A Virtual Approach to Communication:

Augmented Reality and Language Related Episodes

in Second Language Learning

by

Silvana Domaz

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2020 by the
Graduate Supervisory Committee:

Barbara Lafford, Co-Chair
Bryan Smith, Co-Chair
Marta Tecedor

ARIZONA STATE UNIVERSITY

December 2020

ABSTRACT

In the last decade, the educational field, in general, has experienced increasing interest in applying augmented reality (AR) for educational purposes. Studies have shown that when AR is effectively applied in education, it can increase students' learning interest and concentration (Zhang et al., 2014), reduce cognitive overload (Bower et al., 2014, p.1), and provide a more authentic learning experience (Klopfer, 2008). This study uses both cognitive and sociocultural theoretical perspectives to better understand the role of AR in peer interaction by investigating language-related episodes (LREs) during collaborative dialogue. The current study investigates whether mobile-based AR influence the number, nature, outcome, and correction orientation of LREs during two oral and writing-focused activities of ten advanced L2 Spanish dyads using AR and non-AR mobile applications. The results show significant differences in the incidence of LREs in both settings (AR vs non-AR) and modality focus (oral vs writing-focused). Although significant differences were found between mechanical LREs vs. lexical and grammatical LREs, no significant differences were found between lexical and grammatical LREs in both modalities and settings. Likewise, the correction orientation was similar in both modalities, whereas the LRE outcomes were significantly different in both settings. Immediate posttests were administered to determine whether participants retained the results of the LREs based on the LRE outcome types. The posttests showed a strong correlation between the recognition and production scores of the grammatical structures. However, no significant differences were found in the recognition or production of grammatical structures nor the production of lexical items between the two settings.

ACKNOWLEDGMENTS

Many people helped me through this journey and without them I would have not been able to complete this degree. First, my sincere gratitude to my dissertation committee for their invaluable guidance and thoughtful feedback on this paper. I am extremely grateful to Dr. Barbara Lafford for her personal and professional guidance, her time, and detailed feedback on the ever-evolving drafts of this dissertation. I wish every graduate student could have a professor like Dr. Lafford on their committee. I wish also to express my gratitude to Dr. Bryan Smith for his technology wisdom and for teaching me that when we try a new approach to language learning, we do not have to prove that it is going to be better than what we have, but we need to determine if it is equally effective. Finally, I thank Dr. Marta Tecedor for her statistical advice and comments on the final draft of this paper.

Second, I would like to thank my classmates and colleagues in the Department of Spanish and Portuguese at Arizona State University, with whom I shared many successes and laughs. I am very grateful for the opportunities presented to me, the financial and academic support the department has provided me during these years. A special thanks goes to Anne Walton-Ramirez and the students who participated in this study. My deepest gratitude to all graduate and undergraduate students in my courses from whom I have learned so much.

Last but not least, I want to thank my family and friends for their love and encouragement. I want to thank my mom for showing me the meaning of determination and strength. A special thanks to my children for conceding to accompany me in this

journey. My profound appreciation goes to my husband for his patience and support throughout these years. Thank you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The field of second language acquisition (SLA) has traditionally encompassed two prominent perspectives – the cognitive and the sociocultural approaches. The SLA strand of cognitive research explores the processes involved in automatization, information processing, attention resources and manipulation, the functions of working memory capacity in relation to pedagogy, and conditions supporting implicit and explicit linguistic knowledge, among other phenomena. Therefore, one acquires a language through the cognitive process of developing "systems of knowledge (morphological, phonological, and lexical) which make up the target language" (Foster & Ohta, 2005, p. 234). Several researchers explore the cognitive interactionist dimensions of SLA with a focus on information processing within human interaction among learners, native speakers (NS), and teachers. From this perspective, the role of interlocutors is primarily to provide the learner with input that is processed to construct the learner's *interlanguage* (IL). IL is defined as the separate linguistic system displayed when adult learners attempt production of a target language norm (Selinker, 1972).

In contrast, the sociocultural perspective of SLA, inspired by Vygotsky's (1934, 1978, 1986) sociocultural theory of mind (SCT), affirms that what individuals eventually learn to do by themselves, they first learn it collaboratively. Thus, the role of the interlocutor from this perspective is more integral to the construction of new linguistic systems through a collaborative cognitive process. Therefore, it is through interaction with others that learning occurs and it is then integrated into the individual's mind.

Validating this idea two decades later, the Russian developmental psychologists, Stetsenko and Arievitch (1997), claimed that "psychological processes emerge first in collective behavior, in cooperation with other people, and only subsequently become internalized as the individual's own possessions" (p.161). Language is one powerful tool that mediates this cognitive process (Swain, 2001).

In the field of SLA, language learners are viewed to use language both as a tool to learn and as an object (the language itself). This process is carried out by interaction towards the completion of a goal-oriented activity. The concept of *activity* was the basic unit of analysis in Vygotsky's (1978) research. Learning happens when individuals "shift control within activities from the social to the individual, from the external to within self" (Ellis & Barkhuizen, 2005, p. 232). Drawing from Vygotsky's (1978) theory, language learning then happens during an interaction, and not simply as a product of interaction.

Several second language (L2) cognitive interactionist studies examined the practices in which learners used their IL resources either by thinking aloud or when asked by another interlocutor (e.g., teacher, peer, NS) to make their utterances more comprehensible. These strategies, in turn, led learners to restructure their IL toward better language proficiency measured by accuracy and complexity (Gass et al., 2011). From the cognitive perspective of SLA, learners make use of *communication strategies* (CSs) to attend to both form and meaning. CSs are defined as "tools used in a joint negotiation of meaning where both interlocutors are attempting to agree as to a communicative goal" (Tarone, 1980, p. 420).

In comparison, from the SCT perspective of SLA, this cognitive process of negotiating for meaning and producing comprehensible output is defined as *languaging* (Swain & Watanabe, 2013). One example of *languaging* that learners engage in for problem-solving and knowledge-building is *collaborative dialogue* (Swain, 2000). Within this framework, during *a collaborative dialogue* learners use CSs not only when there is a miscommunication that causes a break in the conversation, but also when one or both speakers identify a linguistic problem and try to find solutions for it. In doing so they may boost their knowledge or come to a new or deeper understanding of the issue. Swain (2000) defines *collaborative dialogue* as "dialogue in which speakers are engaged in problem solving and knowledge building" (p.102). These occurrences of linguistic inquiry are called *language-related episodes* (LREs; Swain & Lapkin, 1995, 1998).

Thus, from the SCT perspective of SLA, learners initiate and resolve LREs to build language awareness. LREs are defined as "[…] any part of the dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct" (Swain, 1998, p. 70). These LREs are often composed of CSs that allow learners to bridge communication gaps, confirm meanings, and build knowledge while keeping the conversation flowing. Peer interaction is one of the contexts that allows this socio-psychological process to happen. The episode below, between two Spanish learners, shows a lexical LRE (in bold) composed of six different CSs (for a list of CSs types and examples see Appendix A):

Pat: *Okay, chulo. ¿Hay otras más… opciones más, es todo que necesitamos hacer?* (CS = 28 SR [self-repair])
(Okay, cool. Are there more… options more, is everything that we need to do?

Jenn: *No, hay más, pero, no para cenar, es como para desayunar.*
[No, there are more, but not to dine, is like to have breakfast]

Pat: *Oh, sí, sí, sí.*
[Oh, yes, yes, yes.]

Jenn: *Y* ***¿meriedes?*** (CS = 31 AC [own accuracy check])
[And snacks?]

Pat: *Huh?* (CS = 2 CR [clarification request])
[huh?]

Jenn: snacks? (CS = 33 CW [code switching])
Pat: *¡oh¡* ***¿meriendas?*** (CS = 3 FC [confirmation check])
[Oh! Snacks?]

Jenn: ***Meriendas, sí****. Solamente hay meriendas y desayuno.*
(CS = 15 RC [response-confirm])
[Snacks, yes. There are only snacks and breakfast.]

(data from the current study)

In the example above, Patty and Jenn (pseudonyms) talk about food and meals. Jenn is not sure how to say "snacks" in Spanish and asks for Patty's help. The dyad resolved this communication gap by testing a hypothesis and co-constructing knowledge. This example illustrates the process of learning new vocabulary through peer interaction.

The research on peer interaction under the interactionist perspective (Blake & Zyzik, 2003; Pellettieri, 2000; Smith, 2003, 2004) and the sociocultural perspective (Yanguas & Bergin, 2018; Yilmaz & Granema, 2010) has gained increased popularity in the advent of computer-assisted language learning (CALL), more specifically synchronous computer-mediated communication (SCMC). However, in the field of mobile-assisted language learning (MALL) the concept of using mobile devices to

promote peer collaboration is still relatively new (Kukulska-Hulme & Shield, 2008). More recently, Burston (2014) in his meta-analysis of MALL studies revealed that:

> Ironically, it is precisely in the areas where they potentially have the most to offer – mobility, peer connectivity, oral interactions, and learner collaboration – that the advanced communication features of mobile phone technology have been, and continue to be, the least exploited in MALL. (p.350)

Technology and Peer Interaction

In a world where digital technologies, such as smartphones, are deeply connected to the way we interact with others and the environment, learning experiences that leverage the affordance of mobile applications as tools that students can use to interact and collaborate still remain an exception in the L2 classroom (Godwin-Jones, 2011, 2017). According to the Pew Research Center (2017), about 77% of the U.S population owns a smartphone, up from 35% in 2011, making the smartphone one of the most rapidly adopted consumer technologies in recent history. In addition, smartphones are most common among young adults, 92% of 18 to 29-year-old adults own a smartphone (Perrin, 2017). Today, consumers in the U.S. spend an average of 3.5 hours per day on mobile devices and 90% of this time is spent on web apps (Wurmser, 2018).

In the field of SLA, research shows that students prefer using mobile devices instead of computers (Thornton & Houser, 2005), and some teachers believe that a MALL environment, including social media, offers a more authentic learning experience for students which leads to an increase in students' motivation and engagement in learning (Krueger, 2014).

One mobile technology that can be used with smartphones is augmented reality (AR). AR is a technology that allows a user to see the real world with virtual elements overlapped upon it in real time (Cheng & Tsai, 2013). AR was introduced into sports broadcasting in the 1990s showing the first-down line in football, the path of a golf ball, or the ball placement on a tennis court. The popularity of smartphones over the last decades has contributed to the use of AR. Smartphones include several technologies that make this possible, such as, powerful computer processors, GPS receivers and sensors that can detect when the phone moves, allowing AR graphics to shift as needed. All of this creates the illusion of virtual objects existing in the real world. One example is the mobile game *Pokemón Go* which gained immense popularity (500 million downloads) with its release in the summer of 2016. The computer magazine *Wired* described several examples of AR becoming mainstream – AR apps will instruct people how to use household machines, play a game with virtual zombies in the living room, or scan a street to find the restaurant with the best deals and an available table. The Horizon Report (Johnson et al., 2016) has recognized AR as one of the most important advancements in educational technology. According to the report, AR is expected to reach widespread adoption in the next few years in higher education.

Studies thus far have shown that AR can increase students' learning interest and concentration (Zhang et al., 2015), reduce cognitive overload (Bower et al., 2014), provide a more authentic learning experience (Klopfer, 2008), increase motivation (Bergig et al., *2009*; Dunleavy et al., 2009) and enhance problem solving, critical thinking, and collaboration (Wasko, 2013). Mobile AR is thought to help produce

learning that is customized, socially constructed, and which expands beyond the classroom (Holden & Sykes, 2011). Furthermore, AR in the educational setting has the potential to deliver content in a three-dimensional perspective, to create real-time and collaborative learning opportunities, to link formal and informal learning (Wu et al., 2013). Therefore, it is important to examine how this immersive technology can be integrated into the teaching learning process, more specifically in L2 learning.

Hence, the overarching goal of this study is to examine the role that mobile AR applications can play in enhancing peer interaction and facilitating L2 development. This will be accomplished by examining the LREs produced by advanced learners of Spanish during two collaborative tasks using mobile AR applications vs traditional computer applications. Personalized posttests based on the LREs produced by each dyad will inform the effectiveness of this technology in aiding language learning.

Theoretical Background

The theoretical background of this study draws from both SLA's cognitive interactionist and sociocultural approaches to SLA. The Interactionist Approach (IA) (Long, 1981, 1983b, 1996) and the Output Hypothesis (Swain, 1985, 1997) of SLA originally had a cognitive interactionist perspective. However, Swain (2000) later broadened the definition of output as a communicative cognitive activity mediated by collaborative dialogue. Thus, from a sociocultural perspective, learner's "internal mental activity has its origins in external dialogic activity" (p. 113). In other words, language learning happens during a collaborative dialogue. It is important then to briefly explain both theories.

The IA has its roots in Krashen's (1981, 1982) Monitor Model. According to Krashen language learning has two very distinct processes: *acquisition* and *learning*. The first is an unconscious process that involves meaningful interactions in the target language in natural communication settings, the latter is a conscious process that involves receiving information about the target language (form and rules), transforming it into knowledge and storing it in memory. A good example of second language acquisition is when people live abroad for a year or two, often achieving near native fluency while knowing little about the target language forms or rules. They speak the language fluently without having any concept of phonology, nor any idea of what a present perfect tense is, but instinctively they learn to recognize and know how to use the language structures. These factors led Krashen and other researchers (e.g., Long, 1981, 1983b; VanPatten, 1991, 1996) to believe that input is the most important factor in second language acquisition, while other researchers (e.g., Pica, Holliday, Lewis, & Morgenthaler, 1989; Swain, 1985, 1995, 2000, 2005) believe that output is as important as input, and it is a central aspect of language learning.

According to the Input Hypothesis (Krashen, 1982), language acquisition happens when a learner receives input that is one step beyond his/her existing linguistic knowledge, what Krashen defines as *comprehensible input* (*i* +1). Krashen believes that language acquisition happens when we understand the meaning of a message/input. Taking the process of comprehensible input a step further, VanPatten argues that learners prefer to process lexical items to grammatical items, they prefer processing content words in the input before anything else, thus meaning before form (see Input Processing model,

Lee & VanPatten, 1995).

In contrast, Swain (1995) argues that just comprehensible input is not enough for second language learning, based on her experience with French immersion students who were routinely exposed to the L2 for years and still lacked native-like speaking and written skills. These observations led Swain to propose the Output Hypothesis, which affirms that "the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning" (Swain 2000, p. 471). Swain (2000) believes that by creating output, learners are forced into a more deeply mental effort than when receiving input - "Output may stimulate learners to move from the semantic, open-ended, strategic processing prevalent in comprehension to the complete grammatical processing needed for accurate production" (p. 99).

The importance of comprehensible input and output is further emphasized in Long's (1996) IA. Although comprehensible input is essential to both Krashen's view and Long's IA, they differ in how they think input is made comprehensible. The IA describes the interaction-L2 learning relationship as "*negotiation for meaning*, and specially negotiation work that triggers *interactional adjustments* by the native speaker or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways" (Long, 1996, p. 451). When learners experience difficulties in communicating, they try to overcome those problems by modifying or restructuring their utterances, what is referred to as negotiation for meaning (Long, 1981).

The miscommunication that happens between interlocutors can have two origins: misunderstanding and incomplete understanding (Gass & Varonis, 1991). The former refers to a "simple disparity between the speaker's and hearer's semantic analysis of a given utterance" (Milroy, 1984, p. 15 cited in Gass & Varonis, 1991). In a misunderstanding, interlocutors do not recognize a disparity or do not bother to investigate it, so communication goes on to a different topic or it is terminated all together. Incomplete understanding or nonunderstanding refers to when "one of more participants perceive that something has gone wrong" (Milroy, 1984, p. 15 cited in Gass & Varonis, 1991) and try to remediate the problem by negotiating meaning. Such negotiation involves CSs which may result in the successful transmission of the intended message or a complete lack of understanding.

According to the IA, an interaction may serve as a priming device where learning takes place, with negotiation being the initial step in learning (Gass & Selinker, 2001). Negotiation is part of *apperception,* that is, the process by which learners compare the input with their cognitive knowledge (Smith, 2003). Therefore, interaction offers learners with opportunities to negotiate form and meaning and to "notice the gap" (Schmidt & Frota, 1986) between their production and the target language. When learners perceive difficulties in message comprehensibility, they try to overcome these obstacles by asking, for example, for clarification requests, repetition, or confirmation checks to achieve comprehension. Selinker (1972), Dörnyei and Scott (1997), Lafford (2004), Lyster (2004), and others have examined focus on meaning under the umbrella of CSs, while other researchers (Doughty & Long, 2003; Lyster & Ranta, 1997; Norris & Ortega, 2000;

Smith, 2003, 2004; Swain & Lapkin, 2000) have demonstrated the importance of drawing learners' attention to form during interaction as it facilitates L2 acquisition.

From a SCT perspective, Swain (2000) argues that collaborative dialogue may bring students' attention to both form and meaning to build linguistic knowledge. Swain (2005) argues that "the dialogue that students produce collaboratively provides them both with opportunities to use language, and opportunities to reflect on their own language use. Together their jointly constructed performance outstrips their individual competences" (p. 111).

Unlike the concept of negotiation for meaning where there is a miscommunication between interlocutors that causes a break in the conversation, in a collaborative dialogue, one or both speakers identify a linguistic problem and try to find solutions for it, in doing so they may boost their knowledge or come to a new or deeper understanding of the issue. Furthermore, she argues that it is the act of verbalization initiated through social interaction (e.g., collaborative dialogue) that makes learners aware of linguistic problems (e.g., LREs), monitor their own language and look for solutions. Thus, it is learners' collaborative efforts, mediated by dialogue, that engage them in knowledge building.

According to Swain (1998), LREs "may serve the function of helping students to understand the relationship between meaning, forms, and function in a highly context-sensitive situation" (p. 69). Further explanation of the Output Hypothesis and its functions: noticing, hypothesis testing, and metalinguistic awareness, as well as the construct of LREs can be found in Chapter 2.

Statement of the Problem

Researchers have been investigating the concept of L2 interaction since at least the 1980s. Of the three types of L2 interactions, teacher-learner interaction, NS-learner interaction, and peer interaction (NNS-NNS), the latter has received the least attention (Philp et al., 2013). Nonetheless, peer interaction occupies a considerable amount of time in the L2 student-centered communicative classroom. Despite the research published on peer interaction in the last two decades, only a limited number of studies have empirically shown the effectiveness of peer interaction on L2 development (Sato & Ballinger, 2016). Thus, it is important to determine and test the factors that may affect L2 learning during peer interaction and explore ways to maximize pedagogical potential. One of the goals of this study is to investigate how tasks, mediated by an immersive technology such as AR, affect peer interaction and L2 learning.

An innovative way to increase opportunities for language interaction in and out of the classroom is using technology. Although successful language learning may happen with no technology at all, the use of technology tools offers advantages over the traditional classroom with regard to ease and range of material accessibility, NS, and in some cases, domain experts (Doughty & Long, 2003). However, in many instances, the use of technology is just a reproduction of traditional practices in digitally mediated settings. Educators should use these new learning environments to build knowledge rather than simply transmitting knowledge so students learn *in* the target language rather than just learning *about* the target language (Holden & Sykes, 2011).

In addition, much research (Harmer, 2007; Herrington & Oliver, 2000; Herrington

et al., 2002; Willis, 1996; and others) has shown that language learning, at its best, occurs when the target language is used in real life situations. However, when instructors are hindered from taking students to authentic language environments and situations (e.g., study abroad programs), the use of mixed-reality (XR) technologies such as AR, can *virtually* bring these environments to the students. AR offers location-based technology that can take language learning outside of the classroom or transform a classroom into a more realistic setting, thus giving students the opportunity of situated learning within an *augmented* environment, and encouraging more meaningful interactions.

Situated learning, which stresses the importance of human knowledge and interaction in *situ*, can be approached from a cognitive perspective (e.g., artificial intelligence), or from a SCT perspective (e.g., community of practice). Brown, Collins, and Duguid (1989) argued that "situations co-produce knowledge through activity" (p. 32). They were the first to develop a theory of *situated cognition* or *situated learnin*g and created a model of instruction with implications for classroom practice. Brown et al. (1989) added that many traditional classroom activities would not make sense by the cultures to which they are attributed. In addition, they limit students' access to the cues that arise from context, which offer structure and support to the process of learning. Thus, Collins (1988) defined *situated learning* as "the notion of learning knowledge and skills in contexts that reflect the way the knowledge will be useful in real life" (p. 2).

Both place-based technology and situated learning emphasize hands-on, real-world learning experiences, by leveraging the power of place and technology to personalize learning. Research shows that situated learning not only increases students'

engagement and motivation, but also increases academic achievement (Sobel, 2005). In the field of SLA, situated learning have had an important impact in various learning contexts, including in MALL (see Burton's meta-analysis, 2014).

The mobility and portability of innovative and immersive mobile technologies such as AR can leverage the power of place-based learning and may empower students' learning experience by heightening awareness and interaction with the real world. The GPS and other enhanced capabilities of AR applications such as add-on digital features to explore and expand scenes from the physical world, give students the opportunity to learn and interact in the target language in a localized, contextualized, and meaningful way (Godwin-Jones, 2013), thus combining learning concepts with real life experiences. Gee (2003) described this type of situated learning as *active learning*. He stated that "language is tied to people's experiences of situated action in the material and social world" (Gee, 2004, p. 52) and these experiences are kept in the mind in the form of lively images. Furthermore, Gee argued that when we learn a new domain in a more active way, instead of as passive content, 1) we experience the world in new ways, 2) we form new affiliations, and 3) we gain resources that prepare us for future learning and problem-solving in the domain and in related domains.

Although mobile devices appear to be ubiquitous in our society, the full potential for their value in education, and specifically language learning, is still in its infancy. There is little research which examines the use of mobile devices as an aid for interaction and language learning. Likewise, the research on the use of AR for education is in its early stages, even more in the field of SLA. To date, research on AR in SLA has been

shown to promote peer engagement and problem solving by providing place-based information and simulated dialogues (Holden & Sykes, 2011), motivation (Perry, 2015), and uses of images and videos (Zheng et al., 2017). Other studies have investigated the development of writing skills (Liu & Tsai, 2013), vocabulary (He et al., 2014), and listening and oral skills (Liu & Chu, 2010). Moreover, a limited number of studies compare place-based learning mediated by technology with traditional CALL classroom-based methods (Thorne, 2010).

Likewise, most sociocultural (García Mayo & Azkarai, 2016; Swain & Lapkin, 1998; Yanguas & Bergin, 2018) and cognitive interactionist (Gass & Mackey, 2007; Pica, 2006; Smith, 2004) studies have examined the effects of task type (e.g., jigsaw, dictogloss), modality (oral vs written), or context (e.g., CMC vs face-to-face) in language development. However, only one study (Sydorenko et al., 2019) to date, has examined the effect of AR in peer interaction. Sydorenko et al.'s (2019) study revealed that the small number of LREs produced by the participants in a place-based simulated AR task were only lexical in nature and focused on the particular study's topic of green technology.

Using both cognitive and sociocultural theoretical perspectives, the current study is important in that it will help us understand in general terms how interactions aided by AR are different from or like classroom interactions aided by traditional computer applications. This study will fill this gap by specifically examining the LREs produced by learners of Spanish engaged in place-based, in oral and written interactions in AR and non-AR settings. From a cognitive perspective this study will investigate if participants

focus on form and meaning by repairing and amending their language and constructing their IL. From a sociocultural perspective it will investigate if the social interaction mediated by collaborative dialogue makes learners attend to linguistic problems, monitor their own language, and rely on each other to find the right form or word.

Purpose of the Study

The current study will help us better understand how tasks mediated by AR technology affect language learning. It explores the unique features that AR provides and gathers evidence to determine whether AR can facilitate the types of interactions believed to facilitate L2 acquisition; that is, if learners provide and make use of rich input, negotiate for meaning and form, produce output, and build and co-construct knowledge in the process. AR provides a high level of realism and visual salience that may create a beneficial environment for input processing and more opportunities for "pushed" output in that it provides learners with a more authentic experience. This study will also advance the research agenda on interaction and L2 development through the analysis of LREs in this new, technologically mediated context. An empirical examination is essential to identifying the scope to which AR can effectively be used for language learning.

This study examines four relevant components of AR applications for IL development of L2 learners – (1) the effects that using AR (vs non-AR) applications have on learners' IL development by analyzing the frequency distribution of LREs, and the nature, outcome, and correction orientation of the LREs in these two types of settings, (2) the effects that oral and written modalities have on learners' IL development by analyzing the frequency distribution of LREs, and the nature, outcome, and correction

orientation of the LREs in these two modalities, (3) learners' recollection of the LRE outcomes through the use of posttests two days after their interactions. This study uses a microgenetic analysis, that is, it examines changes as they occur in collaborative dialogues between learners as their interactions unfold utterance-by-utterance towards self-regulation. Excerpts from the dialogues will be used to illustrate the results.

The first goal of the proposed study is to evaluate the effectiveness of AR (vs non-AR) applications as instructional tools to promote peer interaction. It measures the frequency distribution of LREs, and the nature, outcome, and correction orientation of LREs by advanced learners of Spanish in a place-based AR setting and a non-AR computer lab setting.

The second objective of this study is to examine the effectiveness of the modality (oral vs writing-focused) as a tool to foster collaborative dialogue, by measuring the frequency distribution of LREs, and the nature, outcome, and correction orientation of LREs by advanced learners of Spanish during oral and writing-focused interactions in both settings.

These two first goals look deeply into the elements of each problem-solving and knowledge building routine that occurs during peer interaction, classifying each language-related episode according to established models (Swain, 1995, 1998). To analyze this, the recordings from both oral and writing-focused activities were transcribed and coded according to modality and setting.

The third objective of this study is to use posttests to analyze quantitatively the learner data to understand if the lexical and grammatical structures from the positive or

negative resolutions of LREs were remembered and produced in their posttests two days after the interactions. This provides preliminary insight into the effectiveness of AR applications for L2 acquisition.

In sum, this study explores the process of peer interaction and how Spanish learners manage conversations, collaborate toward mutual understanding, manage output, and build knowledge when engaged in place-based AR and non-AR oral and writing-focused activities. It also seeks to determine whether the subsequent recall of target lexical items and grammar structures can be traced directly to the outcomes of the LREs in which they occur.

Overview of the Dissertation

This chapter has provided a preliminary overview of many of the relevant issues related to the current project, including the role of interaction in IL development from the sociocultural and cognitive perspectives of SLA, and the LRE construct being addressed, the benefits and limitations of interactional research, and the potential benefits offered by AR applications as instructional tools. Finally, it presented the objectives of the proposed study.

Chapter 2 will provide an overview of the literature relevant to the project and discuss in greater detail the issues presented in this introduction. First, Chapter 2 outlines the theoretical foundations of the cognitive interactionist and the sociocultural frameworks including the Interactionist Approach and the Output Hypothesis. Second, it discusses the relevant issues related to interaction and IL development, specifically considering three aspects of interaction, from both the cognitive and sociocultural

perspectives of SLA – peer interaction, communication strategies and language related episodes. Thus, the empirical research surrounding LREs is explored specifically in relation to peer interaction, as well as its presence in CALL and MALL research. The next section of Chapter 2 examines the role of AR in SLA in recent research. Finally, the chapter concludes with a statement of the specific research questions being addressed in the current work.

Chapter 3 describes the methodology used in this study. It presents a detailed description of the participant groups and continues with an explanation of the instruments and data collection process. It then discusses the data analysis methods to answer each research question. In addition, the creation of the posttests is discussed in detail as well as the method for scoring them.

Chapter 4 presents the findings for each of the research questions, followed by a detailed description of the results. It concludes with additional observations that emerged during the data analysis related to the research questions.

Finally, Chapter 5 presents the interpretation and discussion of the results for each research question. Finally, in Chapter 6 the conclusions, general implications the results may have for SLA theory and pedagogy, and the limitations of the current study are explained. The chapter ends with suggestions for future research into AR technology and peer interaction for L2 development.

Overall, this study establishes an empirical baseline to allow informed future exploration of AR for L2 development, specifically in peer interaction and IL development.

CHAPTER 2

REVIEW OF LITERATURE

The goal of this chapter is to further discuss the underlying theoretical framework and describe the empirical work that informs the current study of place-based AR task interaction among advanced learners of Spanish. It is important then to define the concept of interaction, more specifically, the notion of *peer interaction* used throughout this study. Peer interaction is defined as any communicative activity (either oral or written) carried out between learners for a common purpose, where there is minimum amount of teacher involvement or none at all (Philp et al., 2013). The structural properties of these conversations are the tools that are argued to contribute to L2 acquisition.

This chapter begins with a description of the Interactionist Approach and the Output Hypothesis, drawn from theoretical foundations upon which the cognitive interactionist and sociocultural perspectives of SLA are based, to explain one specific type of interaction - peer interaction, that is, interaction between learners. The cognitive perspective (Ellis, 1985; Gass, 1997; Long, 1981, 1996) argues that language learning is a psychological intrapersonal process that happens inside the mind when learners interact and try to process comprehensible input and express output. In contrast, the sociocultural perspective (Lantolf, 2000; Vygotsky, 1978) argues that learning is first an interpersonal process that becomes intrapersonal and occurs as learners work together and scaffold each other during a collaborative dialogue. Both perspectives agree that language is the tool that mediates the acquisition process.

This chapter continues with a detailed definition of peer interaction and a description of types of peer interaction. Then, the construct of collaborative dialogue (e.g., LREs) are described in this chapter. Next, the empirical research that investigates LREs relevant to peer interaction is examined, including in SCMC. Finally, this chapter explores AR technology and how it has been used in the educational field and SLA. The chapter concludes with the study's research questions.

The Interactionist Approach – A Cognitive Perspective

Long's (1981, 1982, 1983b, 1996) seminal work on the IA asserted that there is a strong connection between learners' selective attention in conversational interaction and L2 acquisition. Long (2014) stated that when interlocutors negotiate meaning they also may pay attention to linguistic features. Thus, learners are exposed to explicit learning that might improve their own implicit input processing. To solve a problem, learners shift their attention from meaning to form; when this process occurs, they might also notice the new information (White, 1987). The following exchange illustrates this process. Betty and Mark (pseudonyms) are talking about food (tofu). Mark repeats Betty's wrong verb aspect choice and right away notices the problem and reformulates his utterance, first using the incorrect gender and then the correct gender and verb aspect (4). Betty takes notice of it and replies with the correct form also (5).

1. Betty: *¿Te gusta tofu?*
[Do you like tofu?]

2. Mark: *Sí, cuando es bueno.*
[Yes, when it is good.]

3. Betty: *Cuando **es cocinando** bien.*
[When it is cooking well.]

4. Mark: *Hay mucho tofu que no **fue cocinando**… **¿cocinada? Cocinado** muy bien.*
[There is much tofu that was not cooking… cooked (f)? Cooked (m) very well.]

5. Betty: ***Cocinado**, sí.*
[Cooked (m), yes.]

6. Mark: *Pero cuando está hecho bueno es muy bien, pero la mayoría no es muy bien.*
[But when it is done good it is very well, but most are not very well.]

(data from this study)

The IA draws on Hatch's (1978a) key contribution that learners can learn an L2 through the process of interacting. For instance, it is not the learning of L2 structures that leads to communication in L2, rather the learning of the L2 is claimed to evolve out of communication. Hatch (1978b) affirms that "one learns how to do conversation, one learns how to interact verbally, and out of this interaction syntactic structures are developed" (p. 404).

Even though IA applied Hatch's key insight as a starting point, its first version was closely associated with Krashen's (1985) Input Hypothesis which, as mentioned before, stated that exposure to comprehensible input is both necessary and sufficient for L2 learning to happen. However, for Long, input only becomes comprehensible when learners think about and negotiate meaning through interaction with others. Although Long's idea of comprehensible input diverged from Krashen's in the way input is made comprehensible, the earlier versions of the IA were challenged in several ways. Theorists such as Smith (1986), Faerch and Kasper (1986) pointed out that learners might successfully comprehend input (e.g., by drawing on the context and not necessarily attending to linguistic forms) but that does not mean they have acquired the language.

Secondly, learners do not always successfully comprehend a modified input through negotiation for meaning (Hawkins, 1985), and they sometimes 'fake' comprehension. Thirdly, some aspects of language (e.g., inflectional morphology) are typically not subject to negotiation (Sato, 1986) or not noticed by learners (Mackey et al., 2000). Lastly, the claim that comprehensible input is enough for acquisition was challenged by Swain (1985, 1995) who affirmed that comprehensible output is also central to L2 acquisition. Further analysis of the Output Hypothesis will be explained below.

The later version of the IA (Long, 1991) addressed these concerns and suggested that, besides comprehensible input, negative feedback and modified output also contribute to acquisition. In addition, Long (1991) emphasized the importance of *focus on form* defined as: "Focus on form… overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (p. 45-46).

Over the past three decades, several studies have confirmed the link between interaction and L2 acquisition (Ellis, 2008; Gass, 1997, 2003; Gass & Mackey, 2007; Gass, Mackey & Pica, 1998; Gor & Long, 2009; Mackey, 2007, 2012, Mackey, Abbuhl & Gass 2012, Pica 1994; and others). Mackey and Goo's (2007) meta-analysis on interaction concluded that:

> "Interaction plays a strong facilitative role in the learning of lexical and grammatical target items. The 28 interaction studies qualified for the present meta-analysis showed large mean effect sizes across immediate and delayed

posttests, providing evidence of short-term as well as longer-term effects on language acquisition" (p. 405).

In conclusion, the cognitive interactionist perspective of SLA emphasizes negotiation for meaning as the central aspect of interaction. One approach interlocutors use to direct their attention to form and meaning and compensate for communication breakdowns is the use of communication strategies, explained later in this chapter.

The Output Hypothesis

As with the IA, the Output Hypothesis (Swain, 1985) also started from a cognitive perspective. Swain (1995) argued that during an interaction, on occasion, even without implicit or explicit feedback, learners notice a linguistic problem, which in turn triggers a mental process that allows them to modify the output. Thus, output not only helps with fluency, but it also promotes three fundamental functions: noticing, hypothesis testing, and metalinguistic awareness.

Noticing. Noticing a communication problem is a catalyst for change (Schmidt & Frota, 1986). In these moments, when learners are trying to produce meaning they learn something they did not know before or rectify an incorrect grammatical structure or lexical item. Noticing makes a learner focus their attention in his/her output specifically on the "hole" in his/her IL that triggers mental processes to mend the breakdown. When there is a breakdown in communication during an interaction, learners pay more attention to language forms used in theirs or other's utterances to correct the problem.

There are two aspects of noticing, the first one is that learners may become aware of the differences between the target language forms and the forms in their L2 (Schmidt

& Frota, 1986). The second aspect of noticing is that learners might identify frequently used forms (Gass, 1997). Swain and Lapkin's (1995, 2002) study with French immersion students also showed that learners were able to notice language gaps which triggered opportunities to engage in cognitive processes beneficial to IL development. Izumi (2002) adds that just attending to a grammatical form is not enough to encourage language learning. Rather, learners need to engage in deeper, integrative mental processes that are typical in communicative interaction. Leeser (2008) found the same results when comparing a control group who completed a comprehension post-test and an experimental group who had to produce output during a reproduction task. The results showed that the experimental group outperformed the control group.

Hypothesis testing. Learners explore and experiment with language to express themselves in the L2. Swain (1995, 1998, 2000) affirmed that what students are doing is creating hypotheses and testing them out. She argued that "output may sometimes be, from the learner's perspective, a 'trial run' reflecting their hypothesis of how to say (or write) their intent" (Swain, 2005, p. 476). Through interaction, the hypothesis is confirmed or discarded. During the process, IL development occurs with the incorporation of new forms and/or vocabulary. Testing a hypothesis also exposes a learner to a possible communication strategy or *corrective feedback* (CF), that is, feedback "that encourages self-repair involving accuracy and precision and not merely comprehensibility" (Lyster & Ranta, 1997, p. 43). Once a learner recognizes that his/her hypothesis is wrong, he/she then may amend the hypothesis and try again.

Metalinguistic awareness. This third function brings forward the concept of language as a tool and as an object, mentioned above. The process of transforming thoughts into words, referred to as verbalization permits learners to become aware of their IL limits, to predict linguistic needs and to set goals for future learning (Swain, 2005). When students question their own or other's use of the language, they are testing hypotheses about the language while they are "deepening their awareness of forms, rules and their relationship to the meaning they are trying to express" (p. 51). This cognitive process, mediated by language, to find a solution to a linguistic problem and producing comprehensible output is what Swain called languaging (Swain & Watanabe, 2013). One example of languaging that learners engage in for problem-solving and knowledge-building is collaborative dialogue, a sociocultural construct (languaging – ex.: collaborative dialogue).

In sum, from the perspective of the interactionist approach and output hypothesis, the current study will investigate how learners are exposed to explicit learning that might improve their own implicit input processing.

The Sociocultural Perspective of SLA

Swain (2000) later broadened the concept of output from viewing output solely within a cognitive perspective (above) to viewing output within a sociocultural theoretical framework. As mentioned in Chapter 1, the SCT perspective of SLA has Vygotsky's *zone of proximal development* (ZPD) and *scaffolding* (explained below) as its primary concepts. Within this framework, learner's linguistic development occurs during collaboration with others.

Thus, when learners try to produce language, they may recognize a problem in the output. For instance, a word or a structure they have not internalized it yet. They then form hypotheses and test them out, thus building their linguistic knowledge. Unlike negotiation for meaning, this knowledge building often happens not because of a breakdown in the communication but as a result of a linguistic problem in which learners try to find a solution for it. Collaborative dialogue mediates this cognitive process.

Collaborative dialogue – A sociocultural construct. In this approach, interaction is seen as a socially situated activity mediated by collaborative dialogue. Language facilitates this collaborative process, which, in turn, facilitates learning. There are three main mediational tools learners use during collaborative dialogue: scaffolding, the use of L1, and repetition (Swain & Watanabe, 2013).

***Scaffolding*.** The concept of scaffolding comes from Vygotsky's model of ZPD which described how children learn from an expert (adult or older child) as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). The term has been expanded in SLA to describe L2 learning through peer collaboration, regardless of proficiency levels (Donato, 1994; Lantolf & Thorne, 2006). During peer interaction learners take the role of experts and novices, thus providing scaffolding assistance to each other (Ohta, 2000). As other studies of L2 learners' interaction have shown, learners with the same proficiency levels can scaffold each other during collaborative tasks (DiCamilla & Anton, 1997; Donato, 1994; Ohta, 1995).

The use of L1. Researchers from the sociocultural perspective of SLA believe that when learners are learning an L2, they can understand the L2 process better if they are able to use their L1 when solving language-related problems and generating ideas (Vilamil & de Guerreiro, 1996), and by mediating intersubjectivity and externalizing inner speech during problem-solving activities (DiCamilla & Anton, 1997). By using the L1 and L2 as socio-cognitive tools in a collaborative dialogue, learners come to a mutual understanding of the task at hand and can provide aid to each other to complete the task (DiCamilla & Anton, 1997).

Repetition. The same happens when learners make use of repetition to negotiate solutions to language-related problems. Thus, repetition has also the function of creating and maintaining learners' mutual understanding of the task and to mediate the co-construction of linguistic knowledge (DiCamilla & Anton, 1997). In a collaborative dialogue, learners repeat their own utterances (self-repetition) or what others said (allo-repetition). These repeated utterances may be a replica of that has been said or they might have a variation. In an interaction, repetition serves the function of 1) production – it facilitates fluency while one decides what to say next, 2) comprehension – providing less dense discourse, 3) listenership – getting or keeping the floor, 4) interactional coherence – a speaker repeats another interlocutor's utterance to show acceptance (Tannen, 1989).

In conclusion, from the sociocultural approach, the current study will investigate how learners make use of these mediational tools as they seek out and help each other with linguistic problems. As Swain (2006) suggested, by examining the LREs, researchers can find out what learners attend to and how they resolve the language

problems that occur through peer interaction. For example, learners may collaborate to search for a new lexical item or to clarify the meaning of an unknown word or syntax, and in the process accept or reject their own or each other's hypotheses. However, the type of interlocutor (e.g., teacher, NS, learner) influences greatly the type and amount of negotiation for meaning and the LREs that occur during an L2 interaction.

L2 Interaction

Teacher-learner vs peer interaction. Studies that compared peer interaction with teacher-learner interaction have shown that learners when interacting with their peers negotiated meaning at their language developmental level (García Mayo & Pica, 2000), and used more CSs with each other than when interacting with the teacher (Alcón, 2002; Fernández Dobao, 2012; Pica et al., 1996; Toth, 2008). Also, in a class period where the teacher tries to interact with each student, each learner has very little time to speak if they speak at all; in contrast, peer interaction gives learners more opportunities and time to speak (Harmer, 2001). Furthermore, peer interaction does not necessarily contain "the pedagogical intent to assist language development as teacher-learner interaction does" (Sato & Ballinger, 2012, p. 159). Thus, students feel less anxious to perform the language correctly and that gives them greater autonomy to experiment with the language and try new things (Philp et al., 2013). For example, students may revise words or structures in response to one's feedback, or by saying it out loud. Together they find solutions to linguistic problems they encounter while communicating in the language (e.g., collaborative dialogue).

However, despite the many in-class hours learners spend in pair and group work the usefulness of collaborative time vs whole class interaction with the teacher is still questionable. One drawback of peer interaction is that many opportunities for language learning go unnoticed by learners (Williams, 1999). Another disadvantage of peer interaction is the possibility for non-target-like input and feedback. Studies have revealed that when trying to solve a communication problem, students may create "non-target-like" words or structures and miscorrect each other (Fernández Dobao, 2012, p. 248) or negotiate less (Williams, 1999) and have less L2 development (Toth, 2008) than in teacher-learner interactions.

NS-learner interaction vs peer interaction. Research has shown that peer interaction differs from NS-learner interaction in terms of the CF given. The CF literature has suggested that the types of feedback (Pica et al., 1996; Varonis & Gass, 1985), the quantity of feedback moves (e.g., Mackey, Oliver, & Leeman, 2003), and the production of modified output (Sato & Lyster, 2007) in peer interaction contexts are different from NS-learner contexts. Pica et al. (1996) found that NS-learner interaction addresses more modified input and feedback compared to peer interaction. However, while the production of modified output was similar between the two types of dyads, the amount and type of feedback were greater in peer interactions, where mostly repetition segments were used.

Similarly, Varonis and Gass (1985) compared three types of dyads: NS-learner, NS-NS, and learner-learner. The dyads were composed of English NS and ESL students who had not previously met. The dyads were recorded and instructed to converse freely

and informally in English only. The results showed greater amount of CF between L2 learners than NS-learner or NS-NS dyads due to the perception that "as learners they have little to lose by indicating a non-understanding, because they recognize their shared incompetence" (p. 84). Among the NS-NS dyads, the results showed little need for negotiation because NS-NS shared the same background. Furthermore, in a NS-learner interaction, the presence of a language proficiency inequality "discourages negotiation, because it amplifies rather than masks the differences between them" (Varonis & Gass, p. 86), so conversation resumes without negotiation.

In contrast, other studies using the same type of participants (e.g., English NS students and ESL learners) found that in adult NS-learner interaction, the amount of implicit feedback (e.g., clarification requests, recasts, and confirmation checks) was greater than in peer interaction (Mackey, Oliver, & Leeman, 2003). However, the feedback learners provided offered more opportunities for modified output than that from NSs, but there were no significant differences in terms of production of modified output between the two types of dyads.

In contrast, Sato & Lyster's (2007) study showed how differently Japanese ESL learners interacted when the interlocutor was another learner or an English NS student. The results showed that learners provided each other with a significant higher proportion of elicitation feedback (e.g., clarification and confirmation requests) than NSs, while NS in turn provided learners with significant more reformulation (e.g., recasts) feedback than did other learners. However, neither type of CF provided a significant amount of modified output, which suggested that the type of feedback does not affect the amount of

modified output in either types of dyads. When comparing the proportions of modified output in NS-learner and learner-learner dyads in both types of feedback (elicitation and reformulation) there was significant more modified output between learners than NSs. In conclusion, learners felt more comfortable and less pressure when interacting with another learner than a NS.

In the last few decades, the communicative approach has been the focus of L2 instruction. Thus, peer interaction takes up a significant amount of time in the L2 communicative classroom as students collaborate to complete a task. The studies below have suggested that the collaborative dialogue that learners participate in during peer interaction may promote language learning.

Peer interaction. More recent studies have suggested the idea that students feel more comfortable working on their linguistic problems and testing their linguistic hypothesis with another learner (Foster & Ohta, 2005; Leeser, 2004a; Sato & Lyster, 2007). Adams (2007) found that the feedback learners received during peer interaction was retained over time, an indication of L2 learning. However, other CF studies have shown that peers focus more on getting the utterance meaning across than pointing out grammatical errors (McDonough & Mackey, 2000; Mackey, Oliver & Leeman, 2003; Sato, 2007; Philp et al., 2010). Lyster et al., (2013) argue that this happens for three reasons: 1) learners are more interested in completing the task, 2) learners do not feel comfortable pointing out errors when they themselves feel they are not proficient (Foster, 1998), or 3) learners ignore peer's feedback "due to mistrust of each other's linguistic abilities" (p. 28).

In addition, other studies have shown that peer interaction when done in pairs (as opposed to groups of 3 or more) offered more opportunities for practice and individual language use (Fernández Dobao & Blum, 2013). However, L2 development through peer interaction might appear limited because students tend to focus on meaning more than form (García-Mayo & Azkarai, 2016).

To improve peer feedback some researchers are investigating how to train learners to provide feedback to peers (Sato & Ballinger, 2012; Sato & Lyster, 2012). The results showed an increase in peer feedback which led learners to autonomously shift their attention to form during negotiation and consequently improve accuracy and fluency (Sato & Lyster, 2012). These same studies also raised an important issue in peer interaction – peer engagement. For CSs and feedback to happen, learners need to feel comfortable with one another and be willing to collaborate. It is important to examine the patterns of peer interaction and how they affect L2 learning.

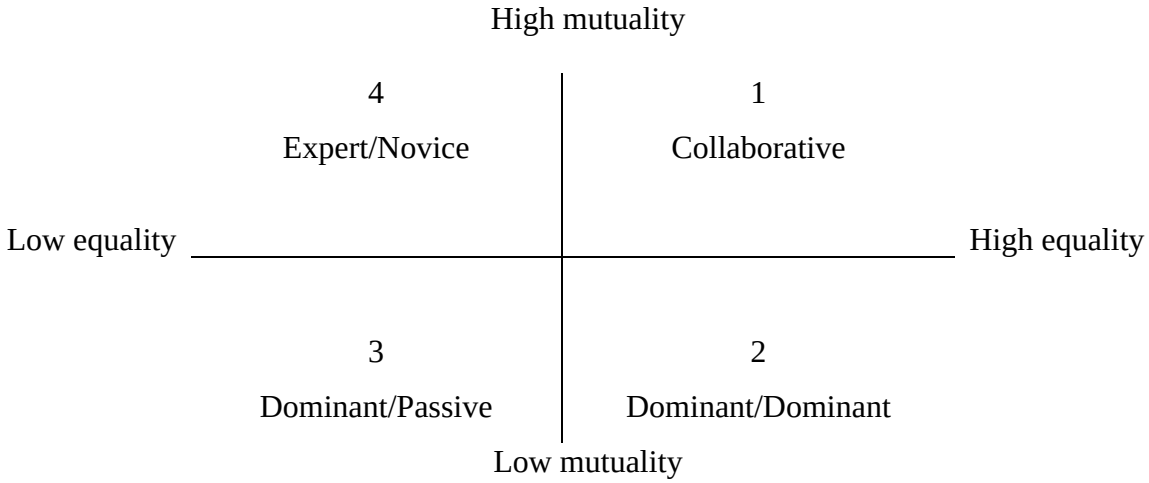
Patterns of peer interaction. According to Storch (2002) four distinct relationships roles occur during an interaction: collaborative, dominant/dominant, dominant/passive, and expert/novice. These roles are determined by two indexes: equality and mutuality. The former refers to "the degree of control or authority over the task, and the latter refers to the level of engagement with each other's contribution" (p.127). The four-role relationship patterns are represented in Figure 1. Each of the two axes represents a continuum (low to high) with the intersection point at a moderate level rather than zero. The following are the relevant descriptors (Storch 2002):

Collaborative. Quadrant 1 refers to a "collaborative" relationship represented by both high equality and moderate to high mutuality. The dyad works together during the entire task, they exchange ideas, they offer and discuss alternative views, and resolutions seem to be acceptable to both participants. Thus, high degrees of both equality and mutuality are present.

Dominant/Dominant. Quadrant 2 refers to a "dominant/dominant" relationship represented by moderate to high equality, but a moderate to low mutuality. In this pattern, both learners contribute to the task but are unwilling or unable to fully accept each other's contribution. The dominant/dominant discourse has a high level of disagreement and an inability to reach consensus. Alternatively, this pattern can also show a high level of equality and a low level of mutuality characterized by a division of labor, equal contribution to the task, but little engagement with each other's contributions.

Figure 1

Model of Dyadic Interaction



Source: Storch (2002)

Dominant/Passive. The third quadrant refers to a "dominant/passive" relationship represented by a moderate to low equality and mutuality where the dominant participant takes an authoritarian attitude and takes over the task while the other participant adopts a passive, subservient role, with few contributions. In this type of interaction, there is evidence of little negotiation because the passive learner does not contribute as much as the dominant learner.

Expert/Novice. The fourth quadrant refers to an "expert/novice" relationship represented by moderate to low equality but moderate to high mutuality. In this type of interaction, one participant takes control of the task, and actively encourages the other participant to participate.

Besides the four types of peer relationship patterns, peer interaction can be classified into four types: cooperative learning, collaborative learning, peer tutoring, and other forms of help from peers (Damon & Phelps, 1989). This project focuses on *collaborative learning*, that is, students in pairs or groups who depend on one another to complete a task. Unlike cooperative learning in which each student is assigned a piece of information and individually contributes information to complete a task, in collaborative learning students must engage in problem-solving and exchange of ideas to complete the task together.

Two additional perspectives on peer interaction worth mentioning but beyond the scope of this study relate to 1) language socialization and identity, and 2) age and developmental factors. There is a large body of literature within educational psychology and L2 learning which investigates the process of becoming a member of a community of

users of the L2 (Miller, 2003; Pavlenko, 2002; Swain & Deters, 2007). Similarly, research on age and developmental factors shows that the social and cognitive development of children, adolescents, and adults manifests very differently in these groups. Consequently, research on the nature of peer interactions among children is quite different from adult peer interactions (Duchesne et al., 2012; Muñoz, 2007).

Tasks and peer interaction. Bygate et al., (2001) defined a task as "an activity which requires learners to use language, with emphasis on meaning, to attain an objective" (p.11). In the 1980s several cognitive interactionist studies (e.g., Varonis & Gass, 1985; Doughty & Pica, 1986; Pica et al., 1989) investigated what type of classroom tasks produced the most negotiated interaction. Information gap tasks were found to give the most opportunities for negotiated meaning (Pica et al., 1993; Schmidt & Frota, 1986).

Other tasks found to promote negotiated interaction in collaborative work in the language classroom are jigsaw, problem-solving, opinion-exchange, and decision-making tasks (Pica et al., 1993). Other studies, from the sociocultural perspective of SLA have also shown that jigsaw and text reconstruction tasks (Kowal & Swain, 1997; Swain & Lapkin, 1998) when completed collaboratively, helped students engage in negotiation for meaning which led to focus on form.

The goal of the current study was not to focus on types of tasks that explicitly elicit negotiation for meaning and form (e.g., information gap, jigsaw) but instead find tasks that have *ecological validity,* i.e., those tasks that could be identified as "real world things people do in everyday life" (Long, 1985, p. 89). Therefore, the tasks for this study had to 1) be something students could do/use outside of the classroom, 2) "whose content

is of high intrinsic interest, that are fun to do, and that are just challenging enough to intrigue learners, but not so challenging as to discourage them" (Long, 2014, p. 249), and 3) have real-world language use (Long, 2014). Furthermore, the tasks should "engage productive or receptive, and oral or written skills, and also various cognitive processes" (Ellis, 2003, p. 16). Thus, the goal was to maximize learning opportunities by finding tasks not specifically designed to elicit LREs but that made learners autonomously attend to form and consequently produce LREs in the process.

Task setting. The impact of task setting in peer interaction has been extensively examined in the last few decades by researchers interested in the pedagogical application of their research, that is, if whether the results obtained in a laboratory setting can be replicated in a classroom setting. Most findings claim that there are no differences in negotiation of meaning in these two settings (e.g., Ellis, 2010; Gass et al., 2005; Philp et al., 2006), whereas other results show that classrooms lack the interaction control that a laboratory setting provides (Foster, 1998; Nunan, 1991). Thus, some researchers have claimed that negotiation of meaning happens more frequently in a laboratory setting where learners attend more to form than in the classroom where students just try to get the meaning across.

Foster's (1998) seminal study comparing negotiated interaction in a classroom vs laboratory setting revealed that the amount of meaning negotiation in the classroom was low with usually one student dominating the group or peer interaction. She concluded that students might perceive peer interaction in the classroom as an informal activity in which the priority is to maintain the flow of the conversation rather than its accuracy,

thus students “miss opportunities to gain comprehensible input and to create modified output” (p. 19).

Motivated by this claim, Gass et al. (2005) suggested that task setting must be analyzed carefully as an “important issue in interaction research” (p. 194). To investigate this matter further, Gass et al. (2005) divided four Spanish classes in the two settings to complete three different tasks - a jigsaw, information gap, and a decision-making task. They examined the most common interactional features of peer interaction: negotiation for meaning (e.g., confirmation checks, clarification requests, and comprehension checks), LREs, and recasts, that is a “correct statement of a learner’s incorrectly formed utterance” (Nicholas et al., 2001, p. 721). Results showed that while task type was an important factor in accounting for the incidence of CSs, it did not affect LREs or recasts. Most importantly, the setting had very little impact on the learners’ interactional patterns, that is, in both settings learners negotiated meaning and form and made use of LREs and recasts to solve linguistic problems.

The importance and applicability of task setting for L2 research is an important issue in interaction research. To this end, the current study takes a novel approach in investigating the differences between a new setting – “an outdoor campus” place-based setting enhanced by AR vs a laboratory setting, by examining the extent to which interaction is present in two different settings. This will be done by comparing a traditional controlled laboratory setting, in which learners work in dyads to carry out tasks using traditional computer applications, with an outside setting, in which learners from the same population, carry out the same tasks but outside using AR applications.

Communication Strategies

Communication strategies, as Canale and Swain (1980) first described, are "verbal and nonverbal strategies that may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence" (p. 30). Despite being part of the learner's IL system, these strategies are triggered by the learners' inability to understand their interlocutors or to convey their ideas in the L2 (Lafford, 2004). The significance of strategic competence in communication, that is "the ability to communicate within restrictions" (Savignon, 1983, p. 43) by using strategies to overcome misunderstandings, has been widely recognized since it became a major component in the model of communicative competence (Canale & Swain, 1980). However, CSs have also been described as ways to "enhance the effectiveness of communication" (Canale, 1983, p.11). CSs are used to help construct LREs, as in the following example:

1. Mark: *y hay menos lluvia ahí, ¿sí?*
   [and there is less rain there, yes?]

2. Betty: *Sí* ***más menos de****...*
   [Yes, more less than…]

3. Mark: ***mucho menos*** *...* (CS=1 EC [explicit correction])
   [much less…]

4. Betty: ***mucho menos****, más menos (laughs)…* ***mucho menos*** *del oeste donde yo vivo.* (CS=15 RC [response confirm])
   [much less, more less… (laughs)… much less than of the west where I live.]

(data from this study)

In this example, Mark and Betty (pseudonyms) assist each other in solving a linguistic problem and co-constructing knowledge in the absence of a breakdown in communication. Betty uses the wrong adverb in Spanish (2) and Mark corrects her using the right form (3). Betty realizes and laughs at her own mistake and agrees with the correct form offered by Mark (4).

In sum, CSs not only serve to prevent or solve communication problems when they occur but also to co-construct knowledge and build one's own linguistic understanding by forming and testing hypotheses about the correct use of the target language (for more CSs examples see Appendix A). Tarone and Liu (1995) suggested that it is in situations "where the learner needs to produce output which the current IL system cannot handle, that pushes the limits of that interlanguage system to make it handle that output" (p.120). Pica et al.'s (1989) study showed that learners changed their output either semantically or morphosyntactically when faced with a clarification request or a confirmation check during negotiating for meaning.

Although it would be interesting to look at the relationship between the LREs and the CSs that comprise them, to limit the scope of this study, this dissertation will focus only on LREs.

Language-Related Episodes (LREs)

Following Swain and Lapkin's (1998) definition of LREs as "[…] any part of the dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct" (p. 70), LREs have different categories and outcomes.

Nature of LREs. The original definition by Swain & Lapkin (1998) categorized LREs as *lexis-based, form-based,* and *orthographic*. Lexis-based or lexical LREs are the ones in which learners may question the meaning of a word (Example 1) or its pronunciation. In a lexical LRE, learners may search for an L2 word or choose among competing lexical items. Form-based or grammatical LREs (Example 2) include examples in which students may question the syntax or morphology of a target word or construction, thus the focus is on grammar. Orthographic or mechanical LREs (Storch, 2008) (Example 3) emphasize the spelling or punctuation of a written target word. For this study, the emphasis on the pronunciation of a spoken word was also classified as a mechanical LRE. All the following examples were taken from the data set from the current study.

Example 1. Lexical LRE (e.g., questioning word meaning)

Mark: ***A la parrilla****… ¿no sé qué es eso?*
[Grilled… I don't know, what is that?]

Betty: *Parrilla es grilled.*
[Parrilla is grilled]

Mark: *Entonces quiero un queso a la parrilla.*
[Then I want a grilled cheese]

Example 2: Grammatical LRE (e.g., questioning the adjective gender form for a noun)

Mark: *Ahí hay muchos bebidos de fruta ricas y otras* ***sabores buenos… ¿buenas?***
[There, there are many drinks of rich fruits and other good (m).... good (f)? flavors]

Betty: *Hay muchos bebidos y otras* ***sabores buenas****.*
[There are many drinks and other good (f) flavors].

Example 3: Mechanical LRE (questioning the accent on a written word)

Pat: *¿**Sofa** tiene acento? No sé…*
[does sofa have an accent? I don't know…]

Jenn: *¿**Sofa**? No sé.*
[sofa? I don't know.]

LREs can be further classified into subtypes. For this study, lexical LREs were classified by parts of speech into subgroups, i.e., *nouns, verbs, prepositions, pronouns, adverbs, adjectives, conjunctions,* and *determiners*. Examples 4, 5 and 6 below show examples of lexical LREs of the type determiners, conjunctions, and verbs, respectively. Similarly, grammatical LREs were subclassified by verb morphology (subject-verb agreement, tense, aspect, and mode), and noun-gender and noun-number. Example 2 above is a grammatical LRE of the type noun-gender. Examples 7 and 8 below are subject-verb agreement and noun-number grammatical LREs, respectively. In all the examples below, participants used self-repair as a CS within their respective LREs.

Example 4: Lexical - Determiner LRE

Mark: *Es lo mismo situación… hay **una**… **dos** amigos, y uno le gusta el carne y el otro no come el carne, ¿sí?*
[It is the same situation… there is one… two friends, and one likes meat and the other doesn't eat meat, yes?)

Example 5: Lexical - Conjunction LRE

Lisa: *no soy vegana, **pero**… **porque** me gusta carne.*
[I'm not vegan, but… because I like meat.]

Example 6: Lexical - Verb LRE

Mark: *Yo he **tratado**, yo he **probado** macha.*
[I have tried (attempt to do something), I have tried (to test/try food or clothes) macha.]

Example 7: Grammatical – Subject-verb agreement **LRE**

Betty: *¿te **gusta** los… te **gustan** los restaurantes de ASU?*
[Do you like the (singular subject) … do you like (plural subject*) the ASU restaurants?]
* the verb *gustar* is always conjugated to match the subject noun (the thing being liked).

Example 8: Grammatical – noun-number LRE

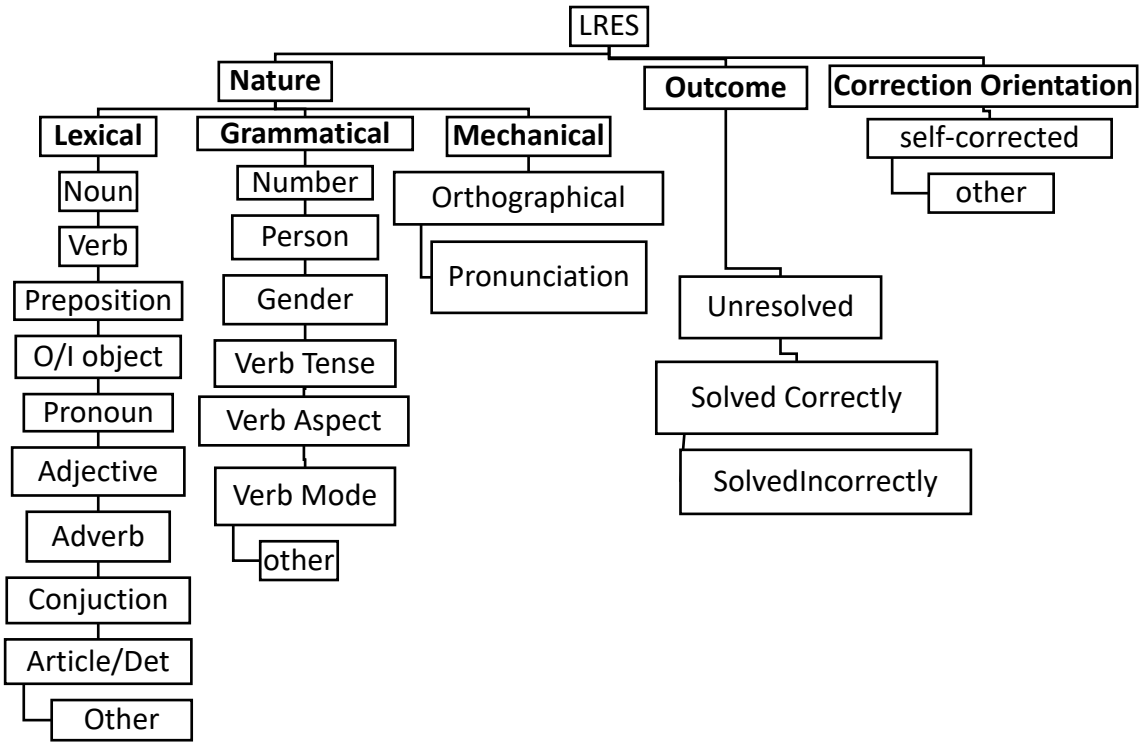
Mark: *Solo **los**… **el** arroz.*
[Just the (plural)… the (singular) rice.]

LREs outcomes. When faced with an LRE (e.g., focusing on a word, form, or structure) the learners may decide to solve it themselves or ask for help from a peer. Previous research (García-Mayo & Azkarai, 2016; Leeser, 2004a; Swain, 1998; Yilmaz, 2011), showed three types of outcomes: LREs that are unresolved, solved correctly, and solved incorrectly. Taking the examples above, example 1 is a correctly resolved LRE, where Mark accepts the right answer from Betty, example 2 is an incorrectly solved LRE where Betty confirms the wrong gender form of the adjective "good", and example 3 is an unresolved LRE where neither interlocutor knew the answer to the orthography problem.

LREs Correction Orientation. LREs were further distinguished between two types of correction orientation, namely LREs that were self-corrected and LREs corrected by the partner. Examples 1 through 3 above show correction orientation of the type *other correction*, whereas examples 4 through 8 are all *self-correction* LREs. The classification of this study's LREs are found in Appendix B. Figure 2 illustrates the categorization of LREs for this study.

Figure 2

Categorization of LREs in the Current Study



Overall, prior studies have shown that lexical LREs tend to be far more prevalent than form based LREs (Choi, 2011, Kim & McDonough, 2008; Walter & Basturkmen, 2010; Williams, 2001). However, research on collaborative, focus on form tasks under the LRE model has shown that several factors can affect learners' attention to their interlanguage, and, thus, their production of LREs.

In sum, this literature review of the cognitive interactionist and sociocultural perspectives of SLA emphasizes interaction as the main element for negotiation and collaborative dialogue. Learners make use of CSs and LREs to solve communication breakdowns and enhance their communication by solving linguistic problems and producing comprehensible output. During this process, learners are exposed to explicit

learning that might improve their own implicit input processing. The LREs learners produce during a collaborative dialogue are the evidence of the cognitive process that happens during peer interaction.

Empirical Research on Peer Interaction and LREs

Swain and Lapkin's (1998) study was the first to investigate focus on form under the LRE model. Their study showed evidence of cognitive processes instantiated during a collaborative task (e.g., jigsaw) between 12 dyads of 8th grade French immersion learners. The customized posttest showed that learners in general retained the language points discussed during LREs produced during the writing task. Since then, many peer interaction studies have focused on several factors that can affect the number, type, and outcomes of LREs, such as task type (Adams, 2006; Ismail & Samad, 2010; Swain & Lapkin, 2001; Yilmaz & Granena, 2010), task mode (e.g., video and audio in CMC) (Hsu, 2019; McDonough and Sunitham, 2009; Shekary & Tahirian, 2006; Yanguas & Berguin, 2018; Yilmaz, 2011; Zeng and Takatsuka, 2009), and different language contexts, foreign and L2 (Amirkhiz et al.,2012; García-Mayo & Azkarai, 2016).

Other studies have examined different factors such as learners' proficiency level (Choi & Iwashita, 2016; Kim & McDonough, 2008; Lapkin et al., 2002; Leeser, 2004b; Watanabe & Swain, 2008; Williams, 2001); pair dynamics (Storch, 2008; Wigglesworth and Storch, 2012; Young & Tedick, 2016), and task modality (written and oral) (Adams, 2006; García-Mayo & Azkarai, 2016; Rouhshad & Storch 2016). Studies have shown that tasks that include a written part, offer learners more opportunities for LREs than just oral tasks (Adams & Ross-Feldman, 2008; Williams, 2008). This is due to the use of different

structures in writing vs speaking (Williams, 2008) and levels of accuracy, as written language has more permanence than spoken language, which is ephemeral in nature. For instance, a conversation has a context and usually includes feedback from the listener, while a written text does not. Instead, the written text requires a higher level of explicitness, a level of linguistic proficiency, and metacognitive knowledge from the writer (Schoonen et al., 2009). Furthermore, spoken language can only be traced back if it is recorded. Therefore, there is more time for the dyad to go over a written text and try to perfect it, as opposed to each member of the dyad holding the interlocutor's message in short term memory and then giving feedback to correct it.

Furthermore, several studies have shown evidence of language acquisition through the development of tailor-made dyad specific posttests to trace both grammar and vocabulary learning in LREs. Adams (2007); Kim (2008), Lapkin et al., (2002), LaPierre, 1994; McDonough & Sunitham (2009), Storch (2002, 2008), Swain and Lapkin (1998, 2002), Watanabe and Swain (2008), Williams (2001), and Zeng and Takatsuka (2009) have found evidence of learner's independent use of the vocabulary and grammatical knowledge that was previously constructed in LREs.

In sum, these studies have shown that there are several factors that can affect the number, nature, and outcome of LREs, while other studies have shown evidence of L2 acquisition through the production of LREs. The current study focuses on two factors - task modality and task setting. Task setting has been examined above (see Peer Interaction section). Studies on task modality will be investigated below along with two other important factors in LRE investigation: Pair dynamics and Task mode.

Task modality: oral vs writing. Adams (2006) and Adams and Ross-Feldman (2008) investigated the effect of task modality on the production of LREs of ESL learners during both task modalities. The tasks focused on two target forms: locative prepositions and past tense. The results from the first study showed that there were more LREs, self-repair and use of target structures in the writing part of the task than the speaking part. The results from the second study showed that most LREs were of the grammatical type in both task modalities and that learners also produced more LREs in the writing task. However, the differences were not statistically significant.

Azkarai and García Mayo (2012) investigated the number and outcome of LREs generated by six dyads of EFL Basque-Spanish learners when they worked collaboratively on four communicative tasks. The authors found that the dyads generated more LREs in the tasks that had a writing component, than in the other two tasks with just the oral component. In a subsequent study, García-Mayo & Azkarai (2016) investigated the impact of task modality and level of engagement on the LREs of 44 Spanish-Basque EFL learners during 2 writing tasks and 2 oral tasks. The results confirmed previous research in ESL settings on the impact of task modality, as participants initiated more LREs in the writing tasks, mainly focusing on form, than in the oral tasks, where LREs mainly focused on meaning. They concluded that tasks that have an element of writing elicit more attention to language than oral tasks alone (Niu, 2009). In addition, task modality did not impact the participant's level of engagement which was similar in the oral and written tasks.

Pair dynamics. The willingness to collaborate, or the learner's level of engagement, is an important factor in determining the quality and quantity of LREs and language development in a collaborative dialogue. Storch (2008) described 'engagement' as "the quality of learners' metatalk" (p. 98), that is, the attention learners pay to language when faced with a linguistic problem. She described two types of engagement – *elaborated engagement* (E LREs) in which both learners engage in the LRE discussion providing explanations, and *limited engagement* where one participant mentions a linguistic problem and the other learner just repeats or acknowledges it (L+L LRE) or ignores the suggestion (L LRE).

Storch (2008) examined the nature of the engagement of 11 intermediate ESL self-selected dyads who worked on a text reconstruction task where learners were given a text containing lexicon and some spaces to be filled for the grammatical structures (e.g., articles, prepositions, verb tense/aspect). She found that learners focused more on grammatical items than lexical items and most LREs were correctly resolved regardless of the level of engagement. Furthermore, the higher the learner's engagement, the more opportunities they had to develop their L2. Specifically, learners who showed an elaborate engagement in their grammar choices learned the target structures. Sato and Ballinger (2012) stressed this factor affirming that "social relationships between learners are a significant variable that facilitates or prevents subsequent L2 development" (p. 173). They argued that "the key element to make the bidirectional language awareness in peer interaction may lie in collaborative patterns" (p. 171).

Task mode. Similarly, Rouhshad and Storch (2016) found that the mode of interaction also affects the pattern of peer interaction. In their study, ESL dyads engaged in 2 written tasks in face-to-face (FTF) and SCMC mode. In FTF mode, learners collaborated and co-constructed the text while solving linguistic problems and providing each other assistance and corrective feedback. In the SCMC mode, participants sat at two different computers away from each other. Each participant had access to the same Google Doc and chat window but only one participant wrote the text. While dyads collaborated to create the text in FTF mode, in the SCMC mode, participants took a cooperative approach, that is, one participant wrote and composed the text while the other edit, made suggestions, and contributed minimally to the content of the text. In general, there was little interaction and substantially fewer LREs between participants in the SCMC mode compared to FTF, even though dyads spent more time on the writing task in the SCMC mode than the FTF mode. Although the mode affected the number of LREs, it did not affect the type (grammar, lexical, and mechanics) or outcome of LREs.

Another SCMC study by Yilmaz' (2011) found a significant difference in the number of lexical LREs in two tasks (jigsaw vs dictogloss) performed by 54 EFL university students. The results revealed that the LREs were more frequent in the written (dictogloss) task, confirming the results of a previous study (Yilmaz & Granema, 2010). In addition, the written task had more resolved LREs than the oral (jigsaw) task. The author concluded that the oral task required more focus on meaning with learners paying more attention to the message content than to message form, which resulted in more unresolved LREs. On the other hand, the written task required learners to reconstruct a

verbal input into a text, which led to more LREs because learners were focused on recreating the correct version of the input, thus increasing attention to form (Yilmaz, 2011). Yilmaz and Granema (2010) also made a comparison between implicit and explicit LREs. In their study the LREs produced during the oral task were implicit, that is, they were reformulations of incorrect utterances of one learner by his/her partner, also known as *recast* CF.

Contrary to Yilmaz and Granema (2010) and Yilmaz' (2011) studies which found significant differences in written SCMC between the same task types, Yanguas & Bergin's (2018) found that learners produced a similar number of LREs on both task types (jigsaw vs dictogloss) and mode (audio SCMC vs video SCMC). They believed that modality- oral (jigsaw) vs written (dictogloss) may influence the way students interact and the types of LREs they produce. Furthermore, the number of lexical LREs were higher in the oral task than the written task, which favored a higher number of grammatical LREs. Regarding LREs outcome produced by learners, there were no significant differences between tasks. However, audio SCMC had a higher number of unresolved LREs than the video SCMC.

Finally, Loewen and Wolff's (2016) study compared 24 dyads in three types of communication modes: face-to-face, oral SCMC, and written SCMC. Negotiation for meaning (confirmation checks, clarification requests, comprehension checks, and recasts only), and LREs were identified amongst 48 ESL learners engaged in three different tasks (jigsaw, information gap, conversation) in one of the three communicative modes (8 dyads per mode). Results showed that written SCMC had fewer confirmation checks and

LREs than FTF and oral, which were similar. There were no statistical differences for clarification requests, comprehension checks or recasts which were infrequent in the data. Comparing this study with Gass et al., (2005), the number of LREs in both studies were very few per task (one to three). This study also confirms what other studies have suggested that in terms of negotiation, oral interaction in F2F and oral SCMC are similar and have more interactional features than written SCMC (García Mayo & Azkarai, 2016, Jepson, 2005). In addition, confirmation checks and LREs hardly occurred in the written SCMC in all three task types.

In sum, the empirical studies presented above suggest that collaborative tasks that require learners also to write besides speaking, provide learners with the opportunity to generate more LREs than speaking tasks alone, and that the writing tasks focus more on form while speaking tasks focus more on meaning. In addition, some studies have suggested that learners' level of engagement also plays a role in the quality and quantity of LREs produced during a collaborative dialogue. Finally, task mode can influence how peers interact.

The current study will be the first to investigate how task modality and setting can affect learners LREs' number, type, and outcome during oral and written collaborative dialogue in AR vs non-AR settings.

Peer Interaction Research from both Cognitive and Sociocultural Frameworks

Foster and Ohta (2005) was the first study to look at peer interaction from the sociocultural and cognitive perspectives of SLA. The authors argued that not all negotiated interactions are signaled by communication problems but also from learners

actively aiding each other to execute the task through co-construction and encouragement. The study comprised two separate data samples. The first sample had 20 ESL learners from different L1 backgrounds studying English in London. The second data sample was 19 learners of Japanese at an American university. Participants interviewed each other using a list of prompt questions about their impressions of studying in England (sample 1) or their plans to study abroad in Japan (sample 2).

All semantic, phonological, lexical and morphosyntactic modifications made by a participant to his or her utterance were identified and divided into 1) modifications due to a communication problem or 2) modifications such as self-repair, and elaborations due to an interlocutor's expression of encouragement, interest, or surprise. In each sample set, the quantitative analysis showed a higher number of modified output produced in the absence of negotiation for meaning than in its presence. The qualitative analysis of the data in both sets found plenty of evidence of learners giving and receiving assistance using co-constructions, self-corrections, and correcting-other in the absence of negotiation for meaning. Furthermore, self-correction and modification of learners' own utterances were more common than corrections initiated by others. In addition, learners incorporated the help received from peers into their own utterances.

More recently Sato & Ballinger (2012) also investigated peer interaction from both SLA theoretical frameworks. Their study described two independent studies, one from the cognitive approach and the other from the sociocultural approach to investigate the effects of instruction designed to maximize learning opportunities during peer interaction by raising L2 learner's awareness of peer interaction and training learners to

increase their language awareness during peer interaction. In the first study, Japanese university students were taught how to use CF during communicative peer interaction activities.

The results showed that learners who were trained to provide CF significantly increased the CF frequency over one semester, thus increasing language awareness. In the second study, Canadian French immersion students from third and fourth grades were given CF instructions to collaborate on-task and language-related problems. The results revealed that factors such as interpersonal relationships or collaboration patterns (operationalized as on-task, partner-focused turns) play a role in language awareness. That is, learners who engaged in unbalanced collaborations provided unbalanced amounts of CF. Thus, Sato and Ballinger's (2012) studies complemented each other to comprehend language awareness in peer interaction. Study 1 investigated the link between language awareness in peer interaction and L2 development showing that training can raise learners' language awareness. Study 2 examined the factors that make language awareness conducive to L2 development showing that collaborative patterns "may mediate the occurrence and effectiveness of CF" (p. 173). The authors concluded that a "collaborative mindset or cooperative atmosphere between learners is a prerequisite for CF to be conducive to L2 development" (p. 172).

In sum, the studies above confirm that peer interaction is an important social process that triggers how the brain process and retrieves information during negotiated meaning interactions as well as other kinds of peer assistance that foster interest and encouragement between students. Furthermore, these studies have shown that learner

feedback depends on a cooperative atmosphere between learners, who might perceive CF as pragmatically inappropriate or face-threatening (Yoshida, 2008). Although Foster and Ohta (2005) and Sato and Ballinger (2012) investigated peer interaction from both SLA perspectives, they did not investigate the number, type and LRE outcomes produced by the learners' negotiated interactions.

The current study will fill this gap, by also investigating peer interaction from both SLA perspectives and examining the LREs produced during collaborative dialogue. It will examine how learners assist each other through co-construction and self-repair of their utterances and use explicit learning to improve their own implicit input processing. No study to date has compared task setting (AR vs non-AR) and its effect on L2 development in peer interaction. Therefore, this study aims to investigate LREs in a mobile-based AR environment vs a non-AR environment (e.g., computer lab) in two different modalities – written and oral. The next section describes AR technology in more detail and reviews prior empirical research carried out on the use of AR in education and SLA contexts.

Augmented Reality (AR)

Azuma (1997) described AR as a variation of virtual reality (VR). VR is a three-dimensional, computer-generated environment where a real person can be immersed into and interact with the imaginary world by manipulating objects and performing actions (Virtual Reality Society). Some examples of recent VR systems are *Oculus Rift*, and *HTC Vive*.

While VR entirely immerses a user in an artificial environment that shuts out the physical world, AR allows a user to see the real world with virtual elements overlapped upon it in real-time (Cheng & Tsai, 2013). Azuma (1997) identified three characteristics of AR: 1) it combines real and virtual, 2) it is interactive in real-time, and 3) it is registered in 3D. In AR, users interact with real and virtual objects which coexist in the same space in real-time. Thus, the user sees the real world *augmented* with virtual objects.

AR elements are not visible to the naked eye; thus, AR relies upon some sort of display such as computers or webcams. This technology has recently become popular through mobile applications which made it simpler to use it and are portable. Through a phone or tablet webcam, users point to an image or object which triggers the augmented media (e.g., a video, audio, a text, or a 3D animation). AR can be image-based or location-based. An example of an image-based AR is a travel poster with images about a vacation place. The images are markers which when detected by the mobile camera triggers a virtual element (e.g., video) generated by the AR software. In contrast, a location-based AR uses position data produced from mobile devices such as a wireless network or global positioning system (GPS), to detect a location, and then superimposes computer-generated information on the users' mobile screens in real-time (Cheng & Tsai, 2013). The current study uses both types of AR applications, an image-based application (IKEA AR) and a location-based application (Bettar AR). More details about these applications on Chapter 3.

AR and cognition theories. Santos et al. (2014) argue that AR can help students learn better by improving 1) perception through real- world annotation, 2) elaboration through *contextual visualization*, and 3) elaboration through *vision-haptic visualization*. Elaboration is defined by the additional associations that help people make sense of and remember information. The juxtaposition of real and virtual symbols and text reduces the cognitive load in the short-term working memory so more of it is used for operating cognitive processes that are stored in long-term memory. AR helps students construct a more elaborate network of knowledge by delivering meaningful signals found in the real environment.

The concept of *contextual visualization* "refers to the presentation of virtual information in the rich context of a real environment" (p. 50). In this sense, AR can be used as a strategy to connect virtual information to an object or setting that a learner is familiar with to provide more effective learning experiences. *Vision-haptic visualization* refers to the integration of two modalities, the sense of sight and sense of touch in perceiving virtual information. AR users can move the object nearer or farther from them and move themselves around the object to see different angles, thus heightening their interaction with the object.

Empirical Research on AR and Education

In the last decade, the educational field, in general, has experienced increasing interest in applying AR for educational purposes. AR is expected to become mainstream within teaching-learning processes, particularly in educational settings (Saltan & Arslan, 2017). AR studies in education have shown positive results such as an increase in

students' learning interest and concentration (Zhang et al., 2015), reduction of cognitive overload by providing students with "perfectly situated scaffolding" (Bower et al., 2014, p.1), authentic learning experiences (Klopfer, 2008), development of problem-solving, critical thinking and collaboration (Wasko, 2013), and increase motivation and collaboration (Bergig et al., 2009; Dunleavy et al., 2009).

Garzón and Acevedo's (2019) meta-analysis of 63 quantitative AR studies between 2010 and 2018 examined the effect of AR on 1) students' learning gains, 2) the learning environment, 3) the level and field of education, and 4) AR compared with other technologies. The selection criteria for studies to be included in the meta-analysis were 1) it had to measure the impact of AR on students' learning gains, 2) it had to have enough information to calculate an effect size, and 3) it had to have a control condition (pretest-posttest or control/experimental groups). The results showed a medium effect size of .64, suggesting that AR had a positive impact on student's learning gains. Furthermore, informal settings involving activities outside of classrooms produced better learning outcomes than formal settings (classrooms and labs).

Regarding the level and field of education, Garzón and Acevedo (2019) showed that AR had a greater impact on students from higher education than students of primary and secondary education. In addition, engineering, manufacturing, and construction were the fields with a very large effect size for the use of AR. Nonetheless, AR also had a large effect size in the field of arts and humanities, more specifically, second language teaching. Although 82% of the studies in humanities in the meta-analysis were SLA studies, Garzón and Acevedo (2019) believed that motivation, considered one of the

biggest factors in language acquisition (Solak & Cakir, 2015; Ushioda & Dornyei, 2017), contributed to the large effect size. Lastly, AR technology also had a greater effect size and thus, more impact on the learning gains of students compared with other technologies (multimedia, games, simulations).

However, other studies show that mobile-based AR games can also present challenges in the L2 classroom, such as cognitive overload (Hsu, 2017). For example, in their study of the AR game *Alien Contact!* with high-school students, Dunleavy et al. (2009) found that students reported feeling "frequently overwhelmed and confused with the amount of material and complexity of tasks that needed to be performed to play the game" (p. 17).

The current study contributes to this literature by investigating the number, nature, outcome, and correction orientation of LREs produced by learners in AR vs non-AR settings in two different modalities (oral vs writing focused) and the factors that might influence language development with the aid of this type of technology.

According to FitzGerald et al., (2014), the challenges associated with the use of AR can be technical, pedagogical, and social.

Technical challenges. The most common, inexpensive GPS systems are likely to be accurate within 10 meters and problematic with local environmental conditions (e.g., the reflection and shadow of skyscrapers). These conditions may produce registration errors that present AR users with nearby locations instead of the actual targeted location. Also, AR requires internet access usually from phone networks that are susceptible to the range and quality of the signal. Therefore, when using phones or tablets to complete

activities, students and instructors must make sure the devices are fully charged, and the network is accessible. Thirdly, maintaining the superimposed information can also be a challenge and may make students feel frustrated (Ji, Tan & Duh, 2018).

Pedagogical challenges. There is always the concern that students might be more interested in the technology than the learning objectives. In situational contexts, learners should engage with the surrounding environment more than the device which should just augment it. As with any other technology, AR should fit the learning objectives and not vice-versa. Other researchers argued that "providing learners with an immediate overlay of information has the potential to reduce observation skills by offering excessive scaffolding and reinforcement" (Fitzgerald et al., 2014, p. 10). In addition, this digital information can cause cognitive overload by delivering more information that one can process or irrelevant information in some cases of user-generated content (Fitzgerald, 2012). Lastly, AR implementation must have the teacher and the administration's interest and support for it to work.

Social challenges. As a new technology AR works on smartphones and mobile devices that run up-to-date versions of operating systems. In addition, some AR apps work only on iPhones and others only on Android devices. These differences may make the digital divide between learners even wider. It is important to create conditions where all learners can benefit from what AR can offer, such as providing them all access to the same type of technology.

Empirical Research on AR and SLA

SLA studies have shown that AR can potentially create optimal foreign language learning because it can transform the notion of a classroom setting into a situated learning environment (Thorne, 2013); this can be accomplished for instance, when AR allows students to go "out of the classroom" to explore the history of cultural artifacts and monuments triggered with images and videos (Holden & Sykes, 2011), and to develop context-awareness (Liu & Chu, 2010; Thorne, 2013; Thorne & Hellerman, 2017).

The ability to provide context-awareness and situational learning is one of the main benefits of AR mobile-based technology which in turn, can provide more effective learning experiences (Santos et al., 2014). Holden and Sykes' (2011) study showed the benefits of place-based AR mobile games to engage Spanish learners with the local environment. In the game *Mentira* (details below) "students use the setting of the game to move their Spanish beyond just the textbook and classroom to a meaningful place and context" (p. 13). Thorne (2013) emphasized "the power and the increasing ubiquity of mobile and GPS enabled devices to engage participants in language-rich experiences outside of the classroom" (p. 17), explaining that movement through the environment affords students language use that illustrates "the significance of context on the form and content of communication" (Thorne & Hellerman, 2017, p.8).

Although the studies above show the value of AR for collaborative interaction in a social context, they have not examined the effect of AR for language learning. Studies that specifically investigate the role of AR for language acquisition are very limited, which can be an indication of a research area that needs further investigation. To date,

research has been carried out on the effect of AR on the development of writing skills (Liu & Tsai, 2013; Wang, 2017), vocabulary (He et al., 2014; Hsu, 2017), pragmatics (Holden & Sykes, 2011), listening and oral skills (Liu & Chu, 2010), and IL development (Sydorenko et al., 2019).

Writing skills. Liu and Tsai (2013) developed an AR-based mobile learning material to help college English learners improve their writing skills. The authors used location-based AR to generate related information about buildings and scenic spots. Students were asked to introduce their campus to new students by visiting specific spots triggered by the AR application. After the trip, participants were asked to describe the observed scenery in their essays. Data collection included students' essays and an open-ended questionnaire about the students' experience with the tool. The researchers used a mixed-method analysis evaluating the words and sentences in the participants' essays and assessing students' perceptions of the experience through the open-ended questionnaire.

The results revealed that the AR-based mobile learning material facilitated the acquisition of linguistic and content knowledge that was later evident in the participants' expanded lexicon in their English compositions toward the end of the course. In addition, the application also helped students elaborate on information based on their own experiences, thus students were able to construct knowledge and produce meaningful essays (Liu & Tsai, 2013). Furthermore, the open-ended questionnaire revealed that the participants felt that the application provided something new and entertaining. Although Liu & Tsai's (2013) case study was one of the first to use location-based AR to expand scenes from a physical place and use writing assignments as a measurement for language

gain, the study had no control group to compare with students who had received the same material but in a different format (e.g., textbook) or context (e.g., classroom).

In contrast, Wang's (2017) study compared an experimental group using AR and a control group using traditional paper-based material to find out if AR could improve students' Chinese writing skills. Thirty students were evenly divided into two groups according to their writing level assessed during writing task 1. For tasks 2 and 3, the experimental group used AR markers in the classroom and around the campus respectively for viewing the AR content during the writing activities, while the control group used only paper-based learning content in the classroom. The results showed that learners in the experimental group performed significantly better than the control group on writing task 4 in terms of content control, article structure and wording. This was especially true for the low-achievement students who benefited from the AR materials which gave them more writing stimulation and helped them start their first paragraph. However, students' writing did not show any significant difference, perhaps because of the treatment's short duration.

Vocabulary. He et al.'s (2014) study used image-based AR to investigate the changes in students' learning interest after using an AR mobile application to learn English vocabulary. Forty pre-school children (ages 4-6) were divided into experimental and control groups equally. Participants in the AR experimental group, with no teacher guidance, aimed at a word on a card with their mobile camera which triggered a corresponding clickable picture with the word pronunciation. The learner then chose the word and connected with the appearing picture by listening to the pronunciation and

repeating it. The non-AR control group received the same type and number of words and corresponding pictures handed out on 8 cards by the teacher who pronounced the words for the students to repeat.

Data was collected from a pre- and post-test match game where the children had to match the English words from column A with the pictures in column B. The statistical analysis showed no significant difference between the control and experimental groups before treatment and a highly significant difference in the number of words retained by each group after the treatment, suggesting that English learners acquire vocabulary better using a mobile-based AR application. He et al.'s (2014) study is one of the first mobile-based AR investigations that show measurable linguistic gains, comparable treatment groups and pre- and post-test data. However, the study presented language outside of context, a limited amount of vocabulary due to the participants' ages, and limited treatment duration.

Similarly, Hsu (2017) also investigated a group of third-grade English learners to find out if different learning approaches influenced students' learning effectiveness, flow state, learning anxiety, and cognitive load. The authors created two AR educational games to help students spell and learn English vocabulary in real-life situations. In one game, learners were free to choose which situational stage or challenge to begin with. The other group of students used a game where they had to solve the situated tasks in sequence, guided by the AR application. Two classes of 20 and 18 students constituted the experimental group and control group, respectively.

The data was collected from pre- and post-tests, and questionnaires regarding learning effectiveness, flow state, learning anxiety, and cognitive load. The results revealed that students who learned with the self-directed AR game experienced a higher flow state. In addition, the study suggested that students' mental efforts are greater when they experience more learning anxiety at the same time. Therefore, the authors concluded that appropriate but not excessive mental effort and L2 learning anxiety are necessary for achieving learning effectiveness.

Hsu's (2017) study was one of the first studies to assess the cognitive load, foreign language learning anxiety and learning effectiveness of the students with different learning styles. However, the sample size was small, and the learning targets were all objects, due to participants' age. In addition, the intervention had a short time. Therefore, future studies with more learning targets and participants, and a longer period of intervention could confirm the results.

Like Dunleavy et al.'s (2009) study, Hsu's (2017) analysis also showed that immersive technology can cause cognitive overload, that is, "AR users may experience certain levels of cognitive overload because AR requires users to process a large amount of information that they encounter in the learning context" (Suh & Prophet, 2018, p. 18).

Pragmatic competence. Holden and Sykes (2011) created a mobile-based AR game called *Mentira,* designed to develop students' pragmatic competence in an intermediate-level Spanish class. In the game, participants had to prove their innocence in a local murder case. To that end, students had to leave their physical classroom and go to a Hispanic neighborhood and interact with local people in the target language. Each

student was assigned a different role along with clues only known by his or her character. The game involved conversations between the player and the game's non-player characters (NPCs) in the form of scripted dialogues in which students had to choose from multiple responses that led to different scenarios. In each level, various pragmatic actions were necessary for successful interactions with the NPCs in the game.

The data was collected from three, four-week implementation sessions that took place over the course of three semesters with a total of 68 participants. The data consisted of gameplay videos and data, in-class observations, interviews, and written assessments. The results revealed that mobile-based learning motivates students. In addition, games can serve as feedback systems to improve practice and a safe place to practice pragmatic behaviors that could have different consequences in real interaction (Holden& Sykes, 2011).

Although *Mentira* is the first mobile-based, AR mobile game for learning Spanish, Holden and Sykes' (2011) lack of quantitative data did not allow them to determine any evidence of gains in pragmatic skills or in aspects of language proficiency in general. However, *Mentira* offers a perspective of the benefits of integrating mobile games in language learning through situated learning to expand students' participation and redirect the curricula towards knowledge building (Holden & Sykes, 2011).

Listening and speaking skills. Liu (2009) examined the use of an AR application called HELLO (The Handheld English Language Learning Organization) with 64 7th grade students assigned to an experimental or control group and 3 high school teachers. The authors created the application/game that involved situated task-based activities

where eight students per group had to take photos of quick response (QR) codes attached to objects around the school for 8 weeks. The control group used traditional learning methods (e.g., printed materials and CDs). The QR codes were converted to data that were used to determine the students' location and to access remote learning material or to display a 3D virtual learning partner to whom students could practice their listening and speaking skills.

An eight-week pilot study was conducted with a pre-test, three tests and a posttest to evaluate students' listening and speaking skills. The results showed that the assessments' average grade of the experimental group exceeded those of the control group in tests 1, 2, 3, and the posttest. A seven-point Likert-scale survey was given to the students at the end of the treatment to evaluate the students' perceived effects of the application on their 1) listening, speaking, and reading abilities; 2) motivation; and 3) context-aware learning. The survey results indicated that most students thought that practicing English in a real-life situation improved their learning and encouraged their creative abilities.

Although Liu's (2009) study is one of the first studies to explore the potential of AR to enhance L2 listening and oral skills, she examined students' improvement by test scores and not each skill (e.g., speaking and listening) separately. Liu and Chu's (2010) study also lacked details about the type of tests given and how speaking and listening skills were measured.

IL development. Sydorenko et al.'s (2019) study is the only study thus far to examine LREs in an AR setting. Twelve students were divided into four groups of three

(two English language learners and one English expert speaker) to play a mobile-based AR game called *ChronoOps*. In the game, players are in the year 2070 and are brought back to the present year to discover green technologies that could help save the future's environmental catastrophe. The groups' interactions during the game were video recorded and LREs identified in each group.

Results showed that lexical LREs were predominant. None of the 32 total LREs addressed form. The CF that initiated the LREs were requests for assistance used by the learners, while the expert speakers used recasts, corrections, and comprehension checks. Both learners and expert used clarification requests. Twenty-two percent of the LREs were unresolved due to poor explanation by the expert speaker or because learners were more interested in progressing through the game. However, 34% of the words that caused the LREs in the first place were used later in gameplay, evidence that students remembered them.

Although Sydorenko et al. (2019) was the first study to explore LREs in an AR setting, the study used a triad where one of the interlocutors was an expert speaker comparable to a teacher or NS. The study did not include a posttest to confirm language gains nor did it have a control group to compare outcomes. However, the study showed that the AR game created "opportunities for just-in-time and situationally driven vocabulary learning" (p. 734) and brought awareness to the topic of 'learning "in the wild" (p. 734) as a context of high relevance for language learning. The authors define the "in the wild" concept as "learning in situated open spaces outside of conventionally structured classrooms" (Sydorenko et al., 2019). Lantolf and Thorne (2006) describe

activities in the wild as an "exogenous activity system… not directly related to education" (p. 225). Although part of the current study uses open spaces for peer interaction, it also uses structured instructions related to education. Thus, the concept of place-based as "the process of using the local community and environment… emphasizing hands-on, real work learning experiences" (Sobel, 2004, p. 6) was more appropriate to use for this study.

In sum, the studies above show the potential of mobile-based AR in the field of SLA. Results from these studies serve as a foundation on which to grow our understanding of the effects of mobile-based AR on L2 acquisition. Three of the studies used image-based AR (He et al., 2014; Hsu, 2017; Liu 2009) while the other three used location-based AR (Holden & Sykes, 2011; Liu & Tsai, 2013, Sydorenko et al., 2019). College students were the subjects of three of the studies (Holden & Sykes, 2011; Liu & Tsai, 2013; Sydorenko et al., 2019) while children were the participants in the other studies (He et al. 2014; Hsu, 2017; Liu, 2009). All studies, except Holden & Sykes (2011), had English as the target language. Qualitative analysis based on interviews, videos, and surveys was the most adopted method in both image-based and location-based AR investigations. He et al.'s (2014), and Hsu's (2017) studies were the only ones to use a quantitative pre- and post-tests method to evaluate the learning outcomes in young children. Additionally, He et al. (2014), Liu (2009), and Sydorenko et al. (2019) were the only studies that investigated linguistic gains among the subjects.

In conclusion, while AR applications have started to receive a lot of attention in SLA studies, many theoretical and methodological questions remain unanswered

regarding the effectiveness of AR technologies for IL development. This study attempts to answer the following questions:

Research Questions

1. Does the task setting (AR vs non-AR) differentially affect the (a) number, (b) type, (c) outcome, and (d) correction orientation of target LREs produced in the two modalities (oral vs writing-focused)?
2. Does the setting and LRE outcomes affect the ability of participants to recognize and produce the correct grammatical and lexical forms on the posttests? Is there a significant correlation between students' ability to recognize the correctness of a grammatical form and their ability to produce that form correctly on the posttests?

CHAPTER 3

METHODOLOGY

This study used a crossover experimental design, in which participants "cross over from one treatment to another during the course of the trial" (Piantadosi, 2005, para. 1). Thus, participants received both treatments and served as their own control groups. This contrasts with a quasi-experimental design in which the experimental group receives a treatment, and another group serves as a comparison. The reason to consider a crossover design is that it could yield a more efficient comparison of treatments, it can be done with fewer participants and have the same level of statistical power or precision as a quasi-experimental design (Piantadosi, 2005). The results are compared between the responses from A vs B (Table 1), in which Group 1 receives treatment A first and treatment B second, while Group 2 receives treatment B first and treatment A second.

The crossover design involved advanced students of Spanish placed in dyads completing two collaborative tasks using AR mobile applications (treatment A) outside of the classroom and traditional computer applications (treatment B) in the language laboratory. These experimental tasks had two components: a production of just oral output (hereafter, oral activities) and the production of oral and written output (hereafter, writing-focused activities). To delimit the scope of this investigation, and for purposes of comparison with other studies on the same topic, only the verbal communication during the oral and writing-focused activities was used for data analysis. No data from the written products of the writing-focused activities was analyzed. The tasks targeted the chapters' grammar structures students were learning at the time. However, the tasks used

different scenarios requiring different vocabulary (food and furniture) not part of the class curriculum.

Table 1

Diagram of Treatment

	Task 1	Task 2
Sequence AB (Group 1)	A (AR)	B (non-AR)
Sequence BA (Group 2)	B (non-AR)	A (AR)

Participants

The subjects in this study were 20 students, 10 males and 10 females, enrolled in the third-year Spanish conversation/composition course at a University in the southwestern region of the United States. To be enrolled in the class, students are required to have taken four lower-division courses: Spanish 101, 102, 201, and 202 or have passed a placement test. Details about students' demographics and technology background can be found below.

The participants interacted in dyads as they performed communicative tasks. The procedure for placing the students in dyads was as follows. The researcher asked the instructor to classify each student based on their performance in class on a three-point scale as "below average (1)," "average (2)," and "above average (3)" in terms of their overall linguistic abilities in Spanish. Seven students were classified as "above average," eleven others were classified as "average," and two students were classified as "below average." Therefore, learners of matching levels of relative proficiency were placed in

dyads based on the teacher's input. Thus, the study had three dyads in the "above average" category, five dyads in the "average" category, one pair in the "below average" category and one mixed dyad ("above" and "average"). Philp et al. (2013), drawing from a Piaget's cognitive perspective, suggested that by matching dyads with the same linguistic competence for the task, students can challenge one another's preexisting conceptions about the target language. This challenge provides the incentive for IL development in a process of "continual construction and reconstruction in response to their experience of language use" (Philp et al., 2013, p. 10). The dyads were split evenly into Groups 1 and 2.

Additionally, every effort was made to make both groups evenly divided based on their linguistic abilities and the phone device they owned**.** The AR applications required an IOS operating system OS X 11 or better (except for *Google Translate AR* which also works on Androids). Therefore, at least one member of each dyad had to have an iPhone7 or higher. Table 2 displays the participants' group and task selection.

Instruments

The instruments used to gather data were the following: a language and technology background questionnaire (pre-questionnaire), audio recordings of task interactions, and two written posttests. These instruments allowed for quantitative analysis of the results and provided the dialogue excerpts that illustrate the findings.

Language and technology background questionnaire (pre-questionnaire). At the beginning of the semester, students completed a pre-questionnaire (see Appendix C) so that the researcher could obtain their demographic data such as age and gender.

Table 2

Participants Group Selection

Dyad No.	Student No.	Pseudonym	Spanish Proficiency	Phone Type	Groups
1	1	Pat	3	iPhoneX	1
	2	Jenn	3		
2	3	Mark	3	iPhoneX	1
	4	Betty	2		
3	5	Karen	2	iPhoneXS	1
	6	Lisa	2		
4	7	Bill	2	iPhoneXR	1
	8	Rob	2		
5	9	Matt	3	iPhone7	2
	10	Ashley	3		
6	11	Kim	1	iPhone8	2
	12	Paul	1		
7	13	Anna	2	iPhone7	2
	14	Nick	2		
8	15	Richie	2	iPhoneX	2
	16	Charles	2		
9	17	Melissa	2	iPhone8+	2
	18	Amy	2		
10	19	Jim	3	iPhoneX	1
	20	John	3		

In addition, this pre-questionnaire was also designed to collect experiential and motivational information from each learner before the beginning of the investigation, such as native language, languages spoken, languages they are studying, major, their reasons for taking Spanish, and any experience with AR technology. The technology questions helped gain additional insight into learners' attitudes toward the use of technological tools involved in the study, specifically toward the use of a virtual technology such as AR.

Audio recordings of task interactions. All conversations during the oral and writing activities were recorded and archived. Each dyad used a recorder and their mobile phones (AR task) or lab computers (non-AR task) to complete the oral activities. In addition, communication during the writing-focused activities in the language lab was also recorded using the same recorder. Since the researcher was not present during the oral activities, it was only possible to observe the interactions through the recordings, thus eliminating the probability of *halo, subject* or *researcher expectancy* effects. All recordings were later transcribed verbatim by the researcher or with the help of the transcription company *Transcription Panda* (http://www.transcription panda.com). The researcher listened to the recordings a second time to make sure all transcriptions matched the recordings verbatim.

Posttest design. The posttests after Task 1 and Task 2 were based on the lexical and grammatical LREs identified in the oral recordings. These LREs were the ones that dyads solved correctly, incorrectly, or were left unresolved. The objective of the posttests was to determine whether learners remembered the resolution to the grammatical and lexical LREs forms that caused the LREs in the first place, and whether students could produce those outcomes. The customized immediate posttest items were developed under considerable time pressure (1 day for a total of 10 dyads – about 10 hours) and administered two days after each task. Thus, the researcher listened to the first 10 minutes of the oral and written interactions of each dyad and tried to choose a balanced amount of lexical and grammatical items that caused most discussions (i.e., LRE turns) during interaction. Thus, participants were not specifically tested on the grammatical structures

they were asked to use for the writing activities. If not enough LREs were present in the first 10 minutes, the researcher continued to listen until there was enough items to make up the test.

The posttests had four questions. Question 1 had 5 grammatical sentences, question 2 and 3 had a lexical or grammatical item within a context. Finally, question 4 had four lexical items. These questions were created to test the participant's grammatical and vocabulary knowledge. Drawing from Swain and Lapkin's (1998) study, question (1) was intended to "capture movement along a continuum of not knowing something or not being certain of something to greater certainty" (p. 326). Thus, in question (1) learners were provided with a certainty scale and had to make a judgment about the grammaticality of five sentences, indicating if the sentences were *definitely wrong, probably wrong, probably correct, definitely correct,* or *they did not now* (Swain & Lapkin, 1998). If the learner indicated the sentence(s) was *wrong* or *probably wrong*, they could then write the sentence(s) in a way they thought was correct.

These sentences in one way or another were the point of focus during task interaction. The episode below is a grammatical LRE in which a participant (pseudonyms used) initiates the episode and his peer tries to help, in this case with the incorrect gender form (masculine = (m), feminine = (f)) of the adjective "good". The Likert-scale test question based on this LRE follows it.

Episode 1: An example of a Likert-scale test item of a grammatical LRE

Mark: *El otro es Nekter… escriba, después de Chop Shop cuando tenemos sed, podemos ir al Nekter para buscar algo a beber. Ahí hay muchos bebidos de frutas ricas y otras* ***sabores buenos… ¿buenas?***

[The other is Nekter… write, after Chop Shop, when we are thirsty, we can go to Nekter to look for something to drink. There are many drinks of rich fruits and good (m) flavors… good (f)?]

Betty: *Hay muchos bebidos y otras **sabores buenas**.*
[There many drinks and good (f) flavors.]

Test question for Mark and Betty: For each sentence below indicate whether the sentence is correct or incorrect by indicating to what extent you are certain of your answer by checking the appropriate box according to the scale below. Please circle any part of the sentence that contains errors and correct them.

5 definitely wrong	4 probably wrong	3 probably correct	2 definitely correct	1 don't know

Ahí hay muchos bebidos de frutas ricas y otras sabores buenas. 5 4 3 2 1
Correct: _____

In question (2), learners had to translate a word or expression in English to Spanish

within a context. This word or expression came from a direct request for help (CS = DA

5) CS or own accuracy check (CS = AC 31) CS present in the interaction.

Episode 2: An example of a translation test item of a lexical LRE

Mark: *El museo, ¿sí? Podemos hablar del museo. Yo de vez en cuando yo tengo ganas de ir ahí, pero siempre es que… ¿**las maleta**?*
(CS = 31 AC [own accuracy check])
[The museum, yes? We can talk about the museum. I sometimes, I feel like going there, but always is that… the bag?]

Betty: The balet? *¿**Baleta**?* (CS = 3 CR [clarification request])
[The balet?] (CS= 20 LT [literal translation])

Mark: *No, la palabra se me olvidó, voy a buscar… **el boleto**. El boleto para ir al Gammage siempre está tan cara, tan caro para comprar.*
[No, I forgot the word, I'm going to search… the ticket. The ticket to go to *Gammage* always is very expensive (f), so expensive (m) to buy.]

Test question for Mark and Betty: Provide the Spanish translation for the English word in the text below:

Tengo ganas de ir al museo, pero siempre… Cómo se dice "*tickets*"?

_____

For question (3) learners provided the translation of a conjugated Spanish verb or expression to English.

Episode 3: An example of a verb definition from Spanish to English

Mark: *No, sabemos que tendrás ganas* ***de tratar*** *la comida americana… no,* ***de probar*** *la comida americana.*
[No, we know that you will like to try the American food… no, to taste the American food.]

Test question for Mark and Betty: Provide the English definition of the boldfaced Spanish word:

Sabemos que tendrás ganas de ***probar*** *la comida americana.* _____

For question (4) learners translated four words from English to Spanish. These words were lexical LREs present during interaction that were present in the first 10 minutes of the recording or caused most discussions (i.e., LRE turns) during interaction.

Episode 4: An example of a vocabulary test item (sugar came from a different episode)

Mark: *El tazón de mantequilla de maní.*
[the peanut butter bowl]

Betty: *¿****maní*** *es* almonds?
[maní is almonds?]

Mark: ***maní*** *es* peanuts.
[Maní is peanuts.]

Betty: *¿que significa* ***almendra****s?*
[What does *almendras* mean?]

Mark: almonds.
Test question for Mark and Betty: Provide the Spanish translation for these English words:

açai bowl: _____ almonds: _____

peanut butter: _____ sugar: _____

A delayed posttest would have shown how much the participants had retained from the LREs produced during their collaborative dialogue and likely would have confirmed the relationship between LREs and L2 development (e.g., Kim 2008; McDonough & Sunitham 2009; Swain & Lapkin 1998). However, for logistical reasons the implementation of a delayed posttest was not possible, which is a limitation of this study. Appendix E shows a sample of a posttest from Task 1.

General Procedures

Treatment timeline. The time frame and tasks for this experiment are shown in Table 3. To summarize, participants first were introduced to the project, completed the pre-questionnaire, and downloaded the app for Task 1. Next, students received their group assignments and completed the tasks. The interaction between learners was recorded, and the recordings were used to create the customized posttests based on the LREs produced by each dyad during the tasks. The posttests were administered two days after each task session.

Table 3

Treatment Timeline

	1	2	3
Phase	**Pre-treatment**	**Treatments**	**Posttests**
Duration	Week 6	Weeks 7 & 12	Weeks 7 & 12
Instrument	- pre-questionnaire - download apps	AR & non-AR Tasks 1 & 2 administered	Same-week custom posttests 1 & 2

One week before the completion of the Task 1, the researcher visited the class and told the students that they were going to do two activities in pairs using AR technology and non-AR applications (e.g., websites). The researcher informed the students that these activities were part of a research study on language learning, but that they were also part of required class assignments. However, they were free to decide whether they wished their assignment data to be used in the study or not. They were informed that their decision would in no way affect their grade in the course. Students who agreed to take part in the study, read and signed a Human Subjects Consent Form. Once they signed the consent form, they then completed the pre-questionnaire.

Finally, for students to get familiar with the AR applications, the researcher showed the applications used in Task 1 and explained the *Bettar* AR and *Google Translate AR* applications. Students then were instructed to download them to their phones and practice using them over the week. Although the applications were easy to use and students were instructed to practice using the applications on their own, they did not have enough time in class to practice with the apps nor ask questions about it before the treatment. This could be considered as a limitation of this study.

All participants met during classroom time to perform the tasks. The following week, the class met at the language lab and the groups and dyads were identified according to the criteria explained above. The researcher then explained to students their task and directed the Group 1 to go outside to areas around the building to complete the oral activity using the *Bettar* AR and *Google Translate AR* apps. Group 2 stayed in the lab to complete the oral activity using the Google Maps and Google Translate websites.

During the task, the researcher briefly visited each dyad from both groups to make sure there were no technical problems or questions related to the oral activity.

After 20 minutes, the Group 1 came back to the language lab and both groups started on the writing activity of Task 1. The researcher showed an example of the email both groups should complete for the assignment, emphasizing the grammar structures they should include. The students were subsequently told to use a specific URL where they would find folders with their dyad's names. They were instructed to try to use the grammar structures in the writing activity and help each other if any language-related problems arose. Both groups completed the writing activity in 20 minutes. Students' conversations were recorded as they interacted in pairs to complete the oral and writing activities of Task 1.

One week before the completion of Task 2, per the researcher's request, the class instructor sent an email to students who were going to use AR for Task 2 asking them to download the *IKEA AR* app and familiarize themselves with it. On week 12, the class once again met at the language lab at regular class time to complete Task 2. The task was explained by the researcher and a video shown to students on how to use the *IKEA AR* app. The researcher took each dyad from Group 2 to designated areas around the campus where, per the oral activity, they could have enough space to "furnish" their "living room". Group 1 stayed in the lab with the Instructor and completed the oral activity using the *IKEA* website.

After 20 minutes, Group 2 came back to the language lab to complete the writing activity along with Group 1. The researcher showed an example of the email to the

students with samples of the grammar structures in Spanish. The students were subsequently told to use a specific URL where they would find folders with their dyad's names. Both groups completed Task 2 writing activity in 20 minutes. Although studies (Sato & Ballinger, 2012; Sato & Lyster, 2012) have shown that training students to provide feedback can improve accuracy and fluency, for this study students were just verbally instructed to help each other if any language-related problems arose. This was done to keep with the study's *ecological validity* (Long, 1985).

The posttest then was administered to the students two days after each treatment. The items on the posttest were chosen according to the LREs produced during peer interaction. Because this study examines the concept of peer interaction, the posttests were customized per dyad and not per individual. The dyads' conversations during the collaborative oral and writing-focused activities of each task were audio-recorded and transcribed verbatim for data analysis. The purpose of analyzing these transcriptions was to study the nature of the collaborative process, and the LREs produced by learners during the collaborative tasks.

The recordings were collected on the two task days and the posttests administered and collected two days after each treatment, during the participants' regularly scheduled class. At the end of the experiment, recordings, posttests, and pre-questionnaire were gathered for analysis.

Due to technical issues, one dyad did not have their conversation recorded during Task 1 and the same happened with two dyads on Task 2, where one did not record their conversation during the oral activity and another did not record their conversation during

the writing-focused activity. For the data analysis, to control for the differences in the quantity of discourse produced across dyads and the number of tasks completed per dyad, the number of LREs were divided by the total number of words produced by each dyad and multiplied by 100 (Yilmaz & Granema, 2010). Although there were 10 dyads, the number of LREs was divided by nine to compensate for the missing data from one dyad who did not complete Task 1 and two dyads who did not complete part of Task 2 (the oral or the writing-focused activity - see Table 14 and 15 for details).

Experimental tasks. Following Swain & Lapkin's (1998, 2001) LRE experimental design and previous research on LREs, this study employed two decision making tasks. These collaborative tasks were created to generate conversation and engage students in linguistic problem-solving. The tasks created for this study were incorporated into the course assignments.

The tasks required dyads to search for restaurants, food, and events in the town surrounding the campus (Task 1) and shop for furniture online (Task 2). The oral activity of each task required each pair of students to engage with either an AR application or the non-AR version (e.g., website) of the same or a similar application. Therefore, for Task 1, Group 1 used the *Bettar AR* and *Google Translator AR* applications on campus while Group 2 used *Google Maps* and the *Google Translator* websites in the language laboratory. Both applications, *Bettar* and *Google Maps*, are navigation applications that use GPS and maps to help users find public places (e.g., restaurants, events). The difference is that *Bettar* uses AR maps and have the functionality of showing name tags of these places when you point the phone camera in a certain direction (Figure 3). Similarly,

Google Translator AR uses Word Lens technology to translate photos of signs, menus, and similar items from one language to another when the user points the phone camera at an item.

For Task 2, the two groups switched settings, the group who previously used AR for Task 1, used the non-AR applications for Task 2 and vice-versa. Group 2 used the *IKEA AR* app outside on campus while Group 1 used the *IKEA* website in the lab. While the AR oral activities took place outside, the non-AR oral activities and all writing-focused activities were conducted in the language laboratory. Both groups used *Google Docs* to complete the writing-focused activities. Once the student was assigned to work with another student, both students stayed in the same dyad for both tasks.

Although the topics, entertainment and shopping, were not part of the class curriculum, they were covered in previous semesters and familiar to the students, which allowed students to focus attention on their language and writing. For the oral activities, students had complete freedom to use the language as they saw fit. For the writing activities, students were instructed to include some specific grammar structures, but they had the freedom to develop their compositions the way they chose. The grammatical structures encouraged by the tasks (e.g., subjunctive, direct and indirect objects) were based on the content of the chapters they were studying at the time, allowing the students to use language forms they had been exposed to in class. This could be considered a possible limitation if the LREs could not be isolated from what participants were learning in class at the time. A detailed analysis will reveal whether the problems prompting the LREs were related to the grammar points they were learning at the time.

The compositions for the writing activities were to be informal in nature - an email to a friend (Task 1) and to parents (Task 2), as opposed to the production of narratives or persuasive essays, keeping with the goal of task authenticity and real-life examples. The topics for each task were intended to draw from learners' life experiences.

Task design. ***Task 1***. Group 1 dyads went outside on campus to search three restaurants and three events in Arizona for the weekend using the *Bettar* AR app. The user points his/her phone to scan the area in a 180-degree motion to see different colored tags showing restaurants (blue), events (green), and people (orange) (Figure 3). After chatting and reviewing three restaurants' menus and events, dyads had to agree on one restaurant and one event to take a fictional friend from Mexico (María), who was coming to visit for the weekend. Once the restaurant was chosen, each student from the dyad chose an item in the menu to eat.

The dyads then used *Google Translate AR* to translate their menu choices to Spanish. To encourage more conversation between learners, each student had a different set of parameters. For example, one student pretended to be a vegetarian and like music while the other loved meat and sports. Figures 3 and 4 show screenshots of *Bettar* and *Google Translate AR*, respectively. All apps chosen for the tasks were free applications and had the features that were required to complete the tasks.

Dyads from Group 2 had to complete the same oral activity in the language lab, using *Google Maps* and *Google Translate* instead. Dyads had the option to use the non-AR apps on their phones or the website on the lab computer. Figures 5 and 6 show screenshots of the two apps. Both groups had 20 minutes to complete the oral activity.

After 20 minutes, Group 1 came back to the language lab and both groups completed the writing-focused activity. Dyads were instructed to write an email to their fictional Mexican friend María telling her the plans for the weekend (see Appendix D for task instructions). Dyads had 20 minutes to complete the writing activity using *Google docs* in the lab computers or personal laptops.

Figure 3

Bettar AR Application Screenshots

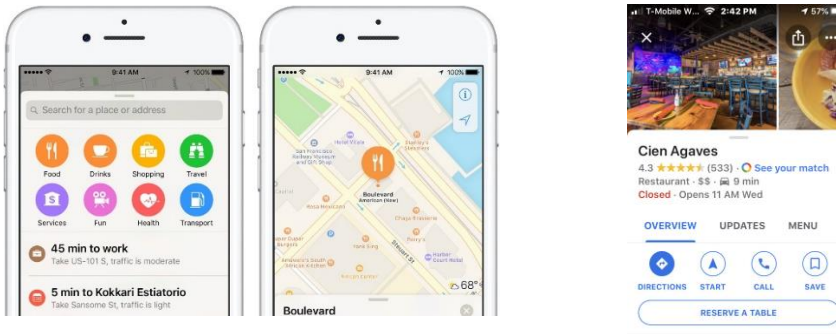
Source: apps.apple.com

Figure 4 ***Google Translate AR Application Screenshot***

Source: cnet.com

Figure 5

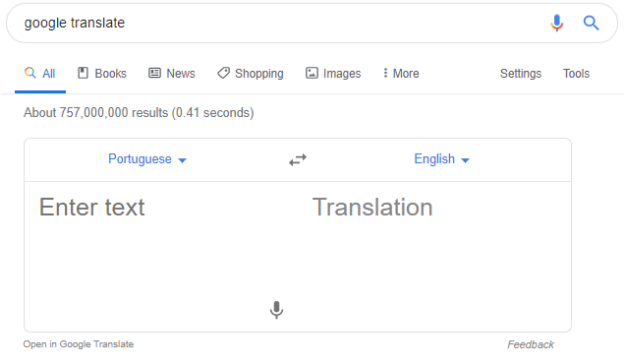
Google Maps (non-AR) Application Screenshot



Source: appleinsider.com

Figure 6

Google Translate (non-AR) Website Screenshot



Source: Google.com

Task 2. For Task 2, the dyads switched settings. The previous AR group was now the non-AR group while the previous non-AR group was now the AR group. Thus, all groups were exposed to the same types of experiment. Each dyad had to furnish a living room using the *IKEA AR* app (Group 2) or the *IKEA* website (Group 1) with a fictional budget of $1,500.00. Dyads went to areas on campus determined but the researcher to

furnish their “living room”. They had to choose a total of 8 pieces of furniture from different categories (*IKEA Collections* or *Back to College*) with up to 2 pieces from each category.

The *IKEA AR* app shows exactly how furniture items would look and fit into a physical space. Users can change the size, rotate, and move pieces of furniture as they choose. Together, participants decided and agreed on which things to buy without exceeding the budget, and how they were going to place them in the real space. Dyads chose and placed the 3D pieces of furniture in the space and saw through their phone camera the physical room as if the 3D furniture pieces were real. They were able to walk around them and interact with them (see Figure 7). The dyads had 20 minutes to complete this oral activity.

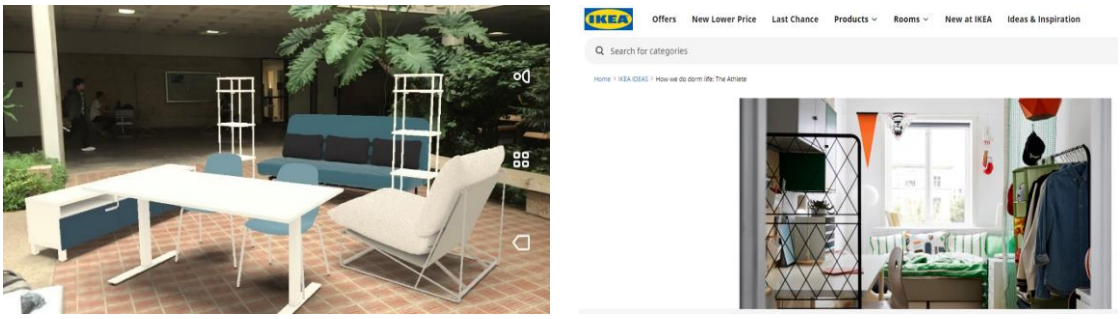
Group 1 used the *IKEA.com* website to complete the same oral activity in the language lab. Once it was completed, both groups completed the writing activity in the lab. Dyads had 20 minutes to write a letter to their parents using *Google Docs* telling them 1) what items they chose and why, 2) how they were placed in the room, 3) how much they spent and 3) why they decided on the items they bought. Dyads used a recorder with a microphone to record oral interactions during the oral and writing-focused activities.

The AR apps (*Bettar*, *Google Translate AR* and *IKEA AR*) and the traditional apps and websites (*Google Maps*, *Google Translate*, and *IKEA.com*) used in the study required only basic technology skills from the students. Thus, all participants were able to use the

applications to complete the tasks, regardless of their technological background knowledge. Students used their iPhones to download the AR applications.

Figure 7

IKEA (AR and non-AR) Applications Screenshots



Source: photo from this study (1) and *IKEA.com* website (2)

In sum, the structure of both tasks was essentially the same in that participants were faced with a situation where they had to narrow down from a number of possibilities which required them to come to a mutually agreeable decision. In this respect, both tasks can be defined as decision-making tasks (Pica et al., 1993) in which participants are expected to work toward a single outcome but have several outcomes available to them. As students choose among different outcomes, they negotiate impasses in their mutual understanding or decision making, thus, offering each other modified input and feedback, and responding with a modified output (Pica, 2005). This type of task may allow for more open discussion than other tasks, however, they do not guarantee that there will be detailed interaction information exchange because participants might not contribute equally, or students might be just focused on reaching a decision without giving elaborate justifications for their opinions (Gass & Mackey, 2011). In addition,

some of the group members or an individual from a dyad may take control over the task to the exclusion of others.

However, convergent tasks such as this, in which learners must agree on a specific outcome (Lowen & Sato, 2018) results in more negotiation of meaning than divergent tasks, in which learners may hold different opinions regarding the task outcome (Gilbert, 2009). Other studies have shown that dyad make-up, as in NNS-NNS was more important to the amount of negotiation than the nature of the task (García-Mayo & Pica, 2000). As mentioned before, the goal of the current study was not to focus on types of tasks that explicitly elicit negotiation for meaning and form (e.g., information gap, jigsaw) but instead find tasks that had ecological validity.

Coding. Because negotiated interaction can happen during any time of the conversation, dyads' collaborative dialogue was used in calculating turns, words, and LREs. Turns were calculated as each time there was a transfer of the "floor" from one participant to the other (Smith, 2003b).

LREs. LREs were coded according to their types - lexical, grammatical, and mechanical; in addition, a total of 17 LREs subtypes were identified and defined (see Appendix B). Each LRE deals with only one linguistic item. If the same LRE was discussed in several turns during a single conversation, that LRE type and outcome was coded only once. However, one larger LRE can have smaller ones embedded in it (Swain & Lapkin, 1998).

Below, is an example of one LRE, coded from the current data (pseudonyms used) with triggers in bold:

Mark: *Sí, esa se ve bien, tiene cebolla,* ***que más****…* yo *quiero algo, ve...* ***como se dice ¿"vegan"?***
[Yes, this looks good, it has onion, what else… I want something, ve… how do you say vegan?]

Betty: ***¿vegano? ¿Algo vegano?***
[Vegan? Something vegan?]

LRE type = lexical, subtype = adjective (vegan), outcome = solved correctly.

As illustrated in the example above, an LRE can involve many CSs and a few conversation-turns focused on one single lexical, grammatical, or mechanical problem. This lexical LRE begins with the participant's awareness that the word he/she wants to communicate is not available in his/her IL and a CS needs to be used. The LRE ends when the dyad reaches a mutual agreement and/or decides to move on with the conversation. A few turns later Mark used the word "*vegano"* confirming the knowledge building that happened through their collaborative dialogue during the LRE. The episode below illustrating a grammatical LRE shows Patty *implicitly* helping Jenn realize that the correct adjective for the noun *día* is not feminine (f), but masculine (m):

Patty: *malgastar* ***toda*** *el día…*
[to waste the whole (f) day…]

Jenn: *el día,* no, ***el*** *día* (CS = 7 IS [inferential strategy])
[the (m) day. No, the (m) day] (emphasizing the noun *day* in Spanish is masculine so it requires a masculine adjective)

Patty: *tú eres correcta.* ***Todo***, ***lo siento***. (CS = 15 [response confirm])
[you are right. All (m), I'm sorry.] (CS = 32 SF [social formula])

LRE type = grammatical, subtype = gender, outcome = solved correctly.

Posttests scores. On the posttests, the number of potentially remembered words for each dyad was counted based on the LREs produced during peer interaction. The

posttests had two sections, one grammatical and one lexical. The grammatical section had two parts – recognition and production. For the recognition part students were scored based on their answers to the Likert-scale prompts. These prompts (drawn from the dyad's dialogue) were either grammatically correct or incorrect (see section "Posttest design" above for examples).

Higher scores show better grammaticality judgments (they reflect a combination of the correctness of the prompt and participant's degree of certainty with the correctness of the prompt). For instance, a score of zero indicated no recollection from the participant ("I don't know"), a score of one indicated that a participant recognized the structure incorrectly, a score of two indicated that a participant recognized the structure incorrectly with a certain degree of doubt, a score of three indicated that a participant recognized the structure correctly but with a certain degree of doubt, and a score of four indicated that a participant recognized the structure correctly. Below the scores with LRE examples drawn from the Likert-scale posttest questions:

5 Definitely correct	4 Probably wrong	3 Probably correct	2 Definitely wrong	1 I don't know

- Grammatically correct posttest prompt: "*En domingo podemos hacer cosas que te interesan.*"

	Answer	Score
o Student marked that it was definitely correct	5	4
o Student marked that it was probably correct	3	3
o Student marked that it was probably wrong	4	2
o Student marked that it was definitely wrong	2	1
o Student marked "I don't know" answer	1	0

- Grammatically incorrect posttest prompt: "*Creo que compráramos muchos muebles*."

	Answer	Score

- Student marked that it was definitely wrong 2 4
- Student marked that it was probably wrong 4 3
- Student marked that it was probably correct 3 2
- Student marked that it was definitely correct 5 1
- Student marked "I don't know" answer 1 0

The second part of the grammatical section was production. Participants had the opportunity to replace the incorrect form with what they thought was the correct form.

Thus, production scores were coded as follows: a score of one indicated that a participant did not try to repair an incorrect prompt or incorrectly repaired a correct prompt, a score of two indicated that a participant repaired a correct prompt with an unnecessary correction (i.e., "*Te voy a mostrar mi ciudad*" with "*¿Voy a mostrarte mi ciudad*") or repaired an incorrect prompt still incorrectly, a score of three indicated that a participant left a correct prompt as is or repaired an incorrect prompt with the correct form. As with recognition scores, higher scores show better productive knowledge (i.e., participants left the correct forms alone and corrected the wrong forms correctly). Below the production scores with the same LRE examples from above:

- Grammatically correct posttest prompt: "*En domingo podemos hacer cosas que te interesan*."
  - Student did not try to correct 3
  - Student "repaired" the correct prompt with an unnecessary correction* 2
  - Student "repaired" the correct prompt incorrectly (bad form provided) 1

  * adding, eliminating words or just copying the original correctly, but the result of that "correction" is still grammatical (e.g. "prompt: ¿Cómo estás tú? Correction: ¿Cómo estás?)

- Original prompt was grammatically incorrect: "*Creo que compráramos muchos muebles*."
  - Student repaired the incorrect prompt correctly 3 answer: compramos
  - Student "repaired" the incorrect prompt incorrectly 2 answer: comprábamos

- Student did not try to repair the incorrect prompt 1 answer: (blank)

The second section of the posttest consisted of lexical items that participants had to translate from English to Spanish or Spanish to English within a context and without. Higher scores show better production knowledge. For instance, for words within a context a score of zero indicated no translation of the lexical item, a score of one indicated that participants produced an incorrect translation of a previously unresolved, correctly solved or incorrectly solved lexical LRE, a score of two indicated that a participant produced the correct translation of a previously unresolved, correctly solved, or incorrectly solved lexical LRE. Below the production scores for with the LRE examples drawn from the posttest questions 2 thru 4:

2. Provide the Spanish translation for the English word in the text below:
   Las luces 'on top' del lago… Answer: encima de (Score: 2)

3. Provide the English definition of the boldfaced Spanish word:
   Es negro y **hecho** de metal. Answer: fact (Score: 0)

4. Provide the Spanish translation for these English words

Rice:	answer: el arroz (2)	roast meat:	answer: carne rosado (0)
Cauliflower:	answer: (blank) (0)	onion:	answer: cebolla (2)

Total score: 4/8

Procedures for Data Analysis

All peer interaction consisting of a total of 10 hours and 22 minutes of talk, were transcribed verbatim, and the total number of words, turns, and LREs were tallied. A quantitative analysis of each of the tasks will be completed to better understand what occurs when students interact using AR vs non-AR, – what activity (oral vs writing) elicits more LREs, and the nature (lexical, grammatical, mechanical), outcome (solved

correctly, incorrectly, unresolved) and correction orientation (self-corrected, other-correction) of these LREs. A quantitative analysis of the posttests will be used to analyze more deeply the outcomes of the LREs and whether they were recalled on the posttests, and with what accuracy.

The study's research design entails a comparison of the LRE number, types, and outcomes by modality (writing-focused vs oral) and task setting (AR vs non-AR). The results of the study are based on the results of regression analysis and non-parametric independent t-tests to measure any possible significant differences in the number, nature, outcomes, and correction orientation of LREs in the oral and writing-focused activities per setting.

Because this study explores a new type of setting for language learning, it seeks to establish whether the collaborative dialogue learners partake using AR outside of classroom is similar or different to using traditional computer applications in a laboratory setting. Moreover, studies have shown that task modality also plays an important role in peer interaction and language development (Adams & Ross-Feldman, 2008; Williams, 2001, 2008). Thus, Research Question 1 asks:

> **RQ#1:** Does the task setting (AR vs non-AR) differentially affect the number, nature, outcome, and correction orientation of target LREs produced in the two modalities (oral vs writing-focused)?

The frequency distribution function of SPSS will be used to find the number and percentage of LRE types and subtypes by setting (AR vs non-AR) and modality (oral and writing focus). The descriptive statistics will show the mean and standard deviations of

these variables per setting and modality. A comparison of the relative incidence of each type (lexical, grammatical, mechanical) and subtype (e.g., noun, verb, etc.) of LREs across the two task settings and modalities will be analyzed. Multinomial regressions will compare LRE types and subtypes in both settings (AR vs non-AR) and modalities (oral vs writing focus) to test for statistical significance. Regression analyses will be performed to find any significant effect of setting on the nature, outcome, and correction orientation of LREs in the oral vs writing-focused activities.

As other studies have shown, learners seem to retain what they learn from peers and self-correction during an LRE. Swain & Lapkin's (1998) study for example, showed that learners move from an incorrect response (pretest) to a correct response (posttest) which suggests that the outcomes of LREs and the posttest scores were positively related. Williams (2001) provided evidence that there is a strong connection between attention to form and subsequent use of those forms. Thus, Research Question 2 asks:

RQ#2: Does the setting and LRE outcome significantly affect the ability of participants to recognize and produce the correct lexical and grammatical forms on the posttests? Is there a significant correlation between students' ability to recognize the correctness of a grammatical form and their ability to produce that form correctly on the posttests?

A quantitative analysis will be performed to compare the number of potentially remembered words for each participant based on the outcomes of correctly solved, incorrectly solved and unresolved LREs (see explanation under Posttest section above). As mentioned before, Question 1 of the posttests examined recognition and production

scores of grammatical forms, questions 2 to 4 examined production of lexical items. Only lexical and grammatical LREs will be examined. Posttest scores will be calculated for items on the posttests that were based on actual LRE outcomes in the peer dialogues. Non-parametric independent *t-tests* will determine if the AR or non-AR settings influence the ability of the participants to remember the LREs.

CHAPTER 4

RESULTS

In this section the quantitative findings of the study are presented in terms of the two research questions investigated. The data were analyzed using the statistical package SPSS 26.0 for Windows. This chapter begins with the data results of the pre-questionnaire followed by the results of the experimental tasks answering the two research questions.

Pre-Questionnaire

The results of the pre-questionnaire are provided in Tables 3 through 7. There was a total of twenty participants, 10 females and 10 males, with ages ranging from 18 to 22 with a mean of 19 years old. Table 4 provides information on the language background of the participants.

Most of the participants were NS of English who have Spanish as their second language. Nineteen participants had English as their native language and one female student had Chinese as her L1. One male student was a heritage learner of Arabic. Fourteen participants reported Spanish as their L2 and had no knowledge of other foreign languages. One female student reported German as her L2. Five participants reported some knowledge of an L3 (Italian, Japanese, Urdu, Portuguese, and Mandarin). Seventeen participants were learning Spanish only and had no other foreign language classes, and three participants were also taking classes in English, German, and Portuguese, respectively.

Table 4

Participants Language Background (N = 20)

Student*	L1	L2	L3	FL courses besides Spanish
1 Pat	Chinese	English	Spanish	English
2 Jenn	English/Arabic	Spanish		
3 Mark	English	Spanish	Italian	
4 Betty	English	Spanish		
5 Karen	English	Spanish		
6 Lisa	English	Cantonese	Spanish	
7 Bill	English	Spanish	Urdu	
8 Richie	English	German	Spanish	German
9 Matt	English	Spanish	Portuguese	Portuguese
10 Ashley	English	Spanish		
11 Kim	English	Spanish		
12 Paul	English	Spanish		
13 Anna	English	Spanish		
14 Nick	English	Spanish		
15vRichie	English	Spanish		
16 Charles	English	Spanish		
17 Melissa	English	Spanish		
18 Amy	English	Spanish		
19 Jim	English	Spanish		
20John	English	Spanish		

*Pseudonyms indicated

Table 5 provides data on the participants' academic majors.

Table 5

Participants Majors

Majors	Totals	Percentage
Arts & Sciences (Sec. Math Ed., Film, Justice Studies & Psychology, Sustainability)	10	50%
Engineering	6	30%
Spanish (English & Spanish, Spanish & Education, German & Spanish)	3	15%
Spanish only	1	5%

Most participants were majoring in areas other than Spanish. Ten participants were arts and sciences major other than Spanish (10/20 [50%]) and 6/20 (30%) participants were majoring in engineering. Three participants had Spanish as part of a double major (3/20 [15%]), with a second major in English, German, or Education. Only one participant (1/20 [5%]) had Spanish as their only major.

Table 6 provides data on participants' reasons for taking Spanish courses.

Table 6

Participant's Reasons for Taking Spanish

Reasons for taking Spanish	Totals	Percentage
Employment/Work-related	14	70%
Tourism/Vacation	13	65%
Degree Requirement	5	25%
Communicate better with family and friends	5	25%
Personal Interest	4	20%

Most participants were taking Spanish for employment/work-related reasons (14/20 [70%]) and/or for use in tourism/vacation activities (13/20 [65%]). Five participants (5/20 [25%]) wanted to satisfy a degree requirement, five (5/20 [25%]) would use it to communicate better with family and friends, and four (4/20 [20%]) were taking Spanish for personal interest. Many of these participants also indicated multiple reasons for taking Spanish classes. For instance, participants wanted to use Spanish for work-related reasons and tourism (4/20 [20%]), for work-related reasons and to satisfy a degree requirement (1/20 [5%]), for work-related reasons, tourism and to satisfy a degree requirement (3/20 [15%]), for work-related reasons, tourism, and to communicate better

with family members and friends (3/20 [15%]), for tourism and personal interest (2/20 [10%]), and tourism and to communicate with family members and friends (1/20 [5%]).

For the technology questions, participants were asked to report what cellular phone they had in order to get to know if at least one student per dyad had a mobile phone that supported AR technology (e.g., iPhone version 6 and above, Android version 8 and above) or if the researcher needed to provide any devices (e.g., iPads). Of the 20 participants, seventeen reported having an iPhone 6 or above, two participants reported having an Android 8 or above, and one student reported having an Android 7 or below. The participants with Android phones were paired with participants who had an iPhone 6 or above (both apps used in the treatment work in IOS systems only).

It was important to know participants' experiences with AR technology before they were introduced to the treatment. Sixteen of 20 participants (80%) reported having used AR technologies such as the *Pokémon Go* app (10), *Snapchat* filters (9), games (3), QR codes (4), *Google Maps AR* (2), and *Google Translate AR* (1). Four participants (20%) reported never having used any AR technology, either because 1) "there are not enough choices for apps and games," 2) "don't know how to use it," 3) "never heard of them," or 4) "no purpose of using this type of technology." Before completing the pre-questionnaire, the researcher briefly explained the technology and described some examples of AR technology that were familiar to the participants, such as the *Pokémon Go* game and *Snapchat* filters.

The next question asked the participants' opinions about how to best use this technology in the language classroom. Table 7 below show the measurable results listing

the percentage scores for this question, based on a five-point Likert scale, with 5 reflecting the highest level of agreement and 2 reflecting the lowest level of agreement. A score of 1 indicated no opinion.

Table 7 shows the measurable results listing the percentage for this question.

Table 7

The Use of AR in the Language Classroom - Percentage Distribution (N = 20)

Question 10: I would use AR in my language class for:	Strongly agree (5)/ Agree (4)	Disagree (3)/ Strongly disagree (2)	No Opinion (1)
Learn about history, culture, artifacts of the target country	70%	15%	15%
Learn vocabulary	65%	20%	15%
Language training	60%	25%	15%
Interact with virtual objects of the target culture	60%	30%	10%
Interact with the Hispanic community	55%	25%	20%

Table 7 shows that most participants would use AR technology to learn about history, culture, artifacts of the target country, followed by learning vocabulary, language training, interacting with virtual objects of the target culture, and lastly, interacting with the Hispanic community. In a Likert-scale from 5 (strongly agree) to 1 (no opinion), 70% (14/20) of the participants strongly agreed (8) or agreed (6) that AR could be used to learn vocabulary, and 15% (3/20) strongly disagreed (1) or disagreed (2), while 15% (3/20) had no opinion (3). Sixty five percent (13/20) of the participants strongly agreed (10) or agreed (3) that AR would be best used to learn about history, culture, and artifacts

of the target country, and 20% (4/20) disagreed (4) while 15% (3/20) had no opinion (3). Sixty percent (12/20) of the participants strongly agreed (9) or agreed (3) that AR technology could be best used to support language training, 25% (5/20) disagreed (4) or strongly disagreed (1), while 15% (3/20) of the participants (3) had no opinion. Sixty percent (12/20) of the participants strongly agreed (7) or agreed (5) that AR would be best used to interact with virtual objects from the target culture, and 30% (6/20) disagreed (4) or strongly disagreed (2), while 10% (2/20) had no opinion (2). Fifty five percent (11/20) of the participants strongly agreed (7) or agreed (4) that AR would be best used to interact with the Hispanic community via an AR game, and 25% (5/20) disagreed (4) or strongly disagreed (1), while 20% (4/20) had no opinion (4).

Table 8 shows the results for the survey question about how participants would use AR technology.

Table 8

The Use of AR - Percentage Distribution (N = 20)

Question 11: How would you use AR technology:	Totals	Percentage
As a requirement (homework/assignments)	14	70%
On my own time to improve language skills	8	40%
On my own time just for fun	7	35%
Other	0	0%

Seventy percent (14/20) participants reported that they would use AR if it were a requirement for the course, 40% **(**8/20) to improve language skills on their own time, and

35% (7/20) indicated they would use AR for fun. Many participants also indicated that they had multiple reasons to use AR: 10% (2/20) as a requirement and for fun, 10 % (2/20) to improve language skills and for fun 10% (2/20), 10% (2/20) as a requirement, to improve language skills, and for fun.

Seventeen participants had no comments about AR technology (open-ended question), but three participants said: "I think actual immersion in a culture is more beneficial and legitimate than virtual reality." (Kim); "I'm not that kind of woman who is interested in technology." (Amy); "I really like the idea of using AR for improving language and I think it would be beneficial." (Betty)

The results from Questions 10 and 11 about AR use show that overall, before the treatment started, participants were willing to use AR for language learning.

Experimental Tasks

One LRE can have several turns. If an LRE was composed of several turns during a single conversation, that LRE type and outcome was coded only once. In the example below, Paul and Kim talk about the correct way of using the reflexive verb *divertirse*. There are eight turns in total, one LRE (in bold), and one outcome (solved correctly):

Paul: *…vamos, vamos al Dave y Busters para****… divertir… divertirnos****.*
[…(we're) going, going to Dave & Busters to… have fun… have fun (ourselves).]

Kim: *para… ¿tenemos divertido?*
[To… (we) have fun (participle)?]

Paul: *no, divertirnos, divertir es el verbo.*
[no, to have fun (ourselves), have fun is the verb.]

Kim: *divertimos…*
[(we) have fun.]

Paul: *No, divertirnos.*
[No, (we) have fun (ourselves).]

Kim: **Divertirmos* [sic]. *Ok, lo siento.*
[(not a word in Spanish). Ok, I'm sorry.]

Paul: *Es sin "mos". Pienso que… porque después de "para" es infinitvo y… pienso que porque es "m" es "n".*
[it's without "mos". I think that… because after "to" is infinitive and… I think that why is "m" is "n".

Kim: *Sí, es "n".*
[Yes, it's "n".]

Table 9 displays the order of tasks and settings per dyad. Dyads who used AR on Task 1 were part of the non-AR group for Task 2 and vice-versa.

Table 9

Counterbalancing of Tasks per Setting and Dyad

	Task 1	Task 2 (dyad #)
AR	Pat/Jenn (1)	Matt/Ashley (5)
	Mark/Betty (2)	Paul/Kim (6)
	Karen/Lisa (3)	Anna/Nick (7)
	Rob/Bill (4)	Richie/Charles (8)
	Jim/John* (10)	Melissa/Amy (9)
Non-AR	Matt/Ashley (5)	Pat/Jenn (1)
	Paul/Kim (6)	Mark/Betty* (2)
	Anna/Nick (7)	Karen/Lisa* (3)
	Richie/Charles (8)	Rob/Bill (4)
	Melissa/Amy (9)	Jim/John (10)

Note: Numbers in parentheses are dyads numbers. *Jim and John did not record Task 1, Karen and Lisa did not record the written part of Task 2, and Mark and Betty did not record the oral part of Task 2.

Descriptive Data

Before addressing the research questions in detail, an overview of the data is provided. Table 10 shows a general picture of the data for the total number of words, and LREs produced in the collaborative dialogue of 10 dyads.

Table 10

Total Frequency of Words, LREs, and Turns

	Total	Mean	*SD*
Words	27,780	7.78	6.71
LREs	695	.20	.38
Turns	3,571		

Note: Mean was calculated by dividing the number of words per turn and LREs per turn.

There was a total of 27,780 words in the data produced across all 10 dyads, combining tasks, settings, and modality. Repetitions (e.g., "más… más opciones") and self-corrected words (e.g., "Tuvo, tuve una clase…"), were included in the number of words; hesitations (e.g., ah, uhm, hmm) and exclamations (e.g., ¡Oh! Ah!) were removed from the data. The average number of words was 7.78 per turn (*SD* = 6.71), with a range of 74 to 1. There were 695 instances of LREs with an average number of .20 (*SD* = .38) per turn. There was approximately one (695/3,571 [.20]) LRE in every 5 turns, with a range of 4 to 0 per turn.

Table 11 shows the overall total number of LREs by type, outcome, and correction orientation. As mentioned in Chapter 2, an LRE can be of a lexical, grammatical, or mechanical nature. Participants produce a lexical LRE when they question the meaning of a word. Grammatical LREs include examples in which

participants may question the syntax or morphology of a target word or construction, thus the focus is on grammar. Mechanical LREs emphasize the pronunciation or, the spelling or punctuation of a written target word.

Table 11

Total number of LREs by Type, Outcome, and Correction Orientation (N = 695)

	LREs	Sum	Percentage
Type	Lexical	364	52.37
	Grammatical	248	35.68
	Mechanical	83	11.94
Outcome	Solved correctly	480	69.06
	Solved incorrectly	133	19.13
	Unresolved	82	11.80
Correction Orientation	Other correction	353	50.79
	Self- correction	342	49.21

In this study, the overall data shows that more than half, 52.37% (364/695) of the LREs were lexical in nature, 35.68% (248/695) were grammatical, and 11.94% (83/695) were mechanical. Sixty-nine per cent (480/695) of LREs were solved correctly, 19.13% (133/695) were solved incorrectly, and 11.80% (82/695) were unresolved LREs. Regarding the orientation of the LRE correction, 50.79% (353/695) were corrected by the peer and 49.21% (342/695) were self-corrected.

Tables 12 and 13 show the total number of words, and LREs during the oral (Table 12) and writing focus (Table 13) activities per setting (AR vs non-AR).

Table 12 shows that during the oral activities the participants using AR produced a total of 7,824 (47.38%) words out of 16,512 total words with a mean of 7.49 (*SD* =

7.32) per turn, whereas in the non-AR setting they produced 8,688 words (52.62% of the total number of words) with a mean of 8.59 per turn (*SD* = 8.67). Out of a total of 254 LREs, 130 (51.18%) were produced in the AR setting with a mean of .12 LREs (*SD* = .36) per turn, whereas 124 (48.82%) LREs were produced in the non-AR setting with the same mean of .12 (*SD* = .36) per turn.

Table 12

Total Frequency of Words and LREs per Setting and Oral Modality

	AR				Non-AR			
	n	%	Mean	SD	n	%	Mean	SD
Words	7,824	47.38	7.49	7.32	8,688	52.62	8.59	8.67
LREs	130	51.18	.12	.36	124	48.82	.12	.36
Turns	1045	50.83			1,011	49.17		

Note: Percentage calculated considering the total number of words and LREs in the oral modality in both AR and non-AR settings. Mean calculated per turn.

Table 13 displays the total number of words and LREs during the writing focus activities per setting (AR vs non-AR).

Table 13

Total Frequency of Words and LREs per Setting and Writing Focus Modality

	AR				Non-AR			
	n	%	Mean	SD	n	%	Mean	SD
Words	4,930	43.75	7.91	7.90	6,338	56.25	7.11	7.68
LREs	210	47.34	.34	.59	231	52.66	.26	.51
Turns	623	41.12			892	58.88		

Note: Percentage calculated considering the total number of words and LREs in the writing-focused modality in both settings. Mean calculated per turn.

During the writing activities, the dyads who used AR produced a total of 4,930 (43.75%) words out of 11,268 total words with a mean of 7.91 words (*SD* = 7.90) per turn, whereas when they used the non-AR applications, they produced 6,338 (56.25%) total words with a mean of 7.11 words (*SD* = 7.68) per turn. Out of 441 LREs, 210 (47.62%) were produced by the AR groups with a mean of .34 LREs (*SD* = .59) per turn, whereas 231 (52.38%) LREs were produced by the non-AR groups with a mean of .26 LREs per turn (*SD* = .51).

To control for the differences in the quantity of discourse produced across participants and the number of tasks completed per dyad (see Table 9), the number of LREs was divided by the number of words per dyad (9) and multiplied by 100. This way the distribution of LREs was standardized across all dyads (Yilmaz & Granema, 2010). This new variable (LREs/100words) was used to analyze the distribution of LREs per setting and modality.

Table 14 displays the number of LREs, number of words, and the number of LREs per 100 words for the oral activities. As seen in Table 14, in the oral focus activities participants had more LREs in the AR setting (130) than in the non-AR (124) setting. This trend is also seen in the average number of LREs, for instance, the average of LREs during the oral activities was slightly higher when participants used AR (14.44 with a mean of 1.58 LREs per 100 words) than the non-AR applications (13.78 with a mean of 1.47 LRES per 100 words).

Table 14

Ratio of LREs per Dyad during Oral Activities

# Dyad	Oral AR			Oral Non-AR			Average
	# Words	LRE Total	LRE/ 100words	# of Words	LRE Total	LRE/ 100words	LRE/ 100words
1 Pat & Jenn	1025	7	.68	1931	24	1.24	.96
2 Mark & Betty	2709	59	2.18	0	0	0	2.18
3 Karen & Lisa	672	5	.74	759	10	1.32	1.03
4 Rob & Bill	1018	10	.98	891	9	1.01	1.00
5 Matt & Ashley	445	8	1.80	542	9	1.66	1.73
6 Paul & Kim	860	22	2.56	826	25	3.03	2.79
7 Anna & Nick	352	8	2.27	804	9	1.12	1.70
8 Rich & Charles	191	3	1.57	471	2	.43	1.00
9 Melissa & Amy	552	8	1.45	859	22	2.56	2.01
10 Jim & John	0	0	0	1605	14	.87	.87
Total	7,824	130	14.24	8,688	124	13.24	13.74
Average	869.33	14.44	1.58	965.33	13.78	1.47	1.53

Table 15 displays the number of LREs, number of words, and the number of LREs per 100 words in the writing activities. As Table 15 illustrates, during the writing focus activities participants had more LREs in the non-AR setting (231) than the AR setting (210). The total number of LREs per 100 words (38.04) and the mean (4.23) in the AR setting were higher than the total number of LREs per 100 words (31.26) and the mean (3.47) in the non-AR setting. That is when looking at the standardized number of LREs, the participants using AR produced more LREs than when they used the non-AR setting.

Table 15

Ratio of LREs per Dyad during Writing Focus Activities

# Dyad	Writing AR			Writing Non-AR			Average
	# of Words	LRE Total	LRE/ 100words	# of Words	LRE Total	LRE/ 100words	LREs/ 100words
1 Pat & Jenn	786	29	3.69	1012	31	3.06	3.38
2 Mark & Betty	604	23	3.81	588	13	2.21	3.01
3 Karen & Lisa	779	24	3.08	0	0	0	3.08
4 Rob & Bill	461	9	1.95	581	12	2.07	2.01
5 Matt & Ashley	480	16	3.33	512	17	3.32	3.33
6 Paul & Kim	581	39	6.71	952	64	6.72	6.72
7 Anna & Nick	265	14	5.28	380	14	3.68	4.48
8 Rich & Charles	335	10	2.99	776	33	4.25	3.62
9 Melissa & Amy	639	46	7.20	805	38	4.72	5.96
10 Jim & John	0	0	0	732	9	1.23	1.23
Total	4,930	210	38.04	6,338	231	31.26	34.65
Average	547.78	23.33	4.23	704.22	25.67	3.47	3.85

In sum, the standardized distribution of LREs (LRE/100words) across all dyads in both settings and modalities shows that participants when using AR applications produced slightly more LREs than when using non-AR applications during the oral and writing-focused activities. Research question #1 will answer if this difference is significant or not.

Research Questions

This study presented two research questions addressing the setting (AR vs. non-AR) and modality (oral vs. writing-focused) and their effect on LREs during peer collaborative dialogue. To rule out any task type effect, both tasks were decision making tasks and were considered equivalent in nature. A non-parametric independent samples *t-test* showed no significant difference in the number of LREs in each task when comparing them in the same setting and modality (e.g., Task 1 AR oral vs Task 2 AR oral). Table 16 shows the summary of the results of the Mann-Whitney U test.

Table 16

Summary of LRE Mean Differences between Tasks in each Setting and Modality

Modality Focus	Setting	Task 1 Mean Rank	Task 2 Mean Rank	z-value	*p* value
Oral	Non-AR	62.13	58.64	-.55	.58
	AR	62.98	69.67	-.98	.32
Writing	Non-AR	115.02	113.16	-.19	.85
	AR	103.31	102.77	-.06	.95

Note: $p < .05$

The findings related to each of the research questions are presented below.

RQ#1: Does the task setting (AR vs non-AR) differentially affect the (a) number, (b) type, (c) outcome, and (d) correction orientation of target LREs produced in the two modalities (oral vs writing)?

LRE numbers. Table 17 displays the distribution of LRE/100 words by setting and modality focus.

Table 17

Distribution of Standardized LREs per Setting and Modality

	AR				Non-AR				Total
	n	%	Mean	SD	n	%	Mean	SD	
Writing	210	6176	.18	.06	231	65.07	.13	.04	441
(per setting)		*47.62*				*52.38*			
Oral	130	38.24	.11	.10	124	34.93	.11	.04	254
(per setting)		*51.18*				*48.82*			
Total	340	100	.15	.09	355	100	.12	.04	695

Note: The vertical (modality) percentages are in normal type. The horizontal (setting) percentages are in italics. Means and standard deviations were calculated using the standardized LRE variable (LRE/100 words).

When comparing the settings in each modality, both settings produced similar LRE/100 words numbers. In the writing focus modality, the AR groups produced 210/441 (47.62%) with a mean of .18 (*SD* = .06) per turn, and the non-AR 231/441 (52.38%) LREs with a mean of .13 (*SD* = .04) per turn. Although the percentage difference of 4.76% favors the non-AR setting, the mean number of LREs/100 words per turn (.18) confirms that the AR groups produced approximately 1/3 more LREs than the non-AR groups (.13). In contrast, in the oral modality, the AR groups produced 130/254 (51.18%) LREs/100 words with a mean of .11 (*SD* = .10) per turn, whereas the non-AR groups produced 124/254 (48.82%) LREs/100 words with the same mean of .11 (*SD* = .04), a difference of 2.32% favoring the AR setting for LRE production in the oral modality.

A generalized linear model (GZLM) with a log link function was conducted to investigate the effect of setting and modality on the numbers of LRE/100 words in the oral and writing activities. GZLM allows for non-normal distributions, non-equal

variances, and no multicollinearity. The predictor variables were tested a priori to verify there was no violation of the assumption of no multicollinearity. Setting and modality (categorical dichotomous variables) were entered as fixed effects and as interaction terms into the model. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. The variables, setting and modality, were found to contribute to the model. This final model significantly predicted the LRE/100 words numbers over and above the intercept-only model, x^2 (3) = 143.72, $p < .001$.

Table 18 displays the results of the GZLM regression predicting likelihood of LRE/100 words based on setting and modality focus.

Overall, the setting had a statistically significant effect on the prediction of the LRE/100 words, Wald x^2 (1) = 58.58, $p < .001$. The number of standardized LREs in the non-AR setting was .96, 95% CI [.94, .97] times less than the number of standardized LREs in the AR setting, a statistically significant effect, x^2 (1) = 24.11, $p < .001$. Likewise, modality also had a statistically significant effect on the prediction of LRE/100 words, Wald x^2 (1) = 104.69, $p < .001$. The standardized LRE numbers in the oral modality were .93, 95% CI [.92, .94] times less than the number of standardized LREs in the writing focus modality, a statistically significant effect, Wald x^2 (1) = 102.79, $p < .001$.

Table 18

Linear Regression Predicting Likelihood of Standardized LRE Numbers based on Setting and Modality

	B	SE	Wald	*df*	*p*	Odds Ratio	95% CI for Odds Ratio	
							Lower	Upper
Oral	-.07	.007	104.69	1	.000	.93	.92	.94
Non-AR	-.46	.006	58.58	1	.000	.96	.94	.97
Setting * Modality	.04	.010	18.92	1	.000	1.04	1.02	1.06
Intercept	.181	.004	1750.29	1	.000	1.20	1.19	1.21

Note: Non-AR setting compared to AR setting, oral modality compared to writing focus modality, intercept is LRE/100 words.

When comparing the interaction between setting and modality, the odds of an interaction between setting and modality having an effect on the distribution of LREs was 1.04, 95% CI [.92, .94] times more likely than if there was no interaction, a statistically significant effect, Wald x^2 (1) = 18.92, $p < .001$. In sum, setting and modality and their interaction had a significant effect on the number of LREs.

Type/Nature of LREs. Table 19 shows the distribution of LREs by types in the two settings and modalities. For these calculations, the total number of LREs is used instead of the standardized number (LREs/100 words).

As Table 19 illustrates, overall, the oral activities produced more lexical LREs (58.27% [148/254]) than grammatical (37.80 % [96/254]) LREs. Similarly, the writing-focused activities produced more lexical (48.98% [216/441) LREs than grammatical (34.47% [152/441]) LREs.

Table 19

Distribution of LRE Types by Setting and Modality Focus

	Oral					Writing				
	AR		Non-AR		Total	AR		Non-AR		Total
	n	%	n	%	n	n	%	N	%	n
Lexical	83	63.85	65	52.42	148	100	47.62	116	50.22	216
(per setting)		*56.08*		*43.92*			*46.30*		*53.70*	
Grammatical	46	35.38	50	40.32	96	78	37.14	74	32.03	152
(per setting)		*47.92*		*52.08*			*51.32*		*48.68*	
Mechanical	1	.77	9	7.26	10	32	15.24	41	17.75	73
(per setting)		*10.00*		*90.00*			*43.84*		*56.16*	
Total	130	100	124	100	254	210	100	231	100	441
(per setting)		*51.18*		*48.82*			*47.62*		*52.38*	

Note: The vertical percentages are in normal type. The horizontal percentages are in italics.

Looking at each modality, the oral activities using AR had more lexical LREs (83/130 [63.85%) than grammatical (46/130 [35.38%]) and mechanical 1/130 [.77%]) LREs. Similarly, the non-AR dyads produced more lexical LREs (65/124 [52.42%] than grammatical (50/124 [40.32%]) and mechanical (9/124 [7.26%]) LREs. The LRE types had the same pattern during writing focus activities. That is, lexical LREs (100/210 [47.62%] in the AR and 116/231 [50.22%] in the non-AR) were more frequent than grammatical (78/210 [37.14%] in the AR and 74/231 [32.03%] in the non-AR) and mechanical (32/210 [15.24%] in the AR and 41/231 [17.75%] in the non-AR setting) LREs.

Looking at settings, overall, both settings produced more lexical LREs (53.82% [183/340] in the AR and 50.99% [181/355] in the non-AR) than grammatical (36.47%

[124/340] in the AR and 34.93% [124/355] in the non-AR) LREs. Looking at each modality per setting, out of a total of 148 lexical LREs initiated during the oral activities, 83/148 (56.08%) were produced in the AR setting, whereas the non-AR setting had 65/148 (43.92%) lexical LREs. From a total of 96 grammatical LREs, the non-AR setting contained 50/96 (52.08%), whereas the AR setting contained 46/96 (47.92%). The least frequent LREs, although with the largest difference, were mechanical (pronunciation) LREs, 1/10 (10%) in the AR setting, and 9/10 (90%) in the non-AR setting. In contrast to the trends found in the oral data, during the writing focus activities dyads produced more Lexical LREs in the non-AR setting (116/216 [53.70%]) than the AR setting (100/216 [46.30%]). In addition, unlike the oral data patterns, grammatical LRES were more frequent in the AR setting (124/248 [51.32%]) than the non-AR setting (74/152 [48.68%]). As during oral activities, there were more mechanical LRES during the writing focus activities in the non-AR setting (41/73 [56.16%]) than the AR setting (32/73 [43.84%]).

A multinomial regression was conducted to investigate the effect of setting and modality on the nature of LREs in the oral and writing activities. The full model with the predictor variables (setting and modality) were found to contribute to the model, x^2 (4) = 31.384, $p < .001$. The findings showed that modality (oral vs. writing focus) had a statistically significant effect on the nature of LREs when comparing lexical and grammatical LREs to mechanical LREs. However, the results of the analysis revealed no statistically significant effect of setting (AR vs. non-AR) on the nature of LREs, x^2 (2) = 2.73, $p = .26$. Table 20 shows the results of the multinomial regression.

Table 20

Multinomial Regression Predicting Likelihood of LRE Types based on Setting and Modality

		B	SE	Wald	*df*	*p*	Odds Ratio	95% CI for Odds Ratio	
								Lower	Upper
Lexical	Non-AR	-.41	.25	2.59	1	.11	.67	.41	1.10
	Oral	1.61	.35	20.55	1	.000	4.98	2.49	9.96
	Intercept	1.31	.20	42.57	1	.000			
Grammatical	Non-AR	-.38	.26	2.15	1	.14	.68	.41	1.14
	Oral	1.51	.36	17.49	1	.000	4.54	2.23	9.27
	Intercept	.95	.21	21.04	1	.000			

Note: Lexical and grammatical compared to mechanical LREs (intercept). Non-AR setting compared to AR setting; oral modality compared to writing focus modality.

As expected, the odds of having lexical rather than mechanical LREs was 4.98, 95% CI [2.49, 9.96] times higher in the oral activities than in the writing activities. Likewise, the odds of having grammatical LREs rather than mechanical LREs was 4.54, 95% CI [2.23, 9.23] times higher in the oral activities than in the writing activities. Both effects were statistically significant, Wald x^2 (1) = 20.55, $p < .001$ and x^2 (1) = 17.49, $p < .001$, respectively. Not surprisingly, the odds of having mechanical LREs rather than grammatical or lexical in the writing activities was also statistically significant .22, 95% CI [.11, .45], .20, 95% CI [.10, .40] times higher respectively than in the oral activities. However, no significant difference was found between lexical and grammatical LREs in either setting (Wald x^2 (1) = .02, $p = .89$) nor modality (x^2 (1) = .30, $p = .58$).

Lexical LREs sub-types. Table 21 displays the distribution of Lexical LRE sub-types that participants focused on during the collaborative dialogue (e.g., asking about the meaning of an L2 word or how to say a particular L2 word that falls under a specific part of speech [noun, preposition, adverb] in the target language) in both settings and modalities.

Table 21

Distribution of Lexical LRE Sub-types per Setting and Modality

	Oral					Writing				
	AR		Non-AR		Totals	AR		Non-AR		Totals
	n	%	n	%		n	%	n	%	
Noun	36	43.37	41	63.08	77	39	39.00	39	33.62	78
(per setting)		*46.75*		*53.25*			*50.00*		*50.00*	
Verb	16	19.28	11	16.92	27	20	20.00	24	20.69	44
(per setting)		*59.26*		*40.74*			*45.45*		*54.55*	
Adjective	9	10.84	6	9.23	15	13	13.00	10	8.63	23
(per setting)		*60.00*		*40.00*			*56.52*		*43.48*	
Preposition	1	1.21	4	6.15	5	13	13.00	11	9.48	24
(per setting)		*20.00*		*80.00*			*54.17*		*45.83*	
Pronoun	7	8.43	0	0	7	12	12.00	20	17.24	32
(per setting)		*100.00*		*0.00*			*37.50*		*62.50*	
Article	6	7.23	3	4.62	9	2	2.00	6	5.17	8
(per setting)		*66.67*		*33.33*			*25.00*		*75.00*	
Adverb	5	6.02	0	0	5	1	1.00	2	1.72	3
(per setting)		*100.00*		*0.00*			*33.33*		*66.67*	
Conjunction	2	2.41	0	0	2	0	0.00	4	3.45	4
(per setting)		*100.00*		*0.00*			*0.00*		*100.00*	
Other	1	1.21	0	0	1	0	0.00	0	0	0
(per setting)		*100.00*		*0.00*			*0*		*0.00*	
Total	83	100.00	65	100.0	148	100	100.0	116	100.0	216
(per setting)		*56.08*		*43.92*			*46.30*		*53.70*	

Note: The vertical percentages are in normal type. The horizontal percentages are in italics.

For these calculations, the total number of Lexical LREs is used instead of the standardized number (LREs/100 words).

As Table 21 illustrates, during the oral AR activities, dyads produced most lexical LREs that dealt with word choice or meaning of nouns (36/83 [43.37%], followed by verb choice (16/83 [19.28%], the use of adjective 9/83 (10.84%), pronouns (7/83 [8.43%]), articles (6/83 [7.23%]), adverbs (5/83 [6.02%]), conjunctions (2/83 [2.41%]) and prepositions (1/83 [1.20%]). In contrast, dyads who used non-AR applications during the oral activities had no instances of pronouns, adverbs, nor conjunctions LREs. More than half (41/65 [63.08%]) of the total LREs were noun issues, followed by verb choice (11/65 [16.92%]), the use of adjectives (6/65 [9.23%]), prepositions (4/65 [6.15%]), and articles (3/65 [4.62%]).

Similarly most of the lexical LREs produced orally during the writing focus activities in both settings dealt with nouns (39/100 [39%] in the AR setting and 39/116 [33.62 %] in the non-AR setting), followed by verb choice (20/100 [20%] in the AR setting and 24/116 [20.69%] in the non-AR setting), prepositions (13/100 [13%] in the AR setting and 11/116 [9.48%] in the non-AR setting), adjectives (13/100 [13%] in the AR setting and 10/116 [8.62%] in the non-AR setting), pronouns (12/100 [12%] in the AR setting and 20/116 [17.24%] in the non-AR setting), articles (2/100 [2.0%] in the AR setting and 6/116 [5.17%] in the non-AR setting), and adverb (1/100 [1%] in the AR setting and 2/116 [1.72%] in the non-AR setting). There were four instances of LREs related to conjunctions by the non-AR group (4/116 [3.45%]) whereas the AR group had none.

As illustrated on Table 21, the number of lexical LREs subtypes during the oral activities were more frequent in the AR setting than the non-AR setting, except for LREs

related to nouns and prepositions which were more frequent in the non-AR setting (41/77 [53.25%] and 4/5 [80%] respectively) than AR (36/77 [46.75%] and 1/5 [20%] respectively). In contrast, in the writing activities, LREs related to nouns had the same frequency in both settings (39/78 [50%]). Unlike the oral activities, all LRE sub-types during the writing activities were more frequent in the non-AR setting than the AR setting, except for LREs related to adjectives and prepositions which were more frequent in the AR setting (13/23 [56.52%] and 13/24 [54.17%] respectively) than the non-AR setting (10/23 [43.48%] and 11/24 [45.83%] respectively).

A multinomial regression revealed that setting did not significantly affect the sub-types of LREs produced. However, modality had a significant effect on nouns and pronouns. The odds of having nouns compared to no nouns in the oral modality was 2.07, 95% CI [1.40, 3.05] times higher than in the writing focus modality. In contrast, the odds of having pronouns compared to no pronouns in the oral modality was .40, 95% CI [.16, .98] less likely than in the writing focus modality.

Grammatical LREs sub-types. Table 22 displays the distribution of grammatical LRE sub-types that participants focused on during the collaborative dialogue (e.g., the use of subject-verb agreement, the gender of a word, etc. in the target language) in both settings and modalities. For these calculations, the total number of Grammatical LREs is used instead of the standardized number (LREs/100 words).

Table 22

Distribution of Grammatical LRE Sub-types per Setting and Modality Focus

	Oral					Writing				
	AR		Non-AR		Total	AR		Non-AR		Total
	n	%	n	%		N	%	n	%	
Gender	17	36.96	18	36.00	35	28	35.90	27	36.49	55
(per setting)		*48.57*		*51.43*			*50.91*		*49.09*	
SVA*	10	21.74	19	38.00	29	17	21.79	11	14.86	28
(per setting)		*34.48*		*65.52*			*60.71*		*39.29*	
Verb tense	7	15.22	6	12.00	13	16	20.51	22	29.73	38
(per setting)		*53.85*		*46.15*			*42.11*		*57.89*	
Number	6	13.04	1	2.00	7	4	5.13	3	4.05	7
(per setting)		*85.71*		*14.29*			*57.14*		*42.86*	
Mood	3	6.52	3	6.00	6	8	10.26	9	12.16	17
(per setting)		*50.00*		*50.00*			*47.06*		*52.94*	
Aspect	2	4.35	3	6.00	5	3	3.85	1	1.35	4
(per setting)		*40.00*		*60.00*			*75.00*		*25.00*	
Other	1	2.17	0	0	1	2	2.56	0	0	2
(per setting)		*100.00*		*0*			*100.00*		*0*	
Total	46	100.00	50	100.0	96	78	100.0	74	100.0	152
(per setting)		*47.92*		*52.08*			*51.32*		*48.68*	

Note: *subject-verb agreement. The vertical percentages are in normal type. The horizontal percentages are in italics.

As illustrated on Table 22 over one third of grammatical LREs during the oral activities using AR were of the gender subtype (17/46 [36.96%], followed by SVA subtype (10/46 [21.74%]), verb tense LRES (7/46 [15.22%]), the use of number

agreement (6/46 [13.04%], mood (3/46 [6.52%]), and verb aspect (2/46 [4.35%] LREs. In contrast, in the non-AR setting, SVA LREs (19/50 [38%]) were more frequent than gender LRES (18/50 [36%]), followed by verb tense LREs (6/50 [12%]), mood LREs (3/50 [6.52%]), verb aspect LREs 3/50 [6%], and number agreement (1/50 [2%]).

Similarly, over one third of grammatical LREs during the writing focus activities using AR were of the gender subtype (28/78 [35.90%], followed by SVA subtype (17/78 [21.79%]), verb tense LRES (16/78 [20.51%]), LREs related to mood (8/78 [10.26%]), use of number agreement (4/78 [5.13%]), and verb aspect (3/78 [3.85%]. Similarly, in the non-AR setting, gender LREs were the most frequent (27/74 [36.49%]), followed by verb tense (22/38 [29.73%]), SVA LREs (11/28 [14.86%]), mood LRES (9/17 [12.16%]), number agreement (3/7 [4.05%]), and verb aspect LREs 1/4 [1.35%].

When looking at settings during the oral activities, gender LREs had a slightly more frequency in the AR setting (17/35 [48.57%] than the non-AR 18/35 [51.43%] setting. The incidences of SVA LREs in the non-AR setting was almost twice as much (19/29 [65.52%] the frequency in the AR (10/29 [34.48%]) setting. Verb tense LRES were slightly more frequent in the AR setting (7/13 [53.85%]) than the non-AR setting 6/13 [46.15%]). Also, number agreement LREs were three times more frequent in AR (6/7 [85.71%]) than in the non-AR setting (1/7 [28.57%]). Mood LREs had no difference in both settings (3/6 [50%]). Verb aspect LREs was the category with the fewer numbers in both settings (2/5 [40%] in the AR and 3/5 [60%] in the non-AR).

Similarly, during the writing focus activities half of the grammatical LREs were of the gender type (28/55 [50.91%] in the AR setting and 27/55 [49.09%] in the non-AR

setting). In contrast to the oral activities, participants produced more SVA LREs during the writing focus activities in the AR setting (17/28 [60.71%]) than the non-AR setting (11/28 [39.29%]). On the other hand, in contrast to the oral activities, verb tense LREs were more prevalent in the writing focused dialogue produced by the non-AR groups (22/38 [57.89%]) than the AR groups (16/38 [42.11%]). Also, like the oral data, number agreement LRE episodes were slightly higher in the AR (4/7 [57.14%]) than the non-AR (3/7 [42.86%]) groups. In contrast, unlike the oral data LREs of the subtype mood were more prevalent in the non-AR (9/17 [52.94%]) than in the AR (8/17 [47.06%] setting. Verb aspect LREs was the category with the lowest frequency in AR (3/4 [75%] and in non-AR (1/4 [25%]).

A multinomial regression revealed that setting did not significantly affect the types of grammatical LREs produced. However, modality had a significant effect on verb tense and verb aspect, that is, the oral modality was .21, 95% CI [.08, .59], .19, 95% CI [.08, .49] times less likely to present these sub-types respectively, than the writing focus modality.

Mechanical LREs. Table 23 shows the distribution of mechanical LREs in both settings and modalities. For these calculations, the total number of mechanical LREs is used instead of the standardized number (LREs/100 words).

Table 23 shows that no orthographical LREs were present during the oral activities. The AR groups produced only one (1/1 [100%]) pronunciation LRE during oral activities whereas the non-AR groups produced 9 (9/9 [100%] LREs related to pronunciation. During the writing-focus activities, the non-AR groups initiated more

orthographic LREs (39/41 [95.12%]) than the AR groups (29/32 [90.63%]). In contrast, the AR group initiated twice as much pronunciation LREs (3/32 [9.37%]) than the non-AR group (2/41 [4.88%]).

Table 23

Distribution of Mechanical LRE Sub-types per Setting and Modality Focus

	Oral					Writing				
	AR		Non-AR		Total	AR		Non-AR		Total
	n	%	n	%		n	%	n	%	
Orthographic	0	0	0	0	0	29	90.63	39	95.12	68
(per setting)		*0*		*0*			*42.65*		*57.35*	
Pronunciation	1	100.00	9	100.00	10	3	9.37	2	4.88	5
(per setting)		*10.00*		*90.00*			*60.00*		*40.00*	
Total	1	100.00	9	100.00	10	32	100.0	41	100.0	73
(*per setting*)		*10.00*		*90.00*			*43.84*		*56.16*	

Note: The vertical percentages are in normal type. The horizontal percentages are in italics.

Looking at the setting, pronunciation issues during the oral activities were more prevalent in the non-AR (9/10 [90%]) setting than the AR (1/10 [10%]) setting. In contrast, during the writing activities, pronunciation issues were more frequent in the AR groups (3/5 [60%]) than the non-AR groups (2/5 [40%]). During the writing activities, the non-AR groups (39/68 [57.35%]) had a higher incidence of orthographic LREs than the AR groups (29/68 [42.65%]).

A binomial logistic regression was performed to understand the effect of setting and modality on mechanical LREs (orthographic and pronunciation). The logistic regression model was statistically significant, x^2 (2) = 42.54, $p < .001$. The model

explained 66% (Nagelkerke R^2) of the variance in the mechanical LREs. However, there was no significant difference between the effect of setting, x^2 (1) = .55, p < .46 or modality, x^2 (1) = .00, p < 1.00 on the production of mechanical LREs.

LRE outcomes. Regarding the outcome of LREs, most of the LREs were resolved correctly, regardless of the setting or modality. Table 24 below displays the distribution of LRE outcomes in the two settings and modalities. For these calculations, the total number of LRE outcomes is used instead of the standardized number (LREs/100 words).

Table 24

Distribution of LRE Outcomes by Setting and Modality Focus

	Oral					Writing				
	AR		Non-AR		Total	AR		Non-AR		Total
	n	%	n	%		n	%	n	%	
Solved	88	67.69	71	57.26	159	142	67.62	179	77.49	321
Correctly		*55.35*		*44.65*			*44.24*		*55.76*	
(per setting)										
Solved	25	19.23	35	28.22	60	34	16.19	39	16.88	73
Incorrectly		*41.67*		*58.33*			*46.58*		*53.42*	
(per setting)										
Unresolved	17	13.08	18	14.52	35	34	16.19	13	5.63	47
(per setting)		*48.57*		*51.43*			*72.34*		*27.66*	
Total	130	100	124	100	254	210	100	231	100	441
(per setting)		*51.18*		*48.82*			*47.62*		*52.38*	

Note: The vertical percentages are in normal type. The horizontal percentages are in italics.

Table 24 shows that during the oral activities, 67.69% (88/130) of the LRES were solved correctly compared to 19.23% (25/130) solved incorrectly and 13.08% (17/130)

unresolved. Similarly, 57.26% (71/124) of the LREs in the non-AR setting were also solved correctly, 28.22% (35/124) were solved incorrectly, 14.52% (18/124) were unresolved LREs. Comparably, most LREs were also solved correctly during the writing focus activities (142/210 [67.62%] in the AR and 179/231 [77.49%] in the non-AR). AR and non-AR groups had a similar frequency of incorrectly solved LREs (34/210 [16.19%] in the AR and 39/231[16.88%] in the non-AR). The AR groups had the same number of unresolved LREs (34/210 (16.19%) as LREs that were solved incorrectly. The non-AR groups had 5.63% (13/231) unresolved LREs.

Looking at settings, LREs that were solved correctly during the oral activities were more frequent in the AR (88/159 [55.35%]) than the non-AR setting (71/159 [44.65%]). In contrast, incorrectly solved LREs were more frequent in the non-AR setting (35/60 [58.33%]) than the AR setting (25/60 [41.67%]). The frequency of unresolved LREs was higher in the non-AR setting (18/35 [51.43%]) than the AR (17/35 [48.57%]) setting. In comparison, the writing activities had a higher incidence of LREs solved correctly (179/321 [55.76%]) and incorrectly (39/73 [53.42%]) in the non-AR setting than LREs solved correctly (142/321 [44.24%]) and incorrectly (34/73 [46.58%]) in the AR setting. In contrast, the AR setting had three times more unresolved LREs (34/47 [72.34%]) than the non-AR setting (13/47 [27.66%]).

A multinomial regression model was conducted to investigate the effect of setting and modality on the outcome of LREs in the oral and writing activities. The full model with the predictor variables, setting and modality, were found to contribute to the model, x^2 (4) = 14.92, p = .005. The results of the analysis revealed that the odds of correctly

solved LREs relative to unresolved LREs in the non-AR setting was 1.77, 95% CI [1.09, 2.86] times higher than in the AR setting, a statistically significant effect, Wald x^2 (1) = 5.35, p = .021. Similarly, the odds of incorrectly solved LREs relative to unresolved LREs was 2.02, 95% CI [1.18, 3.64] times higher in the non-AR setting than in the AR setting. Oral vs. Writing modality focus did not significantly affect the outcome of LREs when comparing unresolved LREs to LREs correctly solved, Wald x^2 (1) = 2.53, p = .112, nor when comparing unresolved LREs to LREs solved incorrectly x^2 (1) = .19, p = .665. Table 25 shows the results of the multinomial regression predicting the likelihood of LRE outcomes based on setting and modality focus.

Table 25

Multinomial Regression Predicting Likelihood of LRE Outcome based on Setting and Modality

		B	SE	Wald	*df*	*p*	Odds Ratio	95% CI for Odds Ratio	
								Lower	Upper
Solved Correctly	Non-AR	.569	.246	5.348	1	.021	1.766	1.091	2.860
	Oral	-.389	.245	2.525	1	.112	.678	.420	1.095
	Intercept	1.659	.186	79.219	1	.000			
Solved Incorrectly	AR	.729	.287	6.437	1	.011	2.072	1.180	3.638
	Oral	.123	.285	.188	1	.665	1.131	.647	1.977
	Intercept	.090	.231	.151	1	.698			

Note: Setting is non-AR compared to AR, modality is oral compared to writing, and intercept is LRE outcome: unresolved.

LREs correction orientation. Table 26 shows the distribution of the correction orientation of LREs in both settings and modalities. For these calculations, the total number of LRE outcomes is used instead of the standardized number (LREs/100 words).

As Table 26 illustrates, during the oral activities, self-correction occurred more often in the AR (87/130 [66.90%]) and non-AR (78/124 [62.90%]) settings compared to other correction (43/130 [33.10]) in the AR and non-AR setting (46/124 [37.10%]). In contrast, during the writing focus activities, more LREs were corrected by the peers (other correction) in both groups (121/210 [57.60%] in the AR and 143/231 [61.90%]) in the non-AR) than self-corrected LREs (89/210 [42.40%]) in the AR and (88/231[38.10%]) in the non-AR groups).

Table 26

Distribution of LRE Correction Orientation per Setting and Modality

	Oral					Writing				
	AR		Non-AR		Total	AR		Non-AR		Total
	n	%	n	%		n	%	n	%	
Self-correction	87	66.90	78	62.90	165	89	42.40	88	38.10	177
(per setting)		*52.73*		*47.27*			*50.28*		*49.72*	
Other correction	43	33.10	46	37.10	89	121	57.60	143	61.90	264
(per setting)		*48.31*		*51.69*			*45.83*		*54.17*	
Total	130	100.00	124	100.00	254	210	100.00	231	100.00	441
(per setting)		*51.18*		*48.82*			*47.62*		*52.38*	

Note: The vertical percentages are in normal type. The horizontal percentages are in italics.

In addition, when looking at settings, the self-corrected LREs were slightly more in the AR setting (87/165 [52.73%]) than the non-AR setting (78/165 [47.27%]). In contrast, LREs corrected by peer were slightly more in the non-AR setting (46/89

[51.69%]) than the AR setting (43/89 [48.31%]). The number of self-corrected LRES during the writing activities were very similar in both settings (89/177 [50.28%] in the AR and 88/177 [49.72%] in the non-AR). Like the oral activities, in the writing focus activities the number of other-correction LRES were more frequent in the non-AR groups (143/264 [54.17%]) then the AR groups (121/264 [45.83]).

A binomial logistic regression was performed to understand the effect of setting and modality on LRE correction (self or other correction). The logistic regression model was statistically significant, x^2 (2) = 41.49, $p < .001$. The model explained 7.7% (Nagelkerke R^2) of the variance in LRE correction. The odds of having LREs corrected by other in the writing focus activities was 2.75, 95% CI [2.00, 3.79] times more likely than having the same type of correction in the oral activity, a statistically significant effect, Wald x^2 (1) = 38.29, $p < .001$. No significant difference was found in the setting, Wald x^2 (1) = 1.29, $p < .256$.

Table 27 shows the results of the binomial logistic regression predicting the likelihood of LRE correction based on setting and modality.

Table 27

Logistic Regression Predicting Likelihood of LRE Correction based on Setting and Modality

	B	SE	Wald	*df*	*p*	Odds Ratio	95% CI for Odds Ratio	
							Lower	Upper
Setting	-.18	.16	1.29	1	.256	.84	.62	1.14
Modality	1.013	.16	38.29	1	.000	2.75	2.00	3.79
Constant	-.528	.153	11.89	1	.001	5.90		

Note: Setting is AR compared to non-AR, and modality is writing compared to oral. Constant is LRE outcome: other correction.

In sum, the findings related to the first question of the study: does the task setting (AR vs non-AR) differentially affect the number, type, outcome, and correction orientation of target LREs produced in the two modalities (oral vs writing) revealed that in terms of numbers, dyads produced a similar number of LREs using AR and non-AR applications.

The results of the standardized distribution of LREs across all dyads in both settings and modalities showed that participants when using AR applications produced slightly more LREs than when using non-AR applications during the oral (mean = 1.58 in the AR and mean = 1.47 in the non-AR) and writing focus activities (mean =4.23 in the AR and mean = 3.47 in the non-AR). The results of the regression analysis showed that the setting had a statistically significant effect on the prediction of the number of standardized LREs produced, Wald x^2 (1) = 20.65, $p < .001$. Likewise, modality also had a statistically significant effect on the prediction of standardized LREs, Wald x^2 (1) = 112.81, $p < .001$.

In terms of the nature/type of LREs, a multinomial regression revealed that setting did not significantly affect the types of lexical LREs produced. However, modality showed some significant differences in the number of LREs related to nouns and pronouns. There were significantly more nouns produced during the oral activities than in the writing focus activities. In contrast, the were significantly more incidences of pronouns in the writing focus activities than the oral activities.

For the grammatical LREs, a multinomial regression also revealed that setting did not significantly affect the types of grammatical LREs produced. However, as with

lexical LREs, modality had a significant effect on two sub-types of grammatical LREs. Verb tense and verb aspect were significantly more frequent in the writing focus activities than the oral activities. For the mechanical LREs, orthographic and pronunciation, a binomial logistic regression showed no significant differences between settings. Not surprisingly, there were significant differences in modality focus when comparing mechanical and lexical LREs, and mechanical and grammatical LREs, but no significant differences between the number of lexical and grammatical LREs in both modalities.

Regarding the LRE outcome, more than half of the total LREs were solved correctly in both settings and modalities. Unresolved LREs had the lowest numbers overall. The setting significantly affected the outcome of LREs, with higher odds of having correctly and incorrectly solved LREs compared to unresolved LREs in the non-AR than the AR setting. Modality did not significantly affect the outcome of LREs. The results of a multinomial regression revealed that the odds of correctly solved LREs relative to unresolved LREs in the non-AR setting was higher than in the AR setting. Similarly, the odds of incorrectly solved LREs relative to unresolved LREs was also higher in the non-AR setting than in the AR setting. Modality focus did not significantly affect the outcome of LREs when comparing unresolved LREs to LREs solved correctly and unresolved LREs to LREs solved incorrectly.

In terms of the LRE correction orientation, both settings had similar numbers of self-correction and other correction LREs. Thus, no significant differences were found between settings. However, modality significantly influenced the type of correction (self

vs. other). There were significant more instances of other-correction LREs in the writing focus activities than in the oral activities.

RQ#2: Does the setting and LRE outcomes affect the ability of participants to recognize and produce the correct grammatical and lexical forms on the posttests? Is there a significant correlation between students' ability to recognize the correctness of a grammatical form and their ability to produce that form correctly on the posttests?

To answer these questions the results of the immediate posttests were analyzed in four ways. First, a recognition analysis was performed to compare the results of the Likert-scale grammatical judgment questions by LRE outcomes and settings. Next, an analysis was performed to compare the production results of the same grammatical questions based on LRE outcomes and settings. Third, the relationship between recognition and production scores was analyzed. Finally, an analysis was performed to compare the production results of lexical items based on LRE outcomes and settings. Posttest scores were only calculated for items on the posttests that were based on actual LREs in the participants' dialogues.

Recognizing grammatical forms. There was a total of 92 individual responses for the grammatical Likert-scale questions, 46/92 (50%) from the AR tasks and 46/92 (50%) from the non-AR tasks. As mentioned in the methodology section, higher scores show better grammaticality judgments (these scores reflect a combination of the correctness of the prompt and their degree of certainty with the correctness of the original prompt). To recap, the prompts in the posttests (drawn from the dyad's dialogue) were either grammatically correct or incorrect. A score of zero indicated no recollection from the

participant (“I don’t know”), a score of one indicated that a participant recognized the structure incorrectly, a score of two indicated that a participant recognized the structure incorrectly with a certain degree of doubt, a score of three indicated that a participant recognized the structure correctly but with a certain degree of doubt, and a score of four indicated that a participant recognized the structure correctly.

Table 28 shows the recognition posttest scores based on the type of LRE outcome and setting. As illustrated in Table 28 a total of 92 prompts originated from LREs produced during AR tasks (46/92 [50%]) and non-AR tasks (46/92 [50%]).

Table 28

Grammatical Recognition Posttests’ Scores per LRE Outcome and Setting (N = 92)

Posttest Scores														
outcome	AR							Non-AR						
Score scale	0	1	2	3	4	Total	*Mdn*	0	1	2	3	4	Total	*Mdn*
unresolved	0	2	1	6	3	12	2.00	1	0	2	2	2	7	2.00
Solved correctly	4	1	4	6	11	26	3.00	0	4	3	8	12	27	3.00
Solved incorrectly	0	1	4	1	2	8	3.00	2	3	2	3	2	12	3.00
Total	4	4	9	13	16	46	3.00	3	7	7	13	16	46	3.00

In terms of the posttest scores, 63% of the prompts in both settings (AR [29/46] and non-AR [29/46]) received a score of 3 or higher (i.e., demonstrating that participants recognized the structure correctly as grammatically correct or ungrammatical). Regarding the outcome types, participants recognized 68.42% of the unresolved LREs, 69.81%

(37/43) of the grammatical LREs that were solved correctly, and 40% of the LREs that were originally solved incorrectly.

The regression model proved to be not statistically better at predicting the dependent variables (recognition and production scores) than a mean model. Thus, a Kruskal-Wallis H test was applied to determine if there were differences in recognition scores between the three types of LRE outcomes: unresolved (n = 19), solved correctly (n = 53), solved incorrectly (n = 20) in the two settings. Distributions of recognition scores were similar for all types, as assessed by visual inspection of a boxplot. Recognition scores increased from unresolved (*Mdn* = 2.00) to solved outcomes (*Mdn* = 3.00). However, there were no differences between solved correctly (*Mdn* = 3.00) and solved incorrectly (*Mdn* = 3) outcomes. Median recognition scores were not significantly different between the three types in the AR setting, x^2 *(2)* = .76, *p* = .68 nor the non-AR setting, x^2 *(2)* = 5.04, *p* = .08.

Producing grammatical forms (student correction of Likert-scale prompts). As part of the posttest, students not only had to recognize the grammatical structures as correct or incorrect, but they also had the opportunity to replace the incorrect form with what they thought was the correct form. To recap, production scores were coded as follows: a score of one indicated that a participant did not try to repair an incorrect prompt or incorrectly repaired a correct prompt, a score of two indicated that a participant repaired a correct prompt with an unnecessary correction (i.e., "¿*Cómo estás*?" with "¿*Cómo estás tú*?") or repaired an incorrect prompt still incorrectly, a score of three indicated that a participant left a correct prompt as is or repaired an incorrect prompt with

the correct form. As with recognition scores, higher scores show better productive knowledge (i.e., participants left the correct forms alone and corrected the wrong forms).

Table 29 shows the production posttest scores based on the type of LRE outcome and setting. As illustrated in Table 29, in terms of the posttest scores, 50% of the grammatical prompts in the AR (23/46) and 63% of the grammatical prompts in the non-AR (29/46) received a score of 3. Overall, 56.52% (52/92) of the LREs received a score of 3, thus demonstrating productive knowledge. In terms of outcome types, participants produced the correct form of 42.11% (8/19) of the LREs originally unresolved, left unchanged 73.58% (39/53) of the LREs that were originally solved correctly and corrected 25% of the LREs that were originally solved incorrectly (5/20).

Table 29

Grammatical Production Posttests' Scores per LRE Outcome and Setting (N = 92)

	Posttest Scores											
Outcome	AR						Non-AR					
Score scale	0	1	2	3	Total	*Mdn*	0	1	2	3	Total	*Mdn*
Unresolved	0	3	4	5	12	2.00	1	1	2	3	7	2.00
Solved correctly	4	0	5	17	26	3.00	0	2	3	22	27	3.00
Solved incorrectly	0	6	1	1	8	1.00	2	5	1	4	12	1.00
Total	4	9	10	23	46	2.50	3	8	6	29	46	3.00

A Kruskal-Wallis H test was run to determine if there were differences in production scores between the three types of LRE outcomes: unresolved (n = 19), solved correctly (n = 53), and solved incorrectly (n = 20) in the two settings. Distributions of production scores were similar for all types, as assessed by visual inspection of a boxplot.

Production scores increased from "solved incorrectly" (*Mdn* = 1.00) to "unresolved" (*Mdn* = 2.00) to "solved correctly" (*Mdn* = 3). Median production scores were significantly different between the three types, x^2 *(2)* = 17.52, $p < 001$. Subsequently pairwise comparisons with a Bonferroni correction for multiple comparisons (statistical significance accepted at $p < 0.15$ level) revealed a significant difference in production scores between the "solved incorrectly" (Mdn = 1.00) and "solved correctly" (Mdn = 3.00) groups in the non-AR ($p = .004$), but not in the AR setting ($p = .023$). Adjusted *p*-values are presented.

Relationship between recognition and production scores. On their posttests, participants were able to recognize (as correct or incorrect) 63% of the grammatical LREs and leave unchanged the correct form or correct the ungrammatical forms of 56.52% of the original grammatical LREs. A Spearman's rank-order correlation was run to assess the relationship between recognition scores and production scores in both settings. Preliminary analysis showed the relationship to be monotonic, as assessed by visual inspection of a scatterplot. The results revealed a significant strong positive correlation between the recognition scores and the production scores in the AR setting, r_s (44) = .63, p < .001, and non-AR setting, r_s (44) = .64, p < .001, that is, an increase in the recognition scores was strongly associated with an increase in production scores in both settings.

***Production of lexical items*.** There was a total of 86 individual responses for the lexical (translation) questions, 40/86 (46%) from the AR tasks and 46/86 (54%) from the non-AR tasks. As mentioned before, higher scores show better production knowledge.

To recap, the lexical items were within a context and without a context (drawn from the dyad's dialogue). A score of zero indicated no translation of lexical item, a score of one indicated that participants produced an incorrect translation of a previously unresolved, correctly solved or incorrectly solved lexical LRE, a score of two indicated that a participant produced the a correct translation of a previously unresolved, correctly solved, or incorrectly solved lexical LRE.

Table 30

Lexical Production Posttests' Scores per LRE Outcome and Setting (N = 86)

	Posttest Scores									
outcome	AR					Non-AR				
Score scale	0	1	2	Total	*Mdn*	0	1	2	Total	*Mdn*
unresolved	1	1	4	6	1.50	1	1	2	4	2.00
Solved correctly	4	1	25	30	2.00	7	4	29	40	2.00
Solved incorrectly	2	2	0	4	1.00	1	0	1	2	.50
Total	7	4	29	40	2.00	9	5	32	46	2.00

The descriptive statistics of the lexical production posttest scores from each setting are displayed in Table 30. In terms of the posttest scores, 72.5% of the prompts from AR tasks (29/40) and 69.57% of the prompts from non-AR tasks (32/46) received a score of 2 (i.e., demonstrating productive knowledge). Overall, on the posttests participants correctly produced 70.93% of items based on lexical LREs. Looking at the types of outcomes, on the posttests, participants produced the correct lexical item for 60% of the originally unresolved LREs, 77.14% of the LREs solved correctly and 16.67% of the LREs solved incorrectly.

A Kruskal-Wallis H test was ran to determine if there were differences in production scores between the three types of LRE outcomes: unresolved (n = 10), solved correctly (n = 71), and solved incorrectly (n = 6) in the two settings. Distributions of production scores were similar for all types, as assessed by visual inspection of a boxplot. There were no significant differences in production scores between the two settings.

However, there were significant differences between outcome types in the AR setting. Production scores increased from "solved incorrectly" (*Mdn* = 1.00) to "unresolved" (*Mdn* = 1.50) to "solved correctly" (*Mdn* = 2.00). Median production scores were significantly different between the three types, x^2 *(2)* = 10.54, *p* = .005 in the AR setting. Subsequently pairwise comparisons with a Bonferroni correction for multiple comparisons revealed a significant difference in production scores between the "solved incorrectly" (Mdn = 1.00) and "solved correctly" (Mdn = 2.00) groups in the AR (*p* = .004) setting. Adjusted *p*-value are presented.

In sum, the findings related to the second question of the study revealed that dyads recognized 63% of the grammatical structures correctly, and were able to produce (or leave unchanged) a correct grammatical form 65.38% of the time in the AR setting and 81.48% of the time in the non-AR setting. No significant differences were found in the recognition posttest scores regarding the setting nor the type of LRE outcome. Similarly, there were no significant differences regarding the production scores of grammatical structures from dialogues using AR and non-AR. Overall, participants correctly produced (or left unchanged the correct forms) of 56.52% of the LREs.

However, when looking at the two settings separately, the LREs solved correctly

and solved incorrectly had significant production scores differences in the AR (17/26 and 1/8 respectively) and in the non-AR setting (22/27 and 4/12 respectively). The Spearman correlation test revealed a strong positive correlation between the recognition and production scores. Lastly, the production scores related to lexical items revealed no significant differences between the two settings. However, looking at each setting separately, more specifically the AR setting, there was a significant difference between the production scores from LREs correctly solved vs LREs solved incorrectly. Overall, on their posttests, participants were able to correctly produce 70.93% of the lexical items related to LREs. Comparing the production of grammatical (56.52%) vs lexical (70.93%) LREs, participants performed better when producing lexical items.

CHAPTER 5

DISCUSSION

The purpose of this study was to explore the impact of task setting (AR vs non-AR) on LREs in oral and writing-focused activities during collaborative dialogue and its effect on the ability of participants to recognize and produce the correct forms on the posttests. Thus, this study took a novel approach in exploring the extent to which interaction is present in two different settings – an outdoor campus setting enhanced by AR applications vs a laboratory setting with similar traditional computer applications. Following the cognitive and sociocultural theoretical approaches, a quantitative analysis was performed to investigate the LREs produced during peer interaction.

The first research question addressed how the setting affected the number, nature, outcome, and correction orientation of LREs during oral and writing activities. A discussion of each category is below.

> **RQ#1:** Does the task setting (AR vs non-AR) differentially affect the number, nature, outcome, and correction orientation of target LREs produced in the two modalities (oral vs writing-focused)?

Number of LREs

The total number of LREs is the first notable difference between the current study and other LRE studies that used oral and oral + writing components in their tasks. In the current study, L2 learners of Spanish produced 695 instances of LREs in AR and non-AR settings. As demonstrated in Chapter 4, Tables 14 an15 showed that the overall average of LREs produced per 10 dyads was 19.30 and the overall average of LREs per 100

words was 2.7. These results showed that, regardless of whether the participants completed the task using AR outside (340) or non-AR applications in the lab (355), dyads produced a similar number of LREs per modality in both settings. However, looking at the average of LREs per 100 words, dyads who used AR produced more LREs in the oral activities and almost 1/3 more LREs during the writing-focused activities compared to dyads who used the traditional computer applications (non-AR group).

The current study produced double the average of LREs compared to Swain and Lapkin's (2001) study with a total of 26 LREs and an average of 9.0 LREs per 31 dyads. The current study's LRE average per 100 words is consistent with Yilmaz and Granema's (2010) study. In their study, which used the same ratio calculation, the average of LREs produced by five dyads was 1.26 LREs, compared to the average of 2.7 LREs produced by 10 dyads in the current study. However, it is difficult to compare these studies with the current study because Yilmaz and Granema's (2010) study was performed in a SCMC context, looking only at the participants' written output, and Swain and Lapkin's (2001) F2F classroom data did not have a LRE ratio. In addition, both studies used different tasks (e.g., dictogloss and jigsaw) and contexts (lab and classroom) than the current study.

Comparing the current study with García-Mayo and Azkarai's (2016) study, which reported an average number of .44 LREs per turn in the writing tasks and .18 LREs per turn in the oral tasks, this study produced .29 LREs per turn in the writing tasks and .12 LRES per turn in the oral tasks. García-Mayo and Azkarai's (2016) had 22 Spanish EFL dyads at different proficiency levels completing 4 collaborative tasks (dictogloss,

text editing, picture placement, and picture differences) whereas the current study had 10 advanced Spanish learners and two collaborative tasks. Consistent with García Mayo and Azkarai's (2016) study, overall, the current study also had more turns during the oral activities but more LREs during the writing activities. Likewise, Adams and Ross-Feldman's (2008) study with a similar task (decision-making) as the current study showed more LREs during writing tasks targeting locatives (64.05%) and past tense (82.76%) than during the oral tasks (35.95% and 17.24% respectively). (LRE ratio: the total of dyad's LREs divided by the total number of LREs).

Nonetheless, it is difficult to compare these studies with the current study because the task types, context, and measurements for calculating the ratio of LREs to learners' language production (e.g., the total number of LREs vs the number of turns vs the number of words per dyad) were different (or nonexistent).

The current study used two different environments (place-based and the lab), as opposed to just one setting (laboratory or classroom). It also aimed to use authentic collaborative tasks aided by technology (AR vs non-AR) as opposed to more structured tasks that explicitly elicit negotiation for meaning and form (e.g., jigsaw, dictogloss).

Comparing this study with the only other study thus far that employed LREs as a unit of analysis using AR technology (Sydorenko et al., 2019), the number of LREs in the current study considerably exceeds the number of LREs in that study. In their study, four groups of three participants (1NS – 2NNS) played a mobile-based AR game called *ChronoOps*. The participants produced a total of 32 LREs while completing one oral task around the university campus. Although their study also used a place-based mobile AR

application, it is difficult to compare both studies because Sydorenko et al. (2019) did not mention the time allocated to complete the task, nor identify the number of words or turns, they used only one oral activity with 12 participants in groups of 3 (1NS – 2NNS) instead of dyads, and no control group for comparison. However, Sydorenko et al. (2019) offered interesting findings of the use of AR and the nature of LREs which are presented in the next section.

Similar to previous studies (Storch, 2001; Watanabe & Swain, 2008), the current study also reported considerable variation in the total number of LREs by dyad. There was a range of 2 to 64 LREs per dyad out of 695 (*SD* = .38) total LREs. However, unlike other studies (Kim & McDonough, 2008), which reported a higher frequency of LREs among dyads with high proficiency students, the results of the current study suggest that pairing students with the same level of proficiency fostered more negotiated interaction regardless of ability level in Spanish. In the current study, the dyad with the lowest proficiency (Paul and Kim) had one of the highest frequencies of LREs across modality and setting (22 to 64 LREs).

One possible explanation might be due to participants' collaborative or non-collaborative orientation to the activity (Swain & Watanabe, 2013) and their level of engagement with the task, which Fernández Dobao (2012) suggests might have a stronger effect on the nature of the interaction and the number of LREs than the overall proficiency of the dyad. Although the relationship between proficiency, the learners' collaborative orientation or task engagement and LREs, was beyond the scope of this study, it is a topic worth mentioning and examining in future research.

Taking into consideration the fact that some of these previous studies had more participants than the current study and used more structured task interactions, the higher average of LREs from the current study, is particularly notable. The results of this study suggest that ecological validity and task settings could have motivated students to talk more and produce more LREs. However, several interactional features and measures must be taken into consideration when investigating peer interaction, as there are a host of variables (e.g., gender, NS vs NNS, task types, level of engagement) that come into play during task completion (Jenks, 2009).

In addition, this study's findings confirm previous research on the impact of task modality (Adams, 2006; Adams & Ross-Feldman, 2008; Garcia-Mayo & Azkarai, 2016), as participants initiated significantly more LREs in the writing tasks than in the oral tasks. Participants had a total of 11,268 total words and an average of .16 LREs per 100 words in the writing-focused activities, whereas in the oral activities there was a total of 16,512 words and an average of .11 LREs per 100 words. Dyads were instructed to use 20 minutes to complete the oral activities and 20 minutes for the writing activities. For the oral activities they were free to use any vocabulary or grammatical forms and they were not told how many words they had to produce. For the writing-focused activities, they needed to write at least 200 words and to continue talking while writing their text. They were also instructed to try to use the grammar forms outlined in Task 1 and Task 2 examples shown previously to them by the researcher. These instructions, besides contributing to the amount of collaborative dialogue, also added to the number of LREs as seen in the examples below.

Example 1 shows dyad discussions focusing on the number of words they had to write for the writing-focused task and the number of minutes they had to talk during the oral activities.

Example 1: (non-AR writing-focused)

Melissa: *probable* like *cien y… cuanta… cuantos palabras tenemos?*
[probably about one hundred and… how many (f/singular) … how many (m/plural) words do we have?]

Amy: *Noventa y dos.*
[Ninety-two.]

Melissa: *Noventa no es muy, hihi mucho.*
[Ninety is not very, (laughs) many.]

Melissa: *Enjoy? Alegrar, maybe? es una palabra, sí? Mucho. Do we have to have 200 words before we leave? This is terrible! It's probably like 100.*

Amy: *Necesito usar el objeto directo y comparativos.*
[I need to use the direct object and comparatives.]

Lisa: *usa más vegetales para las conta de palabras… ¿cuenta? ¿El cuenta? ¿El cuenta de palabras?*
[Use more vegetables for the word count (mispronounced)… count (pronounced correctly)? The (m) count? The (m) word count?]

Lisa: *¿Cuánto de palabras?*
[How much of words?]

Karen: *Oh necesitamos cincuenta más, hihi.*
[Oh we need fifty more, (laughs).

Lisa: *Vamos a poner un conclusión más largo.*
(Let's put a (m) longer conclusion.]

(AR oral)

Bill: *¿Cuántos minutos tenemos? ¿Qué es el tiempo?*
[How many minutes do we have? What is the time?]

Rob: *Tenemos doce minutos, más o menos.*
[We have twelve minutes, more or less.]

This study also adds to the literature regarding the interaction between modality and setting which was also significant. For instance, the average number of LREs in the AR setting during the writing-focused activities was significantly higher (M = .18) than the average of LREs in the non-AR setting during the oral activities (M = .11).

Regarding the setting, although participants produced more words in the non-AR setting (15,026) compared to AR setting (12,754), overall, the average number of LREs per 100 words was higher in the AR setting (.15) compared to non-AR (.12). One possible explanation might be the fact that the immersive technology gave opportunities for participants to use the language in different ways to talk about using new technology and to incorporate the physical context, not possible with traditional classroom materials in a classroom nor with computer applications in a laboratory, as seen in the examples below:

The LREs below originated from using the applications' specific technology features. Example 2 shows dyads using the camera feature from the Google Translate AR application. As dyads place the camera over the menu, they see the menu translated into the target language.

Example 2: (AR oral)

Mark: *y usando cámara vamos a *traduzcar* [sic] *el menú para que podemos leerlo en español. Aquí está…*

[and using camera we are going to translate the menu so we can read it in Spanish. Here it is…]

Betty: Oh my God!
(Surprised)

Example 3 also shows dyads using their mobile phone capabilities to take a picture of their virtual furnished room.

Example 3: (AR oral - IKEA AR app with phone camera)

Matt: *una mesa de café y después tomar una foto para que no perdimos… para que no perdamos la idea.*
[a coffee table and after to take a photo so we don't lose (past tense) … so we don't lose (subjunctive) the idea.]

Matt: *una de azul y una de color blanco y después tomar la foto para… ¿tomar foto?*
[one of blue and one of white color and later take a photo for... take a photo?]

Ashley: uh-huh.
(agreement)

In example 4, dyads use the space awareness feature of the IKEA AR application to move, resize and place the virtual furniture in a physical location.

Example 4: (AR oral - IKEA AR app space awareness feature)

Matt: **Ponería* [sic] *una carpeta si tenemos la espacio para… el espacio pero no tenemos más dinero, así que estamos terminados…*
[I would put a folder (meant to say rug) if we have the (f) space for… the (m) space but we don't have any more money, so we are finished.]

(AR oral)

Paul: S*í, me gusta, sí. ¿Dónde vas a ponerlo?*
[yes, I like it, yes. Where are you going to put it?]

Kim: *¿Cómo? ¿Pone?*
[how? (3rd person) put?]

Paul: *Oprima la pantalla.*
[press the screen.]

Kim: Oooh! Okay!

Paul: *¡Un salón! la televisión puede a ser aquí tener aquí ser aquí estar aquí. ¿Un otro lugar? ¿Te gustar en un otro lugar?*
[A living room! The television can be here, have here, be here, be (location) here. Another place? Do you like it in another place?]

Kim: *oh puedes…*
[Oh you can…]

Paul: *¡muévalo!*
[Move it!]

Paul: *Sí, azul. A la derecha.*
[Yes, blue. To the right.]

Paul: *Sí, la *madura de algo de la cosas son la misma también. Me gusta la locación de las sillas. Necesito dos más.*
[Yes, the wood from some of the things are the same also. I like the location of the chairs. I need two more.]

(AR oral)

Nick: *Dónde puedo…*
[where can I…]

Anna: *Aquí. Wow, muy largo.*
[Here. Wow, too big.]

Nick: *Yo voy a ponerlo aquí y va a ver (inaudible)*
[I'm going to put it here and you will see (inaudible)]

Anna: *¡Muy bonito!*
[Very pretty!]

In example 5, dyads mention the Bettar AR application in their writing-focused activity (email to a friend) as a good tool to look for activities in town.

Example 5: (AR writing-focused)

Karen: *Con muchos verduras, ah… verduras… puedes poder… poder… pero si quieres platos sin carne con muchos verduras hum… tienes muchos opciones también. Es bueno para comer y necesitamos más de karate (Reading). Es Lo podemos usar lo, nosotros lo encontramos en el app Bettar, sí. Es uno que puede buscar por actividades divertidos cuando estás aquí con…. ¿Que es conmigo, pero con nuestras?*
[With many (m) vegetables, ah… vegetables… you can can… can… but if you want dishes without meat with many vegetables hum… you have many options too. It is good to eat and we need more of karate (reading). It is it, we can use it, we found it in the (m) app Bettar, yes. It is one that you can look for fun (m) activities when you are here with… What is with me but with ours (f)?]

Karen: *Curioso. Cuando estás aquí con nosotras. Ah… Debe… Debes buscar en Bettar para otros… o por otros? Para?*
[Curious. When you are here with us. Ah… You (formal) must... must (informal) look for in Bettar for (incorrect form) others… or for others? For (incorrect form)?]

Lisa: *No sé…*
[I don't know…]

Karen: *Para otros actividades que quieres hacer.*
[For (incorrect form) other (m) activities that you want to do.]

The current study's results contribute to previous research on negotiated interaction in different environments (Foster, 1998; Gass et al., 2005, Russel & Spada, 2016). As Garzón and Acevedo's (2019) AR meta-analysis suggested, informal environments involving activities outside of classrooms produce better learning outcomes than formal environments (classrooms and labs). The current study compared a controlled environment (laboratory) vs an uncontrolled place-based environment. That is, participants outside of the classroom engaged with the environment and used the target

language in *situ*. Foster (1998) attributed the relatively low numbers of negotiated moves between students in classroom settings to the “general informality of the classroom and the absence of any strict requirement to fulfill the task inclined them (the students) no to pay close attention to form of their language” (p. 16).

However, taking into consideration that participants in this study completed tasks in an "informal setting," the results of the current study are not consistent with Foster's (1998) results. On average, students had more standardized LREs in this place-based “uncontrolled” environment than the participants using computer applications in the laboratory-controlled environment. On the other hand, other studies (Gass et al., 2005; Russel & Spada, 2016), have found no significant differences in the participants’ number of negotiations of meaning and form in two different environments (e.g., classroom vs laboratory). The results of the current study showed significant differences between this new type of environment vs laboratory. The AR groups (compared with the non-AR groups) had the same average number of LREs in the collaborative dialogue during the oral activities outside of the classroom and higher number of LREs during their dialogues during the writing focused activities.

Three factors may account for these results. First, the new technology motivated students to talk about the experience of using something innovative. Second, the oral communication tasks in an authentic context influence the topics participants choose to communicate about (Jamshidnejad, 2011), giving them a broader variety of topics and more chances to use the target language. The third reason is that mobile AR applications provide context-awareness and situational learning (Example 4), offering more

opportunities for students to question and reflect on their language choices. For instance, the Bettar AR app displays several places and events at once just by positioning the phone camera in front of you, which could have motivated participants to explore more. Participants using the app talked about 6 different events and 5 restaurants on average compared to 4 events and 3 restaurants from participants using the Google Maps computer application (non-AR group). Likewise, when using the IKEA AR app participants were able to interact with the furniture (Example 5), something they could not do by surfing the website. The LRES shown in the examples below originated from the context-awareness and situational learning provided by the environment. In example 6 participants are talking about the color of the furniture which matches the ceiling of their house (blue as in the sky, because they are outside).

Example 6: (AR oral)

Matt: *Se eligieron… para decir por sus por el sentimiento de… no sé. ¿Se eligieron por sus, por el sentimiento que hacen?*
[they selected themselves… to say for their for the feelings of… I don't know. They selected themselves for their, for the feelings they make?]

Ashley: hum… (laughs)

Matt: *del cielo de nuestra casa. ¿Qué más?*
[Of the sky (meant to say ceiling) of our house. What else?]

In example 7 the dyads are in front of a building under construction. The dyads make comments about the building. Later, the same dyad talks about the birds flying nearby. Both examples show how the setting influences the topics discussed and offers students the opportunity to discuss new, unique topics that they would not normally bring up in a classroom setting.

Example 7 (AR oral)

Thomas: *no me gusta este edificio, ellos están construyendo el otro parte y yo creí que iban a mejorarlo, pero se ve lo mismo que antes de ellos empezaron a construir.*
[I don't like this building, they are building the other part and I thought that they would make it better, but we see the same as before they started to build.]

Elise: *Tengo o tuve una clase con este esta profesora en este edificio en el año pasado.*
[I have or I had a class with this (m) this (f) teacher in this building last year.]

Elise: *¡Hay aves!*
[There are birds!]

Thomas: *Sí, hay muchas aves.*
[Yes, there are many birds.]

Elise: *No me gusta…*
[I don't like...]

Thomas: *¿De veras? Me encantan los pájaros.*
[Really? I love birds.]

Elise: *me gusta, pero no cuando hum…*
[I like, but not when hum…]

Thomas: *No están tan cerca.*
[They are not so close.]

Elise: *Yeah, no están tan cerca de mí.*
[Yeah, they are not so close to me.]

Thomas: *No están haciendo nada.*
[They are not doing anything.]

The examples below show that AR provides a high level of realism and visual salience that creates a beneficial environment for input processing and more opportunities

for “pushed” output in that it provides learners with a more authentic experience. In example 8, Rich makes a comment in English about the app's ability to have users interact with the virtual furniture.

Example 8

Rich: Sit on the chair, oh my God that is so funny!

In example 9, the dyads are “moving” the furniture Rich makes a comment in English about the app's ability to have users interact with the virtual furniture.

Example 9 (AR oral)

Matt: *Para sus colores, yeah. ¿Qué más dice el papel? *Poneremos* (sic) *un estante para… al izquierda de la sofa grande y el otro a la derecha.*
[For (incorrect form) their colors, yeah. What else does the paper say? We will put a shelf for… at the (m) left of the (f) big sofa and the other at the right.]

(AR oral)

Kim: *Tengo un ottoman en… ¿que es…*
[I have an ottoman in… what is…]

Paul: *A la *iziquierda.* (wrong pronunciation)
[At the left.]

Kim: *¿Iziquierda? una silla a la…*
[Left? (repeating the same pronunciation)] a chair at the (f)…]

Paul: *Dos sillas a la derecha y un silla cerca de…a la ottoman.*
[Two chairs at the right and one (m) chair next to… at the ottoman.]

These examples show how students used the app’s space awareness feature to place their furniture in the physical location. Interestingly, when completing the writing activity, even though, participants using the IKEA website (non-AR) had a sheet with an

apartment plant drawing to draw the furniture placement, they did not talk nor wrote on their texts about preposition locatives like the AR group.

The examples below show the participant's metalinguistic awareness during collaborative dialogue in an AR setting (example 10) and non-AR setting (example 11). In example 10 below, the two students talk about a game one of them plays outside of school. Even though the topic is not part of the task, they keep their communication in the target language. Charles wonders about his classmate's use of the word *juntar*.

Example 10 (AR oral)

Rich: *yo no sé, es posible que los jugadores van a juntar y practicar, yo no.*
[I don't know, it is possible that the players are gathering and practice, I'm not.]

Charles: *¿*deciste* [sic] *juntar? To gather? Interesante.*
[Did you say to gather? To gather? Interesting.]

In example 11, Anna and Nick collaborate to find the right conjugation and tense for the verb *divertirse* (to have fun).

Example 11 *(*non-AR writing)

Anna*:* *tener divertido… o nos divertimos*
[to have fun (adjective)… or we had fun]

Nick*:* *Vamos a divertirnos…*
[We are going to have fun (ourselves)]

Anna*:* It would be…

Nick*:* It's easier to *divertirse* in the future tense. I can look it up really fast…

Anna*:* Doesn't the second verb have to be in the infinitive when you use *vamos a*?

Nick*:* I don't know how… *dive…*

Anna*:* You're right… *vamos a divertirnos?* Because you still have to add the *nos.*

Interestingly, there were 8 instances of metalinguistic talk during oral activities using AR and no instances in the non-AR setting. In contrast, during the writing-focused activities, the non-AR groups had 61 instances of meta-talk whereas the AR groups had 27 instances of meta-talk. These results confirm the interaction findings between setting and modality previously mentioned (pg. 6). Future research should investigate this interaction more closely to understand these opposing trends in the interaction between modality and setting.

In sum, the current study validates other studies' claims that adding a writing component in oral peer interactive tasks leads to the production of more LREs. Unlike previous studies, there was a high incidence of LREs during dyads' interactions across all proficiency levels. This shows that to complete the tasks, participants at all levels collaborated with each other to resolve the language problems they encountered during the activities. This was even more evident when participants used place-based AR taking advantage of the environment and the applications' features to explore with the language and further use that knowledge during their writing assignments. The examples above show how dyads during collaborative dialogue benefitted from the authentic context and the apps' features to talk about different topics and reflect on their language choices. A detailed discussion about the nature of LREs follows below.

Nature of LREs

Previous studies (Adams & Ross-Feldman, 2008; Garcia-Mayo & Azkarai, 2016, Kim & McDonough, 2008; Lapkin et al., 2002; Leeser 2004; Niu, 2009; Swain &

Lapkin, 1998), showed that participants produced more grammatical LREs in the writing tasks and more lexical LREs in oral tasks. Unlike these previous studies, oral and writing activities in the current study produced more lexical LREs (52.37%) than grammatical LREs (35.68% respectively). Looking at the grammatical LREs only, the writing modality produced more instances of this type of LRE (61.29%) than the oral modality (38.71%), which confirm the findings of other studies' claim that written output draws learner's attention to grammatical knowledge to a greater extent than just oral output does. Adams and Ross-Feldman (2008) claim that participants focus more on form during writing activities due to "the effects of producing output through writing than to having previously discussed the content in a collaborative speaking section of the task" (p. 261).

Looking at the lexical LREs only, the writing modality also produced more lexical LREs than grammatical LREs, suggesting that perhaps was the nature of the task (e.g., talking about foods to eat, furniture to buy) vs a more structured text editing (e.g., dictogloss) task that afforded more opportunities for dyads to talk about vocabulary (e.g., nouns and adjectives) more than grammar during the writing activities. In Sydorenko et al.'s (2019) study using an open-ended AR game activity, participants produced only lexical LREs. The authors suggest that the nature of the game with a time limit to finish it, made participants focus only on the lexical knowledge necessary to complete the task (e.g., to get their message across). In addition, they suggested that AR games are a meaning-oriented task unless focus on grammar is clearly requested via instructions. This study confirms this claim (Sydorenko et al., 2019) that AR may offer a more meaning-oriented opportunity for communication even when grammatical instructions are woven

into the task. Perhaps, because learners were aware that the tasks were part of the class curriculum they also focused on form. However, the fact that the number of grammatical LREs were higher in the AR setting could mean that AR is also a good technology to promote grammatical LREs.

Another interesting result was that the number of lexical LREs (53.82% in the AR and 50.99% in the non-AR) and grammatical LREs (36.47% in the AR and 34.93% in the non-AR) were comparable but slightly higher in the AR setting than the non-AR setting. These results suggest that the AR technology and the place-based environment could have pushed students to focus more on form and meaning. However, as illustrated on Table 20, these differences were not significant in the lexical LREs (p = .11) nor the grammatical LREs, (p = .14).

The examples below show dyads focusing on vocabulary (example 12 and 14) and grammar (example 13 and 15) during oral and writing activities in AR and non-AR settings.

In example 12, dyads question the meaning of the noun "bowl" and the adjective "grilled" in Spanish.

Example 12 (AR Oral – lexical LRE)

Mark: *Tazones de proteínas.*
[Protein bowls.]

Betty: *Tazones, ¡tazón es la palabra!*
[Bowls, bowl is the word!]

Mark: *A la parrilla… ¿no sé qué es eso?*
[Grilled… I don't know what is that?]

Betty: *Parrilla es grilled.*

[Grilled is grilled.]

In example 13, dyads question the correct verb conjugation for the present perfect in Spanish.

Example 13 (Non-AR Oral - grammatical LRE)

Anna: *hum… ha oído que… oí…*
[Hum... (3rd person) heard that… I heard…]
Nick: *Has oído.*
[You have heard.]

Anna: *Has oído… ah… he oído que es muy divertido y hay conciertos, concertos aquí… y ah…tenemos… no, tienen comida de veget…*
[You have Heard… ah… I have heard that it is very fun and there are concerts, concerts (mispronounced) here… and ah… no, they have food of veget…]

In example 14, dyads are looking for the right word in Spanish for "rugs". After looking it up online they come up with the right translation.

Example 14 (Non-AR writing-focused - lexical LRE)

Pat: *Car...carpetas. ¿Qué es?*
[Fol… folders. What is?]

Jenn: *¿*Almofadas?*
[Pillows? (meant to say almohadas)

Pat: huh?

Jenn: *¿*Almohados?*
[Pillows (m)?]

Pat: *¿Enojados? Sí. Ah, no se tiene…no se tiene… muchas… muchas… color... o muebles.*
[Angry? Yes. Ah, there is no… there is no… many… many… color…]

Jenn: *Alfombra.*
[Rugs.]

Pat: Uh?

Jenn: *Alfombra. ¡Oh!*
[Rugs. Oh!]

In example 15, dyads question if they should use a verb in the subjunctive because they are using the verb "esperar" (to hope) which may require they use of the subjunctive.

Example 15 (AR writing-focused grammatical LRE)

Matt: *Esperan que ustedes *puendan ayudarnos, porque ellos están pagando para el dinero para los muebles.*
[(They) hope that you can help us, because they are paying for the money for the furniture.]

Ashley: *no… es subjuntivo.*
[No… it is subjunctive.]

Matt: *¿No es subjuntivo? Estamos usando esperar. ¿Pagar la costa?*
[It is not subjunctive? We are using to hope. Pay for the (f) coast? (meant to say costo – cost)]

The current study also shows that when participants discuss the content prior to writing, it does not suggest that participants do not engage or engage less in collaborative dialogue when writing, as shown by the high frequency of lexical LREs in the writing-focused activities, regardless of setting. These results may have occurred because although students had been exposed to the vocabulary in previous Spanish classes, they did not remember all the items, which prompted them to question their meaning during the oral and writing-focused tasks.

The higher frequency of lexical LREs in both modalities in the current study are consistent with Williams' (1999), Niu's (2009), and Ismail and Samad's (2010) results

which showed a higher focus on (lexical) meaning than (grammatical) form in their LRE data. The current study showed a frequency of 59.34% of lexical LREs in the writing modality and 40.66% in the oral activities. These findings support Niu's (2009) suggestion that written output tasks can draw learners' attention to lexical knowledge to a greater extent than just oral tasks. Ismail & Samad's (2010) justified their higher frequency of focus on meaning due to "learners' stage of developmental readiness," that is, "learner's existing knowledge systems and processing capabilities (Lightbown & Spada, 1999)" (p. 94). Thus, since one of the main goals of oral interaction is to get the meaning across, the majority of LREs in oral tasks are related to lexical issues because participants do not know how to say a certain word in the target language and ask for help. Williams (1999) suggests that this high frequency of lexical LREs can happen across all levels of proficiency. In her study, this happened due to the high number of clarification requests performed by participants to teachers and peers on the meaning of lexical items, and secondarily on their pronunciation. Yanguas and Bergin (2018) investigated LREs in video SCMC and audio SCMC and found that the main source of LRE focus was of a lexical nature.

Therefore, the claim that the nature of the task can encourage learners to focus their attention more on meaning (e.g., information-gap tasks) than on form (e.g., dictogloss task) (García Mayo, 2002a, 2002b) is supported by the current research. However, the current study also suggests that more authentic place-based tasks seem to influence the nature of LREs more than modality. However, other variables come into play during task completion (e.g., age, proficiency level and perception, L1, mode,

setting), which are not necessarily controlled in all studies on this theme. As Yanguas and Bergin (2018) suggested, contextual differences among studies make it challenging to draw conclusions.

As mentioned before, although there were differences in the lexical and grammatical LRES by modality and setting, these differences were not statistically significant. Consistent with previous studies' results (Gass, Mackey, & Ross-Feldman, 2005; Rouhshad & Storch, 2016), neither modality nor setting significantly affected the production of grammatical or lexical LREs in the current study.

LRE sub-types. The data analysis also revealed that participants in both settings and modalities focused on the same subtypes of lexical LREs (e.g., nouns, verbs, adjectives), and grammatical LREs (e.g., gender, subject-verb agreement, and verb tense). The top two lexical LRE categories in both settings and modality were nouns and verb meaning. Consistent with Ismail and Samad's (2010), and Leeser's (2004) results, the most frequent grammatical LREs were LREs related to gender (obviously non-existent in Ismail and Samad's EFL study), subject-verb agreement, and verb tense in both modalities and settings. However, these results were not significant by setting, possibly due to the overall focus on meaning over form in LREs generated by the students (Swain and Lapkin, 2001).

On the other hand, modality showed some significant differences in some of the lexical and grammatical LRE subtypes. For instance, there was significantly more noun focused LREs produced during oral activities than in the writing focused activities. In contrast, there were significantly more incidences of pronoun based LREs in the writing

focused activities than oral activities. One explanation is that dyads were instructed to use direct and indirect object pronouns on their first task. Secondly, it seems that participants used more anaphoric references in writing with the use of personal pronouns, direct and indirect object pronouns. Thirdly, the type of assignments (informal email to a friend and parents) could also have influenced the type of language they used. Other studies (e.g., Pica et al., 1989) have shown that the task type can influence negotiated interaction. Similarly, grammatical LREs of the subtype verb tense and verb aspect were significantly more frequent in the writing-focused activities than the oral activities. One reason could be that in writing, participants focus their attention to more aspects of grammar than when just speaking (Niu, 2009) due to the fact that written text requires a higher level of explicitness, a level of linguistic proficiency, and metacognitive knowledge from the writer (Schoonen, Snellings, Stevenson, & VanGelderen, 2009). In addition, a written text gives writers the ability to review the text and perfect it; giving authors more opportunities to talk about the language they are using and try to resolve linguistic problems, specifically issues related to form.

Regarding mechanical LREs, orthographic LREs in the writing modality comprised 93% of the total mechanical LREs. The non-AR groups (39/68 [57.35%]) had a higher incidence of orthographic LREs than the AR groups (29/68 [42.65%]). However, these differences were not significant between settings (x^2 (1) = .55, p < .46) or modality (x^2 (1) = .00, p < 1.00).

One possible explanation is that one of the dyads produced 80% of the orthographic LREs. Example 16 below is an example of the orthographic LREs produced

by a native speaker of Chinese and a native speaker of American English dyad. (The native speaker of Chinese typed the writing assignment).

Example 16

Amy: *…de la cotura.*
[… of the culture (misspelled)]

Melissa: ul... cul...

Amy: *oh CULtura...*
[oh culture (spelled correctly)]

Melissa: *…de Arizona, o nosotros estado, posible si no quieres estar* redundant. *Pero, ok. Entonces, we have… hacemos pensar en unas cosas divertidas para hacer.*
[… of Arizona, or we (meant to say our) state, possible if you don't want to be redundant. But, ok. Then, we have… we have to think about some fun things to do.]

Amy: *Pensar en…*
[to think about…]

Melissa: *cosas divertidos… divertidas para hacer. No hay "d" allí… T I D A S*
[Things fun (m)… fun (f) to do. There's no "d" there… T I D A S (spelling)]

Amy: *Oh, hihihi.*
[Oh, (laughs)]

Melissa: …para hacer contigo. Primero quier…
[… to do with you. First want…]

Amy: Primero cer…
[First cer (incorrect spelling)…]

Melissa: *We can use a different word… primero queremos… maybe? C e e… wait… no… quer… creo… creemos ah! C R primero atrás del primer ah… delete the first E, la primera E, la primero E, si. Primero creemos que podemos ir a… podemos P O D E M… podemos ir al Postinos para tener comida.*
[We can use a different word… first we want… maybe? C e e… wait… (thinking about the verbs want and believe in Spanish with similar sounds)

> We believe ah! C R, first behind (meant to say before) of the first ah… delete the first E, the (f) first (f) E, the (f) first (m) E, yes. First we believe that we can go to… we can P O D E M (spelling) we can go to the Postinos to have food.]

In sum, few studies (Ismail & Samad, 2010; Leeser, 2004; Niu, 2009) have investigated what specific linguistic items participants focus on during LREs. The current study adds to the research by examining these linguistic items in more detail. Overall, the analysis of the nature of LREs showed that both modalities offered opportunities for participants to collaborate and question lexical and grammatical forms. Besides leading learners to focus more on formal aspects of the language, the presence of a writing-focused activity also interrelates the oral and written aspects closely "enabling learners to talk more and to focus on language more" (Niu, 2009, p. 397).

This study substantiates other studies (Lafford, 2004) that suggest that in order to keep the conversation flowing and "not to intrude on the interlocutor's time" (p. 216) during oral communication, L2 learners avoid taking the time to monitor their speech and stop to focus on their IL. Thus, by adding a writing component to a task gives students more opportunities to focus and discuss their own language when faced with a linguistic problem, highlights the noticing of gaps in their knowledge of the target language, and creates opportunities to work out possible solutions through hypothesis formation and testing (Swain & Lapkin, 2001).

Outcome of LREs

Consistent with previous studies the majority of the LRE outcomes was solved correctly in both modalities and setting. In line with previous studies (Adams & Ross

Feldman, 2008; Azkarai & García Mayo, 2012; García Mayo & Azkarai, 2016; Ross-Feldman, 2007) more LREs were correctly resolved in the writing tasks (66.88%) than in the oral tasks (33.12%). Similarly, the writing activities also had more LREs solved correctly in the non-AR setting (52.08%) than in the AR setting (47.92%). Looking at each type of LRE, the number of grammatical LREs and lexical LREs solved correctly were very similar in both settings and modalities. About 63% of the LRES were solved correctly in the writing-focused activities (AR 62.72% and non-AR 62.87%) and 37% of the LREs were solved correctly in the oral activities (AR 37.13% and non-AR 37.28%). Thus, the setting was not a factor in the outcome of LREs correctly solved.

The LREs that were solved incorrectly or unresolved were more frequent in the writing-focused activities (54.89% and 57.32% respectively) compared to García Mayo and Azkarai's (2016) study, in which more unresolved and incorrectly solved LREs were found in the oral tasks (84.25%) than in the writing tasks. Similarly, in the current study, 19% of the LREs were left unresolved, compared to 22% in García Mayo and Azkarai's (2016) study. This low number of unresolved LREs means that most of the time the participants felt comfortable providing answers or talking to each other until they found a resolution, even if it was non-target-like. However, the AR setting had a significantly higher frequency of unresolved LREs (62.20%) compared to the non-AR setting (37.80%).

One possible explanation is that these unresolved LREs were related to the use of the technology with which neither participant had an experience. For instance, when participants were using the Google Translate AR app to translate the menu to Spanish,

they were faced with words not already part of their IL (Fernandez Dobao, 2012). As seen in example 17 below, when a lexical or grammatical hole is noticed and neither of the two participants has the knowledge to fill it, they focus on the successful communication of the message, or move on without explicitly looking for the missing L2 lexical item using technological resources:

Example 17 (AR oral)

Mark: *Jugos, si… por supuesto… que significa… el desintoxicación de manzana verde? ¿Sólo tienen bebidos, bebidas?*
[Juices, yes… of course… what does it mean… the (m) green apple detox? Do they have only drinks (m), drinks (f)?]

Betty: *¿Qué?*
[What?]

Mark: *¿Dónde está la comida?*
[Where is the food?]

In the example above, Mark asks a question about the Spanish meaning of an item in the menu for which Betty did not have an answer. Although he could have removed the phone/camera with the Google Translate AR app and see the English version, he decided to move on to another topic since knowing the meaning of that lexical item was not crucial to completing the task, leaving the LRE unresolved.

Consistent with Adams & Ross-Feldman's (2008) and Rouhshad and Storch's (2016) results, the oral vs. writing-focused modality did not significantly affect the outcome of LREs. As illustrated on Table 25, unresolved LREs compared to LREs correctly solved (Wald x^2 (1) = 2.53, p = .112), and unresolved LREs compared to LREs solved incorrectly (x^2 (1) = .19, p = .665) were not significantly different.

On the other hand, the setting significantly affected the LRE outcomes, with higher odds of participants correctly solving LREs in the non-AR than the AR setting (Wald x^2 (1) = 5.35, p = .021). One possible explanation is the higher number of words and LREs in the writing activity which led to more focus on form and text structure making students speculate more about vocabulary and grammar forms not yet part of their IL.

In sum, the analysis of LREs outcome in the current study revealed that participants correctly solved more lexical and grammatical LREs in the writing-focused activities than in the oral tasks and more LREs were solved correctly in the non-AR setting. The writing focus tasks also had the most LREs (lexical, grammatical, and mechanical) solved incorrectly and unresolved. However, few grammatical LREs were left unresolved, meaning that participants addressed their form-focused linguistic issues when working in the tasks and came to an agreement. No major differences were found in outcomes between modalities, but the AR setting led to significantly more LRES unresolved (62.20%) than the non-AR setting (37.80%).

Correction Orientation of LREs

Unlike Foster and Ohta's (2005) study, which showed self-correction and modification of learner's own utterances more frequently than correction by others, the current study results showed that the number of other correction (50.79%) and self-correction (49.20%) were very similar overall. In terms of modality, there were slightly more self-corrected LREs in the writing activities (51.75%) than in the oral activities (48.24%). In contrast, LREs corrected by others were more predominant in the writing

activities (74.79%) than in the oral activities (25.21%). Thus, modality significantly influenced the type of LRE correction with significantly more instances of other correction in writing activities (Wald x^2 (1) = 38.29, $p < .001$).

One possible explanation is that the visual representation of a text gives students in the dyad the opportunity to go back and review it and fix any mistakes (made by others more than by themselves). Leeser (2004) suggests that learners' self-initiated LREs most often occur with participant's own perceived errors, which happen more frequently in oral interactions. However, this study shows that even when students are engaged in a writing activity, they are interacting orally. The second explanation is that, since the text was a collaborative text, both participants felt the responsibility of creating a suitable final product created by their dyad.

Regarding the setting, the frequency of self-corrected LREs were very similar in both AR (49.43%) and non-AR (50.56%) settings. Similarly, LREs corrected by others were slightly higher in the non-AR (53.54%) than in the AR (46.46%) setting. However, none of these differences were significant as illustrated on Table 27 (Wald x^2 (1) = 1.29, $p = .26$).

In sum, the results of the current study suggest that correction orientation is modality dependent, with self-corrected LREs predominant in the oral activities and other-correction LREs more frequent in the writing focus activities. In addition, there was a similar number of self-corrected and other correction LREs in the data, which suggests that participants felt comfortable about collaborating and correcting each other. Furthermore, it seems that this cooperative atmosphere was a prerequisite for the

negotiated interaction to be conducive to L2 development (Sato & Ballinger, 2012) as shown by the participants' responses in the posttests.

The discussion of the results of the second research question can be found below.

RQ#2: Does the setting and LRE outcome significantly affect the ability of participants to recognize and produce the correct lexical and grammatical forms on the posttests? Is there a significant correlation between students' ability to recognize the correctness of a grammatical form and their ability to produce that form correctly on the posttests?

Recognition and Production of L2 Vocabulary and Grammatical Forms in the Posttests

The results of several studies have indicated that LREs can lead to L2 development. Adams (2007); Kim (2008), Lapkin, Swain, and Smith (2002), McDonough & Sunitham (2007), Storch (2002, 2008), Swain and Lapkin (1998, 2002), Watanabe and Swain (2008), Williams (2001), and Zeng and Takatsuka (2009) have found evidence of L2 development in tailor-made posttests of learner's independent use of the vocabulary and grammatical knowledge that was previously constructed in LREs.

As described previously, the posttests in the current study were divided into two sections – a Likert-scale grammatical section with recognition and production scores, and a lexical section with production scores. For the grammatical recognition scores, 63% of the prompts in the posttests from both settings (AR and non-AR) received a score of 3 or higher for grammatical LREs, demonstrating that overall participants recognized 63% of the structures correctly. That is, participants recognized an ungrammatical sentence as ungrammatical (incorrect), and a grammatical sentence as grammatically correct.

Regarding the three outcome types, participants recognized 68.42% of the unresolved LREs, 69.81% of the grammatical LREs that were solved correctly, and 40% of the LREs that were originally solved incorrectly. However, these scores showed no significant differences between the three outcome types in the AR setting (x^2 *(2)* = .76, *p* = .68) nor in the non-AR setting (x^2 *(2)* = 5.04, *p* = .08).

These results suggest that participants could remember the linguistic issues they had and the solutions (or lack of solutions) they came up with during the collaborative dialogue. Regarding the LREs originally solved incorrectly, participants still expressed doubt in the posttest as to the correctness of the forms. Further investigations could show if such answers came from LREs that were originally self-corrected or corrected by a peer.

Example 18 below illustrates a case of a dyad discussing and reaching an incorrect decision during their oral AR interaction, which is reproduced in their individual performance in the posttest.

Example 18 (AR oral)

Mark*:* *El otro es Nekter… escriba, después de ChopShop cuando tenemos sed podemos ir al Nekter para buscar algo a beber. Ahí hay muchos bebidos de fruta ricas y otras sabores buenos… ¿buenas?*
[The other is Nekter… write, after ChopShop when we have thirst we can go to Nekter to look for something to drink. Over there, there are many (m) drinks (m) of rich fruits and other (f) flavors good (m) good (f)?]

Betty: *Hay muchos bebidos y otras sabores buenas.*
[There are many (m) drinks (m) and other good (f) flavors.

In the example above, Mark is not sure if he should use the feminine or masculine adjective "good" with the masculine noun *sabor*. Betty gives him the incorrect solution –

a feminine adjective. In the posttest, although Betty correctly fixes the adverb and noun (*muchos bebidos* to *muchas bebidas*), they both give *sabores buenas* as the correct form.

For the production scores in the grammatical section, participants were able to produce correctly 50% of the AR grammatical prompts and 63% of the non-AR grammatical prompts. Overall, 56.52% (52/92) of the LREs received a score of 3, thus demonstrating productive knowledge. In terms of outcome types, participants produced the correct form of 42.11% (8/19) of the LREs originally unresolved, left unchanged 73.58% (39/53) of the LREs that were originally solved correctly, and 55% of the prompts were left as is, whereas they corrected 25% of the LREs that were originally solved incorrectly (5/20). Production scores increased from "solved incorrectly" (*Mdn* = 1.00) to "unresolved" (*Mdn* = 2.00) to "solved correctly" (*Mdn* = 3) with a significant difference between the LREs solved incorrectly and LREs solved correctly in the non-AR ($p = .004$), but and in the AR setting ($p = .023$).

These results suggest that participants were able to reproduce correctly over half of the original LREs. In addition, over 40% of the originally unresolved LREs and 25% of the incorrectly solved LREs were produced correctly, meaning that participants learned the correct structure in some other way (e.g., homework, friends, internet). It could also be an instance of IL variation which shows that different styles (e.g., oral vs written) may produce different results (Dickerson, 1975). For instance, during an oral dialogue participants have less time to think and use their full linguistic knowledge, whereas when writing the answers to the posttest they might not make the same mistake because they have more time to think. This would corroborate Ellis' (1992) claim that

situational factors are a source of variability. That is, when a situational factor (contexts of oral vs. written performance) changes, the learner's performance also changes. Variability can also present itself the other way, that is, grammatical accuracy may be better in spontaneous oral communication than in scores on a written grammar test. Thus, the type of task can affect learner's IL performance (Tarone, 1995).

The situational context covers a whole host of factors. When any of the situational factors changes, the learner's performance will change. For example, when a learner is asked to answer the teacher's question in class, he/she will make mistakes which he/she will never make in grammar tests, for he/she is in an urgent situation and has no time to make full use of his linguistic knowledge.

Regarding lexical issues, although there were no significant differences in production scores between the two settings, there were significant differences between the outcome data in the AR setting (x^2 *(2)* = 10.54, p = .005), more specifically between the production scores from lexical LREs solved incorrectly (*Mdn* = 1.00) and LREs solved correctly (*Mdn* = 2.00). Approximately 2/3 of the prompts in the posttests from both settings (AR and non-AR) received a score of 2 or higher for lexical LREs, demonstrating that overall, participants produced 71% of the vocabulary correctly. Looking at the types of outcomes on the posttests, participants produced the correct form for 60% of the originally unresolved lexical LREs, 77.14% of the LREs solved correctly and 16.67% of the LREs solved incorrectly. Again, these results suggest that participants may have learned or remembered the vocabulary later through other input sources (e.g.,

homework, friends, internet). Again, it could also be an instance of IL variation, mentioned above.

In addition, the opportunity learners had to use Google Translate AR (AR group) or the traditional computer Google Translator application (non-AR group) may have aided their ability to remember correctly solved lexical LREs on the posttest. Interestingly, these applications gave more opportunities for participants to question and collaborate with the language they did not previously know. For instance, with the AR application they saw words in Spanish that they were not exposed to before (or did not remember) which made them question their meanings (Example 19), or a lexical LRE that could not be solved with the help of a peer, was solved with the help of the application (Example 20). In the posttest, these lexical LREs were remembered by both members of the dyad (Example 20) or by at least one of the participants (Example 21), the initiator or the responder.

Example 19: (AR oral)

Mark: *Cómo se dice eso… miel.*
[How do you say this… honey.]

Betty: *¿Miel? Oh hum… ¿Frijoles es el mismo de like frijoles negras? y…*
[Honey? Oh hum… Beans is the same of like black beans? and…]

Betty: *Miel… no es marisco… marisco es fish…*
[Honey… it is not seafood… seafood is fish…]

Mark: *Creo que si… o… eso es marisco o no sé… algo de el océano.*
[I believe so… or… this is seafood, or I don't know… something from the ocean.]

Example 20: (AR oral)

Mark: *¿El museo, sí? Podemos hablar del museo. Yo de vez en cuando yo tengo ganas de ir ahí, pero siempre es que… ¿las maleta? ¿Baletas?*
[The museum, yes? We can talk about the museum. I sometimes feel like going there, but always is that the… suitcase? *baletas?]

Betty: *The balet? *¿Baleta?* [sic]

Mark: *No, la palabra se me olvidó, voy a buscar… el boleto. El boleto para ir al Gammage siempre está tan cara, tan caro para comprar.*
[No, I forgot the word, I'm going to look it up… the ticket. The ticket to go to the Gammage always is so expensive (f), so expensive (m) to buy.]

In this last example (21) below, Nick initiated an LRE about the meaning of seaweed in Spanish. Even though both looked up the meaning of seaweed using Google Translate, Nick was the only one who remembered it in the posttest. The LRE initiated by Anna about the meaning of avocado in Spanish was also remembered only by Nick in the posttest, even though was Anna who initiated.

Example 21: (non-AR oral)

Nick: *Hay opciones de ensalada de "seaweed".*
[There are options for seaweed salad.]

Anna: hum!
Nick: Seaweed in Spanish…
Anna: *Hay un* cucumber roll.
[There is a cucumber roll.]

Nick: *Sí, ese es muy bueno*.
[Yes, this is very good.]

Anna: *Y un* avocado roll.
[and an avocado roll.]

Nick: *okay, y ¿cuáles ingredientes están en ese opción?… ¿se dice?*
[okay, and what are the ingredients in this (m) option? Does it say?]

Anna: *Arroz.*

[Rice.]

Nick: *Algas marinas es* seaweed
[Seaweed (sp) is seaweed.]

Anna: seaweed… *alga*… oh you just said that… alga marina… *it's so random*… *alga marina, arroz, y* avocado… what's that called*?*
[seaweed… seaweed (sp)… oh you just said that… seaweed… it's so random… seaweed, rice, and avocado… what's that called?]

Nick: just avocado.

Anna: *Aguacate.*
[Avocado.]

Nick: okay.

In He et al.'s (2014) study using a mobile-based AR application for vocabulary learning, there was a highly statistically significant difference between the control (non-AR) and experimental (AR) groups in the number of words retained by each group after the treatment, suggesting that English learners acquired vocabulary better using a mobile-based AR application. Unlike He et al. (2012), in the current study, the results of the tailor-made posttests showed no statistically significant differences between posttests results by settings, indicating that both settings facilitated students making measurable gains. In addition, there were numerous incidences of vocabulary learning in which students, as they converse during the oral activities and wrote their texts during the writing-focused activities encountered a linguistic problem, worked towards solving it, and were able to remember and reproduce it in the posttests as confirmed by the high posttest scores (71% lexical) regardless of the setting. One example of these language learning episodes is illustrated in example 22. In this example, the verb *malgastar* (to

waste) is part of Jenn's IL but not part of Pat's IL. On the posttest, they both correctly translate it from Spanish to English. They also use the verb correctly in their email to María. Thus, these results suggest that learning has occurred from the posttest response to the tailor-made item, and from the written text produced by the dyad.

Example 22:

Pat: You can waste the entire day there…

Jenn: *Se puede malgastar…*
[One can waste…]

Pat: *¿Qué es malestar?*
[What is discomfort? (she was trying to say malgastar, but said malestar which means discomfort in Spanish)]

Jenn: *Malgastar*, to waste. I don't know if that's the word… but we'll look.

(Pat looks up the word)

Jenn: Is it right?

Pat: *Si, tú eres correcta.*
[Yes, you are correct.]

Also documented in previous research (Fernandez Dobao, 2012; Storch 2002; Swain, 1998), learning did not always occur in a positive way. In the current study, this was more prevalent in the AR setting (6/8 [75%]) than in the non-AR setting (5/12 [41.67%]). The examples below show that participants also remembered the incorrectly resolved LREs. In example 23, one of the participants was talking to another participant from another dyad (while his partner was typing the email) and inviting himself for dinner to her house that night. He did not know the correct word for fasting in Spanish. First, he tried a literal translation of the word in English. Then his partner gives him the correct verb and the correct form. However, he is not sure that is the correct word. He

looks up and finds the adverb form. In the posttest both participants use the incorrect adverb form **rapidando*. It is unclear why both participants chose the incorrect form; perhaps the last translation was the one that stuck in their minds. However, this incorrect resolution also represents evidence of form retention.

Example 23: (AR Writing)

Charlie: *Lisa, ¿tienes clase después de la clase?*
[Lisa, do you have class after the class?]
(Lisa from another dyad answers)

Charlie: *¿Quieres cocinar pasta *a* [sic] *noche para mí?* (laughs) *No puedo estoy *fasteando* [sic].
[Do you want to cook pasta at night for me? (laughs) I can't, I'm fasting.]

Richie: *Ayunar. Estoy ayunando.*
[To fast. I'm fasting.]

Charlie: *Oh ¿Sí? Muchas gracias, amigo*. *Ayunando.* I don't know if that's the right word.
[Oh, yes? Thank you very much, friend. Fasting.]

Richie: **Rapidando* [sic], *¿dieta?*
[*Fasting (adverb form), diet?]

There were also instances when one of the participants acquired the knowledge some other way before the posttest. Evidence of this behavior could be seen in the posttest when one participant had the correct response, and the other did not (Example 24).

Example 24 (non-AR Writing)

Melissa: *¿Un jardín? Sí, muy grande y bonita. Hum… ¿vas4*
[A garden? Yes, very big and pretty (f). Hum…. You are going to…]

Amy: ¿Vas?
[You are going to…]

Melissa: Sí, like you are going to… *vas hum…*

[You are going to hum…]

Amy: vas a…
[You are going to…]

Melissa: Enjoy? *Alegrar*, maybe? *¿Es una palabra, sí? Mucho.* Do we have to have 200 words before we leave? This is terrible! It's probably like 100.
[Enjoy, to be happy, maybe? It is a word, yes? A lot.]

Amy: 114.

Melissa: 114? No sé, podemos ir por favor it's 2:45 yo quiero ir.
[I don't know, can we go please, it's 2:45 I want to go.]

Amy: *Necesito usar el objeto directo y comparativos.*
[I need to use the direct object and comparatives.]

Melissa: *(inaudible) puedes escribir algo en los *tentos eso porque yo no sé cómo usarlos y yo tengo trabajar.*
[you can write something in the "tantos" (Spanish comparatives), this is because I don't know how to use them, and I have to work.]

Amy: *… cosas *divertísimas* [sic] (Reading)
[… very fun things.]

Melissa: ooh! maybe *vas a alegrarlo* with lo at the end.
[Ooh! Maybe "you are going to be happy it with it at the end.]

In the posttest, these two students were asked to translate the English word within a context (e.g., "*Vas a* enjoy *mucho*"). Amy had the correct answer (*divertirse*) whereas Melissa still used the unresolved LRE from the dyad dialogue (*alegrar*).

In sum, consistent with other studies that used customized posttests to trace learner's retention of lexical and grammatical items discussed during collaborative dialogue, this study has also found evidence that learners managed to remember the correct solutions (69.81%), and accurately reproduce (75.36%) them in the posttests. Comparing the overall production of grammatical (56.52%) vs lexical (70.93%) LREs,

participants performed better when producing lexical items. A possible explanation is that learners view vocabulary as the single most important area of L2 (Saville-Troike, 1984), essentially different from grammar, because a vocabulary error is more likely to lead to a problem in communication than grammatical or pronunciation errors (Simon & Taverniers, 2011).

The high grammatical recognition (63%) and production (lexical 70.93% and grammatical 56.52%) scores indicate that overall LREs were remembered and reproduced during the posttest. The results also show that when students produce LREs they benefit from their own knowledge and that of their peer. The type of activity may also influence the number and nature of LREs. This may be due to participants' perceiving the activity to be a more structured lesson activity, with concomitant elevation of attention to form, instead of a communicative activity in which they are less likely to attend to form (Williams, 2001). Even though in the current study, participants focused primarily on lexical items in both settings, there was a high frequency of grammatical LREs in oral activities. Thus, it is possible that the kinds of activities performed in the current study, such as the use of new technology and taking students outside of the classrooms, actually challenged participants to use LREs to negotiate the meaning of new, unique lexical items related to the AR or non-AR technologies used and to the new place-based environment in which they interacted.

Nevertheless, these scores do not necessarily mean that the meaning or form was "acquired," as there were no delayed posttests to examine L2 acquisition. In addition, other factors could have influenced the results such as IL variation or other sources of

input. Possible explanations include LREs that produced more turns and more discussion, had more impact, and were remembered more often. Lastly, the LRE correction orientation, whether the LRE was self-corrected or corrected by other could also have impacted the results of the posttest. A future investigation could also shed some light on why some LREs were remembered and others not. Moreover, situational factors and context affect learners' performance, which reflects the variable nature of IL competence. That is, focusing on a test question allowed participants to retrieve the right form, which did not happen spontaneously during oral interaction with a peer or vice-versa. Thus, the participant's IL variable competence manifests itself in different contexts and situations, hence the relationship between form and function in SLA is a gradual process that evolves over time.

CHAPTER 6

CONCLUSION

In the current study, interactional data were analyzed by looking at the number, nature, outcome, and correction orientation of LREs in two different settings and modalities. In addition, post-test data was analyzed to gauge students' ability to retain information from their LRE interactions.

The total number of LREs identified in the data represented the total number of linguistic inquiries generated by the dyads during the performance of two collaborative tasks. The results of this study showed a higher frequency of LREs per 100 words compared to other LREs studies using the same standard measurement. A more detailed analysis of each LRE shows that dyads had a level of engagement and interaction that occurred not only when there was a breakdown in communication but also as a form of interest and encouragement to continue the flow of the conversation. In other words, when linguistic issues arose, learners collaborated to get their meaning across, but also to build linguistic knowledge to convey the message with accuracy and precision.

The data also showed great LRE variability between dyads with some dyads producing many more LREs than others, which suggests that participants approached the task differently, had different amounts of difficulty getting their meaning across, and/or had different levels of engagement with the tasks and with the peer. Unlike other studies, participants' linguistic ability did not determine the number of LREs, and the study showed that pairing students at the same linguistic level (high or low) helped one another find solutions to their linguistic challenges.

The comparative analysis of AR and non-AR settings suggests that the technology setting affected the overall number of LREs produced. In the current study, LREs per 100 words were more frequent in the AR setting in both modalities (oral and writing-focused) than in the non-AR setting. The immersive technology gave opportunities for participants to use the language in different ways to talk about using new technology and to incorporate the physical place-based context, not possible with traditional classroom materials in a classroom or laboratory setting. In addition, the novelty of the AR applications could have enhanced their engagement and production of LREs. The mobile AR applications provided a more authentic experience with a high level of realism and visual salience that created a beneficial environment for input processing and more opportunities for output. These findings are encouraging considering current language learning settings (e.g., classroom vs lab) cannot provide many authentic experiences to learners. However, this study is one of the first studies to compare the effectiveness of collaborative work using place-based AR applications on L2 development; as a result, these observations still should be considered suggestive until confirmed by further research.

Although previous research examined the relationship between task type and setting (e.g., Foster, 1998; Gass et al, 2005 - classroom vs laboratory) no research to date had considered the relationship between task modality (oral and written) and setting. The current study's findings showed that the interaction between modality and setting was significant. Consistent with other LRE studies, participants produced more LREs in the writing-focused tasks than the oral tasks, regardless of the setting. Similarly, the nature of

LREs showed that form-focused (grammatical) LREs had a higher frequency in the writing tasks, whereas the oral tasks had a higher frequency of meaning-focused (lexical) LREs.

However, these differences were not as significant as other studies that reported a greater gap between grammatical and lexical LREs. In the current study, the number of lexical and grammatical LREs was very close in each modality. These results suggest that the place-based context could have pushed students to focus equally on form and meaning. Nonetheless, when looking at just the (AR vs. non-AR) setting, the results show a different picture. The number of lexical LREs was higher compared to the number of grammatical LREs in the AR and non-AR settings in both modalities. These results suggest that AR technology might encourage meaning-oriented tasks more than form-focused tasks (Sydorenko et al., 2019), or perhaps the non-structured nature of the tasks, compared to a more structured text editing (e.g., dictogloss) task, afforded more opportunities for dyads to talk about vocabulary. More research in this area will confirm how place-based AR affects LRE types.

Consistent with previous studies, the majority of the LRE outcomes was solved correctly in both modalities and setting. More LREs were correctly resolved in the writing tasks than in oral tasks. Similarly, the writing activities also had more LREs solved correctly in the non-AR setting than in the AR setting. Looking at each type of LRE, grammatical LREs and lexical LREs were equally solved correctly in both settings and in both modalities (in the writing focused activities and oral activities). Thus, neither setting nor modality was a factor in the outcome of LREs correctly solved. In addition,

the relatively low number of incorrectly solved LREs could imply that participants felt comfortable providing answers or talking to each other until they found a resolution, even if it were non-target-like. Unresolved LREs overall were also relatively low but had a higher frequency in the AR setting compared to the non-AR setting, perhaps due to the use of the new technology with which neither participant had an experience. Overall, these LRE outcomes were equally self-corrected or corrected by the peer.

However, this correction orientation was modality dependent and setting dependent. In other words, there was a higher frequency of self-corrected LREs in the oral activities and more LREs corrected by others in the writing-focused activities. In terms of setting, there were more LREs corrected in the non-AR setting than the AR-setting. However, the number of self-corrected LREs in the AR and non-AR were very similar and other correction LREs were slightly higher in the non-AR setting. The results suggest that participants felt comfortable about collaborating and correcting each other in both modalities and settings.

The current study has also found evidence that learners managed to remember the correct LRE solutions, and accurately reproduce them in the customized posttests. In addition, in those posttests, participants performed better when producing lexical items than grammatical forms. VanPatten (1991, 1997) argues that learners prefer to process content words in the input before anything else, thus prioritizing meaning over form. The results also show that when students produce LREs they benefit from their own knowledge and that of their peer. Nevertheless, these results do not necessarily mean that the meaning or form was "acquired," as there were no delayed posttests to examine L2

acquisition. Thus, other factors could have influenced the results such as IL variation or other sources of input.

In conclusion, this study set out to examine whether advanced Spanish L2 learners in a place-based environment would spontaneously focus on form and meaning (i.e., produce LREs) during two collaborative tasks and whether the setting (AR vs non-AR) would influence the number, type, outcome and correction orientation of their LREs in two different modalities (oral vs writing-focused). It explored the unique features that AR provides and gathered evidence to show that AR can facilitate the types of interactions believed to facilitate L2 development. The results showed that learners of Spanish engaged in place-based oral and written interactions, provided and made use of rich input, negotiated for meaning and form, self-corrected or corrected each other, and built and co-constructed knowledge in the process. Thus, it is possible that the innovative activities performed in the current study using this new technology and taking students outside of the classroom actually challenged participants to use LREs to negotiate meaning and form related to the AR tasks and the place-based environment in which they interacted.

This study also explored one of the least exploited features in MALL and showed that place-based tasks using AR mobile applications have a lot to offer in terms of oral interactions and learner collaboration (Burston, 2014). The mobility and portability of mobile AR applications can leverage the power of place-based learning and can empower students' learning experiences by heightening awareness and interaction with the real world, creating social contexts that are not typically found in the classroom.

Thus, the current study also advances the research agenda on interaction and L2 development through the analysis of LREs in this new technologically mediated context. From a cognitive perspective, there was evidence of participants focusing on form by repairing and amending their language and constructing their IL. From a sociocultural perspective, the social interaction mediated by the collaborative dialogue made learners aware of linguistic problems, monitor their own language, and rely on each other to find the right form or word by scaffolding and supporting each other. Overall, this study established an empirical baseline to allow informed future exploration of AR and peer interaction for L2 development.

Pedagogical Implications

The results of the current study indicate that the use of AR for L2 peer interaction is positive and feasible. From a pedagogical perspective, this study supports prior research findings that peer interaction in the L2 classroom can foster L2 development through peer support and assistance in resolving linguistic problems. While other studies (Sykes & Holden, 2011; Thorne, & Hellermann, 2017) have created and implemented AR games specifically for L2 development, this study used commercially available AR applications as the basis for the tasks. This means that teachers can feasibly explore other AR applications to emphasize vocabulary or specific grammar structures on which that they want their students to focus. The instructors can then incorporate these applications in tasks that are developmentally appropriate, in which students feel comfortable talking to each other, and that relate to the task topic. In other words, to maximize learning opportunities communicative tasks must make learners autonomously attend to form.

However, it is also important for instructors to think through solutions and evaluate the technology necessary to carry out AR tasks in class. Instructors must create conditions where all learners benefit from what AR can offer by providing them all access to the same type of technology.

This study confirms what other studies have suggested that it is not the type of task (e.g., jigsaw, information-gap, dictogloss) that promotes negotiated interaction but perhaps the pair dynamics (Kim and McDonough, 2008; Storch, 2008). In this case, the authentic context aided by technology encouraged learners to focus on a variety of language forms.

The LRE variability in the data with some dyads producing more LREs than others suggest that participants approach the task differently and had different levels of engagement. Although in this study pairing students with similar linguistic abilities promoted peer interaction and meaningful negotiation, teachers need to be aware of the level of engagement and learners' ability levels when putting students together in dyads.

The results of the current study also support the idea of using place-based communicative tasks (such as the ones used in this study) outside of classroom hours as an assignment or extra activities for learners to practice the language in more authentic contexts. For foreign language settings where students are exposed to the language only during classroom hours and do not have many opportunities to practice it outside of class, place-based tasks could afford students more practice with the language in authentic contexts. For L2 settings with classrooms with many students, pairing students to do

authentic communicative tasks aided by technology could be a valuable resource in promoting L2 development.

This study and other studies have shown that the majority of LREs produced in peer interaction are correctly solved. Thus, learners can provide effective corrective feedback to each other. In addition, the study also confirms that adding a writing component helps learners focus on form more than just oral interactions. Therefore, teachers could consider adding a writing component to help learners focus their attention on linguistic forms.

As Kim (2008) suggested, it is not whether collaborative tasks are more effective than individual tasks, but how to facilitate the effectiveness of collaborative tasks. To address this concern teachers should take into consideration a variety of factors that can affect the effectiveness of collaborative tasks such as patterns of interaction (age, proficiency, cultural background, L1), and context. These issues should be addressed in future research so teachers can implement them in the most effective way to maximize the benefits of collaborative work.

Theoretical and Methodological Implications

From a research perspective, this study advances the research agenda on the importance of considering peer-peer collaborative dialogue and the context in which it is performed as a mediator of IL development. The data offered in this study provide support for a theoretical orientation toward viewing dialogue aided by place-based technology as both a means of communication and a cognitive tool. This study also showed that technology-enhanced interactions (both in AR and non-AR settings),

promoted collaborative dialogue. As Loewen and Sato (2018) suggested, interactionist research needs to continue to explore the implications of such new modalities for L2 development. The significance and applicability of task setting for L2 research is an important issue that needs to be investigated further in interaction research.

Although in the current study only 12% of LREs were unresolved, future research should investigate how to train learners to provide feedback to peers (Sato & Ballinger, 2012; Sato & Lyster, 2012); more specifically learners should be trained how to pay attention to peer's linguistic inquiries, so no instances of negotiation go unnoticed or unresolved.

In terms of methodology, it would be important in future work to investigate what kinds of LREs are remembered or acquired as results in the posttest scores. In other words, are the LREs that produce more turns and more discussion the ones that have more impact and become part of students' IL? Are these LREs self-corrected or corrected by a peer? Future investigations could shed some light on why some types of LREs were remembered (or acquired) and others not.

Limitations of this Study and Future Research

As with any research, this study contained several factors that presented limitations to the interpretation of its results, such as the number of participants, lack of precise measures of participants' proficiency, novel tasks, the lack of a delayed posttest, and the unfamiliarity and unpredictability of the new technology.

Although this study used a crossover design to account for more efficient comparison of treatments with fewer participants, its results cannot be generalized to

other contexts due to the small number of participants. Further research should consider a larger sample of participants to achieve stronger conclusions. In addition, although realistic in an L2 classroom setting, the operationalization of proficiency carried solely on the instructor's intuition and prior student classroom performance is not an extremely precise way of measuring learners' proficiency levels for a research study.

As Leeser (2004a) indicates research should consider more objective measures (e.g., tests, interviews) to obtain a more accurate evaluation of learners' proficiency level. In addition, would be interesting to find out the relationship between proficiency and types of LREs used by participants, for instance, do higher proficiency level learners produce more grammatical LREs than learners with lower proficiency levels (as examined in Leeser 2004a, Williams 1999, 2001)?

Another limitation of the ability to interpret the data from this study is that it explored new types of tasks and environments, different from the traditionally structured tasks of other LRE studies; thus it was difficult to compare the current results with prior research. Future studies might repeat the procedures from the current study to corroborate or refute these findings. In addition, while some insight into learners' perception about the tasks were obtained through the recordings, future studies could employ surveys and interviews, and the use of video recordings to gain a better understanding of the relationship among learners' perceptions, the nature of interaction, and learning outcomes (Kim & McDonough, 2008). More specifically, future studies should explore how learners' view the use of AR technology in peer interaction. The inclusion of a posttest questionnaire on participants' attitudes toward the AR technology and a comparison to a

pre-questionnaire could reveal if participants' attitude gain scores correlate with their posttest scores and the number, nature, outcome and correction orientation of LREs in the AR and non-AR settings.

The way the data was collected could also have been a limitation. The use of an authentic place-based environment has its advantages, but it also carries the risk of uncontrollable variables such as time spent on tasks, since the instructor was not present when students were interacting with AR technologies outside the lab. The place-based context may have also presented uncontrollable distractions away from the tasks not found in the lab (e.g., people riding by on bikes, friends waving to them across the quad, birds flying close to them). Thus, the use of video recording could have made participants accountable for their time and attention to task.

In addition, the way the data was analyzed could also have been a limitation. This study used Swain's (1998) definition of an LRE as linguistic inquiries that students raise during a dialogue and correct each other or self-correct. However, the current study did not remove or code differently the instances when a learner (self)-initiated an LRE, perceived their own error, and immediately corrected it before any change of turn occurred. In other words, self-correction LREs that were not in the context of collaborative dialogue were not factored out separately. In future research, these immediate self-corrections should be eliminated or coded as a separate type of LRE correction orientation (e.g., immediate self-corrections). In addition, one of the questions (Question #3) on the posttests should have been counted as a recognition score, because it was a translation from Spanish to English which required participants to recognize the

form in Spanish and produce it in English. Future research should keep this distinction in mind while coding.

The use of technology (e.g. smartphones and recorders) also presented potential constraints, such as the lack of reliability of wi-fi network connections, device malfunctions, or problems with the recording process, which all happened in the current study. Furthermore, the use of new technology, such as AR presented new challenges, such as participants' maintaining and working with the superimposed virtual information. Although participants had a brief training session and were instructed to practice with the technology beforehand, more practice sections would have been beneficial to familiarize participants with the technology and the tasks.

Another important limitation of the current study is the absence of a delayed posttest to measure L2 acquisition. Future research should include delayed post-tests to investigate how place-based AR technologies impact L2 acquisition.

The current study also lacked interrater reliability due to the absence of a rater other than the researcher. The advantage of having two or more raters addressing the consistency of the LRE coding would have been important to confirm the results.

Despite these limitations, the results obtained through this exploratory study present important implications for future development and implementation of AR for L2 peer interaction. In addition, the study results can form a baseline of empirical data for additional research in this area, such as a focus on the relationship between nature, outcome and correction orientation of the LREs, and the effects of modality and setting and their interaction on LREs production during collaborative dialogue. Future research

should investigate the relationship between LREs and the CSs they contain to find out what types of LREs utilize what types of CSs.

REFERENCES

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

Adams, R. (2007). Do second language learners benefit from interaction with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 29–51). Oxford: Oxford University Press.

Adams, R., & Ross-Feldman, L. (2008). Does writing influence learner attention to form? In D. Belcher & A. Hirvela (Eds.), *The oral-literate connection. Perspectives on L2 speaking, writing, and other media interactions* (pp. 243–265). Ann Arbor, MI: The University of Michigan Press.

Alcón Soler, E. (2001). Interaction and second language learning in the institutional context of the classroom. *Estudios de Linguistica*, 1, 271-287.

Amirkhiz, S., Bakar, K., Samad, A., Baki, R., Mahmoudi, L. (2012). EFL/ESL learners' language related episodes (LREs) during performing collaborative writing tasks. *Journal of Language Teaching and Research, 4*(3), 473-479.

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

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments, 6*(4), 355–385. Retrieved from www.mitpressjournals.org/loi/pres.

Bergig, O., Hagbu, N., El-Sana, J., & *Billinghurst*, M. *(*2009). In-place 3D sketching for authoring and augmenting mechanical systems. 2009 8Th IEEE International Symposium on Mixed and Augmented Reality – *Arts, Media and Humanities, 1,* 19-28.

Blake, R., & Zyzik, E. C. (2003). Who's helping whom? Learner/heritage-speakers' networked discussions in Spanish. *Applied Linguistics, 24*(4), 519–544.

Bower, M., Howe, C., McCredie, N., Robinson, A., & Grover, D. (2014). Augmented Reality in education – cases, places, and potentials, *Educational Media International, 51*(1), 1-15.

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Burston, J. (2014). MALL: The pedagogical challenges. *Computer Assisted Language Learning, 27*(4), 344-357.

Bygate, M., Skehan, P., & Swain, M. (2001). Introduction. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 1-20). New York: Pearson Education Limited.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J.C. Richards & R. W. Schmidt (Eds.), *Language and Communication* (pp. 2-27). London: Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.

Chen, D. (2010). Enhancing the learning of Chinese with Second Life. *Journal of Technology and Chinese Language Teaching, 1*(1), 14-30.

Cheng, K., & Tsai, C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science, Education, and Technology, 22*(4), 449-462.

Choi, H. (2011). Diverse peer interaction in small group work: Analysis of low proficiency learners' language related episodes, contingent on group member proficiency levels. *Psychology*.

Choi, H., & Iwashita, N. (2016). Interactional behaviours of low-proficiency learners in small group work. *Language Learning and Language Teaching, 45*, 113-134.

Collins, A. (1988). *Cognitive apprenticeship and instructional technology* (Technical Report 6899): BBN Labs Inc., Cambridge, MA.

Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research, 58*, 9-19.

DiCamilla, F. J., & Anton, M. (1997). Repetition in the collaborative discourse of L2 learners: A Vygotskian perspective. *The Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes, 53*(4), 609-633.

Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. TESOL Quaterly, 9(4), 401-407.

Donato, R. (1994). Collective scaffolding in second language learning. In J. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.

Dörnyei, Z., & Scott, M. L. (1997). Communication strategies in a second language: Definitions and taxonomies. *Language Learning, 47*, 173–210.

Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology, 7*(3), 50–80.

Doughty, C. & Pica, T. (1986). "Information gap" tasks: Do they facilitate second language acquisition? *TESOL Quaterly*, *20*(2), 305-325.

Duchesne, S., McMaugh, A., Bochner, S., & Krause, K. (2012). *Educational psychology for learning and teaching*. Melbourne, Australia: Cengage Learning.

Dunleavy, M., Dede, C. & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, *18*, 7–22.

Ellis, R. (1985). A variable competence model of second language acquisition. *IRAL, International Review of Applied Linguistics in Language Teaching*, *23*(1), 47-59.

Ellis, R. (1992). Learning to communicate in the classroom: A study of two language learners' requests. *Studies in Second Language Acquisition*, *14*(1), 1-23.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics, 18*(1), 4-22.

Ellis, R. (2010). Second language acquisition. Teacher education and language pedagogy. *Language Teaching, 43*(2), 182-201.

Ellis, R., & Barkhuizen, G. P. (2005). *Analyzing learner language*. Oxford: Oxford University Press.

Faerch, C., & Kasper, G. (1986). Strategic competence in foreign language teaching. In G. Kasper (Ed.), *Learning, teaching and communication in the foreign language classroom* (pp. 179–193). Aarhus: Aarhus University Press.

Fernández Dobao, A. (2012). Collaborative dialogue in learner-learner and learner-native speaker interaction. *Applied linguistics*, *33*, 229-256.

Fernández Dobao, A., & Blum, A. (2013). Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System 41*(2), 365-378.

FitzGerald, E. (2012). Creating user-generated content for geolocated learning: An authoring framework. *Journal of Computer Assisted Learning, 28*(3), 195–207.

FitzGerald, E., Ferguson, R., Adams, A., Gaved, M., Mor, Y., & Thomas, R. (2014). Augmented Reality and Mobile Learning. *International Journal of Mobile and Blended Learning, 5*(4), 43-58.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics, 14*(1), 1–23.

Foster, P., & Ohta, A. (2005). Negotiation for meaning and peer assistance in second Language classrooms. *Applied Linguistics, 26*(3), 402–430.

García Mayo, M. P. (2002). The effectiveness of two-form focused tasks in advanced EFL pedagogy. *International Journal of Applied Linguistics, 12*(2), 156-175.

García Mayo, M. P., & Azkarai, A. (2016). EFL task-based interaction: Does task modality impact on language-related episodes. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Research agenda and pedagogical potential* (pp. 241–266). Amsterdam: John Benjamins.

García Mayo, M. P., & Pica, T. (2000). L2 interaction in a foreign language setting: Are learning needs addressed. *International Review of Applied Linguistics*, *38*, 35–58.

Garzón, J. & Acevedo, J. (2019). Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review, 27*(1), 244-260.

Gass, S. (1997). *Input, interaction and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell.

Gass, S., & Mackey, A. (2007). Input, interaction and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition. An introduction* (175–199). Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S., & Mackey, A. (2011). *Data elicitation for second and foreign language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition introduction to the special issue. *The Modern Language Journal, 82*(3), 299-307.

Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning, 55*(8), 575-611.

Gass, S., & Selinker, L. (2001). *Second language acquisition, an introductory course* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. & Varonis, E. (1991). Miscommunication in nonnative speaker discourse. In N. Coupland, H. Giles, & J. M. Wiemann (Eds.), *Miscommunication and problematic talk* (pp. 121-145). Newbury Park: Sage Publications.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* New York: Palgrave/Macmillan.

Gee, J. P. (2004). Situated language and learning: A critique of traditional schooling. London: Routledge.

Godwin-Jones, R. (2011). Mobile apps for language learning. *Language Learning & Technology, 15*(2), 2-11.

Godwin-Jones, R. (2013). The technological imperative in teaching and learning less commonly taught languages. *Language Learning & Technology, 17*(1), 7-19.

Godwin-Jones, R. (2017). Mobile learning: Languages, literacies and cultures. *CALICO Journal, 34*(3), 376-379.

Gor, K., & Long, M. (2009). Input and second language processing. In Ritchie, W. C., Bhatia, T. J. (Eds), *Handbook of Second Language Acquisition*, 445-472. New York: Academic Press.

Harmer, J. (2001). *The practice of second language teaching*. New York, NY: Longman.

Harmer, J. (2007). *The practice of English language teaching*. New York, NY: Longman.

Hatch, E. *(*1978a*). Second language acquisition: A book of readings*. Rowley, MA Newbury House.

Hatch, E. (1978b). Discourse analysis and second language acquisition. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 402-435). Rowley, MA: Newbury House.

Hawkins, R. (1985). Errors in the use of French past participles by foreign speakers and their implications for a model of morphology. *Lingua, 67*(2-3), 171-188.

He, J., Ren, J., Zhu, G., Cai, S., & Chen, G. (2014). Mobile-based AR application helps to promote EFL children's vocabulary study. *2014 IEEE 14th International Conference on Advanced Learning Technologies,* 431-433.

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23–48.

Herrington, J., Oliver, R., & Reeves, T. C. (2002). Patterns of engagement in authentic online learning environments. *Australasian Journal of Educational Technology 19*(1):279-286.

Holden, C., & Sykes, J. (2011). Leveraging mobile games for mobile-based language learning. *International Journal of Game-Based Learning, 1*(2), 1–18.

Hsu, H. (2019). Wiki-mediated collaboration and its association with L2 writing development: An exploratory study. *Computer Assisted Language Learning, 32*(8), 945-967.

Hsu, T. C. (2017). Learning English with augmented reality: Do learning styles matter? *Computers & Education, 106*(1), 37-49.

Hung, Y. & Higgins, S. (2016). Learner's use of communication strategies in text-based and video-based synchronous computer-mediated communication environments: opportunities for language learning. *Computer Assisted Language Learning, 29*(5), 901-924.

Ismail, L. & Samad, A. A. (2010). The effects of tasks on language-related episodes (LREs) during focus-on-form instruction. *Language Education in Asia, 1*, 87-98.

Izumi, S. (2002). Output, input enhancement, and noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541-577.

Jamshidnejad, A. (2011). Functional approach to communication strategies: An analysis of language learners' performance in interactional discourse. *Journal of Pragmatics, 43*(15), 3757-3769.

Jenks, C. J. (2009). Exchanging missing information in tasks: Old and new interpretations. *The Modern Language Journal, 93*(2), 185-194.

Jepson, K. (2005). Conversations and negotiated interaction in text and voice chat rooms.

Language Learning & Technology, 9(3), 79–98.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., and Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition*. Austin, Texas: The New Media Consortium.

Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal*, 92, 114–130.

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, *12*, 211-234.

Klopfer, E. (2008). *Augmented learning: Research and design of mobile educational games.* MIT Press, Cambridge.

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

Krashen, S. (1981). The fundamental pedagogical principle in second language teaching. *Studia Linguistica*, *35*(1-2), 50-70.

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition.* Oxford Pergamon Press.

Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.

Krueger, C. (2014). *Exploring the relationship between the instructional use of mobile technology and social media and student foreign language learning* (Order No. 3582098). Available from ProQuest Dissertations & Theses Global.

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL: The Journal of EUROCALL*, *20*(3), 271-289.

Lafford, B. (2004). The effect of context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition, 26,* 201-26.

Lantolf, J. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Lantolf, J. P., & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development*. Oxford, United Kingdom: Oxford University Press.

LaPierre, D. (1994). Language output in a cooperative learning setting: Determining its effects on second language learning. Unpublished master's thesis, Ontario Institute for Studies in Education, University of Toronto, Ontario, Canada.

Lapkin, S., Swain, M., & Smith, M. (2002). Reformulation and the learning of French pronominal verbs in a Canadian French immersion context. *The Modern Language Journal, 86*(4), 485-507.

Lee, J. F., VanPatten, B. (1995). Making communicative language teaching happen. New York: McGraw-Hill.

Leeser, M. J. (2004a). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research, 8*(1), 55–81.

Leeser, M. J. (2004b). The effects of topic familiarity, mode, and pausing on second language learners' comprehension and focus on form. *Studies in Second Language Acquisition, 26*(4), 587-615.

Leeser, M. J. (2008). Pushed output, noticing, and development of past tense morphology in content-based instruction. *The Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes, 65*(2), 195-220.

Liu, T Y. (2009). A context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning, 25*(6), 515–527.

Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*(2), 630–643.

Liu, P. H. E., & Tsai, M. K. (2013). Using augmented-reality-based mobile learning material in EFL English composition: An exploratory case study. *British Journal of Educational Technology, 44*(1), 221-254.

Long, M. (1981). Input, interaction and second language acquisition. *Annals of the New York Academy of Sciences, 379*, 259-278.

Long, M. H. (1982). Adaptation to the learner: The negotiation of comprehensible input in conversation between native language speakers and learners. *LiLi, Zeitschrift Fur Literaturwissenschaft Und Linguistik, 12*(45), 100-119.

Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126-141.

Long, M. (1985). Input and second language acquisition theory. In S. Gass & C. Madden

(Eds.), *Input in second language acquisition* (pp. 377-393). Rowley, Mass: Newbury House.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam, Netherlands: John Benjamins.

Long, M. (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. 413–468). New York, NY: Academic Press.

Long, M. (2014). Psycholinguistic underpinning: A cognitive-interactionist theory of instructed second language acquisition. *Second language acquisition and task-based language teaching* (pp. 30-62). John Wiley and Sons Inc.

Loewen, S., & Basturkmen, H. (2005). Interaction in group writing tasks in genre-based instruction in an EAP classroom. *Journal of Asian Pacific Communication, 15*(1), 171-189.

Loewen, S. & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285-329.

Loewen, S., & Wolff, D. (2016). Peer interaction in F2F and CMC contexts. *Language Learning and Language Teaching, 45*, 163.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focuses instruction. *Studies in Second Language Acquisition, 26*(3), 339-432.

Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition 19*(1), 37–66.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching, 46*(1), 1-40.

Mackey, A. (2007). Introduction: The role of conversational interaction in second language acquisition. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 1-26). Oxford, United Kingdom: Oxford University Press.

Mackey, A. (2012). *Input, interaction and corrective feedback in L2 classrooms*. Oxford, United Kingdom: Oxford University Press.

Mackey, A., Abbuhl, R., & Gass, S. (2012). Interactionist approach. In S. Gass & A. Mackey (Eds.), *The Rouledge handbook of second language acquisition* (pp. 7-

23). New York, NY: Routledge.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*(4), 471-497.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407–472). Oxford: Oxford University Press.

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.

McDonough, K., & Mackey, A. (2000). Communicative tasks, conversational interaction, and linguistic form: An empirical study of Thai. *Foreign Language Annals, 33*(1), 82–91.

McDonough, K., & Sunitham, W. (2009). Collaborative dialogue between Thai EFL learners during self-access computer activities. *TESOL Quarterly, 43*(2), 231-254.

Miller, J. (2003). *Audible difference: ESL and social identity in schools*. Bristol, United Kingdom: Multilingual Matters.

Muñoz, C. (2007). Age-related differences and second language learning practice. In R. DeKeyser (Ed.), *Practice in a second language* (pp. 229-255). Cambridge, United Kingdom: Cambridge University Press.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*(4), 719–758.

Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness*, 18(3-4), 384–402.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–428.

Nunan, D. (1991). Communicative tasks and language curriculum. *TESOL Quarterly, 25*(2), 279-295.

Ohta, A. (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), 93-121.

Ohta, A. (2000). Rethinking recasts: A learner-centered examination of corrective feedback in the Japanese classroom. In J. Hall & L. Verplaeste (Eds.), *The*

construction of second and foreign language learning through classroom interaction (pp. 47–71). Mahwah, NJ: Erlbaum.

Pavlenko, A. (2002). Postcultural approaches to the study of social factors in second language learning and use. In V. Cook (Ed.), *Portraits of the L2 user* (pp. 275-302). Clevedon, Avon, United Kingdom: Multilingual Matters.

Pellettieri, J., (2000). Negotiation in cyberspace: The role of chatting in the development of grammatical competence. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practices* (pp. 59–86). New York, NY: Cambridge University Press.

Perry, B. (2015). Gamifying French language learning: A case study examining a quest-based, augmented reality mobile learning-tool. *Procedia-Social and Behavioral Sciences, 174*, 2308-2315.

Pew Research Center. (2017, June 28). *10 Facts about smartphones as the iPhone turns 10*. https://www.pewresearch.org/fact-tank/2017/06/28/10-facts-about-smartphones/

Philp, J., Adams, R., & Iwashita, N. (2013). *Peer interaction and second language learning*. New York, NY: Taylor & Francis.

Philp, J., Walter, S., & Basturkmen, H. (2010). Peer interaction in the foreign language classroom: What factors foster a focus on form? *Language Awareness, 19*(4), 261–279.

Piantadosi, S. (2005) Crossover designs. In *Clinical Trials: A Methodologic Perspective* (pp. 684-697). Hobaken, NJ: John Wiley and Sons, Inc.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, 44, 493–527.

Pica, T. (2005). Classroom learning, teaching, and research: A task-based perspective. *The Modern Language Journal, 89*(3), 339-352.

Pica, T., Doughty, C., & Young, R. (1986). Making input comprehensible: Do interactional modifications help? *ITL, Review of Applied Linguistics, 72*, 1-25.

Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition, 11*, 63-90.

Pica, T., Kanagy, & Falodun, J. (1993). Choosing and using communication tasks for

second language instruction. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: integrating theory and practice* (pp. 48-65), Clevedon, UK: Multilingual Matters.

Pica, T., Lincoln-Porter, F., Paninos, D. & Linnel, J. (1996). Language learners' interaction: how does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly, 30*(1), 59-84.

Ross-Feldman, L. (2007). Interaction in the L2 classroom: Does gender influence learning opportunities? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 52–77). Oxford: Oxford University Press.

Rouhshad, A., & Storch, N. (2016). A focus on mode: Patterns of interaction in face-to-face and computer-mediated contexts In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 267–289). Amsterdam: John Benjamins.

Russel, J. & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 131-164). Amsterdam: Benjamins.

Saltan, F., & Arslan, O. (2017). The use of augmented reality in formal education: A scoping review. *EURASIA Journal of Mathematics, Science & Technology Education, 13*(2), 503-520.

Santos, C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning Technologies, 7*(1), 38-56.

Sato, C. J. (1986). *The syntax of conversation in interlanguage development.* Available from Linguistics and Language Behavior Abstracts (LLBA).

Sato, M. (2007). Social relationships in conversational interaction: Comparison of learner-learner and learner-NS dyads. *JALT Journal, 29*(2), 183-208.

Sato, M., & Ballinger, S. (2012). Raising language awareness in peer interaction: A cross-context, cross-method examination. *Language Awareness, 21*(1-2), 157-179.

Sato, M., & Ballinger, S. (2016). Understanding peer interaction: Research synthesis and directions. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 1–30).

Amsterdam: John Benjamins.

Sato, M., & Lyster, R. (2007). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 123–142). Oxford: Oxford University Press.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition, 34*(4), 591-626.

Savignon, S. J. (1983). *Communicative competence: Theory and classroom practice.* Reading, MA: Addison-Wesley.

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly, 18*(2), 199-219.

Schmidt, R., & Frota, S. (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.

Schoonen, R., Snellings, P., Stevenson, M., & van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. M. Manchon (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 77–101). Bristol, UK: Multilingual Matters.

Selinker, L. (1972). Interlanguage. *IRAL, 10*, 209–230.

Shekary, M., & Tahririan, M. H. (2006). Negotiation of meaning and noticing in text-based online chat. *The Modern Language Journal, 90*, 557-73.

Simon, E. & Taverniers, M. (2011). Advanced EFL learners' beliefs about language learning and teaching: A comparison between grammar, pronunciation, and vocabulary. *English Studies, 92*, 896-922.

Smith, M. S. (1986). Comprehension versus acquisition: Two ways of processing input. *Applied Linguistics, 7*(3), 239-256.

Smith, B. (2003a). Computer-mediated negotiated interaction: An expanded model. *The Modern Language Journal*, *87*, 38–57.

Smith, B. (2003b). The use of communication strategies in computer-mediated

communication. *System, 31*, 29-53.

Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition, 26,* 365–398.

Sobel, D. (2006). Place-based education: Connecting classrooms & communities (2nd ed.). Great Barrington, MA: Orion Society.

Solak, E., & Cakir, R. (2015). Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning. *Journal of Educators Online, 12*(2), 50-72.

Spada, N, & Lightbown, P. M. (1999). Instruction, first language influence and developmental readiness in second language acquisition. *The Modern Language Journal, 83*(1), 1-22.

Stetsenko, A., & Arievitch, I. (1997). Constructing and deconstructing the self: Comparing post-Vygotskian and discourse-based versions of social constructivism. *Mind, Culture, and Activity, 4*, 160-173.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52*(1), 119–158.

Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness,* 17(2), 95–114.

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior, 86,* 77–90.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H.G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.

Swain, M. (1997). Collaborative dialogue: Its contribution to second language learning. *Revista Canaria de Estudios Ingleses, 34*, 115-132.

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

Swain, M. (2000). The output hypothesis and beyond. In J. Lantolf (Ed.), Sociocultural theory and second language learning (pp. 97–114). Oxford: Oxford University Press.

Swain, M. (2001). Integrating language and content teaching through collaborative tasks. *The Canadian Modern Language Review, 58*(1), 44-63.

Swain, M. (2005). The Output Hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 471–484). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M. (2006). Verbal protocols: What does it mean for research to use speaking as a data collection tool? In M. Chalhoub-Deville, C. Chapelle, & P. Duff (Eds.), *Inference and Generalizability in Applied Linguistics: Multiple Perspectives* (pp. 97-113). Amsterdam: John Benjamins.

Swain, M., & Deters, P. (2007). 'New' mainstream SLA theory: Expanded and enriched. *The Modern Language Journal, 91*(5), 820–836.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371-391.

Swain, M. & Lapkin, S. (1998) Interaction and second language learning: Two adolescent French immersion students working together. *Modem Language Journal, 83*, 320-337.

Swain, M., & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research, 4*(3), 251–274.

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 assessment* (pp. 99–118). London: Pearson Education.

Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research, 37*(3), 285–304.

Swain, M., & Lapkin, S. (2005). The evolving sociopolitical context of immersion education in Canada: Some implications for program development. *International Journal of Applied Linguistics, 15*(2), 169-169.

Swain, M., & Watanabe, Y. (2013). Languaging: Collaborative dialogue as a source of

second language learning. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). Oxford: Wiley-Blackwell.

Sydorenko, T., Hellermann, J., Thorne, S. L., & Howe, V. (2019). Mobile augmented reality and Language-Related episodes. *TESOL Quarterly, 53*(3), 712-740.

Tannen, D. (1989). *Talking voices. Repetition, dialogue, and imagery in conversational discourse.* New York, NY: Cambridge U Press.

Tarone, E. (1980). Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning, 30*, 417–431.

Tarone, E. & Liu, G-Q. (1995). Situational context, variation, and second language acquisition theory. In G. Cook, & B. Seidlhofer (Eds.), *Principle and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson* (pp. 107-124). Oxford: Oxford University Press.

Thorne, S. L. (2010). The 'intercultural turn' and language learning in the crucible of new media. In F. Helm & S. Guth (Eds.), *Telecollaboration 2.0 for language and intercultural learning* (pp. 139-164). Bern: Peter Lang.

Thorne, S. L. (2013). Language learning, ecological validity, and innovation under conditions of superdiversity. *Bellaterra Journal of Teaching & Learning Language & Literature, 6*(2), 1-27.

Thorne, S. L., & Hellermann, J. (2017). Mobile augmented reality: hyper contextualization and situated lLanguage usage eEvents. *Proceedings of the XVIII International CALL Conference: CALL in Context* (pp. 721-730). City: Publisher.

Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning, 21*(3), 217–228.

Toth, P. (2008). Teacher- and learner-led discourse in task-based grammar instruction: Providing procedural assistance for L2 morphosyntactic development. *Language Learning, 58*(2), 237–283.

Ushioda, E., & Dornyei, Z. (2017). Beyond global English: Motivation to learn languages in a multicultural world: Introduction to the special issue. *The Modern Language Journal, 101*(3), 451-454.

VanPatten, B. (1991). The foreign language classroom as a place to communicate. In B. Freed (Ed.), *Foreign Acquisition Research and the Classroom* (pp. 54-73). Boston: D.C. Heath.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition.* Norwood, NJ: Ablex Publishing Corp.

Varonis, E. M., & Gass, S. (1985). Non-Native/Non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6*(1), 71-90.

Vilamil, O. S., & de Guerreiro, M. C. M. (1996). Peer revision in the L2 classroom: Social-cognitive activities, mediating strategies, and aspects of social behavior. *Journal of Second Language Writing, 5*(1), 51-75.

Virtual Reality Society. (n.d.) *What is virtual reality?* Retrieved from https://www.vrs.org.uk/virtual-reality/what-is-virtual-reality.html

Vygotsky, L. (1934). Thinking and speech. In R.W. Rieber & A.S. Carton (Eds.), *The collected works of L.S. Vygostsky, Volume 1: Problems of general psychology* (pp. 39-285). New York: Plenum Press.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Vygotsky, L.S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Wang, Y. H. (2017). Exploring the effectiveness of integrating augmented reality-based materials to support writing activities. *Computers & Education, 113*, 162-176.

Wasko, C. (2013). What teachers need to know about augmented reality enhanced learning environments. *TechTrends, 57*(4), 17-21.

Watanabe, Y., & Swain, M. (2008). Perception of Learner Proficiency: Its Impact on the Interaction Between an ESL Learner and Her Higher and Lower Proficiency Partners. *Language Awareness, 17*(2), 115.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics, 8*(2), 95-110.

Wigglesworth, G., & Storch, N. (2012). Feedback and writing development through collaboration: A sociocultural approach. In R. Manchon (Ed.), *L2 writing development: Multiple perspectives* (pp. 69–100). Berlin : Mouton De Gruyter.

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

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

Williams, J. (2008). The speaking-writing connection in second language and academic literacy development. In D. Belcher & A. Hirvela (Eds.), *The oral-literate connection. Perspectives on L2 speaking, writing, and other media interactions* (pp. 10–25). Ann Arbor, MI: The University of Michigan Press.

Willis, J. (1996). *A framework for task-based learning*. Harlow: Longman.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, 62, 41-49.

Wurmser, Y. (2018, June 18). *Mobile time spent 2018*. eMarketer. https://www.emarketer.com/content/mobile-time-spent-2018

Yanguas, I., & Bergin, T. (2018). Focus on form in task-based L2 oral computer-mediated communication. *Language Learning & Technology, 22*(3), 65.

Yilmaz, Y. (2011). Task effects on focus on form in synchronous computer-mediated communication. *The Modern Language Journal*, *95*(1), 115–132.

Yilmaz, Y., & Granena, G. (2010). The effects of task type in synchronous computer-mediated communication. *ReCALL*, *22*(1), 20–38.

Young, A. I., & Tedick, D. J. (2016). Collaborative dialogue in a two-way Spanish/English immersion classroom: Does heterogeneous grouping promote peer linguistic scaffolding? *Language Learning and Language Teaching, 45*, 135.

Yoshida, R. (2008). Learners' perception of corrective feedback in pair work. *Foreign Language Annals*, *41*(3), 525–541.

Zeng, G., & Takatsuka, S. (2009). Text-based peer–peer collaborative dialogue in a computer-mediated learning environment in the EFL context. *System*, *37*(3), 434-446.

Zhang, J., Liu, T. C., Sung, Y. T., & Chang, K. E. (2015). Using augmented reality to promote homogeneity in learning achievement. *2015 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design*, 1-5.

Zheng, D., Young, M. F., Wagner, M. M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *Modern Language Journal*, *93*(4), 489-511.

Zheng, D., Schmidt, M., Hu, Y., Liu, M., & Hsu, J. (2017). Eco-dialogical learning and translanguaging in open-ended 3D virtual learning environments: Where place, time, and objects matter. *Australasian Journal of Educational Technology*, *33*(5), 107-122.

APPENDIX A

INVENTORY OF CSS WITH DESCRIPTIONS AND SELECTED EXAMPLES FROM THIS STUDY

(Adapted from Dornyei & Scott, 1997; Hung & Higgins, 2016; Lafford, 2004; Smith, 2003b)

APPENDIX A

	Code	Communication Strategy	Description	Example
	1	**Interactional Strategies**		
1	EC	Explicit correction	Explicit correction of learner's utterance	"No *le* vi, *lo* vi."
2	CR	Clarification request	Asking for explanation of unfamiliar terms or messages.	"¿Qué?"
3	FC	Confirmation Check	Repeating the trigger in a rising intonation to ensure one heard something correctly, or using a first language term or asking a full question to ensure the correctness of the input comprehension	"¿Qué comida? ¿Qué quiero comer a Chopshop?"
4	CC	Comprehension Check	Asking questions to ensure one's messages are understood.	"¿Me entiendes?"
5	DA	Direct appeal for help	Asking for assistance by an explicit question concerning a gap of one's knowledge in the target language.	"¿Cómo se dice *on top* en español?"
6	IA	Indirect appeal for help	Trying to elicit help from one's interlocutor by indicating the problems either verbally or nonverbally.	"Es muy ah… no sé cómo se dice…*quiet*."

APPENDIX A

	Code	Communication Strategy	Description	Example
	1	**Interactional Strategies**		
7	IE	Input Elicitation	Expressing explicitly or passing signals to encourage one's interlocutor to continue talking.	"Sí, hum… ¿que más?"
8	FU	Feigning Understanding	Pretending to understand the preceding message to carry on the conversation.	A: Y ahora tenemos que mirar a los menús con la AR, ¿sí? B: ¿Con qué? A: ¿La AR? B: Sí,sí, sí. A: Pónelo en tu celular es AR para que podemos ver el menú. B: Sí. Hum… A: Así… tenemos que hacerlo.
9	NU	Expressing nonunderstanding	Learners express verbally or nonverbally something he/she did not understand	"No comprendo."
10	IS	Inferential Strategies	Asking questions or making comments based on established information to test one's hypothesis of the preceding message, show one's current state of understanding, or gain new information.	A: No lo vi en Gammage. Pero en mi secundaria hay una musical para… B: ¡Ah! En tu escuela. A: Sí, en mi escuela secundaria. B: ¿Y estuviste en parte? ¿estabas parte de este programa? A: Sí. B: ¿Eres un actriz?
11	FR	Framing	Marking the shifts of topics.	"Primero hablamos de eventos."

APPENDIX A

	Code	Communication Strategy	Description	Example
	1	**Interactional Strategies**		
12	SM	Verbal Strategy Markers	Using verbal marking phrases such as "no sé" to indicate the use of strategy or less accurate form.	"Pues nunca habi… o habíamos, sí? No sé…"
13	OM	Omission	Leaving an unknown word as a gap and carrying on as if it has been said with the hope that the interlocutor can fill the gap by context.	A: Yeah, sí, mucho. Pero ¿es posible ponerlo en el piso? Contra, like, tú sabes cómo like tenían… B: Sí, yo sé que estás diciendo.
14	TG	Time-gaining Strategies	Using fillers such as "umm..." or repeating interlocutor's words to fill pauses in order to maintain conversation at times of thinking.	"Creo que la cocina necesita mucho trabajo para ser un lugar perfecto para preparar comidas, pero ahora…uhm, ya tiene una…uhm, una…"
15	RC	Response-confirm (response-repeat)	Confirming what the interlocutor has said or suggested	A: Era... Es muy caro, ¿no? B: Sí, es muy caro, no podemos ir.
16	RR	Response-repair (response rephrase)	Providing initiated self-repair by rephrasing the trigger or other as a response.	A: ¿Qué quieres comer? B: ¿Qué? A: ¿Qué quieres para cenar?
	2	**Compensatory Strategies**		
17	PA	Circumlocution (paraphrase)	Exemplifying, illustrating, or describing the features of the target object or action.	"Creo que sí… o… eso es marisco o no sé…algo de el océano."

APPENDIX A

	Code	Communication Strategy	Description	Example
	1	**Interactional Strategies**		
18	AO	Approximation	Using one single substitute term with which the target term shares semantic features	A: Ah… me gusta la… ¿Cómo se dice smoothie? B: jugo de… ¿bebida de fruta?
19	AW	Use of All-purpose Words	Using a genera; "empty lexical term to a replace a specific term to compensate for vocabulary deficiency.	A: ¿Un cosa a chopshop recomienda tienes? Tiene. B: ¿Cosa? A: Yeah, ¿qué comida?
20	LT	Literal Translation	Translating a first language term literally to a target language term.	"¿Cómo se empieza una letra?"45555 (letter)
21	SP	Self-rephrasing	Paraphrasing, restructuring, or repeating one's own utterance. Sometimes new information may be added to the repetition.	"No sé, creo que necesito hacer mi... construir mi cena."
22	FO	Foreignizing	Using L1 word by adjusting to L2 phonology, morphology, or both.	"El lugar donde performan, ¿performar?"
23	DR	Derivationally related word	Word(s) derivationally related to L2 word by form and meaning	"Los cocinadores (cook) en ChopShop son personas con hum…"
24	SS	Similar sounding words	Use of lexical item whose form is similar to target	*"Hice muchos divertidos? Diportivos?"

APPENDIX A

	Code	Communication Strategy	Description	Example
	3	**Reduction Strategies**		
25	WC	Word coinage	Creating a novel L2 word by L2 rule formation patterns	*"Yo interesanto en otros apartamentos"
26	MA	Message Abandonment	Leaving a message unfinished due to an inability to cope with language difficulty.	"Aperitivos allí, como ah… es bueno, tenemos que hablar sobre que ellos tienen opciones vege…. vegetarianas también."
27	RS	Restructuring	Replacing the original message by a new one because of language difficulties.	"De costa de los Estados Unidos, ¿la otra costa? Como el… no sé cómo se dice… ¿al derecho de aquí? hay norte, sur, y derecha e izquierda, no sé."
	4	**Focus-on-form Strategies**		
28	SR	Self-repair	Making self-initiated corrections in one's own speech.	"Podemos ir uno… un día juntos."
29	MT	Meta-talk	Using the target language to reflect on one's own or interlocutor's use of the target language	¡Ah! ¡Esta es la palabra – tazón!
30	RT	Retrieval	Saying a series of incomplete or wrong forms or structures before reaching the optimal form	"Hay mucho tofu que no fue cocinando… cocinada? Cocinado muy bien…"
31	AC	Own Accuracy Check		"¿Meriendas?"

APPENDIX A

	Code	Communication Strategy	Description	Example
	5	**Sociocultural Strategies**		
32	SF	Social Formula	Using fixed patterns for social purposes such as greetings, leave takings, or apology.	"Eres correcta, lo siento."
33	CS	Code-switching	Using L1 words in the L2 speech for purposes such as to show familiarity or to negotiate or establish intersubjectivity.	"El *cake pop* ¿rosada? no sé."
	6	Recast	A reformulation of all or part of a learner's immediately preceding utterance in which a nontarget-like word is replaced by the corresponding target language form with a focus on meaning not language as an object.	A: Sí, necesita un otro silla. B: ¿Otra silla?

*Examples taken from Hung and Higgins (2016), not found in this study.

APPENDIX B

INVENTORY OF LRES WITH DESCRIPTIONS AND SELECTED EXAMPLES FROM THIS STUDY

APPENDIX B

		Type: Lexical				
Subtype	Name	Example	Outcome	Code	Correction Orientation	Code
1	LXNoun	A: ¿Cómo se dice *on top* en español? B: Encima A: Encima de… del lago son muy bonitas y podemos sacar fotos.	Solved correctly	2	Other	2
2	LXVerb	A: No… ¿no *querrás* hacer nada? B: posible, tienes todo el día a descansar.	Unresolved	1	other	2
3	LXPreposition	"Necesitas un evento que podemos hacer *con* María… *para* María."	Solved incorrectly	3	Self	1
4	LXPhrase	A: *A la parrilla*… ¿no sé qué es eso? B: parrilla es grilled.	Solved correctly	2	Other	2
5	LXPronoun	"*Nos*… ¿*te* daré una tour?"	Solved correctly	2	Self	1
6	LXAdjective	A: ¿Es querida? querida maría. B: Creo… no sé.	Unresolved	1	Other	2
7	LXAdverb	A: ¿what's the opposite of *después?* B: antes… antes de estas actividades podemos cenar a ChopShop.	Solved correctly	2	Other	1
8	LXConjunction	"No soy vegana, *pero porque* me gusta carne."	Solved correctly	2	Self	1
9	LXArticle	"Podemos ir *uno*… *un* día juntos."	Solved correctly	2	Self	1
10	LXOther	(not found)				

APPENDIX B

		Type: Grammatical				
Subtype	Name	Example	Outcome	Code	Correction Orientation	Code
11	GNumber	"Solo los… el arroz."	Solved correctly	2	Self	1
12	GPerson	"Dónde *está* las...dónde *están* las opciones para comer?	Solved correctly	2	Self	1
13	GGender	A: Ahí hay muchos bebidos de fruta ricas y otras sabores ¿buenas? ¿Buenos? B: Hay muchos bebidos y otras sabores buenas.	Solved incorrectly	3	Other	2
14	GVerbTense	A: ¿Que lo construyo? Si, que lo construyo… hizo… B: no… hacer…. A: Sí, pero es un verbo y después de un verbo es un verbo no conjugated. B: Sí, él lo construyó, no sé.	Unresolved	1	Other	2
15	GAspect	A: ¿y estuviste en parte? ¿estabas parte de este programa? B: Sí.	Solved incorrectly		Self	

APPENDIX B

		Type Grammatical				
Subtype	Name	Example	Outcome	Code	Correction Orientation	Code
16	GMood	A: Después de llegar del aeropuerto tiene toda la noche a descansar. B: Sí. ¿Llegas? ¿Después puede ser subjuntivo? A: después de llegar… ¿subjuntivo por qué? B: después es como una palabra que puede indicar subjuntivo. A: ¿sí? ¿porque no es like cierto ahora? B: sí A: llegues, aquí B: creo que es cierto… no sé. Llegas aquí…	Unresolved		Other	
17	GOther	"Además, tenemos una mesa de café para que pongan… para que se pongan tazas."	Solved correctly		Self	1
		Type: Mechanical				
0	orthographical	A: Y también el diseña de este… B: como dis… ñ?	Solved correctly		Other	1
1	pronunciation	A: Sí, ñ pona una A. "El primero actividade… actividad."	Solved correctly		Self	1

APPENDIX C

LANGUAGE AND TECHNOLOGY BACKGROUND (PRE) QUESTIONNAIRE

APPENDIX C

The information you provide here will help to inform us about your language and technology experience. The following information is **not** going to be shared with anyone and it will only be used for research purposes. By completing and submitting this questionnaire, you are agreeing to participant in a research study. Your identity will be kept in secret. Upon completion of the study, all identifying information will be deleted and only the demographic data will be kept. Please answer all questions as fully as possible.

Name: **Male ( )** **Female ( )**
Age: **ASU Email:**

SECTION I: Language Background

1) Are you a native speaker of English ( ) Spanish ( ) Other ( ) Specify: _____

2) What languages do you know? Please rate in order of proficiency in the language.
1=most proficient 5=least proficient

1	2	3	4	5

3) What language(s) are you studying:

Spanish	()	Arabic	()	Romanian	()
German	()	Chinese	()	Portuguese	()
French	()	Japanese	()	Vietnamese	()
Italian	()	Russian	()	Other	() Specify:

_____

4) Major:
() Language: _____
() Arts & Sciences
() Business
() Engineering
() Other Specify: _____

5) Do you plan to take this language beyond the two-year foreign language requirement?

() Yes () No () I do not know

6) Why are you taking this course? (choose all that apply)
   () Communicate better with family members, relatives, or friends
   () Satisfy a degree language requirement
   () Use in present or future business or employment
   ( ) Use when visiting a country for tourism/vacation
   ( ) Other reason (specify):
   _____
   _____

SECTION II: Technology Background

7) Do you have any of the devices below? Please check all that apply.

   ( ) Android 7 or below ( ) Android 8 or above - Specify model:_____
   ( ) iPhone 5 or below ( ) iPhone 6 or above - Specify model: _____
   ( ) iPad or tablet ( ) Other phone: _____
   ( ) Windows phone

XR is a term to refer to augmented (AR), mixed, or virtual reality (VR) technologies. These include virtual reality immersive games (i.e., 3D with a headset) and augmented reality smartphone apps that create layers over real objects (i.e., selfie filters, *Pokémon Go*, wayfinding, QR codes), etc.

8) Have you used any *Augmented Reality* apps (i.e., selfie filters, *Pokémon Go*, wayfinding, QR codes), on your phone? yes () no ()

If so, which ones: _____

9) If you haven't used any of these technologies, please indicate why not: (check all that apply)

   () Never heard of these.
   () I do not know how to use it.
   () The hardware is too expensive.
   () The current quality of the experience is not good enough for me to use.
   () There are not enough choices for apps and games in this space.
   () I do not see the purpose of using this technology.
   ( ) Other: _____

APPENDIX C

10) Please indicate the extent to which you agree or disagree as follows:

5 = Strongly Agree, 4 = Agree, 3 = Disagree, 2 = Strongly Disagree 1 = No Opinion

If this technology (AR) were available to me in my language class, I would like to use it for:

1) support language training (acquire academic/technical language relevant to my major or career via labeling of objects, practicing processes associated with my field) 5 4 3 2 1
2) learn vocabulary (colors, body parts, house items, clothes, etc.) 5 4 3 2 1
3) interact with the Hispanic community via an AR game 5 4 3 2 1
4) Interact with virtual objects from the target culture 5 4 3 2 1
5) Learn about the history, culture, artifacts of the target country 5 4 3 2 1
6) Other:_____

_____

11) How often would you use this technology (check all that apply):

() as a requirement only (homework/assignment)
() at my own time to improve my language skills
() at my own time, just for fun
() other

12) Would you like to add any other comments?

_____

_____

_____

Thank you for your time and participation!

APPENDIX D

COLLABORATIVE TASKS

APPENDIX D

Rules

1) Always stay with your partner.
2) You have 20 minutes to complete each phase of the game (oral and writing).

Tasks

Task #1: "La noche es joven" (The night is young)

Context: Your friend from Mexico is coming to spend this weekend with you. She's never been to Arizona. You are excited to introduce her to your classmate, and you both would like to take her out to eat and have some fun afterwards. However, there is one problem, one of you is a vegetarian and the other loves meat! One of you likes sports and the other is more into music! You have one thing in common – you both don't have much money to spend. Together, you must decide where you are going to take your Mexican friend to eat and where you are going afterwards.
Using the recorder, record your conversation during the entire time.

Environment: AR

Time duration: Phase 1 (oral): 20 minutes, Phase 2 (writing): 20 minutes

Directions - Make sure you talk to each other during the entire time. Make sure your recorder is on and recording. You have 20 minutes to complete Phase 1.

Phase 1: (*Bettar AR app*) a) Login into the app. Click the 3 bars on the top left and choose the "events" tab. Make sure the "camera" icon (bottom left) is selected. Point your phone camera at buildings in front of you. Green tabs will appear on your screen with the names of the events. Based on each other's preferences and budget (above), discuss at least three possible places with your partner and decide on one event nearby to which you both would like to take your friend from Mexico. Click on at least three of these events to check out their detail. Once you decide on one event, click on the event's link and check the day and time of the event. Take a picture of your screen with the name, day and time of the event.

b) Now go back to the main screen and click on the three bars (top left) and choose "Places". Make sure the "camera" icon (bottom left) is selected. Point your phone camera at buildings in front of you. Blue tabs will appear on your screen with names of restaurants. Click on the tabs to find out more details about the restaurants. Again, discuss at least three restaurants and why you would like to take (or not) your friend there.

Based on each other's preferences (above), decide to which restaurant nearby you want to take your friend to eat. Once the restaurant is chosen, click on the restaurant's link.

c) (*Google Translate AR*) Check the restaurant's menu by clicking on the menu tab. Your partner should open the Google Translate AR app on his/her phone. Select the languages English and Spanish. Tap on the camera icon. Using the camera feature, position your phone camera over the menu. The app will translate the menu to Spanish. Choose an item that you would like to eat. What is this dish made of? Do you like everything in it? Your partner should do the same with his/her item. What do you think your Mexican friend would like to eat? Take a picture of these menu items.

Together come back to the language lab and complete Phase 2.

Phase 2: Writing activity (back in the language lab)
Now that you both know where you want to take your Mexican friend to eat and have fun, write an email to your friend in Mexico telling her about your plans for the weekend. Tell her what you and your classmate like to eat and where you plan to go and why. Think about the structures and vocabulary you have learned in this chapter and previous chapters (verbs in present tense – Ar-Er-Ir irregulars and stem-changing). You should write at least 200 words. Write with as much detail as possible. You have 20 minutes to complete this part of the activity. Continue talking while you write (make sure your recorder is still on).

Environment: non-AR
Duration: Phase 1 (oral): 20 minutes, Phase 2 (writing): 20 minutes
Directions - Make sure you talk to each other during the entire time. Make sure your recorder is on. You have 20 minutes to complete Phase 1.

Phase 1: a) Using the Google Maps app check the "Explore nearby" option on the bottom of the screen, click on "events" and choose "this weekend" from the "date" tab. Based on each other's preferences and budget (above), discuss at least three possible places with your partner and decide on one to which you both would like to take your friend from Mexico. Take a picture of your screen with the name and time of the event.

b) Now go back to the main screen and using the app's feature "Restaurants" find restaurants nearby. Again, discuss at least three places and why you would like to take (or not) your friend there.

Based on each other's preferences (above), decide to which place you want to take your friend to eat. Click on the "menu" tab. Look up the menu and choose an item that you would like to eat. What is this dish made of? Do you like everything in it? Your classmate should do the same. Make sure there's something on the menu for your Mexican friend also. Take a picture of the three menu items you, your classmate, and your friend like. Now complete Phase 2.

Phase 2: Writing activity - You have 20 minutes to complete this part of the activity. Continue talking while you write (make sure your recorder is still on).

Now that you both know where you want to take your Mexican friend to eat and have fun, write an email to your friend in Mexico telling her about your plans for the weekend. Tell her what you and your classmate like to eat and where you plan to go and why. If you need, use Google Translate to find out the name of the ingredients in Spanish. Your partner should do the same with his/her dish. Think about the structures and vocabulary you have learned in this chapter and previous chapters (verbs in present tense – Ar-Er-Ir irregulars and stem-changing). You should write at least 200 words. Write with as much detail as possible. You have 20 minutes to complete this part of the activity. Continue talking while you write (make sure your recorder is still on).

Task #2: "Mi habitación de Ikea" (My room from Ikea)

Environment: AR
Time duration: Phase 1 (oral): 20 minutes, Phase 2 (writing): 20 minutes
Directions - Make sure you talk to each other during the entire time. Make sure your Zoom app is on and recording. You have 20 minutes to complete Phase 1.

Phase 1: You and your partner are moving to a house near ASU. Together you have a budget of $1.500,00 to furnish your living room. Open the Ikea Place AR app. Follow the instructions on the app. You may choose items from a Collection or Categories (top menu). However, you must choose 8 items total, but you can only choose up to 2 items from each category or collection. Together you must decide and agree on which things to buy (do not exceed your budget!) and where you are going to place them in your living room. Once you have all the items placed in your space, take a picture of the room. You have 20 minutes to decorate your living room.
Phase 2: Writing activity. You have 20 minutes to complete this part of the task. Continue talking while you write (make sure your recorder is still on).

APPENDIX D

Now that you have your room decorated. Write an email to your parents telling them what items (i.e. a blue sofa bed, a white desk, etc.) you both chose and why these items are the most important ones to have in your room. Describe where you decided to place the items in the room, giving them a mental picture of what the room will look like. Tell them how much you and your partner spent on these items. You should write at least 200 words.

Environment: non-AR
Duration: 20 minutes (oral)/ 20 minutes (writing)
Directions - Make sure you talk to each other during the entire time using the recorder. You have 20 minutes to complete Phase 1.

Phase 1: You and your partner are moving to an apartment near ASU. Together you have a budget of $1.500,00 to furnish the living room. Open the Ikea Store app. Choose from

"Offers, New Products, or Popular". You may choose items from any of these categories. However, you must choose only 8 items total. Together you must decide and agree on which things to buy (do not exceed your budget!). Place the items in your shopping cart. Using the apartment plant drawing your instructor gave to you and the list from your cart, decide where you are going to place them. During your discussion draw the items to visualize them in your apartment. You have 20 minutes to decorate your room.
Phase 2: Writing activity. You have 20 minutes to complete this part of the activity. Continue talking while you write (make sure your recorder is still on).

Now that you have your room decorated. Write an email to your parents telling them what items (i.e. a blue sofa bed, a white desk, etc.) you both chose and why these items are the most important ones to have in your apartment. Describe where you decided to place the items in the room, giving them a mental picture of what the room will look like. Tell them how much you and your partner spent on these items. You should write at least 200 words.

APPENDIX E

POSTTEST SAMPLE (TASK# 1)

APPENDIX E

1. For each sentence below, indicate whether the sentence is correct or incorrect. Indicate to what extent you are certain of your answer by checking the appropriate box according to the scale below. If you can, provide the correct answer.

5 Definitely wrong	**4** Probably wrong	**3** Probably correct	**2** Definitely correct	**1** I don't know

Estamos muy emocionados para tu llegada a los Estados Unidos. Correct:	5	4	3	2	1
Te voy a mostrar lo que tiene mi ciudad a ofrecer. Correct:	5	4	3	2	1
Allí hay muchos bebidos de frutas ricas y otras sabores buenas. Correct:	5	4	3	2	1
Voy a terminar la letra Correct:	5	4	3	2	1
Ahí podemos bailar con otros o ver una película si queremos. Correct:	5	4	3	2	1

2. Provide the Spanish translation for the English word in the texts below:

   Tengo ganas de ir allí pero siempre… ¿la baleta? ¿Cómo se dice '**tickets**'? _____

3. Provide the English definition of the boldfaced Spanish word:

   Sabemos que tendrás ganas de **probar** la comida americana _____

4. Provide the Spanish translation for these English words

   açaí bowl _____ almonds _____

   peanut butter _____ sugar _____

APPENDIX F

IRB EXEMPTION GRANTED

APPENDIX F

On 9/3/2019 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Augmented Reality and Language Related Episodes in Second Language Learning
Investigator:	Barbara Lafford
IRB ID:	STUDY00010562
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• Domaz_HRP-502a-Consent Form.pdf, Category: Consent Form; • Domaz_HRP-503a-TEMPLATE_PROTOCOL_SocialBehavioral.docx, Category: IRB Protocol; • Postquestionnaire.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Posttest.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Background-questionnaire.pdf, Category: Recruitment Materials; • SD tasks.pdf, Category: Participant materials (specific directions for them);

APPENDIX F

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (1) Educational settings, (2) Tests, surveys, interviews, or observation on 9/3/2019.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator