	91%	115	79%
	Incorrect	7	16%	6	16%	5	8%	18	12%
	Unresolved	11	26%	1	3%	1	2%	13	9%
Sub-total		43		37		66		146	
F-LREs	Correct	10	67%	27	90%	37	82%	74	82%
	Incorrect	5	33%	3	10%	7	16%	15	17%
	Unresolved	0	0	0	0%	1	2%	1	1%
Sub-total		15		30		45		90	
M-LRES	Correct	11	92%	31	72%	29	83%	71	79%
	Incorrect	0	0%	9	21%	5	14%	14	16%
	Unresolved	1	8%	3	7%	1	3%	5	6%
Sub-total		12		43		35		90	
Outcome	Correct	46	66%	88	80%	126	86%	260	80%
	Incorrect	12	17%	18	16%	17	12%	47	14%
	Unresolved	12	17%	4	4%	3	2%	19	6%
Total LREs		70		110		146		326	

LI = low-intermediate; I = intermediate; HI = high-intermediate

As can be seen from the table, pairs altogether managed to resolve correctly about 80% of all LREs produced, while resolving incorrectly 14 %, and leaving unresolved only 6 %. As per resolution of LREs according to their type, pairs were most successful at resolving grammatical issues, reaching a correct resolution rate of 82%, followed by lexical and mechanical LREs with a similar correct resolution rate of 79% respectively.

These findings indicate that working in pairs is beneficial in terms of solving language-related problems that emerge while co-authoring a text. Thus, 24 pairs altogether managed to resolve correctly over 80% of the language-related problems raised during the process of text co-construction. On the other hand, only a small proportion of language issues, around 14 %, were resolved incorrectly and even fewer were left unresolved, counting only about 6%. As per resolution of LREs according to their type, the pairs were most successful at resolving correctly grammatical LREs,

followed by lexical and mechanical LREs. Thus, the findings reported above show that pairs were not only able to talk about various language issues while writing together, as reflected by different type of LREs produced, but they also resolved correctly almost all the linguistic issues they encountered during the text co-construction process. Next the results concerning the outcome of LREs for each proficiency level are presented.

5.9 LREs across three proficiency levels

This section addresses RQ3a: Are there any differences between LREs produced by pairs across three levels of proficiency?

To address this sub-question, I report results regarding the number, type and outcome of LREs produced by pairs across three proficiency levels, namely pairs of low-intermediate, intermediate and high-intermediate L2 proficiency.

5.9.1 Number and type of LREs

Table 5.22 presents the descriptive statistics for the number and type of LREs produced across three proficiency levels.

Table 5.22 Number and type of LREs across three proficiency levels

Pairs	L-LREs			F-LREs			M-LREs			LREs
24	N	M	SD	N	M	SD	N	M	SD	N
LI = 8	43	5.37	2.97	15	1.87	2.41	12	1.5	1.06	70
I = 8	37	4.62	2.06	30	3.75	1.98	43	5.37	3.92	110
HI = 8	66	8.25	3.15	45	5.62	2.55	35	4.35	2.38	146
Total	146	6.08	2.73	90	3.75	2.31	90	3.74	2.45	326

As can be seen from the table, low-intermediate (LI) pairs produced overall 70 LREs, 43 of which were lexical ($M = 5.37$, $SD = 2.97$), 15 grammatical ($M = 1.87$, $SD = 2.41$) and 12 mechanical ($M = 1.5$, $SD = 1.06$). A One-Way ANOVA test result revealed a statistically significant difference between the types of LREs produced by LI pairs $F = (2, 21) = 6.92$, $p = .005$. Consequently, a Tukey post-hoc test result showed that the number of lexical LREs was significantly higher than grammatical and mechanical LREs (L-G $p = .016$, $d = 1.52$; L-M $p = .008$, $d = 1.68$) with a large effect size, respectively. However, no significant difference was found between grammatical and mechanical LREs (G-M, $p = .943$, $d = 0.16$) showing thus a small effect size. These

findings indicate that LI pairs dealt more with language issues related to lexical aspects than grammatical and mechanical ones as they completed their task in pairs.

On the other hand, intermediate (I) pairs produced a larger amount of LREs, that is, they produced 110 LREs in total, 43 of which were mechanical ($M = 5.35$, $SD = 3.92$), 37 lexical ($M = 4.62$, $SD = 2.06$), and 30 grammatical LREs ($M = 3.75$, $SD = 1.98$). Further, a One-Way ANOVA test result revealed significant differences between the types of LREs produced by I pairs $F = (2, 21) = 4.23$, $p = .029$. A Tukey post-hoc test result indicated that only the differences between mechanical and lexical LREs were statistically significant (M-L $p = .025$, $d = 1.42$) with a large effect size. However, the differences between lexical and grammatical (L-G $p = .155$, $d = 0.96$), and grammatical and mechanical LREs (G-M $p = .634$, $d = 0.46$) were not statistically significant. These findings indicate that I pair learners discussed far more often about mechanical than grammatical and lexical LREs while co-constructing their texts.

Finally, high-intermediate (HI) pairs produced the highest number of LREs compared to low-intermediate (LI) and intermediate (I) pairs. Thus, HI pairs produced 146 LREs overall, 66 of which were lexical ($M = 8.25$, $SD = 3.15$), 45 grammatical ($M = 5.65$, $SD = 2.55$), and 35 mechanical ($M = 4.35$, $SD = 2.38$). A One-Way ANOVA test result indicated a statistically significant difference between the types of LREs produced by high-intermediate (HI) pairs $F = (2, 21) = 4.23$, $p = .029$. Thus, a Tukey post-hoc test result revealed that the number of lexical LREs was significantly higher than the number of mechanical LREs (L-M, $p = .025$, $d = 1.42$) with a large effect size. However, the test failed to detect any significant differences between the frequency of lexical and grammatical, and grammatical and mechanical LREs (L-G, $p = .155$, $d = 0.96$; G-M, $p = .634$, $d = 0.46$) despite having large and moderate effect size values, respectively. These findings indicate that HI pairs devoted more attention to lexical concerns than to grammatical and mechanical ones while completing their task.

5.9.1.1 Time, turns and LREs

Table 5.23 presents the descriptive statistics concerning the time pairs spent on task, and the number of turns and LREs produced by pairs across three proficiency levels.

Table 5.23 Time and Turns of LREs

	Pairs	L	G	M	LREs	Resolved (v) %	Time	Turns	
Low-intermediate	Led & Sal	3	2	4	9	9	100%	10.22	40
	Alt & Kull	0	0	1	1	1	100%	17.59	3
	Aid & Is	5	2	1	8	5	63%	18.4	31
	Sej & Kqi	10	5	2	17	11	65%	21.4	100
	Hed & Mur	5	0	1	6	2	33%	16.44	24
	Hek & Ram	7	6	1	14	8	57%	13.01	61
	Nar & Gash	6	0	1	7	4	57%	17.45	39
	Orgs & Hab	7	0	1	8	6	75%	26.53	44
Sub-total	43	15	12	70	46	66%	141	342	
Intermediate	Ald & Spah	5	4	4	13	13	100%	18.02	76
	And & Meh	7	7	7	21	19	90%	17.39	114
	Bli & Dak	3	2	6	11	8	73%	11.49	44
	Ble & Vlla	5	4	0	9	4	44%	19.05	35
	Gla & Hajr	4	2	1	7	7	100%	18.58	26
	Hen & Vle	8	5	11	24	18	75%	23.45	115
	Jon & Syl	2	5	10	17	14	82%	14.22	93
	Sar & Zek	3	1	4	8	5	63%	19.58	44
Sub-total	37	30	43	110	88	80%	141.78	547	
High-intermediate	Ani & Lim	6	4	7	17	15	88%	17.56	72
	And & Uki	8	5	4	17	16	94%	20.53	80
	Rio & Ram	11	8	7	26	20	77%	19.28	163
	Din & Jus	6	9	4	19	13	68%	19.4	106
	Ema & Byt	14	8	7	29	27	93%	20.58	191
	Ili & Kla	6	3	1	10	7	70%	18.06	57
	Fio & Pit	5	6	3	14	14	100%	16.1	98
	Dea & Mal	10	2	2	14	14	100%	21.11	79
Sub-total	66	45	35	146	126	86%	152.62	846	
Total	146	90	90	326	260	80%	435.44	1735	

*Note. L = lexical, G = grammatical, M = mechanical.

As can be seen from the table, regardless of their language proficiency pairs spent very similar amounts of time on task. Specifically, LI pairs spent 141 minutes ($M = 17.63$, $SD = 4.95$), I pairs 141.8 minutes ($M = 17.52$, $SD = 3.58$), and HI pairs 152.6 minutes ($M = 19.07$, $SD = 1.72$).

A One-Way ANOVA test result detected no statistically significant differences between the amounts of time pairs across three proficiency levels spent on task $F = (2, 21) = .390, p = .628$. However, different results were revealed when the number of turns of pairs across three proficiency levels was tallied. That is, HI pairs generated the highest number of turns 846 ($M = 105.75, SD = 47.05$), followed by I pairs 547 ($M = 68.37, SD = 35.89$) and LI pairs 342 ($M = 42.75, SD = 28.54$).

One-Way ANOVA test result indicated a statistically significant difference between the number of turns produced by pairs across three proficiency levels $F = (2, 21) = 5.580, p = .011$. Specifically, this difference accounts only for LI and HI pairs (LI-HI, $p = .009, d = 1.61$) with large effect, but not for LI and I (LI-I, $p = .384, d = 0.67$) or I and HI pairs (I-HI, $p = .144, d = 0.96$), respectively. Thus, these findings indicate that the number of turns generated while working together differed across three proficiency levels, suggesting a higher number of turns for HI pairs followed by I and LI pairs, who produced not only the smallest number of LREs but also generated the fewest number of turns, despite spending a similar amount of time on task.

Thus, in line with the time devoted to LREs, LI pairs produced the lowest number of LREs, varying from 1-17, and the number of turns they generated ranged from 3-100. Similarly, there was a variation between the amounts of time LI pairs spent on task, that is, it varied from 10.22 – 26.53 minutes. These data suggest for a considerable variation in terms of frequency of LREs, turns and the amount of time LI pairs spent on task.

Pairs composed of intermediate proficiency learners produced a somewhat higher number of LREs, despite a similar amount of time spent on task. Thus, the number of LREs they produced ranged from 7-24, while the number of turns varied from 26-115, and the time spent on task varied from 11.49 – 23.45 minutes. The trend of variation continued even among pairs of intermediate proficiency.

Finally, pairs of high-intermediate proficiency spent a somewhat higher amount of time on task, and this was also reflected in terms of number of LREs produced. Thus, the number of LREs they produced ranged from 10-29, turns varied from 57-191, whereas the time spent on task varied from 16.10 – 21.11 minutes. Thus, similar to the other two proficiency groupings a variation in terms of number of LREs and turns among pairs of high-intermediate was observed.

Taken together, LI pairs mostly focused on lexical rather than grammatical and mechanical issues. This implies that they prioritized meaning over form. On the other hand, I pairs devoted most of their time on mechanical rather than lexical and grammatical issues. They mostly raised and discussed spelling and punctuation issues, then lexical items, and finally grammatical concerns which grabbed the least attention. Furthermore, as in the case of LI pairs, HI pairs mostly discussed lexical issues, then grammatical and mechanical issues. Pairs of this proficiency grouping prioritized meaning over form, whereas mechanical issues did not compose any significant interest to them.

Finally, the time spent on task varied across pairs of three levels of proficiency. This variation was far more reflected in terms of number of LREs and turns generated while dealing with language issues. Thus, as indicated above, the lowest number of LREs and turns was observed among pairs of low-intermediate proficiency learners. This was also true for the time spent on task. On the contrary, a higher number of LREs and turns was observed among pairs of intermediate proficiency members. Finally, the highest number of LREs and turns generated was observed among pairs of high-intermediate proficiency learners. Thus, this finding suggests for a mediating factor other than L2 proficiency among pairs of similar L2 proficiency.

5.9.2 Resolution of LREs

Table 5.24 displays the descriptive statistics concerning the resolution of LREs reached by pairs across three proficiency levels.

Table 5.24 Outcome of LREs for pairs across three proficiency levels

Type	Outcome	LI = 8		I = 8		HI = 8	
		N	%	N	%	N	%
L-LREs	Correct	25	58%	30	81%	60	91%
	Incorrect	7	16%	6	16%	5	8%
	Unresolved	11	26%	1	3%	1	2%
Sub-total		43		37		66	
F-LREs	Correct	10	67%	27	90%	37	82%
	Incorrect	5	33%	3	10%	7	16%
	Unresolved	0	0	0	0%	1	2%
Sub-total		15		30		45	
M-LRES	Correct	11	92%	31	72%	29	83%
	Incorrect	0	0%	9	21%	5	14%
	Unresolved	1	8%	3	7%	1	3%
Sub-total		12		43		35	
Outcome	Correct	46	66%	88	80%	126	86%
	Incorrect	12	17%	18	16%	17	12%
	Unresolved	12	17%	4	4%	3	2%
Total LREs		70		110		146	

LI = low-intermediate; I = intermediate; HI = high-intermediate

As can be seen from the table, pairs of higher proficiency resolved correctly a higher proportion of LREs than pairs of lower proficiency. Low-intermediate (LI) pairs resolved correctly only 66% of LREs, intermediate (I) pairs resolved correctly 80% of LREs, and high-intermediate (HI) pairs resolved correctly 86% of LREs produced. More specifically, LI pairs managed to resolve correctly over 66%, resolved incorrectly 12%, and left unresolved around 12% of LREs produced. Additionally, the LI pairs were most successful at resolving mechanical LREs, followed by grammatical and lexical LREs. Specifically, they resolved correctly 98% of mechanical LREs, 67% of grammatical LREs and 58% of lexical LREs.

I pairs, on the other hand, resolved correctly over 80%, resolved incorrectly around 14%, and left unresolved about 6% of LREs produced. Regarding the resolution of LREs according to their type, I pairs were more successful at resolving correctly

grammatical LREs than lexical and mechanical ones. Thus, they resolved correctly 90% of grammatical LREs, 81% of lexical LREs, and 72% of mechanical LREs.

Finally, HI pairs managed to resolve correctly a fairly high amount of LREs produced, that is, they resolved correctly over 86%, while the proportion of incorrectly resolved was 12%, and left unresolved only 2% which is very small. In terms of resolution of LREs according to their type, HI pairs managed to resolve correctly over 91% of lexical LREs, over 82% of grammatical LREs, and over 83% of mechanical LREs. Thus, in compliance with the time devoted, they were most successful at resolving correctly lexical LREs, followed by mechanical and grammatical LREs.

5.10 Some observations on the nature of pair talk during the CW process

In this section, I provide six examples of pair talk (2 per proficiency level) to illustrate how learners collaborated to solve their linguistic concerns as they were engaged in the task. Collaboration of learners whose LREs were observed was looked at from the perspective of Storch's (2002a) model on patterns of interaction and of the level of learners' engagement in LREs as described in previous research with young and adult EFL/ESL learners (See, for example, Azkarai & Kopinska, 2020; Storch, 2008). In her work Storch (2002a) identified four patterns of interaction, namely collaborative, expert-novice, dominant-dominant and dominant-passive. Further, research with young EFL and adult ESL learners has analyzed the level of learners' engagement in LREs, distinguishing between elaborate engagement in LREs (E LREs) when both learners of the dyad actively involved in the issue being discussed either by providing suggestions, counter-suggestions, questions or proposals. Limited engagement in LREs (L-LREs), on the other hand, were defined those LREs in which only one of the pair members engaged in discussing, suggesting or asking questions about the linguistic issue being discussed while the other pair member did not engage but simply repeated or acknowledged what his/her interlocutor had already said. Finally limited-limited engagement in LREs (L-L LREs) were identified those instances where only one member of the dyad raised an issue about a linguistic concern, but the other member did not engage at all, as a result the learner who raised the question moved on with the task on his/her own (Azkarai & Kopinska, 2020; Storch, 2008). As analyzing the learners' relationship and their level of engagement in LREs was beyond the scope of the present study, these observations tend to only reflect a general trend on the learners' collaboration, i.e. the patterns of interaction they seemed to form, their level of

engagement in LREs, and their nature of interaction during the collaborative dialogue. Specifically, six examples of pair talk have been looked at to illustrate the points mentioned above, which are presented below according to learners' level of L2 proficiency, namely low-intermediate, intermediate, and high-intermediate.

Example 10 presents data from the conversation of Led and Sal, a low-intermediate pair discussing a lexical issue.

Example 10. Talk of Led and Sal, reflecting an expert-novice pattern of interaction

1. Led: Qysh I thojne qantes bag a o po? [How do we say *bag*, bag don't we?]
2. Sal: Po bag, back, B A C K [Yes, bag, back, B A C K]
3. Led: Bag, jo back eshte permas. [Bag, not back, back is for behind – pointing to the back].

As can be seen from the example, learners (Led and Sal) seek assistance from one another to find an adequate word they need to insert in their writing. In doing this, Led being not sure about the meaning of the word *bag* asks Sal for help. Sal helps by confirming the meaning of the word, but is unsure whether it is *bag* or *back*, a context for testing his hypothesis. Led then draws on his previous knowledge and confirms that it is *bag*, and also provides some explanation (meta-talk) concerning the meaning of the word *bag* and *back*, thus making a clear distinction between the two. In this discussion learners seem to form a more collaborative pattern of interaction, namely expert-novice (Storch, 2002a), where Led seems to act as an expert providing some explanations for the word used to his interlocutor encouraging Sal who holds a novice role in the task. In terms of engagement, both learners seem to engage with the task at hand representing thus an elaborate engagement in their LRE (e.g. Azkarai & Kopinska, 2020; Storch, 2008).

Example 11 features data from the conversation of a low-intermediate pair Hed and Mur struggling to complete their task relying extensively on their L1

Example 11. Talk of Hed and Mur, with a dominant-passive pattern of interaction

1. Mur: Chef p'i thom po chef I thojne t'kuzhines, s'po di qysh m'i thane ketij? [I am calling him chef, but chef is used for the person in the kitchen, I don't know how to refer to him – pointing to the person in the picture whose job was a boss]
2. Hed: arsimtar shefit qysh i thojne? Boss?! [Teacher how do we call the person in charge? Boss?!]
3. Teacher: [boss]

4. Mur: po, boss bravo. Une i kam thane chef, po chef o i kuhines, keshtu
qe...! [Yes, boss...hooray. I called him chef, but chef is the one in the
kitchen, so....!]

As can be seen from the example, learners (Hed and Mur) rely almost entirely on their shared L1 as an attempt to overcome the challenges of finding an English word for the Albanian word 'shëf' [boss]. Here (turn 1) Mur tries to find the appropriate word to refer to the person in charge at the bank, but because of his low English proficiency he struggles to express his intended meaning. He seems to engage in private speech which seems to have helped him in mediating this process, since he is convinced the word he is using is not the one he needs. Hed being passive, most likely because of his low English proficiency, unable to assist his partner asks his teacher for help. The teacher responds to his request (turn 3), and Mur feels excited that they managed to find the word they needed to express their meaning, providing some justifications that he knew he was using the wrong word 'chef' for 'boss'. In this conversation learners seem to form a dominant-passive pattern of interaction, where Mur represents the dominant learner taking control of the task and leading its completion. Hed, on the other hand, takes a more passive role and does not offer any suggestions to his interlocutor's call for assistance. He instead asks his teacher about the linguistic problem they were facing. Hence as there is little engagement between learners, this LRE represents a limited elaboration (e.g. Azkarai & Kopinska, 2020; Storch, 2002a).

Example 13 presents data from the conversation of an intermediate proficiency pair Hen and Vle, discussing a verb form issue using their L1 (Albanian)

Example 13. Talk of Hen and Vle, reflecting a collaborative pattern of interaction

1. Hen: a dog came and start, started to bark, barking a qysh me ba?
[barking or how to say/put it]
2. Vle: a dog came and started barking her. E ka leh ata [because the dog
barked at her]
3. Hen: Mire. [OK].

As can be seen from the example, the intermediate learners deliberate over the word form *bark* to be used in their text. Hen initially uses the verb *start*, then modifies it to *started* adding the infinitive, *to bark*, and finally changes it to gerund (implying *started barking*). In this way Hen tests her hypothesis about the word *bark* until she decides to seek help from her interlocutor. Vle then confirms her hypothesis, offering her positive feedback, and also uses L1 to give some justification why they should use *started*

barking (turn 2). Hen accepts the feedback and agrees on the verb form. In this LRE both learners engaged with each other's contribution and thus formed a collaborative pattern of interaction. In this LRE both learners engage with one another suggestions and contribution to solve the linguistic problem they face, as such it is an elaborate engagement. In solving their linguistic concerns learners used both the L1 and L2.

Example 14 features data from the talk of Ald and Spa, a pair of intermediate proficiency discussing a lexical and punctuation issue.

Example 14. Talk of Ald and Spa, with an expert-novice pattern of interaction

15. Ald: Got the bag and the other stuff.... Qysh me thane kur u qu?
[How can I say when she stood up?]
16. Spa: prit qysh e shkruve qitu? [Hold on, let me see how you worded it]
17. Ald: she got down on knees to get the bag and the other stuff
18. Spa: back bonja ma mire, to get the....um....[put it *back* because it is better, to get the....um....]
19. Ald: the purse
20. Spa: she got down on her knees to get the money and the other stuff back, mi marre apet. [to get them back]
21. Ald: aa back? [is it back]
22. Spa: po de, back m'i marre apet. [Yes exactly, to get them back]
23. Ald: ok, a be ba pike? [OK, shall I put a full stop?]
24. Spa: po. [Yes]

In this LRE, the interlocutors Ald and Spa discuss a lexical issue as they attempt to elaborate a point in their text. Once, Ald attempts to write a sentence she faces with difficulties as she is not able to complete it without seeking help from her peer Spa. Perhaps because of lack of confidence in using the L2 or because she wanted to be more clear, she switched to their shared L1 (Albanian) (turn 15) to ask her peer for help concerning the lexical item [stood up]. To be able to help her peer much easily, Spa wanted to check what had already been written and how it was written (turn 16). Ald reads the text aloud (turn 17), then Spa suggests using the word *back* thus providing explicit corrective feedback and giving some justifications why they should use the word *back* (turn 18), and uses the L1 to elaborate her proposal. Following this, Spa tries to reformulate a sentence on her own perhaps by incorporating the word *back* but as she begins wording the sentence she struggled in finding the appropriate words. To help her, Ald proposes the word *purse* (turn 19), but it seems that Spa wanted a better formulation, which she made (turn 20) correcting her peer's sentence by adding the

pronoun *her*, and used the L1 to convince her interlocutor on her choice. Unsure whether it was *back* Ald had to write, she asked her interlocutor Spa with a clarification request (turn 20) [is it back?]. Spa confirms it (turn 21) with a positive confirmation (feedback) [yes]. Then, Ald wonders whether she should put a full stop or any other punctuation marks (turn 21), where Spa gives her positive confirmation regarding the full stop (turn 22). In this LRE both learners engage in finding a solution for their linguistic concerns (lexical and mechanical issues) to complete their writing task. Although both learners seem to engage in the LRE, it is Spa who takes control over the task and hence the role of the expert, whereas Ald seems to obtain a more subservient role that of a novice learner and does whatever her interlocutor wants her to do. As such, in this LRE there is an elaborate engagement with an expert-novice patterns of interaction between learners.

Example 15 displays data from the talk of Mal and Kel, a pair of high-intermediate proficiency discussing a spelling issue.

Example 15. Talk of Mal and Kel, with a collaborative pattern of interaction

1. Mal: drove away, a shkruhet drove?
2. Kel: drive. Drove
3. Mal: drove away, d-r-o-v-e, drove away
4. Kel: bravo.

As can be seen from the example, learners discuss a spelling issue, more specifically they talk about the spelling of the word *drive* in the past tense. Mal seeks assistance from her interlocutor Kel over the spelling of the word *drove*. Kel supports her interlocutor perhaps relying on her previous knowledge, repeating the word *drive* and *drove* (turn 2). Mal accepts her confirmation (positive feedback) and uses it appropriately. Kel congratulates her partner with an equivalent word in Albanian [bravo – well done]. Again, in this LRE learners form a collaborative pattern of interaction, engaging with each other's contribution by means of suggestions and counter-suggestions representing thus an elaborate engagement in the LRE (Azkarai & Kopinska, 2020; Storch, 2002a; Watanabe & Swain, 2007).

Finally, example 16 derives from the data of Rio and Ram, a high-intermediate pair discussing a grammatical issue, namely the verb form *fall*.

Example 16. Talk of Rio and Ram, with an expert-novice pattern of interaction

- 62 Rio: she got so scared as....
 63 Ram: as her bag fall
 64 Rio: jo vec pak, qysh me thane [hold on for a second, how can I put it]
 such as she.... As her bag fell qysh me than...qysh I thojke [how can I
 put... how can one say it] more fell underground..
 65 Ram: she got...
 66 Rio: she got scared and stressed...
 67 Ram: she got scared and stressed...
 68 Rio: as much as she fell
 69 Ram: as much as she....
 70 Rio: as she fell
 71 Ram: as??
 72 Rio: as her bag fell
 73 Ram: fallt a fell? [fallt or fell?]
 74 Rio: vec fell, m'doket vec fell. [Only fell, I guess, only fell] A fallt? Vete
 B.... a eshte fallt a fell? [Is it fallt? Ask B... is it fallt or fell?]
 75 Ram: po na incizon ky seni [yes, but we're being recorded]
 76 Rio: so gajle....[It doesn't matter]
 77 Ram: oj B... a fell a fallt per qante? Fell [Hey B... is it fell or fallt for
 the bag?]
 78 B... fell (B...responds confirming the verb form fell)
 79 Rio: fell, mire. [Fell, OK].

In this LRE the pair members Rio and Ram discussed a grammatical item to be used in their writing. They discuss with each other for quite a long time but they could not find a solution which satisfied their needs, therefore they decided to ask for help one of their classmates who was not part of their dyad. In this conversation, Rio initially tries wording a sentence by spelling out some words (turn 62), Ram jumps in with another proposal, where Rio partly rejects her peer's suggestion (turn 74) and engages in private speech trying to mediate the process of finding an adequate word/expression for her intended meaning. Interestingly, she uses her L1 as she engages in private speech, and repeats a part of the whole idea by combining both the L1 and L2 as an attempt to facilitate this process. Ram returns the process back to the beginning trying to encourage a solution for their issue where both learners repeat each other's wordings (turns 67-70) other-repetition, confirmation check (turn 71) and comprehension check (turn 73) (Azpilicueta-Martínez & Lázaro-Ibarrola, 2022; Oliver, 2002). Interestingly, as they come close to using the correct verb form of fall *fell*, it is Ram who questions their targeted lexical item to be used in their text. Taking her interlocutor's doubts in mind, Rio got confused and thus asks her peer to seek help from one of the classmates who could confirm or reject one of their options *fallt* vs. *fell*. Being aware they were being

recorded, her interlocutor hesitates to ask for help on the grounds that they were being recorded. Rio acknowledges that she is aware of the recording but encourages her peer to ask the ‘targeted’ classmate about the word. Ram finally asked her classmate about the word who confirmed it was *fell* they should use. In the end, they solved their linguistic problem relying in L1 and L2, but most importantly it was her friend’s help that confirmed their right choice of the word which otherwise became a too complicated issue for both interlocutors. In this LRE learners seemed to form an expert-novice relationship, with an elaborate engagement in the LRE (Azkarai & Kopinska, 2020; Storch, 2002b).

To sum up the examples of pair talk above revealed that the collaborative writing task used in the present study enabled learners to engage in peer interaction and discuss about multiple linguistic issues including lexical, grammatical and mechanical items. While discussing these issues learners overall seemed to form collaborative patterns of interaction and showed elaborate engagement in the LREs discussed. They solved all their linguistic concerns interacting in their L1 and L2, and more importantly their shared L1 facilitated the retrieval of vocabulary items and enabled them to better elaborate their points regarding the issues being discussed. There were times when learners engaged in private speech in the presence of their peers (see examples 11 and 16), which seemed to mediate their learning and solutions for the linguistic problems encountered. Similarly, there were instances where learners relied on their previous L2 knowledge to find solutions for their linguistic uncertainties (example 15). In the same vein, there were instances where learners consolidated their existing and potentially created new knowledge about the target language (examples 10, 11, 15). Finally, despite benefits of collaborative dialogue learners also had to seek help outside their dyads, namely from their teacher and classmates (examples 11 and 16), mostly using their L1, to move the task forward.

5.11 A summary of results

The results for the written texts assessed employing quantitative measures, indicated that pairs produced better texts than individuals in terms of fluency, complexity and accuracy. Specifically, pairs produced longer, linguistically more complex, and grammatically more accurate texts than learners working individually. Nevertheless, only the differences for the accuracy measures and complexity (dependent clauses to total clauses) were statistically significant. In addition, the analysis of the written texts produced by pairs and individuals across three proficiency levels indicated that collaboration did not prove very useful for learners of low-intermediate proficiency. As shown by the results, texts produced by pairs and individuals of low-intermediate proficiency were not statistically significant in any of the analytic rubrics measured. On the other hand, pairs of intermediate and high-intermediate proficiency levels obtained higher scores in terms of fluency, complexity and accuracy than individuals. Additionally, results for the accuracy measures were significantly higher for pairs over individuals from both proficiency levels. High-intermediate pairs also performed significantly better than individuals in terms of complexity (clauses per T-unit). This implies that collaboration and perhaps the L2 proficiency enabled intermediate and high-intermediate learners to produce longer, grammatically more accurate, and linguistically more complex texts than learners with similar language proficiency writing individually.

Similarly, results for pairs and individuals with lower and higher proficiency indicated that both pairs and individuals with higher proficiency outperformed their counterparts with lower proficiency in all quantitative measures, where the differences between low-intermediate and high-intermediate learners were mostly significant regardless of their writing conditions i.e. collaborative or individual. However, the differences in the same measures between intermediate and high-intermediate learners were not always significant, especially in the case of individual writing condition. This suggests that not only the learners' L2 proficiency, but also the working condition (collaborative) had an impact on the outcomes of the written texts composed by adolescent EFL learners.

The results of the written texts using qualitative measures of task achievement, coherence and cohesion, grammar, vocabulary and mechanics indicated that pairs outperformed individuals in all holistic rubrics and the differences in terms of scores

obtained were statistically significant for pairs over individuals in all areas, except for mechanics.

However, the analysis of the written texts produced by pairs and individuals across three proficiency levels revealed a mixture of results. Thus, while low-intermediate pairs outperformed their individual counterparts in all rubrics, they obtained significantly higher scores only in vocabulary and mechanics. Interestingly, pairs and individuals of higher proficiency, performed better in the other areas. Thus, intermediate pairs significantly outperformed individuals of the same proficiency level in all rubrics, except for mechanics. Similarly, high-intermediate pairs significantly outperformed individuals of the same language competence in all rubrics, except for vocabulary and mechanics. These findings indicated that collaboration and L2 proficiency may impact differently on learners' performance across areas of language measured through qualitative rubrics.

With regards to the performance of pairs and individuals within their writing conditions, results showed that both pairs and individuals of higher proficiency outperformed their counterparts of lower proficiency regardless of their writing conditions, collaborative or individual. Thus, HI and I pairs significantly outperformed LI pairs in all qualitative measures. However, the differences between HI and I pairs were not statistically significant in any of the measures employed. Similarly, HI and I individuals significantly outperformed LI individuals in all holistic rubrics. Yet, the differences between the scores obtained by HI and I individuals were not statistically significant in any of the rubrics of assessment.

Regarding the analysis of pair talk, the results indicated that the majority of LREs produced by pairs were lexical, followed by grammatical and mechanical. Additionally, HI pairs produced the largest number of LREs, followed by I and LI pairs who produced the smallest number of LREs. As per the type of LREs produced, it was found that both HI and LI pairs produced more lexical than grammatical and mechanical LREs. I pairs, on the other hand, produced more mechanical than lexical and grammatical LREs.

Regarding the resolution of LREs, pairs across three proficiency levels altogether managed to resolve correctly 80% of all LREs produced compared to incorrectly resolved 14% and left unresolved 6% of LREs. In addition, as pairs' proficiency increases so does their ability to resolve correctly the amount of LREs produced. Thus,

HI pairs resolved correctly 86% of LREs produced, I pairs 80%, and LI only 66%. Additionally, as displayed in table 5.8.1.1 the analysis of oral data revealed a variation in terms of number of LREs and turns produced, amount of time spent on task, and a variation on the rate of resolution of LREs among pairs of similar proficiency. Furthermore and more importantly, learners regardless of their L2 proficiency often relied on their L1 to solve their linguistic or task completion problems. Although pairs of low-intermediate proficiency seemed to make greater use of their L1 (Albanian), which proved beneficial in terms of finding solutions to their linguistic concerns, still there were cases when pairs who could use their L1 sought help beyond their peers to overcome the linguistic challenges they faced. In such cases, they sought help from their teacher indicating that learners, to some extent, consider the teacher's help as more important than that of their peers, avoiding thus their peers' potential help and forming non-collaborative patterns of interaction, i.e. dominant-passive (Storch, 2002b). In light of this, a pair of high-intermediate L2 proficiency, relying mostly on their L1, had to seek help beyond their dyad, namely from one of their classmates to solve the linguistic concern in question. This indicates that when assigning learners to work collaboratively, options beyond pair work, perhaps group work (in groups of three or four) should be considered to enable learners to pool from more resources when faced with language difficulties during task completion (e.g. Bueno-Alastuey & Lizarrondo Larumbe, 2017; Fernández Dobao, 2012, 2014; Villarreal & Munarriz-Ibarrola, 2021). The study also found that learners, whose pair work was observed, seemed to form different patterns of interaction, mostly collaborative (collaborative, expert-novice) but also non-collaborative (dominant-passive). Regardless of the patterns of interaction exhibited, LREs were mostly handled with elaborate engagement, indicating that as learners worked collaboratively they were mutually engaged to solve their linguistic concerns and move ahead with their task. Nevertheless, as indicated earlier, there was a considerable variation in the number and resolution of LREs produced by pairs of similar L2 proficiency. This indicates that not only learners' L2 proficiency, but also the patterns of interaction they seemed to form and potentially the level of their engagement in LREs may account for this variation (Sato & Viveros, 2016). Finally, as mentioned above learners used their shared L1 (Albanian) to discuss the solutions of their language uncertainties, as such the L1 use seemed to have a facilitative role not only in terms of solutions of the linguistic issues encountered, but also to better elaborate their proposals and justifications why a specific option should be incorporated

in their written texts. In a word, all of these variables seemed to play a role in the quantity and quality of collaborative dialogue among learners, namely the type of task learners worked on, their L2 proficiency, the patterns of interaction they formed, and the L1 use to discuss the problematized issues. All of these will be discussed in the next chapter.

Chapter 6: Discussion

6.1 Introduction

The main goal of the present dissertation study was to compare the writing performance of lower-secondary EFL learners working in pairs and individually to complete a story based on a set of photographs. Texts produced by pairs and individuals across three proficiency levels were analyzed in terms of quantitative rubrics of fluency, complexity and accuracy, and in terms of holistic rubrics of task achievement, coherence and cohesion, grammar, vocabulary and mechanics. In addition to comparing the written products, the present study sought to analyze the pair talk for language-related episodes (LREs) to determine the number, type and outcome of LREs produced by pairs across three proficiency levels. Although it was beyond of the scope of the study, six pair talks were also observed for the nature of learner talk i.e. the dyadic patterns of interaction learners form and the L1 use while performing the task.

This chapter will now focus on interpreting and discussing results for each research question and sub-question. First, results concerning the written texts produced by pairs and individuals will be presented and interpreted in line with previous research that measured the writing quality employing quantitative and qualitative rubrics. Then, results for pairs and individuals across three proficiency levels will be discussed and where appropriate interpreted in line with previous studies. Next, results for pairs and individuals within their writing conditions will follow and will be explained how learners' L2 proficiency affected their writing performance across two working conditions (collaborative and individual). Finally, results concerning the number, type and resolution of LREs produced by pairs of three distinct L2 proficiency levels, as well as the nature of learner talk observed will be presented and discussed in line with previous research examining the EFL/ESL learners' pair talk for LREs.

6.2 The performance of pairs vs. individuals on CAF measures

The first research question (RQ1) asks whether there are any differences between the texts produced by pairs and individuals in terms of fluency, complexity and accuracy.

H1: Texts produced by pairs are better in terms of accuracy and fluency than those produced by learners writing on their own.

6.2.1 Fluency

Findings of this study indicated that pairs slightly outperformed individuals on all three fluency measures, that is, words per text, T-units per text, and clauses per text. However, none of these measures reached a statistically significant difference level. The average number of words per text was the only measure to get close to a significant level ($p = .057$), but failed. Overall these findings are in line with those of recent research (e.g. Bueno-Alastuey & Lizarrondo Larumbe, 2017; Soleimani et al., 2017; Villarreal & Munarriz-Ibarrola, 2021) which obtained similar results in favor of pairs over individuals concerning the fluency measures. Nevertheless, findings of the present study do not fully support findings of earlier research (for example, Wigglesworth and Storch, 2009) which revealed a similar performance between pairs and individuals on the fluency measures. The slightly contradictory results between these studies may be related to a number of reasons. While for example, participants in the present study were adolescent intermediate EFL learners from a lower secondary school context, participants in Wigglesworth and Storch's (2009) study were students of advanced proficiency level from an ESL university context.

Learners working in pairs from EFL contexts have been shown to make use of their shared L1 for a range of functions, such as hypothesis testing, initiating and solving language-related problems, offering suggestions and counter-suggestions, and deliberating over language use and task content (Sato and Angulo, 2020). As such, participants of the present study could have used their L1 for the same functions, where its use could have enabled them to discuss more freely over the task content, organization and language use. By making more use of their L1 they could have managed to fully describe the pictures and complete the whole story, producing longer compositions in turn.

Additionally, the writing task differed across the two studies. While in the present study, I used a visual prompt, in their study Wigglesworth and Storch (2009) used an argumentative essay. Writing an argumentative essay collaboratively has been shown to be a more time-consuming activity than composing a story based on a visual prompt. This was reported on a study comparing the writing performance of pairs and individuals across two different tasks (e.g. argumentative essay vs. data commentary report) in terms of CAF measures (Storch & Wigglesworth, 2007). The analysis of oral data in Wigglesworth and Storch's (2009) study revealed that pairs completing an

argumentative essay spent a high proportion of time on brainstorming, generation of ideas, planning and revising, agreeing on content, and language use. This may be different when writing a story based on a set of photographs, in which the generation of ideas might be much easier because learners may pool their language knowledge to describe the pictures and thus build the story while concentrating on the details. These may be the reasons that the results of this study differ from that of Wigglesworth and Storch (2009). That pairs produce longer texts than individuals has been reported in previous research with lower-secondary EFL school learners. Thus in line with findings of the current study, Bueno-Alastuey and Lizarrondo Larumbe (2017) reported longer texts for pairs over individuals in their study with lower-secondary EFL learners aged 13-14 years. Therefore, based on the results obtained, it can be stated that collaboration among adolescent EFL learners benefits their written texts in terms of fluency.

6.2.2 Complexity

Results of the present study revealed that pairs outperformed individuals on the complexity measures of proportion of clauses to T-units and percentages of dependent clauses of total clauses. Specifically, texts produced by pairs had on average two more clauses per text, one more T-unit per text, and about two more dependent clauses per text than texts produced by individuals. However, the differences were not statistically significant in any of these measures. Also the ratio of clauses to T-units and dependent clauses to total clauses was higher in texts produced by pairs compared to those of individuals, but only the ratio of dependent clauses per total clauses was statistically significant. These findings suggest for a positive effect of collaboration in terms of complexity measures. That is, collaboration benefits L2 writing in terms of complexity. That collaboration benefits learners' written products in terms of complexity has been observed in earlier research (Bueno-Alastuey & Lizarrondo Larumbe, 2017; Calzada & García Mayo, 2021; Storch, 2005). Even though pairs produced texts with more complex language than individuals in both studies reported above, no significant differences were found in any of the measures employed. The lack of significant differences may be related to the small sample size of the studies, as claimed by Storch. Yet, another body of research has found limited or no impact of collaboration on the complexity measures (e.g. Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019). The limited impact of collaboration on complexity, which does not fully support the findings of the present study, has been attributed to the type of task used, because

complexity has been claimed to increase only if it is required by the task type or task goals (Pallotti, 2009). Therefore, one reason why findings of the present study are almost completely in line with findings of Bueno-Alastuey and Lizarrondo Larumbe (2017) is that both studies used similar writing prompts. Thus, participants from both studies were of similar age and context (EFL), and had to write a story based on visual prompts. Hence, writing a story based on a set of 18 photographs using the past simple tense, in the case of present study, might have required learners to use more subordinating conjunctions to refer to specific times in the past, and thus might have urged them to use more complex language (sentences) for a better elaboration of their ideas in their compositions. Additionally, interacting in L1 has been shown to affect the complexity of language produced when working collaboratively (Zhang, 2018). Zhang found a correlation between pairs interacting in L1 and the complexity of the language (L2) produced. Thus, the ability of pairs to deliberate over language difficulties in a shared L1 might have enabled them to pool their linguistic resources and construct not only grammatically more accurate but also syntactically more complex sentences to elaborate their points in their written texts.

6.2.3 Accuracy

Findings of the present study indicated that collaboration benefits grammatical accuracy. The present study found that collaboration had a positive effect on the outcomes of the written texts in terms of accuracy measures of error-free T-units to T-units, and error-free clauses to total clauses. Thus, the rate of error-free T-units of total T-units and error-free clauses of total clauses was significantly higher in the collaborative texts than in the individual texts. This finding is in line with the majority of research that compared collaborative work with individual work in terms of the accuracy measures (Sarkhosh & Najafi, 2020b; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009; Zabihi & Bayan, 2020). However, the significant results of the present study make a stronger case than the results of other studies whose statistical results were not significantly different favoring pairs over individuals with regards to the accuracy measures (e.g. Bueno-Alastuey & Lizarrondo Larumbe, 2017; Fernández Dobao, 2012; Storch, 2005). The lack of statistical significance reported in these studies was attributed to the small sample size of the studies. Bueno-Alastuey and Lizarrondo Larumbe (2017), for example, compared only 28 texts (10 written in pairs, 18 individually), Fernández Dobao (2012) compared a somewhat larger sample i.e. 36

texts (15 written in pairs, 21 individually), while Storch (2005) examined only 23 texts (18 written in pairs, 5 individually). Thus the sample sizes of these studies is obviously smaller compared to that of the present study which drew on a corpus of 96 written texts (CW = 48, IW = 48). Therefore, the significant results in the present study may partly be related to the sample size. Nevertheless, when looking at the overall accuracy measures findings of the studies mentioned above and those of the present study point into the same direction, suggesting for a positive effect of collaboration on the accuracy of the written texts. The positive effect of collaboration on the accuracy scores is not surprising given the fact that working in pairs provided learners with opportunities to discuss about the task, organization, deliberate on language use, and pool from their shared L2 knowledge to solve the lexical, grammatical and mechanical issues. It is perhaps because of such opportunities that pairs reached predominantly correct solutions for the majority of problems they encountered (Storch & Wigglesworth, 2007). In the present study, pairs often engaged in collaborative dialogue and reached correct solutions for the majority of the problems (86%) they raised around their writing task. This means that collaborative writing is beneficial for intermediate lower-secondary EFL learners. During the analysis of peer interactions it was revealed that learners involved in LREs provided one another with corrective feedback, discussed lexical, grammatical and mechanical problems, offered alternatives for their problems and resolved correctly almost all language issues they raised. This would not be possible had the learners been working on their own, hence collaborative writing is superior to individual writing in this respect (Storch, 2021). All of these attributes may explain the reason why texts produced by pairs were grammatically more accurate than those produced by individual learners.

In sum, collaboration benefits the learners' written products in all quantitative measures, though to varying degrees. It is clear now that pairs wrote significantly better texts than individuals in terms of complexity (number of dependent clauses per clauses) and accuracy (error-free T-units of T-units, and error-free clauses of clauses). Nevertheless, pairs also performed better in all the other measures, including those concerning the fluency and complexity. Therefore, as hypothesized at the outset, pairs produced better texts than individuals not only in terms of fluency and accuracy but also in terms of complexity. With this in mind, it can be said that the first hypothesis is fully confirmed.

6.3 Pairs vs. individuals across three proficiency levels

The first research sub-question (RQ1a) asks whether there are any differences between the texts produced by pairs and individuals across three proficiency levels.

6.3.1 Pairs vs. individuals of low-intermediate proficiency

Results of the present study showed that low-intermediate (LI) proficiency pairs outperformed individuals of the same language proficiency in all CAF measures. Thus, LI pairs did better in the fluency (words per text, T-units per text, clauses per text), complexity (proportion of clauses to T-units, and proportion of dependent clauses to clauses), and accuracy measures (error-free T-units, and error-free clauses). However, the differences were not statistically significant in any of the above measures. The lack of significant differences in the CAF measures may be related to learners' low L2 proficiency, which may have prevented them to benefit more from collaboration, to perform significantly better than their individual counterparts in any of the areas assessed. Despite the fact that low-intermediate learners working in pairs provided feedback to one another, such feedback was not always accurate (Storch, 1998a), therefore for low-intermediate EFL learners teacher's feedback seems to be necessary to have positive changes in their L2 written production. Additionally, given that the present study measured the writing performance only one time, this may not be enough for low-intermediate learners to benefit from collaboration. A recent study by Sarkhosh and Najafi (2020b) found that low L2 proficiency learners working in pairs significantly outperformed their individual peers in accuracy and fluency measures in the immediate and delayed posttests following six sessions of treatment and provision of unfocused corrective feedback (feedback on all grammatical aspects) to both groups of learners during the study. Hence, based on the results of the present study and those observed in previous research with low L2 proficiency learners (Consolini & Soto-Lucena, 2019; Villarreal & Munarriz-Ibarrola, 2021) it can be said that collaboration does not benefit significantly adolescent low-intermediate EFL learners' written products when measured through quantitative rubrics, at least, not in a single treatment, perhaps because their pace of L2 development is slower when compared with higher L2 proficiency learners. The lack of significant effect of collaboration on the written products of the low-intermediate learners may also be related to the ability of learners to provide feedback to their peers as well as to benefit from the feedback they receive

on their non-target-like utterances. The analysis of the oral data of ESL students receiving feedback on their lexical and grammatical features while engaged in interactive tasks found that learners with low L2 proficiency benefited the least from the feedback offered by their classmates or their teacher (Williams, 2001). In fact, low L2 proficiency learners paid very little attention, if any, to their peers' feedback, which was quite the opposite for the feedback offered by their teacher. This suggests that low proficiency learners do not rely on their partner's help to complete the task, either because they do not trust their peer's L2 language knowledge or because they are unwilling to participate actively in the task, which may lead to non-collaborative patterns of interaction (Watanabe & Swain, 2007). The oral data of Hed and Mur, reported in Chapter 5, section 5.10, example 11, illustrate how these two learners sought help from their teacher instead of pooling their shared language knowledge to overcome a language barrier they faced as they completed the task. Thus, teacher's help seems to be more highly valued by low-intermediate learners than the potential help offered by their peers. As an alternative, assigning low-intermediate learners to collaborative tasks and supplementing with teacher's feedback when necessary seems to be the best option to improve their performance in terms of CAF measures.

6.3.2 Pairs vs. individuals of intermediate proficiency

Results of this study indicated that pairs performed better than individuals of intermediate language proficiency in all CAF measures. However, a closer look of the results showed that in line with previous research (e.g. Sarkhosh & Najafi, 2020b; Storch & Wigglesworth, 2007; Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009) only the differences in the accuracy measures were statistically significant. Furthermore, the findings of the present study suggest that although collaboration benefits mostly the accuracy of the written texts produced by intermediate adolescent EFL learners, it has a positive impact on fluency and complexity as well. This was reported by Sarkhosh and Najafi (2020b) who obtained better scores for pairs over individuals on the fluency and accuracy measures as well as by Zabihi and Bayan (2020) who found that pairs outperformed individuals in all the CAF measures. The positive results for pairs over individuals, especially in terms of accuracy may be due to two possible reasons. First, pairs from this group were of intermediate L2 proficiency, therefore their ability to discuss language issues revolving around grammar and to provide accurate feedback to their peers was higher compared to that of low-

intermediate pairs discussed above (Wigglesworth & Storch, 2012). The analysis of the oral data revealed that intermediate pairs resolved correctly over 90% of grammatical LREs suggesting that as learners' L2 proficiency increases so does their ability to resolve accurately the language concerns they raised (Storch & Aldosari, 2013). Second, learners who worked in pairs could benefit from the collaborative dialogue with their peers not only in discussing the grammatical aspects of the task but also in terms of using the feedback provided by their peers more effectively (Williams, 2001) and thus managed to have a better elaboration of the points discussed by means of producing longer descriptions of their ideas as well as by using more complex language to express their points, including the elaboration of ideas using their shared L1 (Albanian), which has been found to play a facilitative role in terms of L2 task performance (Zhang & Crawford, 2021). This is because pairs, as opposed to individual learners, could assist one another by paying attention to the details of the task, structure of the text, points to be elaborated and language use including attention to words, spelling and capitalization issues (Zabihi & Bayan, 2020). Additionally, their L1 use could have facilitated the discussion of grammatical, lexical and mechanical items (Zhang & Crawford, 2021) incorporated in their texts. All of these elements seems to have advantaged intermediate-level pairs over individuals in reaching a better performance in their written products.

6.3.3 Pairs vs. individuals of high-intermediate proficiency

Findings of the present study showed that high-intermediate pairs outperformed their individual peers in all CAF measures such as fluency, complexity and accuracy. However, only the differences in the accuracy and partly in complexity (clauses per T-units) were statistically significant. That collaboration significantly benefits the written texts in terms of accuracy has been observed in earlier research (Villarreal & Gil-Sarratea, 2019; Wigglesworth & Storch, 2009; Zabihi & Bayan, 2020). However, the effect of collaboration on the other CAF measures such as fluency and complexity has been reported to be more moderate. The present study, in line with Bueno-Alastuey and Lizarrondo Larumbe (2017), and Villarreal and Munarriz-Ibarrola (2021) found that the texts produced by the pairs were longer and more complex in terms of the language produced than those produced by individual learners. This may be due to three possible reasons. Firstly, learners who worked in pairs had the opportunity to benefit from the collaborative dialogue with their peers and to discuss and resolve correctly the majority

of language concerns they raised. In addition, by working in pairs high-intermediate learners could pay more careful attention to all the task content, ideas to be developed, content points to be elaborated and thus could produce written language that was grammatically more accurate, linguistically more fluent, and syntactically more complex than the language produced by individual learners, who had to take care of everything concerning their task completion on their own (Zabihi & Bayan, 2020). Secondly, the study by Williams (2001), reported above, found that higher proficiency learners benefited more from the feedback offered either by their teacher or peers compared to the low proficiency learners. This implies that high-intermediate participants from the present study benefited more from the feedback they received from their peers and used such feedback more effectively, not only in terms of grammatical accuracy, but also in terms of fluency and complexity of the language produced. The study conducted by Calzada and García Mayo (2021) also found that collaboration had a positive effect on the complexity of the texts produced by more advanced young EFL learners (aged 11-12). This suggests that collaboration and the higher L2 proficiency of learners benefit the written texts in terms of complexity measures, which seems to be the case for high-intermediate pairs in the present study. Finally, although all participants, regardless of their L2 proficiency, shared the same L1 (Albanian) and sporadic use of their L1 was observed among pairs, it is not quite clear, based solely on occasional observations, to what extent could their L1 use influence their written text outcomes assessed through CAF measures. Nevertheless, Zhang (2018) found that pairs interacting in L1 produced texts that were syntactically more complex than those interacting in L2. This was attributed to the possibility of pairs to discuss the task content, organization, and language use more freely interacting in their L1 than in L2. Zhang also maintained that interacting in L1 helped learners retrieve much more easily the lexical and grammatical items so that they could create a better elaboration of their ideas using more complex language by means of longer clauses. In a more recent study with adult EFL learners Zhang and Crawford (2021) found that pairs interacting in their shared L1 (Chinese) not only generated but also resolved a significantly higher proportion of LREs than pairs interacting in their target language (English) as they completed a collaborative writing task. Therefore, the high-intermediate pairs in the present study may have used their L1 to find better and more adequate expressions to thoroughly elaborate their points producing longer descriptions using more complex language by means of more clauses to T-units and perform better

than their individual counterparts in all CAF metrics. However, without having a more rigorous analysis over the potential effect of L1 use on the writing performance of adolescent EFL learners assessed through analytic rubrics, the result drawn from simple observations, in particular the result concerning the complexity measures should be interpreted with caution.

6.4 Pairs and individuals within their conditions

The second research sub-question (RQ1b) asks whether there are any differences between the texts produced by pairs and individuals across three proficiency levels within their writing conditions.

Results of the present study showed that both pairs and individuals of higher language proficiency outperformed their peers of lower proficiency in all CAF metrics. In the case of collaborative condition, high-intermediate (HI) pairs significantly outperformed both intermediate (I) and low-intermediate (LI) pairs on the fluency measures of average number of words per text, clauses per text, and T-units per text. However, in the case of T-units per text, significant differences were observed only between HI and LI pairs favoring the former, yet the same did not hold true for the scores between HI and I pairs, and I and LI pairs.

Also, pairs of higher proficiency outperformed their counterparts of lower proficiency in the complexity measures of clauses to T-units, and dependent clauses to clauses. That is, the HI pairs significantly outperformed I and LI pairs in both complexity measures. Similarly, I pairs significantly outperformed LI pairs in both these measures.

Finally, pairs of higher proficiency performed better than those of lower proficiency in the accuracy measures of error-free T-units to T-units, and error-free clauses to clauses. However, results revealed that only the differences between HI and LI pairs were statistically significant favoring the former. Though, with a tendency towards reaching a significant level, the results between HI and I pairs were not significantly different in any of the accuracy measures. These findings indicate that the L2 proficiency of learners had a positive impact on the written text outcomes measured through analytic rubrics of fluency, complexity and accuracy. Previous research that has examined the effect of collaboration and L2 proficiency on the written text outcomes found that learners of higher L2 proficiency outperformed their peers of lower

proficiency in all CAF measures (e.g. Calzada & García Mayo, 2021; Soleimani et al., 2017). Thus findings of the present study, along with those of recent research, suggest that collaboration and language proficiency benefit the adolescent EFL learners' written products in terms of analytic rubrics.

In the case of individual writing condition, slightly different results emerged. On the whole, individuals of higher proficiency outperformed their peers of lower L2 proficiency in all CAF measures. Specifically, HI individuals significantly outperformed LI individuals in the fluency measures of words per text, T-units per text, and clauses per text. However, they did not perform significantly better than I individuals in any of these measures. On the other hand, I individuals significantly outperformed LI individuals in all fluency measures, except for T-units per text.

Concerning the complexity measures, results are very similar. Thus, HI individuals significantly outperformed LI individuals in both the complexity measures of clauses to T-units and dependent clauses to clauses. However, this was not true for the scores between HI and I individuals, although the former performed better, no significant differences were detected. On the other hand, while I individuals significantly outperformed LI individuals in the complexity measure of dependent clauses to clauses, they failed to reach a statistically significant level for the other complexity measure of clauses to T-units.

Finally, learners of higher proficiency in the individual condition outperformed their peers of lower proficiency in the accuracy measures of error-free T-units and error-free clauses. Thus, HI individuals significantly outperformed LI individuals in both accuracy measures. However, no significant differences were detected between HI and I individuals in the same measures, though the former performed better. On the other hand, I individuals significantly outperformed LI individuals in both the accuracy measures.

The better result of higher L2 proficiency learners over their peers of lower L2 is not surprising, because the learners' L2 proficiency has been reported to be an influencing factor affecting the quality of the written products assessed both analytically and holistically. Thus, in a study with lower-secondary school children learning English as a foreign language Martínez (2018) found that intermediate learners produced a significantly higher number of simple, compound, complex, and compound-complex

sentences than their low-intermediate peers. Additionally, texts composed of intermediate learners contained a significantly higher number of both independent and dependent clauses, suggesting thus not only longer but also more complex texts in favor of higher L2 proficiency learners. Interestingly Kyle, Crossley, and Verspoor (2021), who measured the child learners' language growth using complexity indices, found that learners of English as a foreign language with low L2 proficiency (A2) produced more dependent clauses and longer T-units each time their written language was measured during two academic years, indicating that the development of learner language is a gradual process in particular for low proficiency learners. Earlier research has also witnessed that language proficiency can affect the quality of the written texts produced by EFL learners (Sasaki & Hirose, 1996). A study with EFL learners of lower and higher proficiency found that the less proficient learners made more mechanics errors (e.g. spelling and punctuation), grammatical and syntactical errors compared to their more proficient peers (Olsen, 1999). A more recent study by Révész et al. (2022) found that learners' L2 proficiency was the main factor impacting the speed fluency of the texts produced by Chinese ESL university students. That is, as learners' L2 proficiency increased so did their ability to write quickly in the target language and make fewer and shorter breaks during the writing process. Their study also revealed that higher proficiency learners paid more attention to various writing stages such as planning, formulation and revision compared to lower proficiency learners, who were mainly focused on formulation. This implies that more proficient L2 learners can complete their writing task faster, and allocate more time to other writing stages thus reaching a better final product compared to less proficient learners, who most probably due to their lack of linguistic knowledge cannot compete with their peers of higher L2 proficiency. These studies demonstrated that learner L2 proficiency is one important factor affecting the quality of writing, not only as a product but also as a process. In the case of the collaborative condition, the better performance of learners can be attributed to both language proficiency and collaboration among learners, because as indicated in Chapter 5 (Results), the effect size values were higher in the pair condition compared to the individual condition, which seems to suggest that not only the L2 proficiency but also the collaboration among learners played a role in the results of collaborative texts. Whereas in the case of the individual condition, L2 proficiency may be taken as the most influencing factors accounting for the differences in the performance of individual learners across three proficiency levels. However, whether besides L2 proficiency there

are other factors that may affect the quality of writing in L2 remains to be explored (e.g. Manchón & Williams, 2016) in the future. Such inquiry should take into account the learners' L2 proficiency, type of task used, and the writing condition (collaborative vs. individual) employing a longitudinal design to get a fuller picture over the role of these variables in the written text outcomes.

To sum up, results for written texts produced by learners from the collaborative and individual conditions point to the same direction, i.e. learners of higher proficiency outperformed their peers of lower L2 proficiency. Based on these findings it can be assumed that not only the L2 proficiency but also the collaboration among pairs played a role in the adolescent EFL learners' writing performance. Results for pairs and individuals of low-intermediate (LI) proficiency indicated that even though pairs outperformed individuals in all CAF measures, the effect of collaboration was small, as demonstrated by non-significant results between groups. Further, the significant results obtained in the case of collaborative and individual conditions comparing the performance of learners across three proficiency levels suggest that learners' L2 proficiency plays a role in their performance. On the other hand, results obtained from low-intermediate pairs and individuals suggest that despite some positive effects of collaboration on the written products measured through analytic rubrics, collaboration is not very beneficial for low-intermediate learners Wigglesworth and Storch (2012). It thus seems that its effect is greater for intermediate and high-intermediate learners as demonstrated by the results of the present study. Now, the discussion of results obtained from the holistic rubrics follows.

6.5 The performance of pairs vs. individuals on global measures

The second research question (RQ2) asks whether there are any differences between the texts produced by pairs and individuals in terms of holistic rubrics of task achievement, coherence and cohesion, grammar, vocabulary and mechanics.

H2: Texts produced by pairs obtain higher scores for task achievement, coherence and cohesion, grammar and vocabulary.

Findings of the present study revealed that pairs writing a picture story significantly outperformed individuals in all global rubrics of task achievement, coherence and cohesion, grammar, vocabulary, but not in mechanics. In line with previous research (e.g. Chen, 2019; Khatib & Meihami, 2015; Shehadeh, 2011 ; Villarreal & Munarriz-Ibarrola, 2021) pairs obtained higher scores than individuals on global rubrics. Thus, collaboration positively affected the written texts produced by adolescent EFL learners. A comparison of findings of previous studies (e.g. Chen, 2019; Khatib & Meihami, 2015; Shehadeh, 2011; Villarreal & Munarriz-Ibarrola, 2021) showed an overall positive effect of collaboration on the written texts produced by adult and young EFL learners. Findings of these studies, however, are somewhat contradictory. Shehadeh (2011), for example, obtained significantly higher scores for pairs over individuals in the areas of content, organization, vocabulary, but not in grammar and mechanics. Chen (2019), on the other hand, reported a significant effect in favor of pair work on the areas of organization, grammar, vocabulary, but not on content and mechanics. Khatib and Meihami (2015), found a significant effect of collaboration on all measuring rubrics, i.e. content, organization, grammar, vocabulary and mechanics, despite involving participants of low-intermediate L2 proficiency. Villarreal and Munarriz-Ibarrola (2021), who used a similar rubric as the present study, obtained higher scores for collaborating students over non-collaborating ones in task adequacy, coherence and cohesion, grammar, vocabulary and mechanics. However, a significant difference was detected only in the area of vocabulary. Thus taken together, results of the present study confirmed those of previous research. However, the slightly different results of the present study and those reported in Villarreal and Munarriz-Ibarrola (2021) may be related to the smaller sample size of the texts compared and the overall learners' L2 proficiency. Thus, they compared 70 written texts (49 individual, 9 in pairs, 12 in groups). In the current study I compared 96 compositions (CW = 48, IW = 48). The L2 proficiency of their participants was beginner, whereas those in my study

was intermediate. Despite these minor differences, results reported in their study indicated a positive trend in favor of collaboration. Similarly, the contrasting results for grammar reported in Shehadeh (2011) and the present study may be explained by the overall language proficiency of participants involved (low-intermediate vs. intermediate), and the length of the texts they produced. While in Shehadeh's (2011) study, low-intermediate EFL learners were asked to write a paragraph around 100 words in length, in the present study intermediate adolescent EFL learners produced longer texts with an average of 170 words. The length of the texts produced in both studies may also be due to the task differences (short paragraph vs. picture description), which in the case of my study the longer texts produced may have created more opportunities for participants to discuss language issues, pool linguistic resources and achieve higher scores in their performance in terms of grammatical accuracy (Chen, 2019). Finally, the different results in the area of mechanics between Khatib and Meihami (2015) and the present study may be due to the differences in coded errors. In the present study, I counted errors only in capitalization and spelling, while Khatib and Meihami (2015) counted errors in capitalization, spelling, punctuation, and indentation, which represents a wider range of coded errors. As mentioned above, the present study indicates a positive effect of collaboration on the global rubrics with an effect size varying from small to moderate on task achievement ($r = .37$), coherence and cohesion ($r = .36$), grammar ($r = .27$), vocabulary ($r = .47$), but limited or no effect on mechanics ($r = .14$). As in the previous research (e.g. Chen, 2019; Shehadeh, 2011; Villarreal & Munarriz-Ibarrola, 2021) the largest effect of collaboration was observed on vocabulary. This is not surprising, given the fact that participants worked on a meaning-focused task (Storch, 2016) and produced and resolved correctly the highest number of lexical LREs compared to grammatical and mechanical LREs. Finally, the lack of significant or limited effect of collaboration on mechanics may suggest that learners collaborated more on issues related to task achievement, coherence and cohesion, grammar and vocabulary, perhaps because these areas were considered more important than those concerning spelling and capitalization, where the ultimate decision may have been taken by the learner mechanically writing the text on his or her own (e.g. Fernández Dobao, 2012). Thus, taken together, the current findings fully confirm the second hypothesis which proposes that pairs obtain higher scores compared to individuals in the writing dimension of task achievement, coherence and cohesion, grammar and vocabulary. The results reported above account for the performance of pairs and individuals from the

whole dataset. Now, we turn to the performance of pairs and individuals across three proficiency levels.

6.6 Pairs vs. individuals across three proficiency levels

The second research sub-question (RQ2a) asks whether there are any differences between the texts produced by pairs and individuals across three proficiency levels.

6.6.1 Pairs vs. individuals of low-intermediate proficiency

Pairs of low-intermediate (LI) proficiency outperformed their individual peers in holistic rubrics, namely in task achievement, coherence and cohesion, grammar, vocabulary and mechanics. However, significant differences were detected only for scores in vocabulary and mechanics. The lack of significant results for task achievement, coherence and cohesion, and grammar may be related to the pairs' low level of L2 proficiency (e.g. Shehadeh, 2011; Storch, 1998a), as well as to their seemingly greater focus on vocabulary and mechanics which could have absorbed all their cognitive resources, hindering them to focus on other writing dimensions at the same time. It has been suggested that if children focus heavily on one or two particular areas of writing they will be then faced with limited cognitive resources to focus on other higher-level processes of writing such as planning, evaluation and revision (Kojima, In'nami, & Kaneta, 2022). This seems to be one of the key reasons why the pairs obtained significantly higher scores only in vocabulary and mechanics, but not in the other areas. Similarly, it has been claimed that the low level of L2 grammar knowledge, in particular for learners with low L2 proficiency, is a big challenge for writing in L2 (Roca de Larios et al., 2008). Consequently, the learners' low level of L2 proficiency could have resulted in a somewhat similar performance between pairs and individuals in the areas of grammar, task achievement, and coherence and cohesion. Nevertheless, when looking closely at the results concerning the pairs' performance especially in task achievement and grammar, there is a tendency towards a significant level with a moderate effect size in both cases ($r = .37$; $r = .28$), suggesting that although the question whether pairs outperformed individual learners in these areas cannot be answered in terms of statistically significant differences, at least there is a positive trend for the pairs over the individuals in that direction. Despite this, to have a fuller picture over the potential changes towards a greater effect it would be wise to investigate this issue employing a longitudinal design with low-intermediate learners measuring their

performance by means of holistic ratings. On the other hand, in line with previous research (e.g. Khatib & Meihami, 2015) collaboration among pair learners had a positive effect on vocabulary and mechanics with a large effect size in both cases ($r = .57$; $r = .47$). The analysis of oral data uncovered that the LI pairs produced a significantly higher number of lexical LREs than grammatical and mechanical LREs, the number of which was almost identical. This clearly reflects the focus of low-intermediate learners while writing together, that is, they were focused on looking for words to express their ideas. Yet, when looking at the outcome of LREs there is an interesting pattern of resolution across types of LREs. Thus based on the analysis of oral interaction, low-intermediate pairs (LI) resolved correctly only 58% of lexical LREs, but 67% of grammatical LREs, and 92% mechanical LREs. Even at this rate of resolution, it meant that pairs outperformed individuals in the area of vocabulary and mechanics. This means that pairs' focused on those areas they considered easier for them and kept working until they finished their writing task, specifically devoting most of their time and attention to the lexical issues rather than grammatical and mechanical issues, and in this manner their written texts resulted in better quality measured through the global rubrics, especially in vocabulary and mechanics. These findings confirm those of previous research with children aged (11-12) (e.g. Villarreal & Munarriz-Ibarrola, 2021) regarding the effect of collaboration on written text outcomes.

6.6.2 Pairs vs. individuals of intermediate proficiency

Results for pairs and individuals of intermediate (I) proficiency revealed a different picture. Thus, I pairs significantly outperformed their individual peers in almost all holistic rubrics, namely in task achievement, coherence and cohesion, grammar and vocabulary, but not in mechanics. These findings support those of previous research with adult EFL learners (e.g. Chen, 2019) which showed a positive effect of collaboration on organization, vocabulary and grammar, but not on mechanics, despite a sustained and prolonged collaborative experience. Similarly, both earlier and more recent studies have reported non-significant differences between scores obtained by collaborating and non-collaborating students in mechanics (e.g. Fernández Dobao, 2012; Villarreal & Munarriz-Ibarrola, 2021). The positive results of pairs over individuals in the above mentioned areas may be due to two possible reasons. First, working in pairs means having additional resources of working memory compared to working individually (Zabihi & Bayan, 2020). Individual learners can be constrained

when working under the time pressure because they have to pay attention to all the writing task demands of on their own, while pair members can assist one another with such demands. This may be the reason why pairs performed better than individuals in the above-mentioned areas. Second, learners working in pairs can engage in collaborative dialogue (Swain & Watanabe, 2013), which helps them to overcome the challenges of the task completion more easily than working on their own, as pair learners can assist one another with feedback which may be centered on organization, generation of ideas, language use and so on. The feedback exchanged among intermediate pair members seems to be more useful compared to that offered by low-intermediate peers Williams (2001). This seems to be the case with intermediate pairs who could have benefited more from the feedback offered by their intermediate peers, which in turn helped them to outperform their individual learners, who had to rely only on their linguistic resources to complete the task. Conversely, the lack of significant results in the area of mechanics may be due to several reasons. First, the fact that pairs spent a lot of their energy on searching for the meaning of words, expressing their ideas using grammatically accurate language, as well as controlling for the other task dimensions, namely task achievement, and coherence and cohesion of their written texts may have left them with limited cognitive resources (Weigle, 2002) to simultaneously focus on the mechanics of their compositions. Second, prior research (e.g. Fernández Dobao, 2012) has found that collaboration does not boost mechanics, which is the opposite of what I found in the data of low-intermediate pairs reported in the previous section. It seems that low-intermediate adolescent EFL learners perform differently in the area of mechanics not only compared to their intermediate and high-intermediate peers, but also compared to adult EFL learners. This implies that as learners' L2 proficiency increase they tend to focus more on other aspects of writing such as task fulfillment, coherence and cohesion, grammar and vocabulary, hence paying less attention to the spelling, capitalization or punctuation of their compositions. Additionally, it may be that for intermediate and high-intermediate learners, the learner writing the text decides on the final form of the written word (Fernández Dobao, 2012) or the pair members do not put sufficient effort to solve correctly all the mechanical issues they face in the process of text co-construction. To get a fuller picture of this aspect of writing, future studies should incorporate complementary data gathering instruments such as follow-up interviews to get a better understanding as to how persistently pair members work to solve mechanics-focused issues. All in all, what the

current findings indicate is that collaboration benefits intermediate adolescent EFL learners written texts in terms holistic rubrics with a moderate effect on task achievement and grammar ($r = .42$; $r = .39$), with a large effect on coherence and cohesion and vocabulary ($r = .84$; $r = .64$), but with small effect on mechanics ($r = .14$). This indicates that collaboration is more likely to benefit the texts produced by intermediate learners than those produced by low-intermediate learners. This can be said at least for the adolescent EFL learners in the Kosovar context. Now, I discuss the performance of high-intermediate learners.

6.6.3 Pairs vs. individuals of high-intermediate proficiency

Finally, results for pairs and individuals of high-intermediate (HI) proficiency reflected a very similar trend to pairs and individuals of intermediate proficiency. Thus, HI pairs significantly outperformed their individuals in almost all measuring rubrics, i.e. task achievement, coherence and cohesion, grammar, but not in vocabulary and mechanics. The findings indicated that collaboration and perhaps learners' L2 proficiency had a stronger effect on task achievement, coherence and cohesion, and grammar as evidenced by large effect size values in these areas ($r = .53$; $r = .51$; $r = .52$), compared to moderate and small effect sizes in vocabulary ($r = .37$) and mechanics ($r = .08$). These findings support those of Chen (2019), discussed above, who found that students who experienced collaborative writing for a sustained and prolonged period outperformed those who did not in terms of organization, grammar and vocabulary. Nevertheless, the lack of significant results for vocabulary in this study is surprising because even young beginner EFL pairs have been shown to perform significantly better than their individual peers in vocabulary (e.g. Villarreal & Munarriz-Ibarrola, 2021). There are several possible explanations for this result. First, one possible reason may be related to the inability of pairs to find the appropriate words to express their intended meaning, despite reaching correct resolutions for their lexical LREs. The long discussions on lexical LREs, as shown by oral data, suggest that they experienced difficulties in finding appropriate words, and/or agreeing with their partner on which lexical item to use in a given context (c.f. Chapter 4, example 2). Murphy and Roca de Larios (2010) also argue that even advanced L2 learners experience difficulties in finding the appropriate words to express the meaning they have in their minds. As such, despite reaching a correct resolution for an LRE they still may have not found the adequate word to express the meaning they had really intended. The second possible explanation may be related to

the inability of pair members to control for all the task requirements. Thus writing the story while controlling for all the task demands (task achievement, coherence and cohesion, grammar, vocabulary and mechanics) may have consumed most of the pair learners' energy, and may have exerted difficulties in their working memory hindering them to focus on global (e.g. organization, coherence, cohesion) and local errors (errors on words, clauses and sentences) simultaneously (Weigle, 2002). Their inability to control for all types of errors could have led the pairs to produce texts with lexical and mechanical errors, which were statistically very similar to the texts produce by individual learners. Finally, although based on the oral data analysis, it is hard to predict how much time HI pairs devoted to other aspects of writing such as organization, discussion of ideas, moving the task along, revision and the like (McDonough et al., 2016), because it may be the case that pair learners did not spend much time revising their texts with a focus on the lexical and mechanical errors, and thus handed them over to the researcher without checking their texts' quality relating to lexical and mechanical errors. Nevertheless, the overall results favored the pair work over the individual work. The analysis of oral data showed that the overall resolution rate of language-related issues was higher (86%) than in the two previous proficiency groupings, a finding which supports previous research (e.g. Fernández Dobao, 2012; Gallardo del Puerto & Basterrechea, 2021; Storch & Aldosari, 2013) that found that higher proficiency pairs were more successful at resolving language uncertainties than their lower proficiency counterparts. Therefore, it may be assumed that collaboration as well as language proficiency affected positively the written products of adolescent EFL learners assessed through holistic rubrics. Now, we focus on the performance of pairs and individuals within their writing conditions.

6.7 Pairs and individuals within their conditions

The second research sub-question (RQ2b) asks whether there are any differences between the texts produced by pairs and individuals across three proficiency levels within their writing conditions.

Results of the present study indicated that learners of higher L2 proficiency outperformed their peers of lower proficiency in task achievement, coherence and cohesion, grammar, vocabulary and mechanics, regardless of their writing conditions. Thus in the collaborative condition, pairs of high-intermediate (HI) and intermediate (I) proficiency significantly outperformed their peers of low-intermediate (LI) proficiency in all measuring rubrics, except for mechanics. Interestingly, results between HI and I pairs in the same measures were not statistically significant. Similarly, results for the individual condition showed that individuals of high-intermediate (HI) and intermediate (I) proficiency significantly outperformed their peers of low-intermediate (LI) proficiency in almost all qualitative rubrics. Nevertheless, as in the case of pair work, results for HI and I individuals were not statistically significant in any of the rubrics.

Thus, in line with recent research with child EFL learners (aged 11-12) that found that higher L2 proficiency learners significantly outperformed their lower proficiency peers in task adequacy, coherence, cohesion, grammar, vocabulary and mechanics (e.g. Calzada & García Mayo, 2021), results of the present study indicate that higher L2 proficiency learners performed significantly better than their lower proficiency peers in task achievement, coherence and cohesion, grammar, vocabulary and mechanics, regardless of their working condition (collaborative or individual). Nevertheless, the performance of both HI and I pairs, and HI and I individuals was not significantly different in any of the writing dimensions examined. The lack of significant results between high-intermediate and intermediate learners may be due to the potential small gap of L2 proficiency between the learners. That is, as the proficiency level of learners was not determined by means of standardized tests, it is possible that the gap between high-intermediate and intermediate learners is too close to obtain a significantly different result in their performance. The small effect size values observed for the performance between HI and I learners point to the same direction, leading one to speculate on a potential small difference in the L2 proficiency between intermediate and high-intermediate learners.

To sum up, both results of quantitative and qualitative rubrics revealed that higher L2 proficiency learners outperformed their lower L2 proficiency peers, regardless of the working conditions of learners, that is, collaborative or individual. Yet, the performance of HI and I pairs, and HI and I individuals was not statistically significant in any of the rubrics measured. This indicates, as explained above, for a close gap in terms of L2 proficiency between learners. In terms of overall performance, as reported earlier, learners working in pairs performed a bit better than those working individually across all three proficiency levels. Additionally, irrespective of the measuring rubrics, pair work proved to be more effective than individual work, as pairs of higher proficiency obtained higher scores compared to pairs of lower proficiency, as well as compared to individuals of the same proficiency level. This slight difference may be attributed to the collaboration among learners (McDonough et al., 2016; Storch, 2021), as working in pairs created opportunities for learners to pool their linguistic resources and compose texts of slightly better quality than learners working on their own, who had to rely only on their individual linguistic resources to complete the task. It is also possible that pairs regardless of their L2 proficiency formed different patterns of interaction, which could have affected the quality of their written products (Watanabe, 2019). Thus, the pairs with more collaborative orientation have been found to perform better in their tasks (e.g Storch, 2002b; Watanabe, 2019; Watanabe & Swain, 2007) than the pairs with non-collaborative patterns of interaction such as dominant-dominant, and dominant-passive. Although results drawn from observations of the pair talks indicated that learners in the present study formed mostly collaborative patterns of interaction, such results are not enough to draw a robust conclusion concerning the relationship between dyadic patterns of interaction and text quality, especially because no other analysis beyond simple observations on a limited number of pair talks (six) were conducted. Therefore, whether the patterns of interaction learners formed affected the quality of texts produced cannot be directly linked. As such, this issue remains to be looked at more closely employing a more rigorous analysis in the future, while results on the observations of pair talks are presented in a more detailed manner below in section 6.9.2 of the discussion chapter. Now we discuss results obtained from the analysis of pair talk for language-related episodes (LREs).

6.8 The nature and outcome of LREs

The third research question (RQ3) asks whether there are any differences between LREs produced by pairs in terms of number, type and resolution.

Results of the present study indicated that pairs of all proficiency groupings generated a high proportion of LREs, the majority of which were lexical (146) followed by grammatical (90) and mechanical LREs (90), which were at a similar rate. Learners discussed significantly more often about lexical than grammatical and mechanical issues. This finding corroborates those of earlier studies with adult EFL/ESL participants (e.g. Fernández Dobao, 2012; Niu et al., 2018; Storch & Aldosari, 2013; Zabihi & Ghahramanzadeh, 2022), and young EFL learners (e.g. Basterrechea & Gallardo-del-Puerto, 2020; Villarreal & Munarriz-Ibarrola, 2021) that obtained similar results concerning the type of LREs. Nevertheless, the same finding does not fully support earlier research with young, adolescent and adult EFL/ESL and CLIL learners (e.g. Basterrechea & Leaser, 2019; Calzada & García Mayo, 2020a; Leaser, 2004) that found a higher number of discussions about grammatical than lexical and mechanical issues. The slight difference between the results of the present study and earlier studies may be related to the nature of task used. While participants in the present study worked on a picture story task, participants in Basterrechea and Leaser (2019), and Calzada and García Mayo (2020a) completed a dictogloss task. Tasks such as dictogloss, text reconstruction, and text editing are considered as more form-focused, while the jigsaw, picture stories, data commentary report, and composition are perceived as more meaning-based (Storch, 2016). As such, the procedures of task completion differ across the tasks, which in turn may draw learners' attention to meaning and form differently. In a dictogloss task, for example, learners have to follow three procedures to complete the task. First, they listen to a story either read by their teacher or played by an audio player. Second, they listen to the same passage for a second time and take some incomplete notes to facilitate their task completion in the final stage. Finally, they work either individually or in pairs to reconstruct the whole story as accurately as possible (García Mayo, 2018). On the other hand, in a picture story, used in the present study, learners had to write a story based on a visual prompt (a set of pictures) for which they needed to find the appropriate words and expressions to express their ideas and complete the task. Learners completing a dictogloss task are provided with some ideas about the content of the story and the nature of language structure used, where

researchers deliberately embed pre-determined linguistic items and test the effect of such a task on the development/learning of a specific linguistic item (e.g. third person singular –s morpheme). In this manner, learners working on more form-focused tasks such as dictogloss are more likely to pay attention to the specific content and language items they have heard most frequently rather than creating a context for expressing their own words by generating their own ideas about the story they are going to complete (Zabihi & Ghahramanzadeh, 2022). Hence, writing a story based on a visual prompt seems to have triggered learners to talk more about lexical aspects of the target language rather than grammatical and mechanical ones, which helps to explain why adolescent EFL learners produced more lexis-focused than form-focused and mechanics-focused LREs as they co-constructed their stories. Additionally, the number of LREs produced in the present study and those with young and adult EFL learners seem to vary. In Calzada and García Mayo (2020a), for example, the number of LREs produced by 8 pairs was only 88, while in Fernández Dobao (2012) who had 15 pairs was 505, whereas in the present study with 24 pairs was 326. When contrasting the number of LREs produced by a single pair in all three studies, it seems that pairs in Fernández Dobao (2012) produced a greater number of LREs not only compared to pairs in Calzada and García Mayo (2020a) but also to pairs in the present study. This difference may be related to the participants' age, time allocated to complete the task, and to the learners' level of L2 proficiency. Although studies comparing the amount and quality of languaging produced by young, adolescent and adult EFL/ESL learners, to the best of my knowledge, are rare, one recent study that compared the amount of negotiation for meaning between young and adult EFL learners found that adult learners significantly outperformed their younger peers in conversational adjustments operationalized as clarification requests, confirmation checks, and comprehension checks (Azpilicueta-Martínez & Lázaro-Ibarrola, 2022). As such the age difference between participants in the present study and those in Fernández Dobao (2012) and Calzada and García Mayo (2020a) may have affected the amount of LREs produced. Thus in Fernández Dobao's (2012) study, participants were older than those in the present study (adult vs. adolescent) studying a different foreign language in a different learning context (Spanish vs. English; University vs. Secondary School) who completed a similar task. Finally, participants in Calzada and García Mayo (2020a) were, as explained earlier, young EFL learners (aged 11-12 years) of beginner L2 proficiency, who were allocated 15-20 minutes to complete a dictogloss task. Thus when it comes to the differences

between LREs produced by participants in Fernández Dobao and the present study, the findings seem to suggest that the variable of age may have played a role in the quantity of languaging of both groups of learners. As adult learners are thought to have higher level of cognitive maturity compared to younger learners, they are more likely to benefit from their richer L1 literacy skills and probably perform better in L2 tasks compared to young learners who lack such skills at their age (Kojima & Kaneta, 2022). On the other hand, the differences between LREs produced by participants in Calzada and García Mayo (2020a) and the present study may be due to the level of L2 proficiency, time allocated to complete the task, and the nature of task used. Specifically, participants in my study had more time to complete the task (30 minutes), were of higher L2 proficiency, and completed a picture story task. All of these variables may account for the differences between the number and type of LREs produced in these two studies, in particular, the L2 proficiency of learners may have been the most influencing factor, as it has been found to positively affect the quantity and quality of LREs produced by adult EFL/ESL learners (Leeser, 2004; Storch & Aldosari, 2013).

Concerning the resolution of LREs, pairs of all three proficiency levels were successful at resolving correctly the majority of language-related concerns, reaching a correct resolution rate of 80% for all linguistic problems they encountered. These findings support those of previous studies that obtained high successful resolution rates of language-related problems with young and adult EFL/ESL learners (Basterrechea & Leeser, 2019; Fernández Dobao, 2012; Storch & Aldosari, 2013; Villarreal & Munarriz-Ibarrola, 2021) and extend those findings with new evidence drawn from adolescent EFL learners. In fact, Storch and Aldosari (2013) found that their adult EFL learners resolved correctly over 75% of the language-related concerns while completing a composition task collaboratively. Similarly, Fernández Dobao (2012) found that her adult Spanish as a foreign language learners solved correctly around 64% of their language-related uncertainties while completing a picture story collaboratively. In a similar vein, Villarreal and Munarriz-Ibarrola (2021) obtained a high resolution rate (71%) of LREs with young EFL learners completing a narrative essay collaboratively. Therefore, the high resolution rate (80%) of LREs reached in the present study, along with findings of studies referred to above, serves as a sound evidence indicating that learners working in pairs or small groups, regardless of their age, can pool their linguistic resources as well as scaffold one another to reach correct solutions for the

linguistic problems they encounter while working on collaborative tasks. This is because when working collaboratively learners can seek and provide corrective feedback to each other, pool from their collective linguistic resources, propose and implement new ideas, and find new alternative words to express their ideas in a better way (Storch, 2020). Nevertheless despite the benefits offered in collaborative work, it has been found that the success of collaborative work is frequently moderated by the L2 proficiency of learners and relationships they form during collaborative writing. Now we discuss the performance of pairs across three proficiency levels on LREs.

6.9 LREs across three proficiency levels

The third research sub-question (RQ3a) asks whether there are any differences between LREs produced by pairs across three proficiency levels.

Regarding the learners' proficiency as a variable affecting the quantity and quality of LREs produced, results of the present study revealed that pairs composed of high proficiency learners produced a higher number of LREs than pairs of low proficiency learners. Additionally, the majority of LREs produced were lexical especially those belonging to the low and high-intermediate pairs, while the intermediate pairs discussed slightly more about mechanical than lexical or grammatical issues during the task completion process. Regardless of the LRE types, the study found that the higher the L2 proficiency of learners, the greater the number of LREs produced thus confirming results of previous studies with adult EFL/ESL learners (e.g. Kim & McDonough, 2008; Leeser, 2004; Storch & Aldosari, 2013) that found a higher number of LREs for higher proficiency learners. Although studies involving young and adolescent EFL participants investigating learners' L2 proficiency as an influential factor on the amount and outcome of LREs produced are very limited, one study with adolescent EFL learners from a CLIL context that examined the impact of learners' L2 proficiency in the production and outcome of LREs involving the third person singular present tense marker in a collaborative dictogloss task found that the higher proficiency learners produced and resolved correctly more LREs involving the targeted grammatical feature (present tense marker) than the lower proficiency learners (Basterrechea & Leeser, 2019). In this respect, findings of the present study help to counterbalance the scarcity of existing research by bringing new evidence indicating that the L2 proficiency of adolescent EFL learners has a positive impact on the quantity and quality of LREs they

produce. Specifically, higher L2 proficiency learners produce and resolve correctly a higher amount of LREs compared to their lower proficiency peers.

Quite expectedly, and in line with previous research (e.g. Kim & McDonough, 2008; Leiser, 2004; Storch & Aldosari, 2013), the present study found that HI pairs resolved correctly over 86% of LREs, followed by the I pairs who resolved correctly 80% of LREs, and finally the LI pairs who resolved correctly only 66% of LREs produced, suggesting for a positive impact of L2 proficiency on the correct resolution rate of LREs. Interestingly, when the resolution rate of LREs was compared according to their type, pairs of all proficiency levels reached different results. Thus, LI pairs were most successful at resolving correctly mechanical LREs (92%), then grammatical (67%) and lexical LREs (58%). I pairs, on the other hand, were most successful at resolving correctly grammatical LREs (90%) followed by lexical (81%) and finally mechanical LREs (72%). Finally, HI pairs were most successful at resolving correctly lexical LREs (91%) followed by mechanical (83%) and grammatical LREs (82%). The different results concerning the resolution rate of LREs across their types may be related to the level of learners' L2 knowledge and L1 use in such LREs and potentially also the patterns of interaction they formed, as well as the level of learners' engagement in LREs during the collaborative work. As such, LI pairs, for example, may have been more successful with mechanics-focused LREs because both pair members may have been more familiar with the spelling, punctuation and capitalization issues of the L2 words than with L2 vocabulary and grammar rules used to express their ideas while co-constructing the story. Additionally, as low-intermediate learners seemed to have made the greatest amount of L1 use while co-constructing their texts, this could have helped them discuss the mechanics-focused and lexis-focused issues more elaborately and thus reach higher correct resolution rates for the same. Previous research with adult EFL participants found that pair learners interacting in a shared L1 resolved correctly a higher amount of mechanics-focused and lexis-focused LREs than interacting in L2 (Zhang & Crawford, 2021). Additionally, previous research has found that when learners discussing LREs show high level of engagement with one another they are more likely to produce and reach more accurate solutions for such LREs than when they show low level of engagement (e.g. Azkarai & Kopinska, 2020; Storch, 2008; Zabihi & Ghahramanzadeh, 2022). Thus, it may be that the low-intermediate pairs in the present study had a higher level of L2 knowledge on mechanical aspects and thus felt more

willing to discuss more elaborately such LREs than lexical and grammatical LREs, in which their engagement could have been of a more limited nature, and hence their accuracy resolution rate lower. In the same vein, the intermediate pairs could have benefited from their higher L2 proficiency and the ability to discuss and resolve correctly a higher number of grammatical LREs compared to lexical and mechanical ones which were correctly resolved at a somewhat lower rate. Earlier research has shown that as learners' L2 proficiency increases so does their ability to discuss and resolve correctly their language-related problems revolving around grammar (Basterrechea & Leiser, 2019; Leiser, 2004; Wigglesworth & Storch, 2009). Hence, the intermediate pairs in the present study possibly due to their higher L2 knowledge about grammatical aspects could have discussed and engaged more elaborately in LREs revolving around grammar than in the ones centered on lexical and mechanical aspects, and thus reached a higher resolution rate for the former compared to the latter type of LREs. Finally, the higher resolution rate of lexical LREs compared to grammatical and mechanical LREs, which were of similar rate in the data of HI pairs, seem to suggest that the HI pairs were more focused in retrieving key vocabulary items to express their intended meaning and complete their task (García Mayo & Imaz Agirre, 2019) without paying enough attention to the appropriateness of their choice of words or expressions. This in fact could have resulted in a higher amount of discussion and resolution of such language-related problems but those solutions were not the best match for the context in question. Additionally, the L1 use among HI pairs could also account for the different outcomes of LREs across their types. As with LI pairs, HI pairs made sporadic use of their shared L1 (Albanian) while discussing the linguistic concerns, in particular, they used their L1 to elaborate the choice of words and their justifications why a given option should be used. As such, the interaction in their shared L1 could have facilitated the solution of lexical and mechanical LREs as they completed the task (Zhang & Crawford, 2021). Furthermore, previous research examining the patterns of interaction and the level of learner engagement in LREs has found that both these variables affect the quantity and quality of LREs produced in collaborative task (e.g. Azkarai & Kopinska, 2020; Storch, 2008; Zabihi & Ghahramanzadeh, 2022). However, whether these variables affected the different outcomes of LREs across their types is still not clear, because it was beyond the scope of the present study to investigate specifically whether the patterns of interaction learners formed and the level of learner engagement have any impact on the outcome of LREs according to their type. In light of this, it can

be said that based on the observations of a limited number of pair talks reported earlier, it was found that in the majority of pair talks observed, learners seemed to exhibit collaborative patterns of interaction (collaborative, expert-novice) and to show elaborate engagement in their LREs. However, such observations were only made on specific LREs produced by pairs across three proficiency levels, zooming in on one particular LRE per each pair without looking specifically whether a given pair showed elaborate engagement in all the LREs it produced, and whether the same pair kept the same pattern of interaction throughout their discussion. All in all, when analyzing the overall outcome of LREs across three proficiency levels, it was found that all pairs as a group reached higher correct resolution rate for grammar-focused than lexis-focused and mechanics-focused LREs, a finding which supports Basterrechea and Gallardo-del-Puerto (2020), and Villarreal and Munarriz-Ibarrola (2021) with young EFL learners (aged 10-11 and 11-12, respectively). Also, the results of the present study indicated that as learners' L2 proficiency increases so does their ability to discuss and reach accurate solutions for the grammatical aspects of their target language, adding new evidence to the existing body of research (Basterrechea & Leiser, 2019; Leiser, 2004; Wigglesworth & Storch, 2009) which reported higher discussion of form-focused LREs in favor of increased L2 proficiency. This implies that similar to adult EFL peers, adolescent EFL learners of higher L2 proficiency not only produce a higher number of LREs but also resolve correctly a higher amount of such LREs, when they work on a collaborative task. Additionally, although the differences in number, type and outcome of LREs may be primarily attributed to the learners' L2 proficiency, the variation in quantity and quality between LREs produced by pairs of similar proficiency seems to suggest that proficiency cannot be accounted as the sole explanatory factor on its own. As previous studies with young and adult EFL/ESL learners have shown, the relationship learners form and the level of engagement with one another when working in pairs or small groups affect the number, type and outcome of LREs produced while in collaborative tasks (Wenxue Chen, 2017, 2018; García Mayo & Imaz Agirre, 2019; Storch, 2002b, 2008; Watanabe & Swain, 2007). Thus the variation in number, type and outcome of LREs in the present study may be related to similar factors as well. As such, due to limited research with adolescent EFL learners examining the effect of learner relationship and the level of learner engagement on the amount and outcome of LREs, the present study highlights the need for more research with adolescent EFL participants examining a range of variables in collaborative tasks including the patterns of

interaction and the level of learner engagement on the quantity and quality of languaging during collaborative writing tasks. As mentioned above, although it was beyond the scope of the present study to analyze the effect of such variables on learner talk, some observations made on learners' interaction while discussing LREs indicated that noticeable differences in terms of number, type and outcome of LREs produced by pairs composed of similar proficiency learners exist. Finally, these observations suggest that peer interaction, in L1 or L2, helped learners raise, discuss and resolve the language-related problems during the collaborative writing process. In addition, learners not only discussed their linguistic concerns in their shared L1 or L2, but they also sought help beyond their peers when faced with difficulties to move the task forward. Therefore, in the next section I discuss some observations of peer interactions featuring the nature of learner talk while assigned to collaborative work.

6.10 Some insights on the nature of pair talk during the CW process

The main goal of this section is to discuss the nature of learner talk observed during the collaborative writing task used in this study. Thus based on observations of six pair talks (2 per proficiency level), adolescent EFL learners seemed to form three distinct patterns of interaction, namely collaborative, expert-novice and dominant-passive. In the case of low-intermediate pairs, the study found that low-intermediate pairs formed dominant-passive and expert-novice dyadic relationships, respectively. The intermediate pairs, on the other hand, formed collaborative and expert-novice patterns of interaction. Similarly, the high-intermediate pairs formed collaborative and expert-novice relationships during their talk. Unlike the low-intermediate pair with a dominant-passive orientation, who seemed to show limited engagement in their LRE, all the other pairs, regardless of their L2 proficiency, showed elaborate engagement in their LREs by means of suggestions, counter-suggestions, proposals, agreements and disagreements with their pair members in an attempt to find solutions for their linguistic concerns. This finding supports earlier research with child EFL learners (aged 11-12 years) who completed a similar task, but that included an oral plus an oral and a written component as well (García Mayo & Imaz Agirre, 2019). In their study, García Mayo and Imaz Agirre asked child learner participants to interact with their peers to put the cartoon strips in order (oral task), while for the oral plus written component learners were asked to give a written answer regarding their opinion about the potential person who might have broken the glass in the kitchen. Additionally, findings of the present

study generally support those reported in Azkarai and Kopinska (2020) with child EFL learners (aged 10-11), who found that the second major pattern of interaction in their study was collaborative, with the majority of LREs handled with elaborate engagement. In line with their findings, findings obtained in our study indicate that the majority of LREs were handled with elaborate engagement, and the majority of pairs were collaborative (collaborative and expert-novice). The more collaborative nature of pair talk in the present study may be related to the pre-existing friendship between learners as they belonged to the same class and had known and worked with each other in language activities for more than 6 years by the time of data collection. The pre-existing friendship among learners has been found to have a positive impact on their collaboration in collaborative task performance (Gallardo del Puerto & Basterrechea, 2021). Secondly, the nature of the task learners in our study performed could have fostered the collaboration among them. This is because building a story based on a visual prompt could have helped learners to get a common understanding of what should be written based on the pictures given. On the other hand, when learners work on a dictogloss task, which includes a listening component, they may have different or even contradictory opinions concerning the content points or ideas they have heard. Therefore, collaboration in a dictogloss task may be much more complex than on a picture story task. This may explain the difference between findings reported in Azkarai and Kopinska (2020), who also do acknowledge that the task features may impact the learners' patterns of interaction. Furthermore, regardless of the L2 proficiency, the majority of learners seemed to be actively in the task and to contribute in the solution of their mutual linguistic concerns while completing the task. This reflects a kind of collaborative mindset exhibited among adolescent EFL learners (Sato, 2017), which has been claimed to be a stronger mediator than the L2 proficiency of learners in terms of raising and resolving language-related problems during task completion (Watanabe, 2008). Previous research has found that the more collaborative the learners are, the greater the number of LREs they generate and resolve correctly (e.g. Storch, 2002b; Watanabe & Swain, 2007). With that in mind, it can be assumed that the more collaborative patterns of interaction exhibited in this study (collaborative, expert-novice) may account to a certain degree for the large number of LREs produced and correctly resolved during the collaborative task completion. Moreover, compared to other patterns of interaction i.e., dominant-dominant, dominant-passive, the collaborative patterns of interaction (collaborative, expert-novice) are more likely to

lead to L2 learning (Storch, 2002b). This implies that the pairs exhibiting collaborative dyadic relationships in the present study may have benefited more from the collaborative task in terms of L2 learning. On the other hand, pairs showing a non-collaborative patterns of interaction do not benefit much from the pair work (Watanabe & Swain, 2007), this is because learners forming non-collaborative patterns of interaction do not engage or assist each other while performing their task. As such, the data observed in our study revealed that one pair of learners with low-intermediate proficiency showed no or limited collaboration in their task. More specifically, the oral data of Hed and Mur (Chapter 5, section 5.10, example 11) indicated that these learners seemed to form a dominant-passive pattern of interaction, where Mur obtained a leading role while Hed a subservient role while completing their task. It is very likely that the emerged pattern of interaction may have arisen due to the possible differences in the L2 proficiency between learners. Previous research with adult EFL learners has found that pairing learners with different L2 proficiency may negatively affect their collaboration (Storch & Aldosari, 2013), because the low proficiency learner may feel intimidated in working with a high proficiency learner, while the highly proficient learner may feel reluctant to work with a lower proficiency learner and thus their engagement in collaborative dialogue and task completion may be lowered (Zabihi & Ghahramanzadeh, 2022). In this manner, the dominant-passive pattern of interaction appeared in the data of Hed and Mur may have emerged as a result of both learners' perception towards each other's English competence, where Hed having known his peer might have perceived him (Mur) as being unable to assist him in solving the lexical problem being discussed (the word *boss*) and thus has completely ignored him. Instead, he engaged in private speech as a language tool to mediate his cognitive process in order to find a solution for their concern (Suzuki & Storch, 2020). Mur, on the other hand, knowing his low English proficiency may have felt intimidated to make any suggestions or propose any idea to solve the problem they encountered in front of his interlocutor perceived as a more competent L2 user and thus chose to be passive until he decided to ask his teacher for help, who assisted them with the lexical item being discussed. What this conversation seems to indicate is that despite the fact that both learners were assessed as being of low-intermediate L2 proficiency, the fact that their English competence was not determined by means of any standardized tests may have resulted in pairing learners with heterogeneous L2 proficiency instead of exact matched-proficiency. Thus being familiar with one another's competence in the target language

or at least having a general perception towards it could have resulted in the learners' lack of willingness to collaborate with each other in their given task. Previous research with adult ESL learners has documented that when learners do not trust in each other's L2 knowledge, they may assume a non-collaborative pattern of interaction in pair work activities (e.g. Watanabe, 2008). As such, our finding seems to suggest that the learners' perceptions (Hed and Mur) towards each other's L2 proficiency resulted in a dominant-passive pattern of interaction with no or very limited mutual engagement in the task echoing the findings reported in Watanabe (2008) and extending them further with data from adolescent learners from an EFL context. Furthermore, and more importantly, the fact that they decided to seek help from their teacher instead of maximizing the learning opportunities that peer interaction offers (Philp et al., 2013) suggests that adolescent EFL learners with low L2 proficiency assume the assistance of their teacher as more reliable than that of their peers. From this perspective, the assistance of the teacher in an EFL classroom is almost indispensable, not only to help learners with their linguistic concerns, but also to help their collaboration in language tasks. Nevertheless, whether the learners' L2 proficiency played any role in their collaboration, reflected through patterns of interaction formed, it cannot be claimed with confidence as these results are based on simple observations on a limited number of pair talks. As such, this finding should be interpreted with caution until a larger number of pair talks are analyzed employing a more rigorous analysis, which is planned to be conducted in the future.

Despite of the differences in the relationships among learners, the majority of pairs across all proficiency levels seemed to engage in interaction as a way to solve their linguistic concerns. When learners are faced with difficulties related to the task performance they engage in collaborative dialogue (Suzuki & Storch, 2020; Swain & Watanabe, 2013), as a way of finding solutions for their linguistic issues. Thus, the collaborative writing task used in the present study gave learners many opportunities to discuss about their language-related problems, provide suggestions and counter-suggestions, provide and receive feedback (positive and negative) during their task completion (c.f. Chapter 5, section 5.10) (Suzuki & Storch, 2020). It was through the interaction that learners pooled their linguistic resources and created collective scaffolding (Donato, 1994), which resulted in reaching correct resolution rates in about 80% of their mutual problematized linguistic issues. Moreover, data from the pair talks (Chapter 5, section 5.10, examples 10, 11, 14, 16) also revealed that learners across

proficiency levels, especially those of low-intermediate proficiency often relied on their L1 (Albanian) to discuss the language issues, and to provide solutions to their linguistic concerns, suggesting a facilitative role of their shared L1 use both in terms of task completion and solutions reached for their L2 (DiCamilla & Antón, 2012). In particular, learners used their shared L1 to retrieve key vocabulary items in L2, seek and provide assistance during the task completion, and to better elaborate their points or choice of words proposed for a specific context. In this manner, the L1 use facilitated their task completion. The use of L1 has been reported to play a key role in retrieval and elaboration of L2 words. More importantly, adult EFL learners who interacted in their shared L1 (Chinese) were found not only to produce but also to resolve correctly a higher number of language-related problems compared to those interacting in L2 (Zhang & Crawford, 2021). Findings of the present study, obtained from the observation of the pair talks, seem to suggest that the L1 shared among learners helped them to discuss and resolve correctly a high proportion of their LREs. Additionally, these findings generally indicate that the lower the L2 proficiency of learners the greater the L1 use among them, echoing those of previous research which found that pairs composed of adult EFL learners of higher L2 proficiency made less use of their shared L1 (Arabic) than those composed of lower L2 proficiency learners (Storch & Aldosari, 2010). This is not surprising given the lack of their L2 knowledge, compared to intermediate and high-intermediate learners who may have a higher level of knowledge about target knowledge, and thus used their L2 to discuss the lexical and grammatical aspects of their L2. Interestingly, as reported in DiCamilla and Antón (2012) with adult EFL learners, high-intermediate learners in the present study relied less on their L1, in fact, there were some pairs whose L1 use was kept at minimum (c.f. Chapter 4, section 4.9.2.1, example 2 data of Lim and Byt). This suggests that the higher the L2 proficiency of learners, the lower their need to rely on their shared L1. Although, it was beyond the scope of the current study to look for any possible correlation between L1 and L2 use and the task performance in terms of written text outcomes, an earlier study with adult EFL learners found that interacting in L1 enabled learners to produce more complex texts than interacting in L2 (Zhang, 2019b). With this in mind, the present study highlights the need for future studies that investigate the potential link between participants' language of interaction L1 vs. L2 and the quality of the text they produce. This would help EFL teachers, in particular those working in secondary schools, to be less concerned that L1 use may dominate among learners working in pairs and the same

can negatively affect their L2 development, which is the opposite of what has been reported in the previous studies with adult EFL learners (e.g. DiCamilla & Antón, 2012; Storch & Aldosari, 2010; Zhang, 2019b; Zhang & Crawford, 2021). In line with previous studies, our findings suggest for a positive role of L1 use, and the same should serve as encouraging evidence for EFL teachers, in particular those working with learners of lower secondary schools in Kosovo, not to deprive their learners of low L2 proficiency from L1 use because such a deprivation may negatively affect their task performance and potentially collaboration with their peers. Nevertheless, despite the ability to use their shared L1 not all pairs worked in a collaborative manner suggesting for different patterns of interaction learners seem to exhibit while working in pairs (c.f. Chapter 5, section 5.10, example 11 data of Hed and Mur). Additionally, these data suggest that when facing with difficulties in solving linguistic problems, either because of non-collaborative patterns of interaction formed or because of learners' low level of L2 knowledge, learners seek help beyond their peers, namely from their teacher, indicating that the teacher's presence in the classroom is almost irreplaceable. Therefore, while peer interaction helps learners to discuss and find solutions for their linguistic issues, the patterns of interaction they form seem to have an important role both in terms of task completion and resolution of language-related problems.

6.11 Summary

The results discussed in this chapter have, on the whole, shown positive effects of collaboration on the quality of the written texts measured through quantitative and qualitative measures. As mentioned above, drawing on a sample of 144 adolescent EFL learners, the present study found that texts produced by learners working in pairs were linguistically longer, syntactically more complex, grammatically more accurate, and better in terms of task achievement, coherence and cohesion, grammar, vocabulary, and slightly better in mechanics than those composed by learners working on their own. This implies that collaboration benefits adolescent EFL learners' written texts. Similar benefits were reported in previous research with young and adolescent EFL learners working under the same conditions and with similar tasks. Nevertheless, as discussed above, studies with adult EFL/ESL learners have often produced inconclusive results, and/or failed to show any effect of collaboration especially in terms of fluency and complexity, and to some extent in terms of holistic measures of grammar and vocabulary. The positive effects of collaboration on the fluency and complexity measures in the present study has been attributed to the type of task (a picture story) learners worked on, which required them to write a story based on a set of photographs using the past tense. Thus, the picture story task (narrated in the past) has been shown to affect the learners' written texts in terms of fluency and complexity. For example, when learners work in pairs an argumentative essay, they spend a wealth of their time on task-related issues, such as dealing with organization, content, generation of ideas, and the like. In contrast, when they work on a picture story task, their focus is more likely to be on how to find the best words and make accurate language use to describe the photographs rather than engage on generation of ideas or organization of task content, which demands a larger amount of time as a process. Hence, the time-consuming process of generation of ideas and perhaps organization is far more shortened when writing a story based on a visual prompt. Nevertheless, to fully resolve the contradicting results between the present study and earlier research with adult EFL/ESL learners, more research is needed with adolescent and adult EFL learners using a similar task. Additionally, learners' use of their shared L1 (Albanian) while in the process of text co-construction is another factor which seems to have positively affected the collaborative texts in terms of fluency, accuracy, complexity, task achievement, coherence and cohesion, grammar and vocabulary. This is because

learners who worked in pairs used their L1 to discuss about the pictorial prompt, ideas to be developed in their written texts and their linguistic concerns related to grammar, vocabulary and mechanics. In this manner, the L1 use helped pair learners to discuss their ideas more freely and elaborately and thus produced longer descriptions of their content points employing more complex language (sentences). Their shared L1 not only helped them to talk about grammatical, lexical and mechanical aspects of their L2 in a more elaborate way, but it also facilitated the process of finding solutions for their concerns, and to some extent, to convince their partners why a specific choice of a linguistic item was better than the other. Furthermore, the study found that collaboration had a positive impact on task achievement, coherence and cohesion, grammar and vocabulary, but not very much so in the area of mechanics when looking at the results drawn from the whole dataset. This is because of two possible reasons: (a) the learner mechanically writing the text may ultimately decide on the final word form, and (b) the area of mechanics is not considered equally important as the other areas of language by adolescent EFL learners. In a word, results drawn from the whole dataset indicated that collaboration mostly benefited the adolescent EFL learners' written products in terms of accuracy, complexity, task achievement, coherence and cohesion, grammar, and vocabulary, but with slight benefits in fluency and limited effect on mechanics.

The impact of collaboration may also vary according to the learners' L2 proficiency. Thus in the present study, collaboration mostly affected the performance of pairs with intermediate and high-intermediate L2 proficiency. This means that the higher the L2 proficiency of learners, the better their performance in terms of fluency, complexity, accuracy, task achievement, coherence and cohesion, grammar, and vocabulary. While pairs of intermediate proficiency did better than their individual peers in all analytic and holistic rubrics, they reached statistically significant results only in the analytic rubric of grammatical accuracy, and holistic rubrics of task achievement, coherence and cohesion, grammar and vocabulary. Pairs of high-intermediate proficiency, on the other hand, reached significantly higher scores than their individual peers in accuracy, complexity (e.g. clauses per T-unit), task achievement, coherence and cohesion, and grammar. The positive effects of collaboration in the above-mentioned rubrics has been attributed to the learners' L2 proficiency, L1 use, and the effectiveness of feedback received from their pair members while completing the task. Finally, low-intermediate pairs, reached statistically

significant results only in the area of vocabulary and mechanics, despite demonstrating a better performance in all the other areas assessed. This implies that collaboration benefits low-intermediate learners written texts only in terms of vocabulary and mechanics. It seems that because of their low L2 proficiency, LI pairs were mostly interested in searching for words to convey their intended meaning simultaneously paying attention to the spelling of words used in their texts. Thus based on these results, it can be said that collaboration benefits adolescent EFL learners' written text outcomes differently, namely the higher their L2 proficiency the higher their performance in the analytic and holistic rubrics, although some variations in terms of vocabulary and mechanics emerged in the data of intermediate and high-intermediate pairs when comparing pairs vs. individuals of the same L2 proficiency. By the same token, the L2 proficiency of learners affects the written products of both pairs and individuals within their writing conditions. Results of both analytic and holistic rubrics revealed that the higher the L2 proficiency of learners, the higher their performance in both types of measures. However, regardless of the measuring rubrics and working conditions (collaborative vs. individual) no significant differences were observed between learners with intermediate and high-intermediate L2 proficiency in any of the areas assessed. This is most likely due to the close gap in L2 proficiency between intermediate and high-intermediate learners. On the other hand, both pairs and individuals of high-intermediate L2 proficiency significantly outperformed their counterparts of low-intermediate proficiency in all measuring rubrics. This implies that the writing performance of pairs and individuals with intermediate and high-intermediate L2 proficiency is closer compared to that of pairs and individuals with low-intermediate proficiency, which is more distant. It is also interesting to note that the performance of pairs in all areas measured was a bit higher compared to that of individuals when looking at learners' performance within their writing conditions across three proficiency levels.

Finally, results regarding the analysis of pair talk revealed that pairs talked significantly more often about lexical than grammatical and mechanical aspects of their L2. Their collaborative dialogue was successful because the majority of LREs were resolved correctly, and only a small number of linguistic issues discussed were resolved incorrectly and even a smaller one remained unresolved. When comparing the results of pairs across three proficiency levels, our findings support earlier studies which found

that the higher the L2 proficiency, the higher the number of LREs produced and resolved correctly. This implies that adolescent EFL learners of higher L2 proficiency are not only able to discuss more about lexical, grammatical and mechanical aspects of their target language, but also to provide accurate solutions to their linguistic concerns encountered while in the process of text co-construction. Likewise, the higher the learners' L2 proficiency, the higher the number and correct resolution rate of grammatical LREs. This suggests that as learners' L2 proficiency increases so does their ability to engage and resolve correctly language-related problems revolving around grammatical norms. Overall, learners' L2 proficiency impacted positively on the number, type and resolution of language-related problems. Nevertheless, the oral data indicated that there were some variation in terms of number, correct resolution rate, amount of time spent on task, and number of turns generated among pairs of similar L2 proficiency across three proficiency groupings. As such, this indicates that not only their L2 proficiency, but also other factors, supposedly the patterns of interaction learners form when working in pairs may impact on the successful production and resolution of LREs, as well as on the quality of the written texts (Watanabe, 2019). In addition, learners used their shared L1 (Albanian) to discuss about their problematized linguistic issues during the task completion. The L1 use among pairs has been found to affect positively the amount and quality of their discussion centered on linguistic issues when assigned to collaborative writing. Therefore, the L1 use among adolescent EFL learners could account to a certain degree, along with other factors such L2 proficiency, patterns of interaction, and the level of engagement in LREs, in the frequency and solution of LREs during collaborative writing tasks. Finally, findings obtained from the observation of several pair talks revealed that learners formed distinct patterns of interaction, the majority of which were collaborative (collaborative, expert-novice), a similar finding reported by previous studies with young EFL learners. Furthermore, to solve any difficulties encountered during the task completion pairs, regardless of their L2 proficiency, engaged in languaging both in the form of collaborative dialogue (more often) and private speech, to facilitate the solution of their dilemmas about the L2 use. More importantly, the interaction among pairs was frequently done in their shared L1 (Albanian) which helped them to discuss more freely and elaborately about different linguistic issues, including discussions about lexical and grammatical aspects of the target language. Thus, their shared L1 helped learners not only to talk about their linguistic dilemmas but also to find accurate solutions for the majority of their linguistic

uncertainties. Although, learners of all proficiency groupings made use of their shared L1 while interacting with their pair members, it is the low-intermediate learners who seemed to make the greatest amount of L1 use during collaborative work. This implies that the lower the L2 of learners, the greater the L1 use among them.

Finally, despite benefits of peer interaction for L2 learning, there were some cases when pairs sought help beyond their dyads, namely from their teacher and other classmates to resolve their dilemmas about particular L2 words or expressions to move forward with the task at hand (c.f. Chapter 5, section 5.10, examples 11 and 16). This indicates that the teacher's role in assisting students not only with their linguistic concerns but also with their collaboration is hugely important. Last but not least, in situations when learners seem reluctant to work with their assigned peers, the teacher's presence is irreplaceable not only to help learners to overcome difficulties in relation to the target language use, but also to help them to get adapted to working in a collaborative manner and to maximize the learning opportunities that peer interaction offers. This is particularly true for low-intermediate adolescent EFL learners in Kosovo, who may form non-collaborative patterns of interaction while assigned to work in pairs.

Chapter 7: Conclusion

7.1 Synthesis of findings

The goal of this study is to compare the writing performance of adolescent EFL learners writing in pairs and individually to complete a picture story task. To assess the written products, the present study employs both quantitative and qualitative measures, that is, the written texts were rated for fluency, complexity and accuracy, as well as for task achievement, coherence and cohesion, grammar, vocabulary and mechanics. The study also examines the peer interaction of learners working in pairs for language-related episodes (LREs), considering the frequency, type and outcome of LREs produced. Furthermore, a limited number of pair talks (2 per proficiency level) are observed to get some insights into the nature of pair talk as well as the amount of L1 use during the collaborative task.

Following the data analysis, the study found that collaboration benefits the written texts in terms of fluency, complexity and accuracy, as well as in terms of task achievement, coherence and cohesion, grammar, vocabulary, but not very much so in mechanics. What is more, the effect of collaboration seems to be larger on the written texts produced by adolescent EFL learners with intermediate and high-intermediate proficiency level, regardless of the assessment rubrics employed. Accordingly, collaboration among learners with low-intermediate proficiency is less beneficial because LI pairs obtained significantly higher scores than their individual peers only in vocabulary and mechanics. This finding indicates that although LI pairs performed generally better than their individual peers in all measuring rubrics, their performance was significantly different only in vocabulary and mechanics. This implies that the effect of collaboration on the adolescent EFL learners' written texts is to a great extent moderated by learners' L2 proficiency, and potentially other mediating factors such as learners' relationships and the L1 use during the collaborative work. In brief, a very noticeable trend observed in our findings is that collaboration benefits more the written products of learners with higher L2 proficiency than those of learners with lower proficiency. Findings also indicated that the deeper the proficiency gap between learners, the bigger the difference on their writing performance, this is especially true for learners within the collaborative condition.

But since pairs outperformed individuals on all proficiency levels (although to a smaller degree with LI pairs) language proficiency cannot be the sole explanatory factor. It was for this reason that the present study also analyzed the peer talk which occurred between the collaborative writers while they were on-task. The analysis of the pair talk found that the higher L2 proficiency learners not only produced a larger number of LREs, the majority of which were lexical, but also resolved correctly such LREs compared to their lower proficiency peers. In other words, the higher proficiency learners identified more language issues around their writing and found it easier to talk about them. Across the whole dataset, pairs talked about lexical issues more often than about grammatical or mechanical issues, which they did in about equal measure. When comparing results for pairs across three proficiency levels, it was found that LI and HI pairs discussed more often about lexical than grammatical or mechanical issues, whereas I pairs talked more frequently about mechanical than grammatical or lexical issues. On the whole, the findings showed that as learners' L2 proficiency increases so does the number and correct resolution rate of grammatical LREs, which is similar to what has been reported in previous research on collaborative writing.

Finally, the analysis of pair talk revealed that the collaborative task used in the present study provided adolescent EFL learners with ample opportunities for languaging both as collaborative dialogue and as private speech, which mediated the solution of their linguistic uncertainties encountered while writing in pairs. Specifically, learners engaged in peer interaction and talked about language use, provided and received corrective feedback, tested their hypotheses about L2 and modified their output. All of these features have been shown to be hugely important for the development of L2. Furthermore, and more importantly the interaction between learners was often done in their shared L1 (Albanian), which facilitated the retrieval of key vocabulary items in L2, and the discussion of their points relating to lexical meanings, choice of words, and grammatical and mechanical aspects of their L2. Last but not least, although pairs across three proficiency levels made frequent use of their L1, it was the LI pairs who made the greatest amount of L1 use while they were on-task. Regardless whether interacting in L1 or L2, the majority of pairs were collaborative and seemed to show elaborate engagement in their LREs. This reflects a kind of collaborative mindset exhibited among adolescent EFL learners. Nevertheless, despite relying on a limited number of observations, the study found that one pair composed of low-intermediate

learners sought help beyond their peers, namely from their teacher to find solutions for some language issues they considered as key to move forward with the task completion. Interestingly, this pair formed a dominant-passive pattern of interaction during the writing process, which may pose challenges to maximize the learning opportunities that peer interaction offers, and which may have emerged due to the potential differences in L2 proficiency between learners. As such, the teacher's presence in L2 classes is hugely important both to help learners with non-collaborative patterns of interaction to get adapted to working collaboratively, and to overcome their linguistic barriers while in the text co-construction process.

7.2 Implications of the Study

Theoretical, methodological and pedagogical

Theoretically, findings of this study help to understand the role of languaging in mediating learning and peer scaffolding. First, results of the present study suggest that writing in pairs is more beneficial than writing individually for adolescent EFL learners in terms of written product outcomes. These findings contribute to the existing body of empirical research (Bueno-Alastuey & Lizarrondo Larumbe, 2017; Fernández Dobao, 2012; Villarreal & Gil-Sarratea, 2019; Villarreal & Munarriz-Ibarrola, 2021; Wigglesworth & Storch, 2009; Zabihi & Bayan, 2020) which found that collaboration benefits the learners written products in terms of accuracy and complexity (Storch, 2005; Zabihi & Bayan, 2020) and overall text quality (e.g. Chen, 2019; Shehadeh, 2011; Villarreal & Munarriz-Ibarrola, 2021), and extends it further by providing empirical evidence from a more rarely explored population of adolescent EFL learners in Kosovo, for whom collaborative writing was a novel set up, because prior to data collection they had mostly been used to writing individually. The positive effect of collaborative work has often been attributed to the possibility of learners to pool their linguistic resources and thus create collective scaffolding (Donato, 1994). In this vein, learners in the current study helped each other to solve their language-related problems by pooling their linguistic resources and thus created collective scaffolding which in turn resulted in a performance that was beyond their individual capacities.

Second, the analysis of oral data revealed that participants in the present study often engaged in languaging both as collaborative dialogue and as private speech, in an

attempt to find solutions for their language and task-related issues. The concept of languaging is deeply rooted in the sociocultural theory of mind (Vygotsky, 1978, 1987) where language plays a critical role in mediating learning. Swain (2006) defined languaging as “the process of making meaning and shaping knowledge and experience through language” (p.98). Hence, in line with the above claim, adolescent EFL learners often engaged in collaborative dialogue and less frequently in private speech as a way to mediate their thinking and learning and to co-construct meaning. The action of co-constructing meaning is a source for language learning (Swain & Watanabe, 2020). By analyzing learners’ oral interactions through language-related episodes, it was found that there were cases when learners in this study consolidated their existing language and created new L2 knowledge (c.f. Chapter 5, section 5.10, examples 10, 11 and 15).

Finally, the analysis of oral data revealed that collaborative writing provided learners with opportunities for peer interaction through which they focused their attention to meaning and form while seeking solutions for the language problems around their writing task. To find solutions for their language-related problems, learners often deliberated over the language use, provided and received feedback, tested their hypotheses about L2 and produced modified output. During such interaction learners used their L1 and L2 as mediational tools to help them think through a solution for their problems (Suzuki & Storch, 2020), and managed to resolve correctly the majority of LREs they raised during the task completion process, regardless of their L2 proficiency level. Furthermore, based on the observational data the majority of pairs observed exhibited collaborative patterns of interaction while they were on-task, and showed elaborate engagement in their LREs (Azkarai & Kopinska, 2020). Therefore, findings of this study seem to support the role of peer interaction for language learning, because from the point of view of sociocultural theory (Vygotsky, 1978) interaction is a context for all human cognitive development including language learning. Similarly, findings of this study seem to support the interactionist approaches (Mackey, 2012; Mackey et al., 2012) to language learning which highly encourage the use of tasks that promote peer interaction because interaction facilitates the process of acquiring a second language, as it provides learners with opportunities to receive modified input, to receive feedback, both explicitly and implicitly, which in turn may draw learners’ attention to problematic aspects of their interlanguage and push them to produce modified output (Gass & Mackey, 2020). Similarly, Long’s (1996) interaction hypothesis posits that

opportunities to receive comprehensible input and interactional feedback while negotiating for meaning and form benefits L2 learning. As such, the amount of LREs produced by learners demonstrated that the collaborative task used in this study created a favorable context for peer interaction in particular for learners whose L2 proficiency was above low-intermediate.

Methodologically, this study showed that although adolescent EFL learners were allocated to one of the proficiency groupings (i.e. low-intermediate, intermediate, and high-intermediate) based on their teacher's assessment, only the performance of high-intermediate and low-intermediate learners reached a significant level of difference in their writing results, while the difference between high-intermediate and intermediate learners, in the majority of cases, was non-significant. This suggests that the gap of L2 proficiency might have not been equally determined from level to level, thus leading to non-significant differences in terms of writing quality across three proficiency levels in turn. Therefore, determining the learners' L2 proficiency level employing standardized tests, in addition to the teacher's assessment, would have enabled us to get better insights into the writing performance of I and HI adolescent EFL learners.

Pedagogically, findings of this study provide several pedagogical implications. First, results of this study provided empirical evidence that collaborative writing offers ample opportunities for learners to engage in interaction whereby they create solutions to their language and task-related concerns. Additionally, assigning learners to work in pairs on a composition task was a novel set up for them, which they received well, indicating that this can be done with secondary EFL school learners. Since the type of task used in the study is meaning-focused, it is no surprise that the number of lexical LREs produced was higher compared to grammatical or mechanical LREs (Storch, 2013, 2016; Zhang & Crawford, 2021). As such depending on the teaching goals, EFL instructors might want to use picture description or composition tasks (e.g. data commentary report, narrative essay) if they want to promote peer-peer interaction revolving around language use in particular with a greater focus on lexical aspects of the L2, task content and organization. On the other hand, if the teaching goals are to promote peer interaction among adolescent EFL learners revolving around L2 grammatical features like present tense markers (-s) morpheme or past tense markers (-ed), they might want to use a dictogloss task, which is considered form-focused and has been found to elicit more interaction about grammatical issues (Alegría de la Colina &

García Mayo, 2007; Basterrechea & García Mayo, 2014; Basterrechea & Leaser, 2019; Calzada & García Mayo, 2020a). Nevertheless, one might also want to set different teaching goals even when using the same task. For example, EFL instructors might want to use a pictorial prompt, as in the present study, and ask learners to build a story once using the present tense and once using the past tense, if the teaching goals are to help adolescent EFL learners identify and practise the use of both these tenses. Hence, depending on the teaching goals one might choose the type of task to be used and the teaching goals to be achieved.

Second, as noted in Chapter 5, adolescent EFL learners of high-intermediate and intermediate proficiency level were able to generate and correctly resolve a higher proportion of language-related problems compared to learners of low-intermediate proficiency. This was observed in studies with adult ESL/EFL learners as well (e.g. Basterrechea & Leaser, 2019; Kim & McDonough, 2008; Leaser, 2004; Storch & Aldosari, 2013). Therefore, collaborative tasks seem to be more useful for learners with higher proficiency levels (e.g. intermediate and above). Nevertheless, the success of pair work does not only depend on the learners' L2 proficiency level, but also on the patterns of interaction learners form when assigned to work in pairs or small groups (e.g. Storch, 2002b; Watanabe & Swain, 2007). In this vein, even pairs of low L2 proficiency can be successful in their task completion, provided that they form collaborative, or expert/novice relationships (c.f. Chapter 5, section 5.10, example 10 data of Led and Sal). Because as Storch, and Watanabe and Swain suggest learners' proficiency is not always a decisive factor for learners' success in their performance, rather the relationships they form when working collaboratively may determine their success or failure in a given task. As such for useful patterns of interaction to be established, teachers need to train their learners how to collaborate with their peers by providing pre-task modeling activities (e.g. Kim & McDonough, 2011) so that they have some prior information as to how to work better collaboratively, or even help learners with their language barriers to facilitate the collaboration among them.

Third, though analyzing the amount of learners' L1 or L2 use during the collaborative task was beyond the scope of the present study, a brief consultation of the oral data seems to suggest that there was a tendency towards a higher proportion of L1 use among adolescent EFL learners, in particular among those belonging to low-intermediate proficiency group. Researchers, working under the framework of

sociocultural theory, maintain that L1 use in L2 classes is not detrimental because it serves important functions, such as searching for words in L2, dealing with task content, organization and language use, which enables learners to complete their task (e.g. Storch & Aldosari, 2010; Storch & Wigglesworth, 2003; Swain & Lapkin, 2000). A similar function of L1 use emerged in the present study as well. Hence, EFL teachers might want to allow their students to use their L1 in L2 classes when working on collaborative tasks because it can benefit their task performance both in terms of language use and task completion. This is most likely for low-intermediate proficiency learners. Additionally, the teacher's help and assistance for low-intermediate proficiency learners is to a certain degree necessary because the assistance needed to complete writing tasks, similar to the one used in this study, may at times require L2 knowledge which is beyond their own and their partner's ability. Finally, collaboration among learners overall benefits the written text outcomes in terms of analytic and holistic assessment rubrics. In an EFL context, collaborative writing tasks as suggested by Storch (2013) are generally implemented from the perspective of writing-to-learn the language (e.g. Manchón, 2011). Therefore, writing in pairs might help adolescent EFL learners to certainly improve their grammatical accuracy, and very likely linguistic complexity and fluency, as well as the overall text quality. This is because learners assigned to work in pairs or small groups can pool their linguistic resources, and thus achieve higher results than when working alone.

7.3 Limitations of the study and future directions

As with any study, the present dissertation project had several limitations that deserve further research in the future. First, the results drawn from this study are based solely on a single treatment comparing the performance of writing in pairs with writing individually. Therefore, had this dissertation project employed a longitudinal design, it is very likely that different results would have emerged. It is interesting to see how the writing performance of adolescent EFL learners changes or progresses after being involved in several collaborative writing sessions, and to what extent collaboration affects the learners' written products measured quantitatively and qualitatively across areas. As we saw, the results of this study indicated that the effect of collaboration was stronger in some areas and weaker in some others and it would be interesting to find out how this does or does not change under a more long-term treatment.

Second, the present study did not investigate whether the learner relationships had any impact on the quality of the written texts measured both quantitatively and qualitatively, and the number, type and outcome of LREs produced. But since there was a variation in the quantity and quality of LREs produced by pairs of similar proficiency, this indicates that apart from the L2 proficiency of learners are other potential factors that may account for this difference. Therefore, whether the patterns of interaction learners form affects the quantity and quality of LREs, as well as the quality of writing performance is an issue that merits to be addressed in future studies using the same data.

Third, participants of this study shared Albanian as their L1, and they used it for many functions such as looking for lexical items in L2, searching the meaning of words, discussing the solutions of their linguistic issues and elaborated their points on their suggestions/proposals. However, it is not clear whether the interaction of pairs in L1 or L2 affected their writing performance in terms of scores they obtained across rubrics measured. Therefore, this issue needs more attention in future studies. Examining the benefits of L1 and L2 use in terms of writing performance would enable lower-secondary EFL teachers to make informed decisions regarding their potential concerns over learners' L1 use when assigning adolescent EFL learners to write in pairs or small groups.

Fourth, as in the present study learners with similar L2 proficiency were assigned to work in pairs, it would be wise in future studies to investigate whether pairing set up, for example, pairing students with higher and lower proficiency (low-low, low-high, high-high) affects their written products measured both quantitatively and qualitatively, and to what extent does the pairing set up affect peer interaction. Examining the pairing conditions would be useful for lower-secondary EFL teachers in terms of providing research evidence on the issue of how to best pair students so to maximize their learning outcomes in writing tasks in particular and foreign language classes in general.

References

- Abrams, I. (2019). Collaborative writing and text quality in Google Docs. *Language Learning*, 23(2), 22–42. <https://doi.org/10.125/44681>
- Adams, R. J., & Oliver, R. (2019). *Teaching through peer interaction*. New York, NY: Routledge.
- Alegria de la Colina, A., & García Mayo, M. P. (2009). Oral interaction in task-based EFL learning: The use of the L1 as a cognitive tool. *IRAL - International Review of Applied Linguistics in Language Teaching*, 47(3–4).
<https://doi.org/10.1515/iral.2009.014>
- Alegria de la Colina, A., & García Mayo, M. P. (2007). Attention to Form Across Collaborative Tasks by Low-Proficiency Learners in an EFL Setting. In García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 91–116). Clevedon [England] ; Buffalo: Multilingual Matters.
- Aliu, L. (2019). *Analysis of Kosovo's Education System*. Friedrich Ebert Stiftung, Pristina. Retrieved from <https://library.fes.de/pdf-files/bueros/kosovo/15185-20190220.pdf>
- Amiryousefi, M. (2016). The differential effects of two types of task repetition on the complexity, accuracy, and fluency in computer-mediated L2 written production: A focus on computer anxiety. *Computer Assisted Language Learning*, 29(5), 1052–1068. <https://doi.org/10.1080/09588221.2016.1170040>
- Ammar, A., & Spada, N. (2006). ONE SIZE FITS ALL?: Recasts, Prompts, and L2 Learning. *Studies in Second Language Acquisition*, 28(4), 543–574.
<https://doi.org/10.1017/S0272263106060268>

- Antón, M., & Dicamilla, F. J. (1999). Socio-Cognitive Functions of L1 Collaborative Interaction in the L2 Classroom. *The Modern Language Journal*, 83(2), 233–247. <https://doi.org/10.1111/0026-7902.00018>
- Appel, G., & Lantolf, J. P. (1994). Speaking as Mediation: A Study of L1 and L2 Text Recall Tasks. *The Modern Language Journal*, 78(4), 437–452. <https://doi.org/10.2307/328583>
- Atkinson, David. (1987). The mother tongue in the classroom: A neglected resource? *ELT Journal*, 41(4), 241–247. <https://doi.org/10.1093/elt/41.4.241>
- Atkinson, Dwight. (2003). L2 writing in the post-process era: Introduction. *Journal of Second Language Writing*, 12(1), 3–15. [https://doi.org/10.1016/S1060-3743\(02\)00123-6](https://doi.org/10.1016/S1060-3743(02)00123-6)
- Azkarai, A., & García Mayo, M. P. (2017). Task repetition effects on L1 use in EFL child task-based interaction. *Language Teaching Research*, 21(4), 480–495. <https://doi.org/10.1177/1362168816654169>
- Azkarai, A., & Kopinska, M. (2020). Young EFL learners and collaborative writing: A study on patterns of interaction, engagement in LREs, and task motivation. *System*, 102338. <https://doi.org/10.1016/j.system.2020.102338>
- Azkarai, A., & García Mayo, M. P. (2012). Does gender influence task performance in EFL? Interactive tasks and language related episodes. *Discourse and Language Learning across L2 Instructional Settings*, 249–278. https://doi.org/10.1163/9789401208598_013
- Azpilicueta-Martínez, R., & Lázaro-Ibarrola, A. (2022). A matter of age: Negotiation in child and adult interactions. *Vigo International Journal of Applied Linguistics*, (19), 9–41. <https://doi.org/10.35869/vial.v0i19.3758>

- Bagheri, M. S., Rahimi, F., & Riasati, M. J. (2012). Communicative Interaction in Language Learning Tasks among EFL Learners. *Journal of Language Teaching and Research*, 3(5), 948–952. Retrieved from <http://www.proquest.com/docview/1040775351/abstract/472B2B4E4AA5419APQ/1>
- Ball, P., Kelly, K., & Clegg, J. (2015). *Putting CLIL into practice*. Oxford: Oxford University Press.
- 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. (2014). Dictogloss and the production of the English third person “-s” by CLIL and mainstream EFL learners: A comparative study. *International Journal of English Studies*, 14(2), 77–98. <https://doi.org/10.6018/j.177321>
- Basterrechea, M., & Leeser, 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>

- Berk, L. E., & Garvin, R. A. (1984). Development of private speech among low-income Appalachian children. *Developmental Psychology*, 20(2), 271–286.
<https://doi.org/10.1037/0012-1649.20.2.271>
- Berlin, J. (1988). Rhetoric and Ideology in the Writing Class. *College English*, 50(5), 477–494. <https://doi.org/10.2307/377477>
- Bickhard, M. H. (2005). Functional scaffolding and self-scaffolding. *New Ideas in Psychology*, 23(3), 166–173.
<https://doi.org/10.1016/j.newideapsych.2006.04.001>
- Bigelow, M., Delmas, R., Hansen, K., & Tarone, E. (2006). Literacy and the Processing of Oral Recasts in SLA. *TESOL Quarterly*, 40(4), 665.
<https://doi.org/10.2307/40264303>
- Bivens, J. A., & Berk, L. E. (1990). A Longitudinal Study of the Development of Elementary School Children's Private Speech. *Merrill-Palmer Quarterly*, 36(4), 443–463. Retrieved from <http://www.jstor.org/stable/23087309>
- Brooks, F. B., & Donato, R. (1994). Vygotskian Approaches to Understanding Foreign Language Learner Discourse during Communicative Tasks. *Hispania*, 77(2), 262–274. <https://doi.org/10.2307/344508>
- Brooks, L., & Swain, M. (2009). Language in Collaborative Writing: Creation of and Response to Expertise. In S. M. Gass, A. Mackey, & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 58–89). New York: Routledge.
- Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research*, 20(4), 436–458.
<https://doi.org/10.1177/1362168814563200>

- Bruffee, K. A. (1984). Collaborative Learning and the “Conversation of Mankind.” *College English*, 46(7), 635. <https://doi.org/10.2307/376924>
- Bueno-Alastuey, M. C., & Lizarrondo Larumbe, P. (2017). *Collaborative writing in the EFL Secondary Education classroom: Comparing triad, pair and individual work*. Retrieved from <https://academica-e.unavarra.es/xmlui/handle/2454/28503>
- Calzada, A., & García Mayo, M. P. (2020a). Chapter 1. Child EFL grammar learning through a collaborative writing task. In W. Suzuki & N. Storch (Eds.), *Language Learning & Language Teaching* (Vol. 55, pp. 20–39). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/llt.55.01cal>
- Calzada, A., & García Mayo, M. P. (2020b). Child EFL learners’ attitudes towards a collaborative writing task: An exploratory study. *Language Teaching for Young Learners*, 2(1), 52–72. <https://doi.org/10.1075/ltyl.19008.cal>
- Calzada, A., & García Mayo, M. P. (2020c). Child learners’ reflections about EFL grammar in a collaborative writing task: When form is not at odds with communication. *Language Awareness*, 1–16. <https://doi.org/10.1080/09658416.2020.1751178>
- Calzada, A., & García Mayo, M. P. (2021). Effects of proficiency and collaborative work on child EFL individual dictogloss writing. *Language Teaching for Young Learners*, 3(2), 246–274. <https://doi.org/10.1075/ltyl.20003.cal>
- Carless, D. (2007). Student use of the mother tongue in the task-based classroom. *ELT Journal*, 62(4), 331–338. <https://doi.org/10.1093/elt/ccm090>
- Chaudron, C. (1977). A Descriptive Model of Discourse in the Corrective Treatment of Learners’ Errors ¹. *Language Learning*, 27(1), 29–46. <https://doi.org/10.1111/j.1467-1770.1977.tb00290.x>

- Chen, W. (2019). An Exploratory Study on the Role of L2 Collaborative Writing on Learners' Subsequent Individually Composed Texts. *The Asia-Pacific Education Researcher*, 28(6), 563–573. <https://doi.org/10.1007/s40299-019-00455-3>
- Chen, W., & Hapgood, S. (2019). Understanding knowledge, participation and learning in L2 collaborative writing: A metacognitive theory perspective. *Language Teaching Research*, 1362168819837560. <https://doi.org/10.1177/1362168819837560>
- Chen, W. (2017). The effect of conversation engagement on L2 learning opportunities. *ELT Journal*, 71(3), 329–340. <https://doi.org/10.1093/elt/ccw075>
- Chen, W. (2018). Patterns of pair interaction in communicative tasks: The transition process and effect on L2 teaching and learning. *ELT Journal*, 72(4), 425–434. <https://doi.org/10.1093/elt/ccy015>
- Choi, S.-Y., & Li, S. (2012). Corrective Feedback and Learner Uptake in a Child ESOL Classroom. *RELC Journal*, 43(3), 331–351. <https://doi.org/10.1177/0033688212463274>
- Clark, I. L. (2012). *Concepts in composition: Theory and practice in the teaching of writing* (2nd ed). New York: Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed). London ; New York: Routledge.

- Consolini, C. H., & Soto-Lucena, I. (2019). Low-proficiency L2 Collaborative writing to enhance individual writing and grammatical accuracy. *Coyote Papers*, 22(2019), 33–42. Retrieved from <http://hdl.handle.net/10150/641484>
- Corder, G. W., & Foreman, D. I. (2009). *Nonparametric statistics for non-statisticians: A step-by-step approach*. Hoboken, N.J: Wiley.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Modern Languages Division, Strasbourg. Retrieved from <https://rm.coe.int/1680459f97>
- Cumming, A. (1989). Writing Expertise and Second-Language Proficiency*. *Language Learning*, 39(1), 81–135. <https://doi.org/10.1111/j.1467-1770.1989.tb00592.x>
- Daiute, C., & Dalton, B. (1993). Collaboration between Children Learning to Write: Can Novices Be Masters? *Cognition and Instruction*, 10(4), 281–333. Retrieved from <http://www.jstor.org/stable/3233656>
- Dale, H. (1994). Collaborative Research on Collaborative Writing. *The English Journal*, 83(1), 66–70. <https://doi.org/10.2307/820963>
- DiCamilla, F. J., & Antón, M. (2012). Functions of L1 in the collaborative interaction of beginning and advanced second language learners. *International Journal of Applied Linguistics*, 22(2), 160–188. <https://doi.org/10.1111/j.1473-4192.2011.00302.x>
- DiPardo, A., & Freedman, S. W. (1988). Peer Response Groups in the Writing Classroom: Theoretic Foundations and New Directions. *Review of Educational Research*, 58(2), 119–149. <https://doi.org/10.2307/1170332>

- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 35–56). Norwood, N.J: Ablex Pub. Corp.
- Duff, P. A., & Polio, C. G. (1990). How Much Foreign Language Is There in the Foreign Language Classroom? *The Modern Language Journal*, 74(2), 154–166. <https://doi.org/10.1111/j.1540-4781.1990.tb02561.x>
- Ede, L. S., & Lunsford, A. A. (1990). *Singular texts/plural authors: Perspectives on collaborative writing*. Carbondale: Southern Illinois University Press.
- Elbow, P. (1998a). *Writing with power: Techniques for mastering the writing process* (2nd ed). New York: Oxford University Press.
- Elbow, P. (1998b). *Writing without teachers* (2nd ed). New York: Oxford University Press.
- Elbow, P. (1999). Individualism and the teaching of writing: Response to Vai Ramanathan and Dwight Atkinson. *Journal of Second Language Writing*, 8(3), 327–338. [https://doi.org/10.1016/S1060-3743\(99\)80120-9](https://doi.org/10.1016/S1060-3743(99)80120-9)
- Elbow, P. (2012). *Vernacular eloquence: What speech can bring to writing*. Oxford ; New York: Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching* (7. print). Oxford: Oxford Univ. Press.
- Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford ; New York: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner Uptake in Communicative ESL Lessons. *Language Learning*, 51(2), 281–318. <https://doi.org/10.1111/1467-9922.00156>

- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar. *Studies in Second Language Acquisition*, 28(2), 339–368. <https://doi.org/10.1017/S0272263106060141>
- Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, 21(1), 40–58. <https://doi.org/10.1016/j.jslw.2011.12.002>
- Fernández Dobao, A. (2012a). Collaborative Dialogue in Learner–Learner and Learner–Native Speaker Interaction. *Applied Linguistics*, 33(3), 229–256. <https://doi.org/10.1093/applin/ams002>
- Fernández Dobao, A. (2014). Attention to Form in Collaborative Writing Tasks: Comparing Pair and Small Group Interaction. *Canadian Modern Language Review*, 70(2), 158–187. <https://doi.org/10.3138/cmlr.1768>
- Fernández Dobao, A. (2016). Peer interaction and learning: A focus on the silent learner. In S. G. Ballinger & S. Masatoshi (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 33–62). Amsterdam ; Philadelphia: John Benjamins Publishing Company.
- Fernández Dobao, A., & Blum, A. (2013). Collaborative writing in pairs and small groups: Learners’ attitudes and perceptions. *System*, 41(2), 365–378. <https://doi.org/10.1016/j.system.2013.02.002>
- Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge ; New York: Cambridge University Press.

- Ferris, D., & Hedgcock, J. (2014). *Teaching L2 composition: Purpose, process, and practice* (Third Edition). New York ; London: Routledge, Taylor & Francis Group.
- Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th edition). Thousand Oaks, CA: SAGE Publications.
- Flower, L., & Hayes, J. R. (1980). The Cognition of Discovery: Defining a Rhetorical Problem. *College Composition and Communication*, 31(1), 21–32.
<https://doi.org/10.2307/356630>
- Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365–387.
<https://doi.org/10.2307/356600>
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–375.
<https://doi.org/10.1093/applin/21.3.354>
- Foster, P., & Ohta, A. S. (2005). Negotiation for Meaning and Peer Assistance in Second Language Classrooms. *Applied Linguistics*, 26(3), 402–430.
<https://doi.org/10.1093/applin/ami014>
- Foster, P., & Skehan, P. (1996). The Influence of Planning and Task Type on Second Language Performance. *Studies in Second Language Acquisition*, 18(3), 299–323. <https://doi.org/10.1017/S0272263100015047>
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3), 215–247. <https://doi.org/10.1177/136216889900300303>
- Fotos, S., & Ellis, R. (1991). Communicating about Grammar: A Task-Based Approach. *TESOL Quarterly*, 25(4), 605–628. <https://doi.org/10.2307/3587079>

- Freebairn, I., Bygrave, J., & Copage, J. (2015). *Live Beat 3—Teacher's book*.
- Gallardo del Puerto, F., & Basterrechea, M. (2021). The role of proficiency and pair formation method in language-related episodes: A study of young CLIL learners' interaction. *Language Teaching Research*, 13621688211024364. <https://doi.org/10.1177/13621688211024363>
- García Mayo, M. P. (2018). Child task-based interaction in EFL settings: Research and challenges. *International Journal of English Studies*, 18(2), 119–144. <https://doi.org/10.6018/ijes/2018/2/319731>
- García Mayo, M. P. (2018). Dictogloss Technique. In J. I. Lontas, T. International Association, & M. DelliCarpini (Eds.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–5). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118784235.eelt0181>
- García Mayo, M. P., & Lázaro Ibarrola, A. (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System*, 54, 40–54. <https://doi.org/10.1016/j.system.2014.12.001>
- García Mayo, M. P. (2002). Interaction in advanced EFL pedagogy: A comparison of form-focused activities. *International Journal of Educational Research*, 37(3), 323–341. [https://doi.org/10.1016/S0883-0355\(03\)00008-9](https://doi.org/10.1016/S0883-0355(03)00008-9)
- García Mayo, M. P. (Ed.). (2017). *Learning foreign languages in primary school research insights*. Bristol ; Blue Ridge Summit: Multilingual Matters.
- García Mayo, M. P. (2019). Learning in foreign language contexts. *Language Teaching Research*, 23(1), 5–8. <https://doi.org/10.1177/1362168818822387>
- García Mayo, M. P., & Alcón Soler, E. (2013). Negotiated input and output / interaction. In J. R. Herschensohn & M. Young-Scholten (Eds.), *The*

- Cambridge handbook of second language acquisition* (pp. 205–208).
Cambridge ; New York: Cambridge University Press.
- García Mayo, M. P., & Zeitler, N. (2017). Lexical language-related episodes in pair and small group work. *International Journal of English Studies*, 17(1), 61–82.
<http://dx-doi-org.uaccess.univie.ac.at/10.6018/ijes/2017/1/255011>
- García Mayo, M. P., & Hidalgo, M. 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. 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. (2003). Input and interaction. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 224–255). Malden, MA: Blackwell Pub.
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research* (1. publ). Mahwah, NJ: Lawrence Erlbaum.
- Gass, S. M., & Mackey, A. (2015). Input, Interaction, and Output in Second Language Acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (Second Edition, pp. 180–206). New York: Routledge.
- Gass, S. M., & MacKey, A. (2020). Input, Interaction, and Output in L2 Acquisition. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second*

- language acquisition: An introduction* (Third edition, pp. 192–222). New York, NY: Routledge.
- Gass, S. M., & Varonis, E. M. (1985). Variation in Native Speaker Speech Modification to Non-Native Speakers. *Studies in Second Language Acquisition*, 7(1), 37–57. Retrieved from <https://www.jstor.org/stable/44488545>
- Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-Based Interactions in Classroom and Laboratory Settings. *Language Learning*, 55(4), 575–611. <https://doi.org/10.1111/j.0023-8333.2005.00318.x>
- Gassner, O., Mewald, C., Brock, R., Lackenbauer, F., & Klaus, S. (2019). *Testing Writing for the E8 Standards—Technical Report 2019*. Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens, Salzburg. Retrieved from www.bifie.at
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Granott, N. (2005). Scaffolding dynamically toward change: Previous and new perspectives. *New Ideas in Psychology*, 23(3), 140–151. <https://doi.org/10.1016/j.newideapsych.2006.07.002>
- Guerrero, M., & Villamil, O. S. (2000). Activating the ZPD: Mutual Scaffolding in L2 Peer Revision. *The Modern Language Journal*, 84(1), 51–68. Retrieved from <http://www.jstor.org/stable/330449>

- Guk, I., & Kellogg, D. (2007). The ZPD and whole class teaching: Teacher-led and student-led interactional mediation of tasks. *Language Teaching Research*, 11(3), 281–299. <https://doi.org/10.1177/1362168807077561>
- Han, Z. (2002). A Study of the Impact of Recasts on Tense Consistency in L2 Output. *TESOL Quarterly*, 36(4), 543–572. <https://doi.org/10.2307/3588240>
- Harmer, J. (2007). *The practice of English language teaching* ((4th ed.)). London.
- Harris, T., & Rowe, A. (1995). *Exploring English: Workbook*. White Plains, NY: Longman.
- Haswell, R. H. (2008). Teaching of writing in higher education. In C. Bazerman (Ed.), *Handbook of research on writing: History, society, school, individual, text* (pp. 405–424). New York: L. Erlbaum Associates.
- Hayes, J. R., & Flower, L. (1980). Identifying the organization of Writing Processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing*. Hillsdale, N.J: L. Erlbaum Associates.
- Hidalgo, M. Á., & García Mayo, M. P. (2019). The influence of task repetition type on young EFL learners' attention to form. *Language Teaching Research*, 1362168819865559. <https://doi.org/10.1177/1362168819865559>
- Hidalgo, M. Á., & Lázaro-Ibarrola, A. (2020). Task repetition and collaborative writing by EFL children: Beyond CAF measures. *Studies in Second Language Learning and Teaching*, 10(3), 501–522. <https://doi.org/10.14746/ssllt.2020.10.3.5>
- Hillocks, G. (2008). Writing in Secondary Schools. In C. Bazerman (Ed.), *Handbook of research on writing: History, society, school, individual, text* (pp. 381–404). New York: L. Erlbaum Associates.

- Hinkel, E. (2011). What research on second language writing tells us and what it doesn't. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 523–538). Mahwah, N.J: L. Erlbaum Associates.
- Hirvela, A., & Belcher, D. (2001). Coming back to voice: The multiple voices and identities of mature multilingual writers. *Journal of Second Language Writing*, 10(1), 83–106. [https://doi.org/10.1016/S1060-3743\(00\)00038-2](https://doi.org/10.1016/S1060-3743(00)00038-2)
- Hirvela, A., Hyland, K., & Manchón, R. M. (2016). Dimensions in L2 writing theory and research: Learning to write and writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 45–64). Boston: Walter de Gruyter Inc.
- Holton, D., & Clarke, D. (2006). Scaffolding and metacognition. *International Journal of Mathematical Education in Science and Technology*, 37(2), 127–143. <https://doi.org/10.1080/00207390500285818>
- Hunt, K. W. (1966). Recent Measures in Syntactic Development. *Elementary English*, 43(7), 732–739. Retrieved from <http://www.jstor.org/stable/41386067>
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Hyland, K. (2009). *Teaching and researching writing* (2. ed). Harlow: Longman.
- Hyland, K. (2010). *Teaching and researching writing* (2. ed). Harlow: Longman.
- Hyland, K. (2016). *Teaching and researching writing* (Third edition). New York: Routledge.
- Iwashita, N. (1999). Tasks and Learners' Output in Nonnative-Nonnative Interaction. In K. Kanno (Ed.), *Language Acquisition and Language Disorders* (Vol. 20, p. 31). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/lald.20.06iwa>

- 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)
- Jamovi. (2021). *The Jamovi Project (2021)* [[Computer Software]]. jamovi (Version 1.6). Retrieved from <https://www.jamovi.org>
- Jang, H., & Cheung, Y. L. (2020). Impacts of dyadic interaction on second language writing: A study with eight bilingual primary school students in Singapore. *Education 3-13*, 48(5), 527–540. <https://doi.org/10.1080/03004279.2019.1624804>
- Kasumi, H. (2015). Communicative Language Teaching and Its Impact on Students' Performance. *Journal of Educational and Social Research*. <https://doi.org/10.5901/jesr.2015.v5n1s1p155>
- Kessler, G., Bikowski, D., & Boggs, J. (2012). *Collaborative Writing Among Second Language Learners In Academic Web-Based Projects*. 16(1), 91–109. Retrieved from <http://dx.doi.org/10125/44276>
- Khatib, M., & Meihami, H. (2015). Languaging and Writing Skill: The Effect of Collaborative Writing on EFL Students' Writing Performance. *Advances in Language and Literary Studies*, 6(1). <https://doi.org/10.7575/aiac.all.v.6n.1p.203>
- Kilbey, E., Freebairn, I., Bygrave, J., & Copage, J. (2015). *Live beat*. 3, 3,.
- Kim, Y. (2013). Effects of Pretask Modeling on Attention to Form and Question Development. *TESOL Quarterly*, 47(1), 8–35. Retrieved from <http://www.jstor.org/stable/43267772>

- Kim, Y. (2008). The Contribution of Collaborative and Individual Tasks to the Acquisition of L2 Vocabulary. *The Modern Language Journal*, 92(1), 114–130. JSTOR. Retrieved from <http://www.jstor.org/stable/25172996>
- 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>
- Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15(2), 183–199. <https://doi.org/10.1177/1362168810388711>
- Knouzi, I., Swain, M., Lapkin, S., & Brooks, L. (2010). Self-scaffolding mediated by languaging: Microgenetic analysis of high and low performers. *International Journal of Applied Linguistics*, 20(1), 23–49. <https://doi.org/10.1111/j.1473-4192.2009.00227.x>
- Kojima, M., In'nami, Y., & Kaneta, T. (2022). Chapter 6. L2 writing and its external correlates: A meta-analysis. In E. H. Jeon & Y. In'nami (Eds.), *Bilingual Processing and Acquisition* (Vol. 13, pp. 159–211). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.13.06koj>
- Kojima, M., & Kaneta, T. (2022). Chapter 5. L2 writing and its internal correlates: A meta-analysis. In E. H. Jeon & Y. In'nami (Eds.), *Bilingual Processing and Acquisition* (Vol. 13, pp. 109–158). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.13.05koj>
- Kopinska, M., & Azkarai, A. (2020). Exploring young EFL learners' motivation: Individual versus pair work on dictogloss tasks. *Studies in Second Language*

- Learning and Teaching*, 10(3), 607–630.
- <https://doi.org/10.14746/sslt.2020.10.3.10>
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3(2), 73–93.
- <https://doi.org/10.1080/09658416.1994.9959845>
- Kowal, M., & Swain, M. (1997). From semantic to syntactic processing: How can we promote it in the immersion classroom? In R. K. Johnson & M. Swain (Eds.), *Immersion Education* (1st ed., pp. 284–310). Cambridge University Press.
- <https://doi.org/10.1017/CBO9781139524667.022>
- Kozulin, A. (1990). *Vygotsky's psychology: A biography of ideas*. Place of publication not identified: Harvard Univ Press.
- Krashen, S. D. (1981). *Second language acquisition and second language learning* (Reprinted). Oxford: Pergamon Pr.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition* (1st ed). Oxford ; New York: Pergamon.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London ; New York: Longman.
- Kucer, S. B. (2014). *Dimensions of literacy: A conceptual base for teaching reading and writing in school settings* (Fourth edition). New York: Routledge.
- Kuiken, F., & Vedder, I. (2002). The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research*, 37(3), 343–358. [https://doi.org/10.1016/S0883-0355\(03\)00009-0](https://doi.org/10.1016/S0883-0355(03)00009-0)
- Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring Longitudinal Writing Development Using Indices of Syntactic Complexity and Sophistication.

Studies in Second Language Acquisition, 43(4), 781–812.

<https://doi.org/10.1017/S0272263120000546>

- Lahuerta Martínez, A. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing*, 35, 1–11. <https://doi.org/10.1016/j.asw.2017.11.002>
- Lantolf, J. P. (2005). Sociocultural theory and L2 learning: An exegesis. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 335–354). Mahwah, N.J: L. Erlbaum Associates.
- Lantolf, J. P., & Beckett, T. G. (2009). Sociocultural theory and second language acquisition. *Language Teaching*, 42(4), 459–475.
<https://doi.org/10.1017/S0261444809990048>
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford ; New York: Oxford University Press.
- Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2015). Sociocultural Theory and Second Language Development. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (Second Edition, pp. 207–226). New York: Routledge.
- Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R* (Second Edition). London ; New York: Routledge.
- Lázaro Ibarrola, A., & Hidalgo, M. Á. (2017). Benefits and Limitations of Conversational Interactions among Young Learners of English in a CLIL Context. In García Mayo, M. P. (Ed.), *Learning foreign languages in primary school research insights* (pp. 86–102). Bristol ; Blue Ridge Summit: Multilingual Matters.

- Lázaro-Ibarrola, A., & Hidalgo, M. Á. (2022). Collaborative writing among young EFL learners in a school context: Product and process. *The Language Learning Journal*, 0(0), 1–16. <https://doi.org/10.1080/09571736.2022.2141828>
- Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In R. DeKeyser (Ed.), *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 111–138). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667275>
- 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>
- Levine, G. S. (2003). Student and Instructor Beliefs and Attitudes about Target Language Use, First Language Use, and Anxiety: Report of a Questionnaire Study. *The Modern Language Journal*, 87(3), 343–364. Retrieved from <http://www.jstor.org/stable/1192959>
- Li, M., & Zhu, W. (2017). Good or bad collaborative wiki writing: Exploring links between group interactions and writing products. *Journal of Second Language Writing*, 35, 38–53. <https://doi.org/10.1016/j.jslw.2017.01.003>
- Li, S. (2010). The Effectiveness of Corrective Feedback in SLA: A Meta-Analysis. *Language Learning*, 60(2), 309–365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Lin, O. P., & Maarof, N. (2013). Collaborative Writing in Summary Writing: Student Perceptions and Problems. *Procedia - Social and Behavioral Sciences*, 90, 599–606. <https://doi.org/10.1016/j.sbspro.2013.07.131>

- Littlewood, W. (2014). Communication-oriented language teaching: Where are we now? Where do we go from here? *Language Teaching*, 47(3), 349–362.
<https://doi.org/10.1017/S0261444812000134>
- Loewen, S. (2012). The role of feedback. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 24–40). London ; New York: Routledge.
- Loewen, S., & Philp, J. (2006). Recasts in the Adult English L2 Classroom: Characteristics, Explicitness, and Effectiveness. *The Modern Language Journal*, 90(4), 536–556. Retrieved from <http://www.jstor.org/stable/4127042>
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329.
<https://doi.org/10.1017/S0261444818000125>
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input1. *Applied Linguistics*, 4(2), 126–141.
<https://doi.org/10.1093/applin/4.2.126>
- Long, M. H. (1981). Input, Interaction, and Second-Language Acquisition. *Annals of the New York Academy of Sciences*, 379(1 Native Language), 259–278.
<https://doi.org/10.1111/j.1749-6632.1981.tb42014.x>
- Long, M. H. (1983). Linguistic and Conversational Adjustments to Non-Native Speakers. *Studies in Second Language Acquisition*, 5(2), 177–193.
<https://doi.org/10.1017/S0272263100004848>
- Long, M. H. (1991). Focus on Form: A Design Feature in Language Teaching. In K. de Bot, R. B. Ginsberg, & C. J. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (Repr, pp. 39–52). Amsterdam: Benjamins.

- Long, M. H. (1996). The Role of the Linguistic Environment in Second Language Acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–454). San Diego: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Mahwah, N.J: L. Erlbaum Associates.
- Lyster, R., & Izquierdo, J. (2009). Prompts Versus Recasts in Dyadic Interaction. *Language Learning*, 59(2), 453–498. <https://doi.org/10.1111/j.1467-9922.2009.00512.x>
- Lyster, R., & Ranta, L. (1997). Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms. *Studies in Second Language Acquisition*, 19(1), 37–66. <https://doi.org/10.1017/S0272263197001034>
- Lyster, R., & Saito, K. (2010). ORAL FEEDBACK IN CLASSROOM SLA: A Meta-Analysis. *Studies in Second Language Acquisition*, 32(2), 265–302. <https://doi.org/10.1017/S0272263109990520>
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40. <https://doi.org/10.1017/S0261444812000365>
- Macaro, E. (2009). Chapter 2. Teacher Use of Codeswitching in the Second Language Classroom: Exploring ‘Optimal’ Use. In M. Turnbull & J. Dailey-O’Cain (Eds.), *First Language Use in Second and Foreign Language Learning* (pp. 35–49). Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781847691972-005>
- Mackey, A. (1999). Input, Interaction, and Second Language Development: An Empirical Study of Question Formation in ESL. *Studies in Second Language Acquisition*, 21(4), 557–587. <https://doi.org/10.1017/S0272263199004027>

- Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning* (1. publ). Oxford: Oxford Univ. Press.
- Mackey, A., Abbuhl, R., & Gass, S. M. (2012). Interactionist approach. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 7–23). London ; New York: Routledge.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the Relationship Between Modified Output and Working Memory Capacity. *Language Learning*, 60(3), 501–533. <https://doi.org/10.1111/j.1467-9922.2010.00565.x>
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.
- Mackey, A., In Park, H., & M. Tagarelli, K. (2016). Errors, corrective feedback and repair Variations and learning outcomes. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 499–512). London ; New York: Routledge.
- Mackey, A., & Oliver, R. (2002). Interactional feedback and children's L2 development. *System*, 30(4), 459–477. [https://doi.org/10.1016/S0346-251X\(02\)00049-0](https://doi.org/10.1016/S0346-251X(02)00049-0)
- Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional Input and The Incorporation of Feedback: An Exploration of NS-NNS and NNS-NNS Adult and Child Dyads. *Language Learning*, 53(1), 35–66. <https://doi.org/10.1111/1467-9922.00210>
- Mackey, A., & Philp, J. (1998). Conversational Interaction and Second Language Development: Recasts, Responses, and Red Herrings? *The Modern Language Journal*, 82(3), 338–356. <https://doi.org/10.2307/329960>

- Manchón, R. M. (Ed.). (2011). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam ; Philadelphia: John Benjamins Pub. Co.
- Manchón, R. M., & Williams, J. (2016). 26. L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of Second and Foreign Language Writing* (pp. 567–586). De Gruyter.
<https://doi.org/10.1515/9781614511335-029>
- Mascolo, M. F. (2005). Change processes in development: The concept of coactive scaffolding. *New Ideas in Psychology*, 23(3), 185–196.
<https://doi.org/10.1016/j.newideapsych.2006.05.002>
- Matsuda, P. K. (2003a). Process and post-process: A discursive history. *Journal of Second Language Writing*, 12(1), 65–83. [https://doi.org/10.1016/S1060-3743\(02\)00127-3](https://doi.org/10.1016/S1060-3743(02)00127-3)
- Matsuda, P. K. (2003b). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the Dynamics of Second Language Writing* (1st ed., pp. 15–34). Cambridge University Press.
<https://doi.org/10.1017/CBO9781139524810.004>
- McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32(2), 207–224.
<https://doi.org/10.1016/j.system.2004.01.003>
- McDonough, K. (2005). Identifying the Impact of Negative Feedback and Learners' Responses on ESL Question Development. *Studies in Second Language Acquisition*, 27(01). <https://doi.org/10.1017/S0272263105050047>
- McDonough, K., Crawford, W., & De Vleeschauwer, J. (2016). Thai EFL learners' interaction during collaborative writing tasks and its relationship to text quality. In M. Sato & S. G. Ballinger (Eds.), *Peer interaction and second*

- language learning: Pedagogical potential and research agenda* (pp. 185–208). Amsterdam ; Philadelphia: John Benjamins Publishing Company.
- McDonough, K., De Vleeschauwer, J., & Crawford, W. (2018). Comparing the quality of collaborative writing, collaborative prewriting, and individual texts in a Thai EFL context. *System*, 74, 109–120.
<https://doi.org/10.1016/j.system.2018.02.010>
- McDonough, K., & Fuentes, C. G. (2015). The effect of writing task and task conditions on Colombian EFL learners' language use. *TESL Canada Journal*, 32(2), 67–80. Retrieved from
<http://go.gale.com/ps/i.do?p=AONE&sw=w&issn=0826435X&v=2.1&it=r&id=GALE%7CA439271391&sid=googleScholar&linkaccess=abs>
- MEST. (2016a). *Korniza Kurrikulare e Arsimit Parauniversitar të Republikës së Kosovës (e rishikuar) (The Core Curriculum of Pre-university Education of Republic of Kosovo)—(Revised)*. Ministria e Arsimit Shkencës dhe Teknologjisë (Ministry of Education, Science and Technology). (Retrieved from
<https://masht.rks-gov.net/arsimi-parauniversitar>).
- MEST. (2016b). *Kurrikula Bërthamë e Arsimit të Mesëm të Ulët të Kosovës (Klasat VI, VII, VIII dhe IX, e rishikuar) The Core Curriculum of Lower-Secondary Education of Kosovo (Grades VI, VII, VIII and IX, revisited)*. Ministria e Arsimit Shkencës dhe Teknologjisë (Ministry of Education, Science and Technology). Retrieved from <https://masht.rks-gov.net/kurrikula-berthame-e-arsimit-te-mesem-te-ulet-te-kosoves-klasa-vi-vii-viii-dhe-ix/>
- MEST. (2019). *Kurrikulat Lëndore/ Programmet Mësimore—Klasa e Tetë (Subject Curricula/Subject Syllabi- Grade Eight)*. Ministria e Arsimit Shkencës dhe

- Teknologjisë (Ministry of Education, Science and Technology). (Retrieved from <https://masht.rks-gov.net/arsimi-parauniversitar>).
- MESTI. (2021). *Administrative Instruction for Normative on Professional Staff of the General Education*. Retrieved from <https://masht.rks-gov.net/udhezimi-administrativ-mashti-nr-05-2021-per-normativin-mbi-kuadrin-profesional-te-arsimit-te-pergjithshem/>
- 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. (2014). Exploring young learners' foreign language learning awareness. *Language Awareness*, 23(1–2), 24–40.
<https://doi.org/10.1080/09658416.2013.863900>
- Murphy, L., & Roca de Larios, J. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL writers. *Journal of Second Language Writing*, 19(2), 61–81. <https://doi.org/10.1016/j.jslw.2010.02.001>
- Nassaji, H. (2015). *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. London ; New York: Bloomsbury Academic, An imprint of Bloomsbury Publishing Plc.
- Nassaji, H., & Fotos, S. (2004). 6. Current Developments in Research on The Teaching of Grammar. *Annual Review of Applied Linguistics*, 24.
<https://doi.org/10.1017/S0267190504000066>
- Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, 14(4), 397–419. <https://doi.org/10.1177/1362168810375364>

- Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness*, 18(3–4), 384–402.
<https://doi.org/10.1080/09658410903197256>
- Niu, R., Jiang, L., & Deng, Y. (2018). Effect of Proficiency Pairing on L2 Learners' Language Learning and Scaffolding in Collaborative Writing. *The Asia - Pacific Education Researcher*, 27(3), 187–195. <http://dx-doi-org.uaccess.univie.ac.at/10.1007/s40299-018-0377-2>
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47(3), 203–210. <https://doi.org/10.1093/elt/47.3.203>
- North, S. M. (1987). *The making of knowledge in composition: Portrait of an emerging field*. Upper Montclair, N.J: Boynton/Cook Publishers.
- Ohta, Amy. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51–78). Oxford: Oxford Univ. Press.
- Ohta, Amy. S. (1995). Applying Sociocultural Theory to an Analysis of Learner Discourse: Learner-Learner Collaborative Interaction in the Zone of Proximal Development. *Issues in Applied Linguistics*, 6(2). Retrieved from <https://escholarship.org/uc/item/0zs5j7ps>
- Ohta, Amy. S. (2001). *Second language acquisition processes in the classroom: Learning Japanese*. Mahwah, N.J: Lawrence Erlbaum Associates.
- 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., & Mackey, A. (2003). Interactional Context and Feedback in Child ESL Classrooms. *The Modern Language Journal*, 87(4), 519–533. Retrieved from <http://www.jstor.org/stable/1192801>
- Olsen, S. (1999). Errors and compensatory strategies: A study of grammar and vocabulary in texts written by Norwegian learners of English. *System*, 27(2), 191–205. [https://doi.org/10.1016/S0346-251X\(99\)00016-0](https://doi.org/10.1016/S0346-251X(99)00016-0)
- Pallotti, G. (2009). CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics*, 30(4), 590–601. <https://doi.org/10.1093/applin/amp045>
- 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>
- Payant, C., & Reagan, D. (2018). Manipulating task implementation variables with incipient Spanish language learners: A classroom-based study. *Language Teaching Research*, 22(2), 169–188. <https://doi.org/10.1177/1362168816669742>
- Philp, J., Adams, R., & Iwashita, N. (2013). *Peer Interaction and Second Language Learning*. Routledge.
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. New York: Viking Press.
- Pica, T. (1988). Interlanguage Adjustments as an Outcome of NS-NNS Negotiated Interaction. *Language Learning*, 38(1), 45–73. <https://doi.org/10.1111/j.1467-1770.1988.tb00401.x>
- Pica, T. (1992). The Textual Outcomes of Native Speaker-Non-Native Speaker Negotiation: What Do They Reveal About Second Language Learning? In C.

- J. Kramsch & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 198–237). Lexington, Mass: D.C. Heath.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and Using Communication Tasks for Second Language Instruction. In G. Crookes & S. M. Gass (Eds.), *Tasks and Language Learning: Integrating Theory and Practice* (pp. 9–34). Clevedon [England]: Multilingual Matters. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=23427&site=ehost-live>
- Poehner, M. E. (2008). Both Sides of the Conversation: The Interplay between Mediation and Learner Reciprocity in Dynamic Assessment. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 35–56). London ; Oakville, CT: Equinox Pub.
- Polio, C., & Williams, J. (2009). Teaching and Testing Writing. In Michael H. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 486–517). Chichester, U.K. ; Malden, MA: Wiley-Blackwell.
- Révész, A., Michel, M., Lu, X., Kourtali, N., Lee, M., & Borges, L. (2022). The relationship of proficiency to speed fluency, pausing, and eye-gaze behaviours in L2 writing. *Journal of Second Language Writing*, 58, 100927. <https://doi.org/10.1016/j.jslw.2022.100927>
- Rhonda, O. (2009). How Young is too Young? Investigating Negotiation of Meaning and Feedback in Children Aged Five to Seven Years. In S. M. Gass, A. Mackey, & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 135–156). New York: Routledge.

- Richards, J. C., & Schmidt, R. W. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed). Harlow: Longman.
- Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research*, 17(4), 368–392. <https://doi.org/10.1191/026765801681495877>
- Robinson, P. (2003). Attention and Memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 631–678). Oxford, UK: Blackwell Publishing Ltd.
<https://doi.org/10.1002/9780470756492.ch19>
- Robinson, P., Mackey, A., Gass, S. M., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247–267). London ; New York: Routledge.
- Roca de Larios, J., Manchón, R. M., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, 17(1), 30–47.
<https://doi.org/10.1016/j.jslw.2007.08.005>
- Roth, W.-M., & Radford, L. (2010). Re/thinking the Zone of Proximal Development (Symmetrically). *Mind, Culture, and Activity*, 17(4), 299–307.
<https://doi.org/10.1080/10749031003775038>
- Rouhshad, A., & Storch, N. (2016). 10. A focus on mode: Patterns of interaction in face-to-face and computer-mediated contexts. In M. Sato & S. Ballinger (Eds.), *Language Learning & Language Teaching* (Vol. 45, pp. 267–289). Amsterdam: John Benjamins Publishing Company.
<https://doi.org/10.1075/llt.45.11rou>

- Rusinovci, X. (2020). *EFL students' perceptions of the process-genre approach and their practices in writing compositions (Unpublished doctoral dissertation)* (University of Graz). University of Graz, Austria. Retrieved from <https://unipub.uni-graz.at/obvugrhs/content/titleinfo/5856622/full.pdf>
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Amsterdam ; Philadelphia, Pa: John Benjamins Pub.
- Sample, E., & Michel, M. (2015). An Exploratory Study into Trade-off Effects of Complexity, Accuracy, and Fluency on Young Learners' Oral Task Repetition. *TESL Canada Journal*, 31, 23. <https://doi.org/10.18806/tesl.v31i0.1185>
- Sarkhosh, M., & Najafi, S. (2020a). *Collaborative and Individual Writing: Effects on Accuracy and Fluency Development*. Retrieved from <https://digibug.ugr.es/handle/10481/62801>
- Sarkhosh, M., & Najafi, S. (2020b). Collaborative and Individual Writing: Effects on Accuracy and Fluency Development. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras*, (33), 27–42. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=7588121>
- Sasaki, M., & Hirose, K. (1996). Explanatory Variables for EFL Students' Expository Writing. *Language Learning*, 46(1), 137–168. <https://doi.org/10.1111/j.1467-1770.1996.tb00643.x>
- Sato, M. (2017). Interaction Mindsets, Interactional Behaviors, and L2 Development: An Affective-Social-Cognitive Model. *Language Learning*, 67(2), 249–283. <https://doi.org/10.1111/lang.12214>

- Sato, M., & Angulo, I. (2020). Chapter 2. The role of L1 use by high-proficiency learners in L2 vocabulary development: A quasi-experimental study of L1 languaging. In W. Suzuki & N. Storch (Eds.), *Language Learning & Language Teaching* (Vol. 55, pp. 42–66). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/llt.55.02sat>
- Sato, M., & Viveros, P. (2016). 3. Interaction or collaboration? Group dynamics in the foreign language classroom. In M. Sato & S. Ballinger (Eds.), *Language Learning & Language Teaching* (Vol. 45, pp. 91–112). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/llt.45.04sat>
- Schmidt, R. (1993). Awareness and Second Language Acquisition. *Annual Review of Applied Linguistics*, 13, 206–226.
<https://doi.org/10.1017/S0267190500002476>
- Schmidt, R. (2001). Attention. In Peter Robinson (Ed.), *Cognition and second language instruction* (Nachdr., pp. 3–32). Cambridge: Cambridge Univ. Press.
- Schmidt, R., & Frota, S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Rowley, MA: Newbury House.
- Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning1. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Scott, V. M., & Fuente, M. (2008). What's the Problem? L2 Learners' Use of the L1 during Consciousness-Raising, Form-Focused Tasks. *The Modern Language Journal*, 92(1), 100–113. Retrieved from <http://www.jstor.org/stable/25172995>

- Shaw, P., & Ting-Kun Liu, E. (1998). What Develops in the Development of Second-language Writing? *Applied Linguistics*, 19(2), 225–254.
<https://doi.org/10.1093/applin/19.2.225>
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263–300. <https://doi.org/10.1191/1362168804lr146oa>
- Sheen, Y. (2011). *Corrective Feedback, Individual Differences and Second Language Learning*. Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-94-007-0548-7>
- Sheen, Y., & Ellis, R. (2011). Corrective Feedback in Language Teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*. Vol. 2: ... (pp. 593–610). New York, NY: Routledge.
- Shehadeh, A. (1999). Non-Native Speakers' Production of Modified Comprehensible Output and Second Language Learning. *Language Learning*, 49(4), 627–675.
<https://doi.org/10.1111/0023-8333.00104>
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20(4), 286–305.
<https://doi.org/10.1016/j.jslw.2011.05.010>
- Soleimani, M., Modirghamene, S., & Sadeghi, K. (2017). Peer-mediated vs. individual writing: Measuring fluency, complexity, and accuracy in writing. *Innovation in Language Learning and Teaching*, 11(1), 86–100.
<https://doi.org/10.1080/17501229.2015.1043915>
- Storch, N. (2001). Comparing Esl Learners' Attention To Grammar On Three Different Classroom Tasks. *RELC Journal*, 32(2), 104–124.
<https://doi.org/10.1177/003368820103200207>

- Storch, N. (1998a). A classroom-based study: Insights from a collaborative text reconstruction task. *ELT Journal*, 52(4), 291–300.
<https://doi.org/10.1093/elt/52.4.291>
- Storch, N. (1998b). Comparing Second Language Learners' Attention to Form Across Tasks. *Language Awareness*, 7(4), 176–191.
<https://doi.org/10.1080/09658419808667108>
- Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research*, 5(1), 29–53.
<https://doi.org/10.1177/136216880100500103>
- Storch, N. (2002a). Patterns of Interaction in ESL Pair Work. *Language Learning*, 52(1), 119–158. <https://doi.org/10.1111/1467-9922.00179>
- Storch, N. (2002b). Relationships formed in dyadic interaction and opportunity for learning. *International Journal of Educational Research*, 37(3), 305–322.
[https://doi.org/10.1016/S0883-0355\(03\)00007-7](https://doi.org/10.1016/S0883-0355(03)00007-7)
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153–173.
<https://doi.org/10.1016/j.jslw.2005.05.002>
- Storch, N. (2008). Metatalk in a Pair Work Activity: Level of Engagement and Implications for Language Development. *Language Awareness*, 17(2), 95–114. <https://doi.org/10.1080/09658410802146644>
- Storch, N. (2009). *The nature of pair interaction: Learners' interaction in an ESL class: its nature and impact on grammatical development*. Saarbrücken: VDM Verlag.

- Storch, N. (2011). Collaborative Writing in L2 Contexts: Processes, Outcomes, and Future Directions. *Annual Review of Applied Linguistics*, 31, 275–288.
<https://doi.org/10.1017/S0267190511000079>
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol ; Buffalo: Multilingual Matters.
- Storch, N. (2016). Collaborative writing. In R. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 378–406). Boston: Walter de Gruyter Inc.
- Storch, N. (2017). Sociocultural Theory in the L2 Classroom. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 69–84). New York: Routledge.
- Storch, N. (2019). Collaborative writing. *Language Teaching*, 52(1), 40–59.
<https://doi.org/10.1017/S0261444818000320>
- Storch, N. (2020). Collaborative writing: Promoting languaging among language learners. In García Mayo, M. P. (Ed.), *Working collaboratively in second/foreign language learning* (1 st edition, pp. 13–34). Boston: De Gruyter.
- Storch, N. (2021). Collaborative writing: Promoting languaging among language learners. In García Mayo, M. P. (Ed.), *Working Collaboratively in Second/Foreign Language Learning* (pp. 13–34). De Gruyter.
<https://doi.org/10.1515/9781501511318-002>
- Storch, N., & Aldosari, A. (2010). Learners’ use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research*, 14(4), 355–375.
<https://doi.org/10.1177/1362168810375362>

- Storch, N., & Aldosari, A. (2013). Pairing learners in pair work activity. *Language Teaching Research*, 17(1), 31–48. <https://doi.org/10.1177/1362168812457530>
- Storch, N., & Wigglesworth, G. (2003). Is There a Role for the Use of the L1 in an L2 Setting? *TESOL Quarterly*, 37(4), 760. <https://doi.org/10.2307/3588224>
- Storch, N., & Wigglesworth, G. (2007). Writing Tasks: The Effects of Collaboration. In García Mayo, M. P. (Ed.), *Investigating tasks in formal language learning* (pp. 136–156). Clevedon [England] ; Buffalo: Multilingual Matters.
- Suzuki, W., & Storch, N. (2020). Introduction. In W. Suzuki & N. Storch (Eds.), *Language Learning & Language Teaching* (Vol. 55, pp. 2–15). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/llt.55.int>
- Swain, M. (2006). Linguaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London ; New York: Continuum.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing* (pp. 99–118). Longman.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In C. G. Madden (Ed.), *Input in second language acquisition* (pp. 235–253). Rowley, Mass: Newbury House Publishers.
- Swain, M. (1993). The Output Hypothesis: Just Speaking and Writing Aren't Enough. *Canadian Modern Language Review*, 50(1), 158–164. <https://doi.org/10.3138/cmlr.50.1.158>

- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & H. G. Widdowson (Eds.), *Principle & practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford Univ. 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 Univ. Press.
- Swain, M. (2006). Linguaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning. The contributions of Halliday and Vygotsky* (pp. 95–108). Oxford, England: Oxford University Press.
- Swain, M. (2010a). Talking-it-through: Linguaging as a source of learning. In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 112–130). Oxford ; New York: Oxford University Press.
- Swain, M. (2010b). Talking-it-through: Linguaging as a source of learning. In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 112–130). Oxford ; New York: Oxford University Press.
- Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). 9. Peer-Peer Dialogue as a Means of Second Language Learning. *Annual Review of Applied Linguistics*, 22, 171–185. <https://doi.org/10.1017/S0267190502000090>
- Swain, M., Kinnear, P., & Steinman, L. (2015). *Sociocultural theory in second language education: An introduction through narratives* (Second edition). Bristol ; Buffalo : Multilingual Matters.
- 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., & Lapkin, S. (2000). Task-based second language learning: The uses of the first language. *Language Teaching Research*, 4(3), 251–274. <https://doi.org/10.1177/136216880000400304>
- Swain, M., & Lapkin, S. (2013). A Vygotskian sociocultural perspective on immersion education: The L1/L2 debate. *Journal of Immersion and Content-Based Language Education*, 1(1), 101–129. <https://doi.org/10.1075/jicb.1.1.05swa>
- Swain, M., Lapkin, S., Knouzi, I., Suzuki, W., & Brooks, L. (2009). Languaging: University Students Learn the Grammatical Concept of Voice in French. *The Modern Language Journal*, 93(1), 5–29. <https://doi.org/10.1111/j.1540-4781.2009.00825.x>
- Swain, M., & Watanabe, Y. (2013). Languaging: Collaborative Dialogue as a Source of Second Language Learning. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 3218–3225). Oxford, UK: Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405198431.wbeal0664>
- Swain, M., & Watanabe, Y. (2020). Languaging: Collaborative dialogue as a source of second language learning. In C. A. Chapelle (Ed.), *The concise encyclopedia of applied linguistics* (pp. 667–674). Hoboken, NJ: John Wiley & Sons.
- Taguchi, N., & Kim, Y. (2016). Collaborative Dialogue in Learning Pragmatics: Pragmatic-Related Episodes as an Opportunity for Learning Request-Making. *Applied Linguistics*, 37(3), 416–437. <https://doi.org/10.1093/applin/amu039>
- Tognini, R., & Oliver, R. (2012). L1 use in primary and secondary foreign language classrooms and its contribution to learning. In E. Alcón Soler & M.-P. Safont-

- Jordà (Eds.), *Discourse and language learning across L2 instructional settings* (pp. 53–78). Brill | Rodopi. <https://doi.org/10.1163/9789401208598>
- Van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective*. Boston ; Dordrecht : Norwell, Mass: Kluwer Academic ; Distributors for North America, Central and South America, Kluwer Academic.
- Veramuthu, P., & Shah, P. (2020). Effectiveness of Collaborative Writing among Secondary School Students in an ESL Classroom. *Creative Education*, 11(01), 54–67. <https://doi.org/10.4236/ce.2020.111004>
- Villarrreal, I., & Gil-Sarratea, N. (2019). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research*, 1362168819829017. <https://doi.org/10.1177/1362168819829017>
- Villarrreal, I., & Munarriz-Ibarrola, M. (2021). “Together we do better”: The effect of pair and group work on young EFL learners’ written texts and attitudes. In García Mayo, M. P. (Ed.), *Working Collaboratively in Second/Foreign Language Learning* (pp. 89–116). De Gruyter. <https://doi.org/10.1515/9781501511318-005>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), *The Concept of activity in Soviet psychology* (pp. 144–188). Armonk, N.Y: M.E. Sharpe.
- Wagner, E., & Toth, P. (2013). 5. Building explicit L2 Spanish knowledge through guided induction in small group and whole class interaction. In K. McDonough & A. Mackey (Eds.), *Language Learning & Language Teaching*

- (Vol. 34, pp. 89–108). Amsterdam: John Benjamins Publishing Company.
<https://doi.org/10.1075/llt.34.08ch5>
- Wajnryb, R. (1990). *Grammar dictation* (8. impression). Oxford: Oxford Univ. Press.
- Wang, Y.-C. (2015). Promoting collaborative writing through wikis: A new approach for advancing innovative and active learning in an ESP context. *Computer Assisted Language Learning*, 28(6), 499–512.
<https://doi.org/10.1080/09588221.2014.881386>
- Watanabe, Y. (2008). Peer-Peer Interaction between L2 Learners of Different Proficiency Levels: Their Interactions and Reflections. *Canadian Modern Language Review*, 64(4), 605–635. <https://doi.org/10.3138/cmlr.64.4.605>
- Watanabe, Y. (2019). 4. The Role of Language in Collaborative and Individual Writing: When Pairs Outperform Individuals. In M. Haneda & H. Nassaji (Eds.), *Perspectives on Language as Action* (pp. 63–79). Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781788922944-007>
- 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>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- White, A. H. (2007). A Tool for Monitoring the Development of Written English: T-Unit Analysis Using the SAWL. *American Annals of the Deaf*, 152(1), 29–41.
<https://doi.org/10.1353/aad.2007.0016>
- Wiegle, S. C. (2014). Considerations for Teaching Second Language Writing. In M. Celce-Murcia, D. Brinton, & M. A. Snow (Eds.), *Teaching English as a*

- second or foreign language* (Fourth edition, pp. 222–237). Boston, Mass.: National Geographic Learning [u.a.].
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445–466.
<https://doi.org/10.1177/0265532209104670>
- Wigglesworth, G., & Storch, N. (2012). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing*, 21(4), 364–374.
<https://doi.org/10.1016/j.jslw.2012.09.005>
- Williams, J. (1999). Learner-Generated Attention to Form. *Language Learning*, 49(4), 583–625. <https://doi.org/10.1111/0023-8333.00103>
- Williams, J. (2001). The effectiveness of spontaneous attention to form. *System*, 29(3), 325–340. [https://doi.org/10.1016/S0346-251X\(01\)00022-7](https://doi.org/10.1016/S0346-251X(01)00022-7)
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21(4), 321–331.
<https://doi.org/10.1016/j.jslw.2012.09.007>
- Williams, J. D. (2003). *Preparing to teach writing: Research, theory, and practice* (3rd ed). Mahwah, N.J: Lawrence Erlbaum.
- Wilson, K., & Devereux, L. (2014). Scaffolding theory: High challenge, high support in Academic Language and Learning (ALL) contexts. *Journal of Academic Language and Learning*, 8(3), A91–A100. Retrieved from <https://journal.aall.org.au/index.php/jall/article/view/353>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, & Complexity*. Second Language Teaching & Curriculum Center, University of Hawaii at Manoa.

- Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100.
<https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Yoshida, R. (2008). Teachers' Choice and Learners' Preference of Corrective Feedback Types. *Language Awareness*, 17(1), 78–93.
<https://doi.org/10.2167/la429.0>
- Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching*, 49(4), 461–493.
<https://doi.org/10.1017/S0261444816000161>
- Zabihi, R., & Bayan, M. (2020). Are Two Voices Better Than One? Comparing Aspects of Text Quality and Authorial Voice in Paired and Independent L2 Writing. *Written Communication*, 37(4), 512–535.
<https://doi.org/10.1177/0741088320939542>
- Zabihi, R., & Ghahramanzadeh, B. (2022). Proficiency pairing and engagement with language in L2 peer interaction. *System*, 105, 102725.
<https://doi.org/10.1016/j.system.2022.102725>
- Zamel, V. (1982). Writing: The Process of Discovering Meaning. *TESOL Quarterly*, 16(2), 195–209. <https://doi.org/10.2307/3586792>
- Zamel, V. (1983). The Composing Processes of Advanced ESL Students: Six Case Studies. *TESOL Quarterly*, 17(2), 165–187. <https://doi.org/10.2307/3586647>
- Zamel, V. (1987). Recent Research on Writing Pedagogy. *TESOL Quarterly*, 21(4), 697–715. <https://doi.org/10.2307/3586990>
- Zhang, M. (2018). Collaborative writing in the EFL classroom: The effects of L1 and L2 use. *System*, 76, 1–12. <https://doi.org/10.1016/j.system.2018.04.009>

- Zhang, M. (2019a). Towards a quantitative model of understanding the dynamics of collaboration in collaborative writing. *Journal of Second Language Writing*, 45, 16–30. <https://doi.org/10.1016/j.jslw.2019.04.001>
- Zhang, M. (2019b). Understanding L1 and L2 interaction in collaborative writing: A lexico-grammatical analysis. *Language Teaching Research*, 1362168819859911. <https://doi.org/10.1177/1362168819859911>
- Zhang, M., & Crawford, W. J. (2021). Attention to form in collaborative writing: Language-related episodes in L1 and L2 use conditions. *Language Awareness*, 0(0), 1–19. <https://doi.org/10.1080/09658416.2021.1998087>

Deutsche Zusammenfassung

Forschung zu kollaborativem Schreiben hat in den letzten Jahren stetig an Bedeutung gewonnen. Die Mehrzahl an Studien mit erwachsenen EFL/ESL Lernenden hat demonstriert, dass kollaboratives Schreiben viele Vorteile für das Erlernen von Zweit- und Fremdsprachen haben kann, in etwa durch die Schaffung von Übungsmöglichkeiten in der L2 (z.B. Wigglesworth & Storch, 2009), im Hinblick auf die Erlernung von L2-Vokabular (z.B. Kim, 2008), und durch die Entwicklung von L2 Schreibkompetenzen (z.B. Chen, 2019). Forschungserkenntnisse haben auch gezeigt, dass kollaboratives Schreiben zu Verbesserungen in der Textproduktion von Lernenden führt, sowohl in qualitativer als auch quantitativer Hinsicht. Diese Vorteile konnten bis jetzt allerdings nur in Studien mit erwachsenen EFL/ESL Lernenden nachgewiesen werden (z.B. Fernández Dobao, 2012; Storch & Wigglesworth, 2007); wie sich die Anwendung kollaborativen Schreibens auf Kinder und jugendliche Lernende auswirkt, wurde hingegen noch kaum beforscht. Ebenso wenig wurde in der vorhandenen Forschung auf die Korrelation zwischen der L2-Kompetenz der Lernenden und das sprachliche Niveau der schriftlichen Textproduktion sowie auf die Qualität und Quantität des „Languagings“ (Swain, 2006) eingegangen.

Ausgehend von diesen Forschungslücken wurden in dieser Arbeit die Schreibkompetenz von 144 jugendlichen (13-14-jährigen) EFL-Lernenden an einer kosovarischen Sekundarschule verglichen, die verschiedene Leistungsniveaus (niedrig – durchschnittlich – überdurchschnittlich) aufwiesen. Die Jugendlichen hatten den Schreibauftrag, auf Basis eines Bildimpulses Geschichten zu schreiben, wodurch 48 (n=48) individuelle und 48 (n=48) in kollaborativer Partnerarbeit verfasste Texte entstanden. Diese 96 Texte wurden anhand von quantitativen Aspekten wie fließende Sprachbeherrschung, Komplexität, Korrektheit und Genauigkeit, sowie anhand von qualitativen Aspekten wie Aufgabenbewältigung, Kohärenz und Kohäsion, Grammatik, Vokabular und Sprachmechanik analysiert. Weiters wurde der Diskurs der kollaborierenden Paare in Hinsicht auf „Language Related Episodes“ (LREs) untersucht. Dabei wurde die Anzahl, Art, und das Resultat der LREs von 24 Paaren unterschiedlicher Leistungsniveaus festgestellt, um zu ergründen, ob die L2 Kompetenz der Lerner eine Auswirkung auf quantitative und qualitative Aspekte der LREs hat. Diese Analyse lieferte vier wesentliche Erkenntnisse: (a) im Gegensatz zu alleine arbeitenden Lernenden produzierten Paare hochwertigere Texte in Hinblick auf

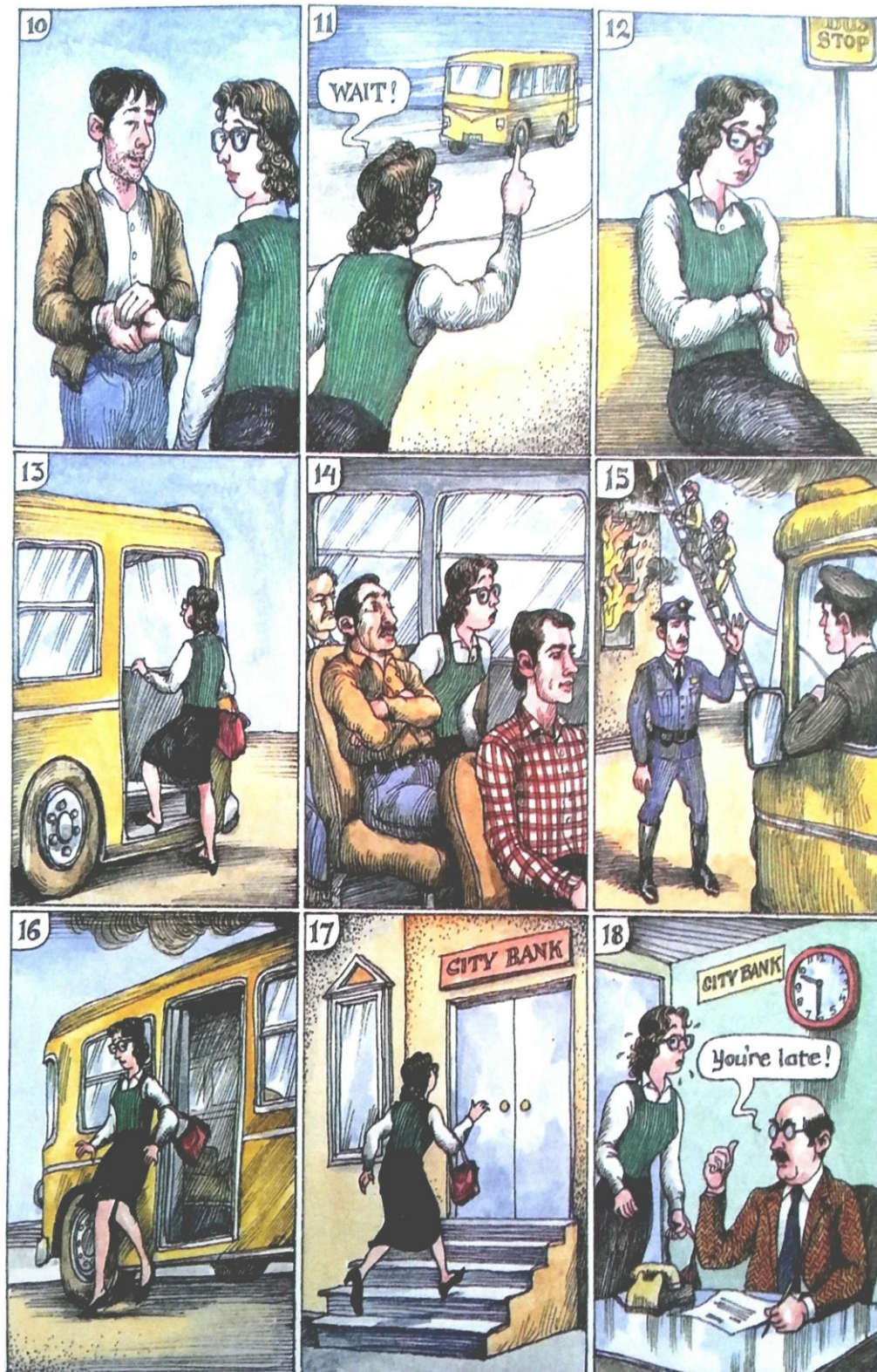
fließende Beherrschung der Sprache, Komplexität, und genauem Sprachgebrauch, wobei jedoch nur in den Bereichen Komplexität und genauem Sprachgebrauch signifikante Unterschiede festzustellen waren (Relation von Nebensätzen zu Sätzen); (b) im Gegensatz zu alleine arbeitenden Lernenden erzielten Paare erzielten signifikant höhere Punktzahlen in den Bereichen Aufgabenbewältigung, Kohärenz und Kohäsion, Grammatik, und Vokabular, jedoch nicht in Sprachmechanik; (c) eine höhere L2 Sprachkompetenz der Lernenden wirkt sich durchwegs positiv auf die produzierten Texte von Paaren und alleine arbeitenden Sprachlernenden aus; (d) die Anzahl an lexikalischen LREs war im Gegensatz zu grammatikalischen und mechanischen LREs am höchsten, und eine höhere L2 Sprachkompetenz entsprach einer größeren Anzahl an produzierten und korrekt ausgeführten LREs. Weiters bestätigt diese Arbeit interaktionistische Perspektiven und die Annahme, dass Sprachlernende sich durch Interaktionsanpassung begünstigende Lernkontexte für den L2 Erwerb schaffen (Long, 1996; Mackey, 2012). In pädagogischer Hinsicht bieten die Ergebnisse dieser Studie eine Hilfestellung für L2-Lehrkräfte, die überlegen, wann und wie sie kollaboratives Schreiben in ihrem L2-Unterricht mit größtmöglichen Erfolgen einsetzen können.

Appendices

Appendix A: Visual writing prompt (obtained from Harris & Rowe, 1995, p. 110-111)



Appendix A: Visual writing prompt (Continued)



Appendix B: The Global Rating Rubric – developed by Gassner et al., (2019)

E8 Writing Rating Scale					
	Task achievement	Coherence and cohesion	Grammar	Vocabulary	Mechanics
7	Complete TA all 4 content points mentioned & 1 or more elaborated „all 6 content points mentioned & 3 or more elaborated	clear and coherent text cohesive devices used successfully to produce a fairly fluent text cohesion on both sentence and paragraph level	good range of structures relatively high degree of control, with few inaccuracies which do not impair communication	good range of vocabulary generally accurate vocabulary formulations sometimes varied to avoid repetition	good knowledge of spelling and capitalization almost no errors in spelling and capitalization
6					
5	Good TA all 4 content points mentioned & 1 weakly elaborated 5 content points mentioned & 2 or 3 elaborated OR all 6 content points mentioned & 1 or 2 elaborated	„mostly clear and coherent text good sentence level cohesion as a linear sequence on a simple level some paragraph level coherence and cohesion	sufficient range of structures occasional inaccuracies which do not impair communication	sufficient range of vocabulary occasionally inaccurate vocabulary major errors possible when expressing more complex ideas	sufficient knowledge of spelling and capitalization occasional errors in spelling and capitalization
4					
3	Sufficient TA 3 content points mentioned & no elaboration OR 2 content points mentioned & 1 elaborated „4 content points	text often lacks clarity and/or coherence some simple sentence level cohesion using simple connectors like ‘and’, ‘but’ and ‘because’ frequent lack of coherence and	limited range of simple structures some inaccuracies which can impair communication	limited range of vocabulary some inaccurate vocabulary, generally without causing breakdown of communication tendency to use phrases from the	limited knowledge of spelling and capitalization some errors in spelling and capitalization

	mentioned & 1 or 2 elaborated OR all 6 content mentioned and no elaboration			prompt	
2					
1	Some TA 2 content points mentioned & no elaboration 3 content points mentioned & no elaboration	text not coherent basic linear connectors (‘and’, ‘then’) on word or word group level	extremely limited range of simple structures limited control, with frequent inaccuracies sometimes causing breakdown of communication	extremely limited range of vocabulary mostly inaccurate vocabulary causing frequent breakdown of communication several chunks lifted from the prompt	extremely limited knowledge of spelling and capitalization frequent of errors in spelling and capitalization
0					

Retrieved from: <http://seminar.koe-at.com/#WRITING%20E8> (last accessed 24 May 2022)

***Note.** The area of mechanics was added to the rubric for the purpose of the study.

Appendix B (Continued)

Text length:

- Short prompt: 40-70 words
- Long prompt: 120-180 words

Answers containing fewer than 20 words/60 words are band 0.

Downgrading:

Task Achievement: It is expected that text type requirements are met. If they are not met (missing/inappropriate salutation and/or closing formula), downgrade by one band.

Coherence and Cohesion: In long texts, paragraphing is mandatory for *bands* 3–7. If there are no visual paragraphs, downgrade by one band.

Task Achievement - Short Prompts

CPs mentioned	CPs elaborated	Band
4	1/2 (1 excellent or 2 good)	7
4	1/2 (1 good or 2 weak)	6
4	1 weak	5
4	0	4
3	2	5
3	1	4
3	0	3
2	1 good	3
2	1 weak	2
2	0	1

Task Achievement – Long Prompts

CPs mentioned	CPs elaborated	Band
6	3 or more	7
6	3 (1 or 2 weak)	6
6	2 good	6
6	1 or 2	5
6	1 weak	4
6	0	3
5	2 or 3 (all good)	6
5	2 or 3	5
5	1 good	4
5	1 weak	3
5	0	2
4	1 or 2 good	4
4	1 or 2	3
4	0	2
3	1	2
3	0	1
2	1	1

Appendix C: Writing Prompt for Pairs

Now, I would like the two of you to work together and write the story of Ana. What happened to Ana this morning? You can start like this: **This morning Ana woke up at seven o'clock. The sun was shining through the window.**

- a. Use the past simple tense whenever possible. Remember to switch on the voice recorder before you start writing the story
- b. If you have any questions please feel free to ask your teacher for help.
- c. You have 30 minutes to finish the task.

Appendix D: Writing Prompt for Individuals

Now, I would like you to write the story of Ana. What happened to Ana this morning?

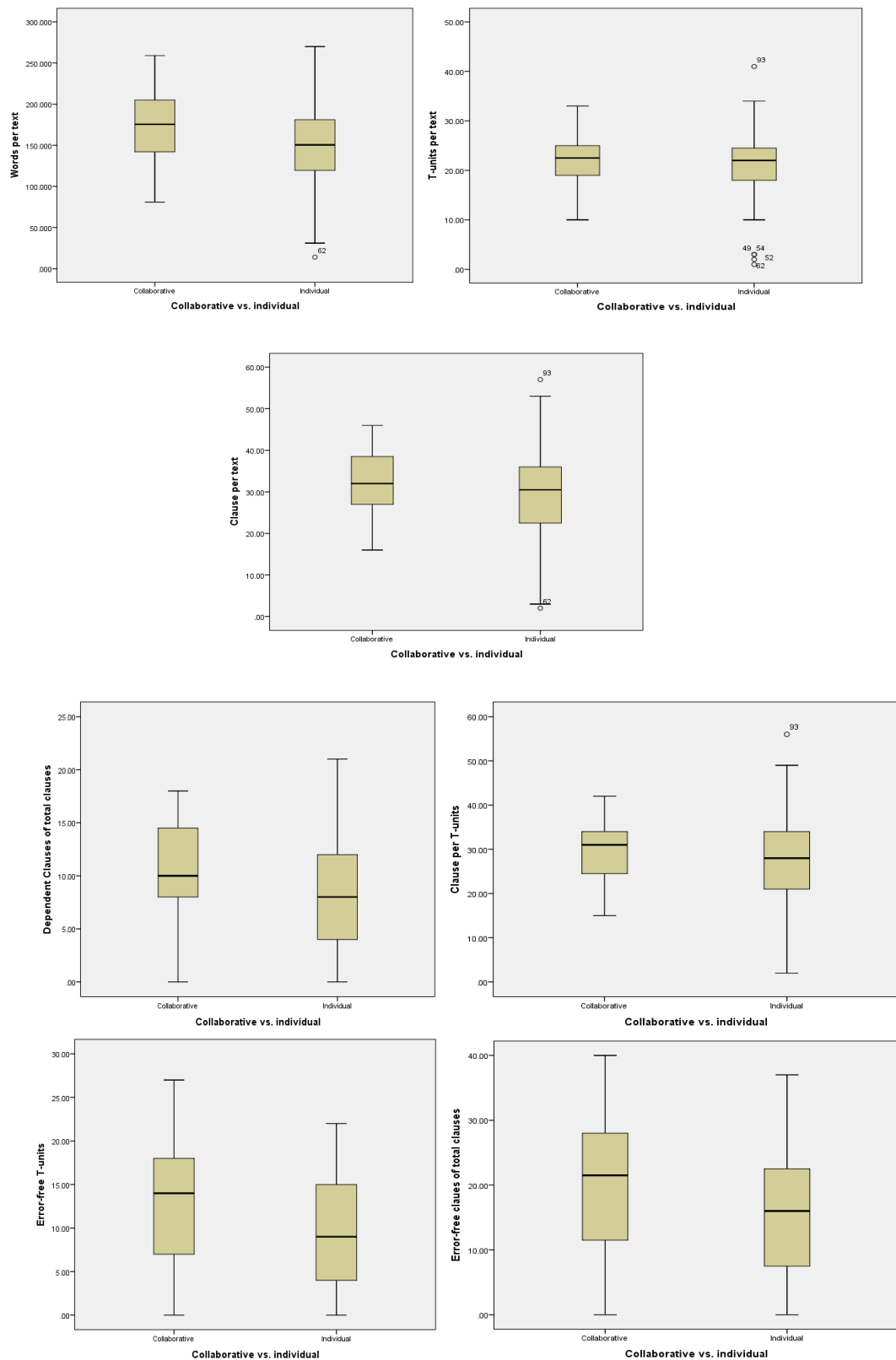
You can start like this: **This morning Ana woke up at seven o'clock. The sun was shining through the window.**

- a. Use the past simple tense whenever possible.
- b. If you have any questions please feel free to ask your teacher for help.

Appendix E: Distribution of quantitative scores

Condition	Measures	Sample	Shapiro-Wilk Test	Boxplots
Quantitative measures				
CW	Words per text	48	W (48) = 0.98, $p = .633$	Figure 1
IW		48	W (48) = 0.97, $p = .460$	
CW	T-units per text	48	W (48) = 0.98, $p = .692$	
IW		48	W (48) = 0.92, $p = .004$	
CW	Clauses per text	48	W (48) = 0.97, $p = .471$	
IW		48	W (48) = 0.96, $p = .193$	
CW	Clauses per T-units	48	W (48) = 0.97, $p = .364$	Figure 1
IW		48	W (48) = 0.96, $p = .133$	
CW	Dependent Clauses per clauses	48	W (48) = 0.96, $p = .113$	
IW		48	W (48) = 0.95, $p = .080$	
CW	Error-free T-units	48	W (48) = 0.97, $p = .469$	Figure 1
IW		48	W (48) = 0.95, $p = .043$	
CW	Error-free clauses	48	W (48) = 0.96, $p = .162$	
IW		48	W (48) = 0.96, $p = .118$	

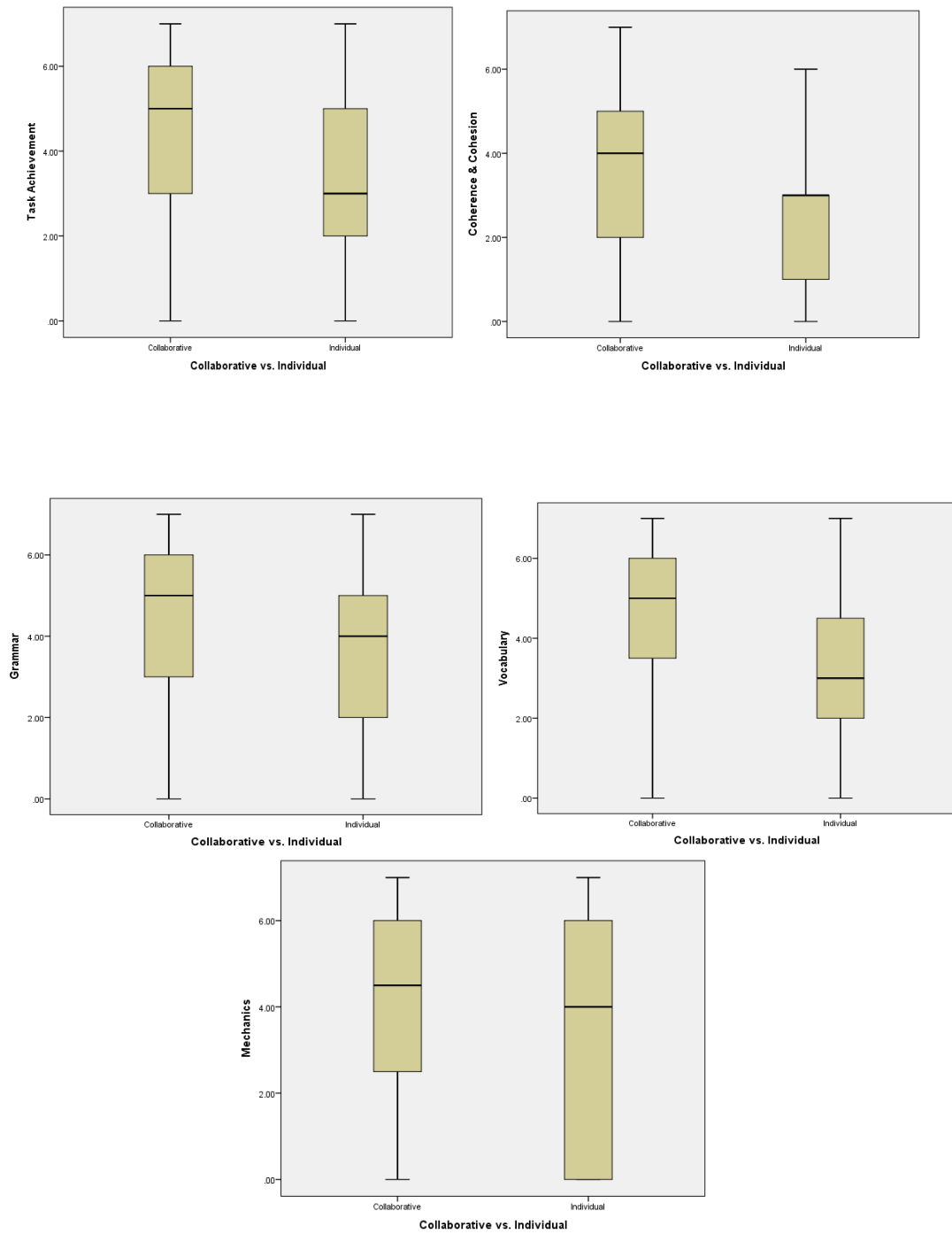
Figure 1. Boxplots - Distribution of quantitative scores



Appendix F: Distribution of holistic scores

Condition	Measures	Sample	Shapiro-Wilk Test	Boxplots
Qualitative measures				
CW	Task achievement	48	W (48) = 0.92, $p = .004$	Figure 2
IW		48	W (48) = 0.95, $p = .057$	
CW	Coherence & cohesion	48	W (48) = 0.94, $p = .038$	
IW		48	W (48) = 0.93, $p = .011$	
CW	Grammar	48	W (48) = 0.92, $p = .004$	
IW		48	W (48) = 0.92, $p = .006$	
CW	Vocabulary	48	W (48) = 0.89, $p = .001$	
IW		48	W (48) = 0.94, $p = .038$	
CW	Mechanics	48	W (48) = 0.88, $p = .001$	
IW		48	W (48) = 0.83, $p = .001$	

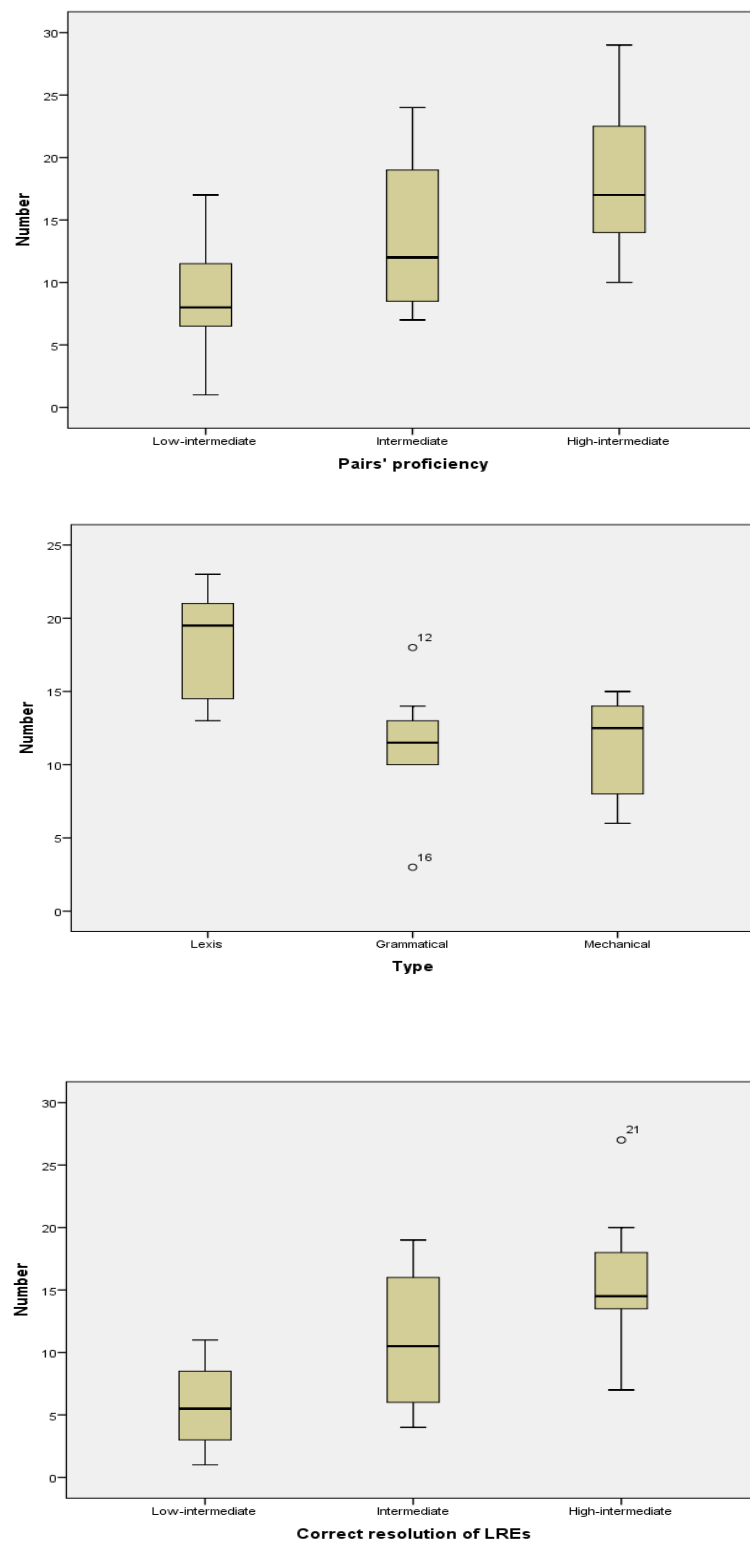
Figure 2. Boxplots - Distribution of holistic scores



Appendix G: Distribution of LRE data

Condition	Measures	Sample	Shapiro-Wilk Test	Boxplots
LREs				
LI	Number	8	W (8) = 0.94, $p = .631$	Figure 3
I	Number	8	W (8) = 0.91, $p = .371$	
HI	Number	8	W (8) = 0.92, $p = .446$	
Type	Lexical	8	W (8) = 0.90, $p = .294$	
Type	Grammatical	8	W (8) = 0.92, $p = .433$	
Type	Mechanical	8	W (8) = 0.84, $p = .206$	
LI	Outcome	8	W (8) = 0.97, $p = .932$	
I	Outcome	8	W (8) = 0.91, $p = .392$	
HI	Outcome	8	W (8) = 0.92, $p = .433$	

Figure 3. Boxplots - Distribution of LRE data



Fletëpëlqim për pjesëmarrje në hulumtim

Titulli: Affordances and challenges of collaborative writing in lower EFL classrooms in Kosovo

Hulumtuesi: Nuhi Bllaca (student i doktoratës në Universitetin e Vjenës, Austri.)

Email: nuhibllaca@gmail.com

Hyrje

Ftesë për pjesëmarrje në hulumtim

Me këtë hulumtim, ne do të hetojmë ndikimin e shkrimit të përbashkët në mësimin e gjuhës angleze nga nxënësit e shkollave të mesme të ulëta në Kosovë. Në këtë fletë pëlqim shpjegohet qëllimi i studimit dhe të drejtat e nxënësve në hulumtim. Vendimi për të marrë pjesë ose jo në studim ju takon juve (nxënësve dhe prindërve si përfaqësues të fëmijëve të tyre). Nëse vendosni që fëmija juaj të marrë pjesë në këtë hulumtim, ju lutem nënshkruani fletëpëlqimin, në rreshtin e fundit të fletës dhe shkruani datën e nënshkrimit në zbrazëtirë.

Shpjegimi i hulumtimit

Përmes këtij hulumtimi ne synojmë të hetojmë dobritë dhe sfidat e mundshme të shkrimit të përbashkët nga nxënësit e shkollave të mesme të ulëta në Kosovë. Shkrimi i përbashkët është të shkruarit e një teksti të vetëm nga dy e më shumë nxënës. Kur nxënësit angazhohen të shkruajnë në mënyrë të përbashkët, ata komunikojnë mes vete, ndajnë ide dhe marrin vendime, dhe diskutojnë shumë për fjalët, formën dhe rregullat gjuhësore që do t'i përdorin gjatë shkrimit. Hulumtimet kanë treguar se këto lloje diskutimesh janë shumë të dobishme për avancimin dhe mësimin e gjuhës së huaj nga nxënësit që shkruajnë ose punojnë në dyshe/çifte. Prandaj ky hulumtim tenton të vë në pa dobritë dhe sfidat e shkrimit të përbashkët për mësimin e gjuhës angleze në Kosovë. Në këtë hulumtim do të përfshihen 148 nxënës nga klasat e 8-ta. Hulumtimi është i ndarë në dy pjesë. Pjesa e parë e hulumtimit do të realizohet në shkurt (2020) dhe do të përfshihen një numër i vogël i nxënësve, pasi që kjo pjesë shërben vetëm si fazë testuese

(Pilot Study). Ndërsa, pjesa e dytë, që është pjesa kryesore e hulumtimit do të realizohet në maj (2020) dhe do të përfshihen 144 nxënës.

Procedura e hulumtimit

Nxënësit e përfshirë në këtë hulumtim do të ndahen në dy grupe. Grupi i parë i nxënësve do të shërbej si grup kontrolli dhe do të angazhohet të shkruajë një temë ose tregim bazuar në një set fotografish. Pra, nxënësit do ta ndërtojnë tregimin bazuar në fotografi. Kurse, grupi i dytë i nxënësve do të shkruaj për të njëjtën temë, duke u bazuar në fotografi të njëjta, porse dy nxënës do ta shkruajnë temën bashkërisht. Biseda e nxënësve në grupin e dytë do të incizohet, me pajisje incizues (diktafon, audio incizues ose pajisje tjetër incizuese). Pas përfundimit të shkrimit/studimit, biseda e incizuar do të analizohet për qëllimet e hulumtimit.

Konfidencialiteti

Të gjitha të dhënat e mbledhura në këtë hulumtim do të jenë konfidenciale dhe do të përdoren vetëm për qëllime të hulumtimit. Identiteti i pjesëmarrësve do të jetë anonim dhe çdo e dhënë do të trajtohet me kujdes të shtuar. Përveç hulumtuesit, askush nuk do të ketë qasje në të dhënat e grumbulluara.

Pjesëmarrja në hulumtim

Pjesëmarrja në hulumtim është krejtësisht vullnetare dhe nxënësit kanë të drejtën që të tërhiqen nga hulumtimi kurdo që e konsiderojnë të arsyeshme. Pjesëmarrja ose mospjesëmarrja në hulumtim nuk ndikon aspak në vlerësimin e nxënësve në lëndën e gjuhës angleze. Prandaj nxënësit kanë të drejtën e tyre të vendosin sipas vullnetit. Meqë hulumtimi bëhet për të mirën e përgjithshme, pjesëmarrja e nxënësve në këtë studim do ta bënte të mundur zbardhjen e dobive dhe sfidave të shkrimit të përbashkët për mësimin e gjuhës angleze në Kosovë. Në secilin rast që e konsideroni të nevojshme për të bërë pyetje shtesë, ju si prindër mund të më shkruani në e-mail adresën e dhënë dhe do t'ju përgjigjem me kënaqësi.

Deklarata e hulumtuesit

Unë Nuhi Bllaca studiues e mësimdhënës i gjuhës angleze, ua kam shpjeguar dhe qartësuar nxënësve qëllimin e hulumtimit, procedurat dhe detyrat e tyre dhe u jam përgjigjur atyre në secilën pyetje lidhur me këtë hulumtim.

Nënshkrimi i hulumtuesit _____ data _____

Pëlqimi i prindit/ës

Unë e kam lexuar më vëmendje çdo informatë në këtë fletë pëlqim dhe të gjitha paqartësitë/pyetjet më janë qartësuar. Unë me vullnet të plotë jap pëlqimin që fëmija im të marrë pjesë në këtë hulumtim.

Nënshkrimi juaj _____ data _____

Appendix H: Form of consent – in English (Continued)

Form of Consent

Title: Affordances and challenges of collaborative writing in lower EFL classrooms in Kosovo

Researcher: Nuhi Bllaca (Ph.D. Student, University of Vienna)

Email: nuhibllaca@gmail.com

Introduction

Invitation for participation in research

In this research, we aim to investigate the benefits of collaborative writing for English language learning in lower EFL classrooms in Kosovo. In this form of consent, we explain the purpose of the study and the rights of prospective participants. The decision whether to participate in this study or not belongs to you and your parents. Therefore, I would kindly ask you to deliver this form of consent to your parents to sign only if you would like to participate in this study. So if you agree to participate in the study, your parents can sign it at the end of the paper.

Explaining the research

In this study we aim to investigate the affordances and challenges of collaborative writing in lower-secondary EFL classrooms in Kosovo. Collaborative writing is defined

as writing of single text by two or more authors (e.g. learners). When learners work in pairs to write a piece of text, it is most likely that they will interact with one another to resolve all the contested issues that might be related to task content, task organization, generation of ideas, language use and the like. Prior research has shown that interaction among peers while writing together benefits L2 learning in different ways. Therefore, the present study aims to examine the affordances and challenges of collaborative writing for L2 learning in Kosovo. Participants of this study will be 144 from grade 8. The study is divided in two stages. The first stage planned to be carried out in February (2020), will include only a small number of participants, because its purpose is to test the instruments to be used in the main study. While, the main study will be conducted in the second stage, and will include around 144 participants.

Research procedures

Participants that will take part in this research will be assigned to work in two groups. One group will serve as a control group, whose main job will be to write a story based on a visual prompt (e.g. a set of pictures). While the second group will be as an experimental group and will complete the same task in pairs. Thus, participants of the experimental group will carry out the task in pairs and their pair talk will be audio-recorded. After the task completion, compositions produced by both groups of learners, and the pair talks will be analyzed only in compliance to the research purposes.

Confidentiality

All the data gathered in this research will be confidential and utilized only for the research purposes. Participants' identity will be anonym and every single piece of information will be treated carefully. Apart from the researcher/s no one else will be able to access the data gathered.

Participation in the study

Participation in this research is completely voluntary and you are entitled the right to withdraw at any stage of the research without consequences. Additionally, your decision to participate or not to participate in this study will in no way affect your school grades in English class. However, your participation in this research will make it possible for me as a researcher to get more insights over the potential benefits of writing in pairs for lower-secondary EFL learners in Kosovo. Finally, whenever you and your parents have

any questions or concerns, you may contact me via email and will get all the information you need.

Statement of the researcher

I, Nuhi Bllaca, teacher of English, have informed and clearly explained to the students the purpose of the study, tasks to be carried and have responded to all questions of participants planed to be involved in the study.

Signature of the researcher _____ data _____

Consent of parent/guardian

I have attentively read every piece of information in this form of consent and every question of mine has been clearly answered. As such, I whole heartedly give my consent that my child be included in this study.

Your signature _____ date _____

Appendix I: A collaborative text written by Dua and Ber (#1 high-intermediate)

This morning Ana woke up at seven o'clock//. The sun was shining through the window//. She stood up// and got ready for work//. Then went out of her apartment//, she was going down the stairs/ when a dog came/ barking of* her//. Ana got scared// and *acidantly droped** her purse on the *graund**//. Her stuff went out of the purse//. She sat down/ collecting her things/ while in* the other side of the street the owner of the dog was yelling at him,/ for scaring the lady//. When Ana stood up/ she saw /that the bus has* arrived/ and people where already geting* in//, she tried to catch the bus// but the bus* already left//. She was waiting at the bus stop/ when a homeless man showed up at the bus stop//. Ana felt sorry for him// so she *offerd** him some money//. While she was giving the money to him/ she missed the bus again//so she sat down on* the bus stop/ until the other bus came//. She got in/ and while she was in the bus// she saw a fire//, so she got out of the bus //and run* to her work//. When she got in /th* boss told her/ she was late//.

CAF		Errors	
Words	206	Grammar	7
T-units	23	Lexis	0
C/T	35	Mechanics	5
DC	15	Total	12
C	38		
EFT	14		
EFC	27		

Note. Among independent and dependent clauses only the latter are underlined. Language errors are marked with asterisk (*). The asterisk (*) is put next to an erroneous word or in case of missing elements (e.g. articles, prepositions, verbs, etc.) it is placed next to, before, or after the word where such an element should have been placed in the text. The boundary of a T-unit is marked with //, while that of a clause with /. When there is more than one T-unit, one after the other, they are not always marked with //, but instead separated with /.

Figure 4: A collaborative text written by Dua and Ber (hand-written)

Now I would like the two of you to work together and write the story of Ana. What happened to Ana this morning? You can start like this: **This morning Ana woke up at seven o'clock. The sun was shining through the window.** If you have any questions, please feel free to ask your teacher.

This morning Ana woke up at seven o'clock. The sun was shining through the window. She stood up and ~~and~~ get ready for work. Then went out of her apartment, she was going down the stairs when a dog came barking at her. ~~She~~ Ana got scared and (accidentally) (dropped) her purse on the (ground). Her stuff went out of the purse. She sat down collecting her things while in the other side of the street ~~the~~ owner of the dog was yelling at him for scaring the lady. When Ana stood up she saw that the bus (has) arrived and people were already (getting) in, she tried to catch the bus but ~~the~~ bus already left. She was waiting at the bus stop when a homeless man showed up ~~at~~ (at the bus stop). ~~She~~ Ana felt sorry for him so ~~she~~ she (offered) him some money. While she was giving the money to him ~~she~~ ~~she~~ she missed the bus again, so she sat down ~~on~~ the bus stop until the other bus came. She got in and while she was in the bus ~~she~~ saw a fire, so she got out of the bus and (ran) to her work. When she got in (th) boss told her she was late.

Appendix J: An individual text composed by Mul (#49 – high-intermediate)

Ana woke up at seven o'clock// she had to go to work.// She opend* the outside door/
 and she saw a dog /barking at hër*// She threw all the work stuff/ bc* she was scared.//
 The owner of the dog calmed the dog,// Ana got her stuff /and start* /going hër* way.//
 She went to the bus station*// and saw/ her bus was about to go away.// She waited at
 the bus stop/ bc* her bus ran away.// She saw a homeless* guy nearby / and gave him
 some changes.// She waited hër* bus/ and the bus* came again /and * saw the bus /she
 said WAIT!// so the bus can* stop.// She got into the bus.// Down the road she saw an
 accident with the police and fierfigters*// she got out of the bus/ bc* it was taking a lot
of time //and ran at* the city bank /the spot she worked.// When she came in/ her boss
 told her/ she was late //and Ana couldn't belive*/ she was that late// /it was about to be
 10 o'clock.//

CAF		Errors	
Words	175	Grammar	4
T-units	25	Lexis	2
C/T	35	Mechanics	10
DC	11	Total	16
C	36		
EFT	9		
EFC	20		

Note. Among independent and dependent clauses only the latter are underlined. Language errors are marked with asterisk (*). The asterisk (*) is put next to an erroneous word or in case of missing elements (e.g. articles, prepositions, verbs, etc.) it is placed next to, before, or after the word where such an element should have been placed in the text. The boundary of a T-unit is marked with //, while that of a clause with /. When there is more than one T-unit, one after the other, they are not always marked with //, but instead separated with /.

Figure 5: An individual text composed by Mul (hand-written)

Now I would like you to write the story of Ana. What happened to Ana this morning? You can start like this: **This morning Ana woke up at seven o'clock. The sun was shining through the window.** If you have any questions, please feel free to ask your teacher.

Ana woke up at seven o'clock, she had to go to work. She opened the outside door and she saw a dog barking at her. She threw all the work stuff because she was scared. The owner of the dog calmed the dog, she got her stuff and started going her way. She went to the bus station and saw ~~that~~ her bus was about to go away. She waited at the bus stop because her bus ran away. She saw a homeless guy near by and gave him some change. She waited for her bus and the bus came ~~and~~ again and saw the bus she said WAIT! so the bus can stop. She got into the bus and down the road she saw an accident with the police and firefighters. She got out of the bus because it was taking a lot of time and ran at the city bank. The spot she worked. When she came in her boss told her she was late and she couldn't believe she was that late it was about to be 10 o'clock.

175 words

Appendix K: A collaborative text composed by Alu and Gash (#17 – intermediate)

This morning Ana woke up at seven o'clock.// The sun was shining through the window.//Ana got ready/ to go on* her job*.// While she was going to the bus station*/ a dog scared her //and a man yelled to* that dog /and Ana was getting the things/ that fell from the bag.// After she got her bag,/ the bus was already on* the bus station*./ Ana tried /to stop the bus /but he* was gone already.// Ana sitted* on* the bus station*,/ to wait* the* next bus.// While she was waiting*/ a poor man didn't have money// and Ana gaved* him some.// The poor man thanked her so much/ and after that another bus got* away.// Ana still waited for another one.// Ana got on the bus/ and while she was on her way// the police stopped the bus /and told them*// that they cant* pass this way/ because a building was on fire.// Ana was running/ to get on* her Job*.// After she got there /Mr. Belegu yelled at her/ because she got* late.//

CAF		Errors	
Words	176	Grammar	11
T-units	20	Lexis	5
C/T	29	Mechanics	3
DC	12	Total	19
C	32		
EFT	10		
EFC	19		

Note. Among independent and dependent clauses only the latter are underlined. Language errors are marked with asterisk (*). The asterisk (*) is put next to an erroneous word or in case of missing elements (e.g. articles, prepositions, verbs, etc.) it is placed next to, before, or after the word where such an element should have been placed in the text. The boundary of a T-unit is marked with //, while that of a clause with /. When there is more than one T-unit, one after the other, they are not always marked with //, but instead separated with /.

Figure 6: A collaborative text composed by Alu and Gash (hand-written)

Now I would like the two of you to work together and write the story of Ana. What happened to Ana this morning? You can start like this: **This morning Ana woke up at seven o'clock. The sun was shining through the window.** If you have any questions, please feel free to ask your teacher.

This morning Ana woke up at seven o'clock. The sun was shining through the window. Ana got ready to go on her job. ~~She~~ While she was going to the bus station a dog seized her ~~and things~~ and a man yelled to that dog and Ana was getting the things that fell from the bag. After she got her ~~bag~~ bag the bus was already on the bus station. Ana tried to stop the ~~bus~~ bus but he was gone already. Ana sat on the bus station to wait the next bus. While she was waiting a poor man didn't have money and Ana gave him some. That poor man thanked her so much and after that another bus got over. Ana still waited for another one. ~~At that time~~ Ana got on the bus and while she was on her way the police stopped the bus and told them that they can't pass this way ~~because~~ because a building was on fire. Ana was running to get on her job. After she got there Mr. Belegu yelled at her because she got late.

Appendix L: An individual text composed by Haz (#65 – intermediate)

This morning Ana woke up at seven o'clock.// The sun was shining through the window.// She got out of the* home* /and when she got out in the street/ shi* saw a dog/ and she was scared /and she dropped the bag,// she was taking the money/ that she dropped from the bag.// Ana was later* for the bus /and the bus goes*./// Later Ana saw one men*,/ he ask* her/ does* she have* a* money./ Ana give* him money*./ And when she was talking with* the man/ the bus goes* again.// She was too...The bus comes* /Ana goes* with bus.// She was going in* the* work,/ but she was too late.//

CAF		Errors	
Words	112	Grammar	12
T-units	18	Lexis	3
C/T	21	Mechanics	1
DC	3	Total	16
C	21		
EFT	6		
EFC	7		

Note. Among independent and dependent clauses only the latter are underlined. Language errors are marked with asterisk (*). The asterisk (*) is put next to an erroneous word or in case of missing elements (e.g. articles, prepositions, verbs, etc.) it is placed next to, before, or after the word where such an element should have been placed in the text. The boundary of a T-unit is marked with //, while that of a clause with /. When there is more than one T-unit, one after the other, they are not always marked with //, but instead separated with /.

Figure 7: An individual text composed by Haz (hand-written)

Now I would like you to write the story of Ana. What happened to Ana this morning? You can start like this: This morning Ana woke up at seven o'clock. The sun was shining through the window. If you have any questions, please feel free to ask your teacher.

This morning Ana woke up at seven o'clock. The sun was shining through the window. She got out of the house and when she got out in the street she saw a dog and she was scared and her bag she dropped the bag she was holding the money that she dropped the bag from the bag. The man takes the bus and the bus goes. Later Ana saw one man, he took her bag she gave a money. Ana give him money. And when she was talking with the man the bus goes again. It was too bad. The bus comes. Ana get on the bus with her. She was going in the work, but she was too late.

Appendix M: A collaborative text composed by Led and Sal (#34 – low-intermediate)

This morning ana woke up at seven o'clock.// The sun was shining through the window.// After that she was* starting/ to go at* work/ when one dog scare* her/ and she drop* her bag.// One man stop* the dog// and help* the laidy*.// After that Ana collect* all (of) the stuff// and go* to the bus station*.// She wasn't there/ when the bus came //so she have* to wait 10 minutes/ or maiby* more for (the) next bus // and she paid the guy/ for helping her.// She was talking to the guy/ when the bus came,// she wasn't looking /so she missed the bus again.// She wait* 10 or more minutes for another bus.// When the bus came/ Ana got in to the bus/ right after she saw a house (was on fire) with fire// and the* policeman stoped* the bus// she have* to take* the* drive away.// She walk* to her job* //and the boss told her/ that she was late/ and fired her for (having) being late.//

CAF		Errors	
Words	169	Grammar	14
T-units	21	Lexis	2
C/T	29	Mechanics	3
DC	9	Total	19
C	30		
EFT	10		
EFC	16		

Note. Among independent and dependent clauses only the latter are underlined. Language errors are marked with asterisk (*). The asterisk (*) is put next to an erroneous word or in case of missing elements (e.g. articles, prepositions, verbs, etc.) it is placed next to, before, or after the word where such an element should have been placed in the text. The boundary of a T-unit is marked with //, while that of a clause with /. When there is more than one T-unit, one after the other, they are not always marked with //, but instead separated with /.

Figure 8: A collaborative text composed by Led and Sal (hand-written)

Now I would like the two of you to work together and write the story of Ana. What happened to Ana this morning? You can start like this: **This morning Ana woke up at seven o'clock. The sun was shining through the window.** If you have any questions, please feel free to ask your teacher.

This morning Ana woke up at seven o'clock. The sun was shining through the window. After that she was starting to go to work when one leg scare her and she drop her bag. One men stop the boy and help the lady. After that ~~the~~ Ana collect all the stuff and go to the bus station. She wasn't there ~~when~~ the bus came so she have to wait for the next bus. She wait 10 minutes or maybe more for next bus ~~and~~ and she paid ~~the~~ the guy for helping her. She was talking to ~~the~~ the guy when the bus came, she wasn't looking so she missed the bus again. She wait 10 or more minutes for another bus. When the bus came Ana got into a the bus. ~~the~~ right after she saw a house with fire and the policeman stoped the bus she have to take the drive away. She walk to her job and the boss told her ~~the~~ that she was late and tired her for being late.

Appendix N: An individual text composed by Byt (#82 – low-intermediate)

One woman in* geting up*/ and he* is* going in* the city Bank,// Bat* 1 dog is* here*/
 and woman go* waith* one go* to bus.// In the moment 1 mean* sit* here*/ and woman
 talk* with* him// and in this moment *bus is* here*/ and woman missing* * bus,// but
 fire fight* and police stop* * bus/ because 1 hause* is* firing*/ and womat* going* to
city Bank with* speed*// bat* her chef* is* nervos*/ because he* is* late//

CAF		Errors	
Words	78	Grammar	20
T-units	10	Lexis	8
C/T	13	Mechanics	7
DC	3	Total	35
C	13		
EFT	0		
EFC	0		

Note. Among idependent and dependent clauses only the latter are underlined. Language errors are marked with asterisk (*). The asterisk (*) is put next to an erroneous word or in case of missing elements (e.g. articles, prepositions, verbs, etc.) it is placed next to, before, or after the word where such an element should have been placed in the text. The boundary of a T-unit is marked with //, while that of a clause with /. When there is more than one T-unit, one after the other, they are not always marked with //, but instead separated with /.

Now I would like you to write the story of Ana. What happened to Ana this morning? You can start like this: This morning Ana woke up at seven o'clock. The sun was shining through the window. If you have any questions, please feel free to ask your teacher.

One woman is getting in and he is going in the city bank, but 1 day is here and woman go with speed. Bus stop in the moment 1 man sit here and woman talk with him, and in this moment bus is here, and woman up in bus but fire light and police stop bus because 1 hour is firing and woman going to city bank with speed but her chef is ~~not~~ ~~not~~ married because he is late.

77