The Impact of Collaborative Computer-mediated Communication on the Emergence of Voice and the Improvement of Overall Writing Quality of L2 Learners

by

Jessica Turetken

A thesis submitted in conformity with the requirements for the degree of Masters of Arts
Graduate Department of Curriculum, Teaching and Learning
University of Toronto

© Copyright by Jessica Evans Turetken 2015
The Impact of Collaborative Computer-mediated Communication on the Emergence of Voice and the Improvement of Overall Writing Quality of L2 Learners

Jessica Turetken
Master of Arts
Graduate Department of Curriculum, Teaching and Learning
University of Toronto
2015

Abstract

This study explored the effects of collaborative Computer-mediated Communication (CMC) on acquiring voice and improving writing quality in English as a second language (L2) writing. This thesis is conceptually grounded in applications of socio-cultural theory to language learning (Vygotsky 1978, 1981). Data were collected from 35 L2 students studying English for Academic Purposes as part of a bridging program in an Ontario university. The data included writing samples over three iterations during 13 weeks of both face-to-face collaborative writing and collaborative CMC, interviews with a subsample of students, and a Technology Satisfaction Questionnaire. The focus of analysis was on changes in the quality of voice in the students’ English writing over time. There was a change in voice intensity in their writing over the course of the study; the use of the collaborative CMC was related to this change. Implications are recommended for teachers, students, and further research.
Acknowledgments

I would first like to thank Dr. Alister Cumming for his continued patience and guidance throughout this process. Without him, I would not have been able to jump the many hurdles that needed to be jumped. As well, I would have not learned the many great things that I have learned throughout the writing of this thesis.

I need to say thank you to Dr. Clare Brett. Not only is she my second reader on this thesis, but when taking her class, I went to her with this idea. She encouraged me to pursue this line of research as a worthwhile endeavor.

Finally, I would like to thank my husband Dr. Ozgur Turetken. It is with his tremendous support and encouragement that I have been able to achieve the goals that I had set for myself. For this, I am ever grateful.
Dedication

For my sons

Sean Turetken and Aydin Turetken

Anything is achievable through hard work and perseverance.
# Table of Contents

Acknowledgments ........................................................................................................ iii
Dedication ....................................................................................................................... iv
Table of Contents .......................................................................................................... v
List of Tables .................................................................................................................. vii
List of Appendices ......................................................................................................... viii
Chapter 1. Introduction ..................................................................................................... 1
  Purpose .......................................................................................................................... 1
  Research Stance .......................................................................................................... 2
  Organization of the Thesis ............................................................................................ 3
Chapter 2. Collaborative CMC in L2 Classrooms ............................................................ 5
  Socio-cultural Theory .................................................................................................. 5
  Voice in L2 Writing ...................................................................................................... 16
  Gap in Research and the Present Study ...................................................................... 25
Chapter 3. Method ........................................................................................................... 27
  Design-Based Research .............................................................................................. 28
  Context of the Research ............................................................................................. 31
  Research Instruments and Procedures ........................................................................ 35
  Analyses for Voice and Overall Writing Quality ....................................................... 41
  Analyses of Questionnaires ....................................................................................... 44
  Analyses of Interviews ............................................................................................... 45
Chapter 4. Results .......................................................................................................... 46
  Results of Analyses of Writing .................................................................................. 46
  Correlational Analysis of Voice Intensity and Overall Writing Quality .................... 55
  Results from the Technology Satisfaction Questionnaire ......................................... 55
  Findings from Interviews ........................................................................................... 63
Chapter 5. Discussion ..................................................................................................... 69
  Summary of Findings ................................................................................................. 69
  Implications ................................................................................................................ 75
  Limitations .................................................................................................................. 80
References ...................................................................................................................... 83
Appendix A: Participant Consent Letter .................................................................................. 89
Appendix B: Screenshot from Etherpad© Session .................................................................. 93
Appendix C: Writing Diagnostic: Baseline Sample, F2F Iteration Sample, and Etherpad©
Iteration Sample ................................................................................................................ 94
Appendix D: Collaborative Writing Tasks for F2F Iterations .................................................. 95
Appendix E: Collaborative Writing Tasks for Etherpad© Iterations ....................................... 96
Appendix F: Sample of Student Writing Taken after the Etherpad© Iteration that Was Deemed
to Have High Voice Intensity and to Have High Overall Quality ...................................... 97
Appendix G: Sample of Student Writing Taken after the Etherpad© Iteration that Was Deemed
to Have Low Voice Intensity and to Have High Overall Quality ......................................... 99
Appendix I: REEP Writing Rubric (adapted from Monsoor & Grant, 2002) ...................... 103
Appendix J: Rater Background Questionnaire (Format adapted from Zhao, 2010) ............ 104
Appendix K: Manual for Voice Rating (Format adapted from Zhao, 2010) ......................... 105
Appendix L: Manual for Overall Quality Rating ................................................................. 107
Appendix M: Technology Satisfaction Questionnaire ......................................................... 108
Appendix N: Individual Interview Questions ........................................................................ 110
Appendix O. Excerpts from Interviews ............................................................................... 111
List of Tables

Table 1: Descriptive Statistics for the Baseline Writing Samples
Table 2: Descriptive Statistics of Samples Taken after the F2F Sessions
Table 3: Descriptive Statistics of Writing Samples after the Use of Etherpad©
Table 4: Change in Voice Intensity among the Three Iterations\textsuperscript{a}
Table 5: Change in Overall Writing Quality among the Three Iterations\textsuperscript{a}
Table 6: Change in the Number of Content Points among the Three Iterations\textsuperscript{a}
Table 7: Change in the Number of Directives among the Three Writing Iterations\textsuperscript{a}
Table 8: Change in the Number of Hedges among the Three Writing Samples\textsuperscript{a}
Table 9: Change in the Number of Boosters among the Three Writing Samples\textsuperscript{a}
Table 10: Change in the Number of Attitude Markers among the Three Writing Samples\textsuperscript{a}
Table 11: Change in the Number of Authorial Self-Mentions among the Three Writing Samples\textsuperscript{a}
Table 12: Change to the Number of Reader References among the Three Writing Samples\textsuperscript{a}
Table 13: Descriptive Analysis of Responses to Technology Satisfaction Questions
Table 14: Chi-Squared Test of Etherpad© Satisfaction Questions 1 to 8
Table 15: Chi-Squared Test of Collaboration Questions 1 to 8
List of Appendices

Appendix A: Participant Consent Letter
Appendix B: Screenshot from Etherpad© Session
Appendix C: Writing Diagnostic: Baseline Sample, F2F Iteration Sample, and Etherpad© Iteration Sample
Appendix D: Collaborative Writing Tasks for F2F Iterations
Appendix E: Collaborative Writing Tasks for Etherpad© Iterations
Appendix F: Sample of Student Writing Taken after the Etherpad© Iteration that Was Deemed to Have High Voice Intensity and to Have High Overall Quality
Appendix G: Sample of Student Writing Taken after the Etherpad© Iteration that Was Deemed to Have Low Voice Intensity and to Have High Overall Quality
Appendix H: Voice Rubric (Format adopted from Zhao, 2012)
Appendix I: REEP Writing Rubric (adapted from Monsoor & Grant, 2002)
Appendix J: Rater Background Questionnaire (Format adapted from Zhao, 2010)
Appendix K: Manual for Voice Rating (Format adapted from Zhao, 2010)
Appendix L: Manual for Overall Quality Rating
Appendix M: Technology Satisfaction Questionnaire
Appendix N: Individual Interview Questions
Appendix O: Excerpts from Interviews
Chapter 1. Introduction

This chapter introduces the thesis, providing background information regarding the use of computer-mediated communication in second language classrooms as well as the role of voice in writing. Three research questions are outlined. In addition, I explain my stance on the current research topic as well as outline the organization of this thesis.

Purpose

The aim of this study was to gauge the influence of computer-mediated communication (CMC), specifically synchronous writing tools, on the overall uptake of written language among a sample of adult English as second language (L2) learners. CMC as an area of research is a rather new one within the entire body of work in applied linguistics. Scholars have recently begun to see the possible advantages of using CMC during classroom instruction, in various contexts, including that of language learning. It has been argued that the socio-affective elements and the interactive discourse of CMC may enable learners to communicate more freely and purposively than they might in conventional media, thus promoting their L2 learning (Arnold, 2007; Lam, 2004). This study incorporated the use of an online source called Etherpad© (http://openetherpad.org/last viewed July 22, 2015), which allows users to have a chat session while collaborating on a single document (both synchronous and asynchronous collaboration). The focus of the analyses was on learner's performance after and during CMC collaborative sessions to see if their engagement in the CMC prompted a greater uptake of language and more specifically allowed for the emergence of an authorial voice as determined through the comparison of a diagnostic writing sample, an individual composition written after
collaborative sessions without the tool, and an individual composition written after the use of the tool. The data collection and the analysis attempted to address the following research questions:

RQ1  Is the regular iterative use of CMC associated with an overall improvement in the quality of students’ English writing?

RQ2  Is the regular iterative use of CMC associated with the emergence of voice in their writing?

RQ3  What facets of CMC might be responsible for this association?

**Research Stance**

My interest in the use of collaborative CMC to elicit voice in writing began through my own teaching practices and observations in classroom as well as after reading Vygotsky’s (1978, 1981) theories on the socio-cultural elements of language learning. Specifically, I was intrigued by the notion of the Zone of Proximal Distance (ZPD) presented by Vygotsky (1978, pg. 86), in which he argued that learners’ current levels of development are measured by the distance between their own independent problem solving skills and their potential level of development through the influence of an expert other, either an adult guide or in collaboration with a more capable peer. The question of whether, through collaborative writing sessions, learners could improve their overall writing quality and increase the salience of voice in their English writing become paramount for me. Furthermore, CMC has become a pervasive means of communication and curricula delivery in the L2 classroom. Deft and Lengel (1996) in their Media Richness Theory (MRT) posited that the richest forms of media are those that most closely replicate face-to-face (F2F) communication. In many instances during my years as a language instructor and as a student, I observed, learners and teachers communicating with each other in asynchronous discussion threads. Through these discussions it was noticeable that after
several iterations, the student’s writing improved. One way to interpret this phenomenon was to view it through a socio-cultural lens where participants developed a static model from the teacher’s comments, and were then able to reproduce a better version of their own writing. However, asynchronous writing does not closely replicate F2F communication. Synchronous chat sessions seemed to be a better fit in relation to MRT. Therefore, I began to search for a CMC application that would allow for synchronous writing. A colleague introduced me to Etherpad ©, which offers a platform that allows for both synchronous and asynchronous collaboration on an editable document as well as a sidebar for synchronous chat sessions. Using this structure I developed an activity that allowed students to collaborate on a single document over a period of two weeks. I was able to follow their progress because Etherpad© (an open-source software) allows the instructor to create a single URL for group members to use to log on to their pad. I observed students informally, and began to notice that during their collaboration processes, students were modeling writing for each other, and through revisions and editing, their writing quality seemed to improve. As well, it appeared that the level of salience in voice in their writing was greater. Voice has typically been understood in western culture as the individual authorial stance, and this definition and its limitations will be explored later in the thesis. This led me to create the above research questions and to develop the study documented in this thesis.

**Organization of the Thesis**

Chapter 1 includes an introduction with research questions, an explanation for the choice of research topic, and an outline of the thesis and study. Chapter 2 focuses on pertinent literature pertaining to the study. Specifically, the chapter addresses published literature about Socio-cultural theory, the Output Hypothesis, and Constructivist theory. These theories and concepts
are collectively considered within the context of the use of CMC for the purpose of language learning. In addition, in this chapter is a review of published literature surrounding collaborative writing and voice in writing. Chapter 3 describes the methodology of the study. This chapter explains Design Based Research and its appropriateness for this study. As well, Chapter 3 describes the context of the study and the materials used. The Etherpad© tool is also described in this chapter as well as the processes of data collection and analyses. Chapter 4 describes the findings of the study. Finally, Chapter 5 further discusses the findings of the study and limitations with the study. This chapter further describes the implications for both teachers and L2 learners, and then proposes possibilities for future research.
Chapter 2. Collaborative CMC in L2 Classrooms

In this chapter, I review ideas and publications discussing the use of collaborative CMC in L2 classrooms in order to elicit students’ voice in their L2 writing. I also describe the theoretical frameworks adopted in the research that inform the research questions addressed by this thesis.

Socio-cultural Theory

One of the most often referenced theories related to L2 learning is Sociocultural Theory (SCT). SCT, following from ideas first developed by Vygotsky (1978), suggests that social interaction plays a major role with regards to cognitive and linguistic development (Lantolf, 2007). SCT argues that learning takes place through socially mediated events and the use of tools (e.g. language) or artifacts (Knouzi et al., 2010; Lantolf, 2007; Lantolf & Beckett, 2009; Lapkin et al., 2010; Nabei & Swain, 2002; Nassaji & Swain et al., 2002a, 2002b, 2007; Watanabe et al., 2007). Lantolf (2007) argued that this Vygotskian notion that human mental activity specifically arises as a result of internalized social relationships, culturally organized activity and symbolic and physical artifacts such as language. The natural environment of humans is comprised solely of society and culture, whose elements are not limited to relationships and activities, but also encompass symbols and artifacts which include a variety of media (e.g., books, computers, furniture, cities, etc...).

Likewise, Gibbs (1999) encouraged researchers to use caution when assuming that cognition only be represented by what is in the human head, but rather suggests that what we perceive as being cognitive is itself inherently the product of culture. Gibbs gives the example of children in different cultures learning math using the artifacts of their culture (e.g., abacus vs. pencil and paper). In this view, the idea of cognition being culturally created is apparent in the
types of mistakes children make. These mistakes depend on the type of artifacts used within the child's own culture. For example, children using pencil and paper methods will often make mistakes with regards to borrowing numbers when doing mental calculations, whereas children who have been taught to use an abacus will often lose track of the 5s when calculating mentally (Gibbs, 1999). This example suggests that the artifacts of our culture heavily influence our cognitive development. If we were to apply the significance of culture to considerations of our linguistic knowledge, Lantolf's ideas become relevant, particularly, the question of what artifacts are important. From this perspective we postulate that our linguistic cognitive development is deeply connected with the artifacts that our cultures consider important for communication of pertinent information. Our modern global culture, for example, deems CMC to have great value. Therefore, CMC should be a fertile environment for supporting cognitive advances in the linguistic development of L2 learners.

With regards to cognitive development, Vytgosky (1978) saw this development as dynamic and unpredictable and able to move in any direction, including previous stages of development (Lantolf 2007). In developing SCT, Vygotsky created the notion of the Zone of Proximal Development (ZPD) (1978, p. 86), defining it as:

the distance between the actual developmental 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.

That is, the ZPD can be simply defined as the current level of problem solving skills held by a learner and the distance to their potential level of problem solving skills with assistance from others. From this concept arose the ideas of scaffolding and modeling (Vygotsky 1978). Scaffolding can be understood as creating temporary cognitive supports through modeling of behavior through which learners are supported to a higher level of cognitive functioning. For example, parents model language for their infants by talking with the child at a level which is
just above their current linguistic and cognitive developmental level. This, in theory, allows the child to begin to mimic and ultimately internalize those utterances for their own use, raising them to a more sophisticated level of linguistic and cognitive development. Thus, if a child is babbling and the parent begins to return two syllable utterances like "ma-ma" and "da-da" while at the same time pointing to the appropriate parent, the child in turn begins to internalize those utterances, add meaning to them, and begins to produce utterances with that same structure, evidence that they have internalized them as their own.

**ZPD in the Classroom**

The ZPD has become an important topic of research for language classrooms. One key element that needs to be noted with regard to the ZPD and learning, especially with regard to writing, is that it is a complex and ongoing process that takes time to establish and develop (Sasaki, 2004). One cannot expect immediate results, but rather over time there is a "gradual release of responsibility" from teachers or other mentors (Kong & Pearson, 2003). As stated before, development is a dynamic process that moves cyclically between stages at times moving forward and then slightly back and then forward again. In instructional settings it is up to the expert to slowly allow the novice to take over the reins with regards to the next phase of development. In a pedagogical approach involving "gradual release of responsibility” the teacher first demonstrates the task, then gradually begins to require students to take on a more active role in the task through scaffolding and modeling, and finally gives the execution of similar tasks over to the students to perform autonomously (Kong & Pearson, 2003). For example, a teacher cannot simply walk into a classroom, model a few grammar points and expect students to be able to use the forms properly and accurately by the end of class. Rather,
the teacher and students must spend time together revisiting uses and applications of that form through different tasks that model the same grammar point in different ways.

**ZPD and the use of CMC**

The idea of computers as artifacts to be used as a medium through which human mental activity can take place supports the widespread use of CMC and the robust possibilities that it has for L2 learning and teaching. Warschauer (2005) discussed the implications of SCT with regards to the use of Computer Assisted Language Learning (CALL) as a means for teaching a second language and more specifically with regards to teaching writing. He uses the three major components of Vygotskian SCT theory: mediation, social learning, and microgenetic analysis to illustrate the relationships between humans and the tools they use (Vygotsky, 1981).

Warschauer described mediation in terms of all human activity being mediated by tools or signs (Vygotsky, 1981; Warschauer, 2005; Wertsch, 1991). He further argued that the role of such tools is determined by how they intrinsically alter human actions (Warschauer, 2005). According to Vygotsky (1981), the concept of mediational means implies that tools should be considered as part of human actions because of their importance in manipulating and transforming the movement and formation of cognitive functions (Warschauer, 2005). Warschauer (2005) also discussed Leontiev’s (1979) expansion of the idea of mediation in his activity theory. Activity theory argues that in order to accurately analyze human cognition and behavior one must not limit the scope simply to the individual, or even the individual and the tools they use, but rather consider the activities that people carry out with the aid of such tools. Warschauer (2005) argued that if we use writing and computers as examples of an action and a tool, then according to the above theories, we can move beyond seeing these two as separate entities (i.e., computers aid writing, but writing is still a completely independent activity).
Rather, he suggested that through the use of computers to write, we have developed new forms of writing that require their own forms of instruction (Shetzer & Warschauer, 2000, 2001; Warschauer, 1999, 2005). Therefore, systematic investigations are needed to examine the implications of using CMC in L2 classrooms.

The second tenet of Vygotsky’s SCT is social learning. The concept of social learning assumes that all functions of a person’s cultural development occur within a binary dialectic between social and individual processes. At the social (interpsychological) level children develop culturally through interaction with others in their world. At the individual (intrapsychological) level children develop culturally on an introspective plane (within the child) (Vygotsky, 1978). This distinction led Vygotsky to the notion paramount in his conceptualization of how children’s cultural development occurred, apprenticeship learning (Vygotsky, 1978; Warschauer, 2005). Apprenticeship learning describes the relationship between a novice and a more skilled other person (e.g., student or teacher). From this view, Warschauer (2005) noted that Bakhtin (1986) built on Vygotskyan notions with regard to how we perceive linguistic interactions. According to Bakhtin, we learn through interpreting the language of others and responding to others. As such, verbal and non-verbal cues of the others are a necessary component of language learning. Therefore, social learning needs to be an integral component of CMC research primarily with respect to how learners integrate linguistic chunks of others’ uses of language into their CMC (i.e. phrases, collocations, etc.) as well as how the influence of an authentic audience encourages writers to refine their work for, and as a response to, the input of others (St. John & Cash, 1995; Warschauer 2002, 2005; Warschauer & Lepintre, 1997).

The final component of Vygotsky’s SCT (1978, 1981; as presented in Warschauer, 2005) is microgenetic analysis, which traces the gradual development of individuals’ mental
functioning through their origins and historical processes. According to Vygotsky, these origins can be categorized as the following: microgenesis (the way specific events develop and occur), ontogenesis (the way in which the individual develops and grows), socio-cultural history, and finally, phylogenesis (referring to the evolution of a species) (Vygotsky, 1978, 1981; Warschauer, 2005). Warschauer (2005) related this aspect of SCT theory to CALL by noting that in order to understand CALL, we have to consider it within the broader context of history, society, and culture. Accordingly, investigations into learners’ uses and motivations toward using a particular technology must consider contributory contextual elements such as present-day economies, cultural norms for interaction, and society (Warschauer, 2005).

Asynchronous and Synchronous CMC

CMC tends to involve one of two modes of communication: asynchronous (ACMC) or synchronous (SCMC). Both types of CMC can potentially lend themselves to prompting the development of cognitive process in L2 learning. Ware (2004) argued that there is a metalinguistic element to the use of CMC for the purpose of learning to write in a second language. Ware considered that both AMC and SCMC allow students the opportunity to review not only their written words but also the words of others. The former is not limited to the use of peer feedback, but includes collaborative settings in which students can model the language of their group members. Writing electronically, students create a written record that they can analyze and from which they can internalize appropriate forms and style. Ware (2004) specifically argued that chat language (in SCMC) creates an environment where learners can become aware of the process of learning various linguistic forms and functions, which further emphasizes the cognitive dimensions of CMC. Parks, Hamers, and Lemmonier (2005) also documented the awareness of metalinguistic processes that L2 learners develop when learning to
write using computer-mediated technologies. Students not only have the opportunity to model the proper forms of the language through the use of CMC, but they also gain insight to the cultural norms surrounding writing within their target language (Ware, 2004).

**The Output Hypothesis**

The Output Hypothesis (Swain, 2001, 2005) adds an SCT perspective to Krashen’s (1985) Input Hypothesis, which had argued that language learning occurs through comprehensible and meaningful input. Swain (2001, 2005) argued that output in the form or speech or writing is integral to the processes of communication and language acquisition; indeed, language learning occurs through the production of language. According to Swain there are three functions of language output: (1) the noticing/triggering function, (2) the hypothesis testing function, and (3) the metalinguistic (reflective) function. The noticing/triggering functions can be simply described as the process of producing the target language through which L2 learners may realize that they do not know the proper form they wish to produce (Swain, 2001, 2005). Lai and Zhao (2006) argued that the use of text-based chat sessions in the L2 classroom would encourage this kind of noticing among learners. Their study found more noticing within the treatment group, which Lai attributed to the longer processing time text-based chat sessions give learners, allowing them to notice the gap between proper language forms and any improper forms they have produced. The second of the three functions, the hypothesis function, is sort of a "trial-run" according to Swain (2001, 2005). In this, learners create hypotheses or guesses about the language and test them out during episodes of discourse, which typically lead to a modification of the output. The final function is the metalinguistic (reflective) function. Here, language is used to reflect back on output (language produced) by
the learner or others around them (Swain 2001, 2005). This function then becomes the tool to mediate L2 learning (Lantolf, 2007; Swain, 2001, 2005; Vygotsky, 1978).

Parks et al. (2005) conducted a longitudinal study with a group of francophone students in Quebec who were learning English. The students participated in a class where the teacher used networked computers to allow students to view websites specific to the English language, to chat back and forth in class, and to perform peer-feedback sessions on their writing (using the track changes function in Word along with chat sessions). The assumption was that through the use of CMC the students would be more inclined to be involved in the feedback sessions and thereby increase their written proficiency. Parks et al. found that these automated feedback sessions allowed the students to incorporate what they "noticed" to be good in each other's writing as well as helped them monitor and edit their own work in an autonomous fashion. Hyland and Hyland (2006) came to a similar conclusion in that such automated feedback devices gave learners a tangible view of their own writing, either when they looked at the edits of others or when they implemented a self-editing tool in their own writing. Both of these examples could be viewed as demonstrating metalinguistic functions in that learners thought about the language, internalized aspects that they felt to be proper, and modified future output. Moreover, Swain, Brooks, and Tocalli-Beller (2002), Swain and Lapkin (2002 & 2007) and Lapkin, Swain, and Psyllaskis (2010) discussed the metalinguistic process of "languaging" in which learners use language (spoken or written) to assist them in cognitively complex activities. Through the use of languaging as a mediator, learners are taking part in critical psychological and cultural activity--learners articulate and transform their abstract thoughts into artifacts that are able to be used for further reflection. Both Parks, Hamers, and Lemmonier (2005) and Hyland and Hyland (2006) also documented students’ uses of feedback and peer-revisions as a
source of languaging, observing that CMC creates a context that helps turn people’s thoughts into permanent artifacts for later use and reflection.

**Peer-feedback and Collaboration**

Peer-feedback and collaboration can be viewed under the umbrella of SCT, playing an integral part in L2 learning processes. Both of these activities lend themselves nicely to Swain's Output Hypothesis (2001, 2005) in that peer-feedback and collaboration can involve aspects of all three processes outlined in the Output Hypothesis. Much research has been dedicated to peer-feedback and collaboration both for L2 learning in general and with regards to CMC. Swain et al. (2002) argued that collaborative writing and peer revisions warranted investigation since the collaborative dialoguing that emerges during collaboration sessions mediates L2 learning. In their 2002 study they measured the "language-related episodes" (LRE) of two anglophone students of French working on a collaborative writing piece. LREs are an aspect of metalinguistic processes in that the LREs are the instances of collaborative dialogue between learners in which they discuss the language they are producing, question their use of the language, and/or engage in other self-correcting language production (Swain et al., 2002). Further to this point, Swain et al. noted that during the study a dialogue between the two students occurred, in which the students began to understand why certain changes were made to their story through reformulated feedback and thus incorporated into their stories those changes they chose to accept. This example shows learning occurring between students through the means of collaboration. With regards to peer feedback on students’ writing Hedgecock and Lefkowitz (1992) observed that such feedback should not simply lie with error correction, but that it should be a procedure for discovering content, structure, and voice. They argued that between-draft interventions or feedback sessions actually led to significant improvements in students’ L2
composing performance. They also pointed to the idea that peer-feedback sessions limited teacher marking while increasing learner writing, whereas written teacher feedback alone had little or perhaps negative effects on students’ production (Hedgecock & Lefkowitz, 1992). As pointed out previously, both Parks, Hamers, and Lemmonier (2005) and Hyland and Hyland (2006) argued that peer-feedback and peer collaboration sessions are strongly supported through the use of CMC in that these sessions that include dialoguing allow learners to have a permanent record of the discourse between peers. This written record is, of course, in contrast to simple oral feedback or collaboration. The dialoguing holds elements of both the written and spoken forms.

Collaborative Writing

Storch (2013) expanded upon much of the prior research on SCT related to the notion that social interactions and communicative practices are key to language learning. She argued that collaborative writing more than collaborative speaking activities are suitable for providing language-rich environments in which learners are better able reach a level of communicative competence beyond just linguistic competence while also achieving a level of discourse competence. This potential for learning is due to several factors, the first being that in order to truly collaborate on a document beyond simple peer review tasks, learners must negotiate meaning in order to achieve a level of mutual understanding within the target language. According to Storch (2013), a process writing approach is most suitable to cultivate such communication. Instructors and teachers should design collaborative activities that take learners through the many stages of writing such as: brainstorming, creating a first draft, editing, giving and receiving peer feedback, and finally incorporating both peer and instructor feedback. In order to go through these processes, learners must rely on the use of their target language in
order to reach mutual agreement and understanding with regards to the document they are writing. In these ways, students not only improve their linguistic competence, but also achieve a higher level of discursive competence. Storch (2013) also argued that the written text allows learners more opportunities than does spoken communication for acquiring language and proper grammatical structures as it gives learners more opportunities and time to ponder word choices and breakdowns in communication. Finally, Tan, Wigglesworth, and Storch (2010) suggested that CMC is a feasible domain for encouraging learners to participate in collaborative writing activities. Unlike F2F communication, CMC can allow learners to have more processing time as they are not expected to always have immediate responses to comments and ideas. Learners can ruminate over linguistic decisions, allowing them to make more appropriate choices. Accordingly, the online domain may very well lead to more language learning than is possible in completely oral interactions.

**Constructivist Theory**

SCT is one type of Constructivist learning theory. Constructivism is based on the idea that knowledge is socially constructed, rather than the act of an individual receiving or discovering new types of knowledge (Senior, 2010). Pegrum (2009, p. 27) defined constructivism as:

Constructivism puts learners, rather than a given body of knowledge, at the centre of the learning process, and aims to build on the pre-existing knowledge and perspectives they bring into the classroom. As students actively engage in educational experiences based on authentic tasks, with guidance provided by teachers or more expert peers, they collaboratively build new understanding. Facilitated by competent teachers, this can be an empowering strategy that encourages the sharing of ideas, giving everyone a voice without unnecessarily imposing particular points of view.

Senior (2010) argued that there is a link between the interactions afforded by the use of Web 2.0 and CMC and social constructivist approaches. Lee (2010) demonstrated this notion with a
study where Wikis were used as a tool for collaborative writing in an elementary Spanish class. Lee found that the Wiki format engaged previously apprehensive writers partly by moving their writing from material that only a teacher read to an interactive social medium. Lee's study also emphasized another tenet of constructivist theory in that it was task based and student centered. Lee argued that task choice was highly important and found that open-ended and problem-oriented tasks offered students opportunities for collaboration and allowed them to not only negotiate meaning but also to strike a balance between fluency and accuracy (Lee, 2010).

**CMC Theoretical Frameworks**

Much of the current CMC research within the field of applied linguistics is informed by the ideas of SCT, the Output Hypothesis, and Constructivist theory. The research also goes a step further to suggest that the use of CMC allows learners who were otherwise apprehensive about language production and participation the opportunity to fully participate and at times take a leading role. Arnold (2007) explained that learner anxiety is a prevalent aspect of L2 learning and that this anxiety often hinders production and thus limits the potential for learning. Arnold suggested that the use of CMC in language classrooms can offer anxious learners opportunities where they do not feel afraid of making mistakes and that it also limits the amount of teacher talk, putting the focus back on the learner. Lam (2004) claimed that CMC can be an instrumental for giving learners a venue where they are free from ridicule or harsh looks for making a mistake. This freeing element of CMC coupled with the socio-cultural benefits that it offers learners makes this an important area for further research and pedagogical development.

**Voice in L2 Writing**
Since the 1960s and 1970s, researchers have begun to pay attention to the role of “voice” in writing. This is an important point to study not only within the L2 writing context, but also within the L1 English writing context. From an English speaking or perhaps Western cultural standpoint, the notion of an authorial voice is a strong undercurrent as to what readers would expect to see in “quality writing”. It is often described as hearing the author themselves speaking as you are reading the words on the page. It is even more critical within scholarly writing. Writers are expected to develop assertions and arguments as well as comment on and critique the work of others. These discourse functions, however, can be challenging for those writers who have been educated or raised in a non-Western culture, yet may be most critical to acquire for those writers who gained their literacy experiences in a collectivist society. This is because the notion of voice is a rather individualistic one and collectivist cultures may not value that notion of individualistic voice encouraged within Western education and societies.

Much of the early research focused primarily on “voice” in L1 writing in a vague, rather abstract way. More contemporary research is beginning to place value on considering in detail how “voice” functions in L2 writing and how this affects the overall writing quality of English language learners. Typically, in an L1 composition context, “voice” is married with the idea of individualism and often considered to reflect the writer’s own personal identity (Stewart, 1992). Accordingly, it was believed that authorial voice was something that would differentiate the author from all other human beings regardless of any shared experiences (Ramanathan & Atkinson, 1999; Stewart 1972). Elbow (1981, 1999) described authorial voice as lending itself metaphorically to the notion of individualism and thus was difficult to separate from the identity of the author. The central idea among these definitions is that of the “self” being interjected into the writing and/or the writing simply being an extension of “the self”.

In contrast, recent scholars have viewed “voice” and language to be social constructs developed through social interactions and shared social experiences, from which a voice emerges that is representative of a collaborative identity. Ede (1992) argued that writers have a “situational voice”, which writers use for specific types or genres of writing. This aspect of voice in writing is likened to the way in which we choose clothes depending on the events of our day. Cummins (1994) developed the notion of “voicing”, which he described as a continual process of development, evolution, and understanding of both the internal and external identities that we carry as writers, all within the boundaries of language, culture, and discourse. Likewise, Matsuda (2001) described “voice” within the context of an ever-changing repertoire of socially available language features that language users utilize either deliberately or subconsciously through both discursive and non-discursive elements.

In an attempt to further refine the definition of “voice”, Hyland (2008) argued that the manner in which writers make choices leads to the construction of their authorial voice. This is most evident in academic writing. Hyland (2008) extended this point by noting that, as writers, we do not sacrifice the individual voice, but rather through negotiated representations of ourselves and the discursive nature of our writing community we achieve an authorial voice. Zhao (2010) elaborated on this idea of voice proposed by Hyland (2008) to devise a scheme for evaluating the relative quality of voice evident in the writing of L2 learners. This interactional model of voice contains two systems. The first system, dubbed the “stance”, refers to writer-oriented features which denote the language students use to present themselves, their opinions, and their arguments. This is accomplished through the use of four linguistic elements: hedges, boosters, attitude markers, and self-mention (Hyland, 2008; Zhao, 2010). The second of the two systems is referred to as “engagement” and primarily focuses on reader-oriented elements of writing accomplished using the following discourse features: reader pronouns, personal asides,
references to shared knowledge, directives, and rhetorical questions (Hyland, 2008; Zhao, 2010). This model consists of social constructivist elements and views voice as a social feature of writing, putting forward a collaborative and dialogic view of voice which has been important to many researchers, specifically with regard to L2 writing (Farmer, 1995; Prior, 2001; Zhao & Llosa, 2008: Zhao, 2012). However, this model does not exclude the individual which is evident through the “stance” system. Hyland’s model seems to marry both research camps that surround voice. The “engagement” system allows for a certain level of interaction or dialogue between the reader and the writer, which in turn allows for the further development of the communal voice (Zhao, 2012). In this thesis, I have adopted this model as a foundation for my empirical study on the development of voice in English learners’ writing.

**Teaching Voice**

Through the development of the above theories surrounding voice in writing, researchers and educators alike have begun to question the necessity for teaching voice in writing and whether it has any impact on overall writing quality. A number of researchers, primarily those in the rhetoric and composition fields, believe that the teaching of an authentic voice in the writing classroom should be an essential part of all writing instruction in order to develop sound writing on the part of the learner (Zhao, 2012). Another camp of researchers, primarily in the L2 field, believe that the teaching of voice to students who do not share a cultural or linguistic background can actually be harmful (Ramanathan & Kaplan, 1996).

Ramanathan and Kaplan (1996) argued the latter view in a study in which they analyzed Freshman Composition textbooks used widely throughout many U.S. colleges and universities. In a general Freshman Composition course the notion of an authorial voice is a prevalent topic. However, their study revealed that the textbooks focussed on topics and social contexts that
would not be either linguistically accessible or culturally available to L2 learners. Similarly, Ramanathan and Atkinson (1999) argued that the U.S. notion of individualism is central in L1 writing courses when teaching voice, peer feedback, critical thinking, and textual ownership, which can all be problematic for L2 learners of English who may not hail from overtly individualistic societies. Likewise, there have been many anecdotal studies that point to the problematic nature of teaching voice in composition or language classes. Shen (1989) explained in his narrative about his own experiences learning how to write in an English context how the notion of an individualistic voice was troublesome for someone who comes from a communitarian culture which emphasises the “we” rather than the “I”. Shen (1989) described how he had to virtually create a new identity to write within the North American context. These types of analysis have pushed researchers to argue for removing voice as a pedagogical tool and shifting the focus of voice to a means for understanding L2 writers (Ramanathan & Atkinson, 1999).

**Voice and Positioning**

Building upon the former body of work, Ivanic and Camps (2001) analyzed the writing samples of six Mexican graduate students. From this analysis, they developed three major types of L2 ‘positioning ‘that mold the voice of L2 students’ writing. These are based on Halliday’s (1985) semiotic theory of micro-functions of language and include: the ideational positioning, interpersonal positioning, and textual positioning. Ideational positioning can be described as the devices that a writer uses to describe the world. Interpersonal positioning considers the relative authoritativeness of writers with respect to their understanding of their target audiences. Textual positioning examines how a writer’s preferred uses of textual features conveys a given meaning to the audience. These three positioning types are used by the authors to argue that a multi-
voiced essence begins to appear in the writings of their subjects, leading them to believe that the negotiation of identity and voice is an essential element to all writing. Based on this notion, Ivanic and Camps (2001) recommended that writing curricula should include a critical awareness of the concept of voice and at the same time allow for learners to acknowledge the difficulties associated with the representation of voice when writing within a different cultural context and a different language.

**Voice and Culture**

Similar to Ivanic and Camps (2001), Matsuda (2001) argued against the idea that voice is an intrinsically Western manifestation of individualistic beliefs, but that L2 writers could, in fact, access an authorial voice. In a case study that closely examined a blog written in Japanese by a Japanese female, Matsuda found strong evidence of an individual voice through the author’s use of discourse features that are exclusive to Japanese. The study counters the notion that an individualistic voice is foreign to writers of non-Western communitarian cultures. Therefore, based on these findings it appears that not only is voice learnable, but it is also a potentially teachable construct.

**Voice and Writing Quality**

Adding to these ideas, researchers have begun to query not only how to teach voice, but also to what extent voice plays in the overall quality of writing and how to assess this. Helms-Park and Stapleton (2003) was one of the first studies to investigate the relationship between voice and overall quality of writing. For this study, they devised a *Voice Intensity Rating Scale* which included four major components: assertiveness, self-identification, reiteration of the central point, and authorial presence and autonomy of thought. This scale was used to assess the
writing of 63 ESL students enrolled at a major Canadian university. They also incorporated Jacob`s et al.’s (1981) *Composition Profile* in order to measure the overall quality of the students’ writing. Although, their inter-rater reliability measures were tested and shown to be high, the results of the study showed no significant correlations between voice intensity and the overall quality of the writing, nor did the study yield any significant results between the overall writing quality and any of the four elements of voice. Therefore, it was argued that there may be very little significance with regard to the role of *voice* in academic writing, thus begging the question as to whether or not this should be a core tenant for teaching composition.

**Zhao’s Studies**

Although Helms-Park and Stapleton’s (2003) study was innovative, Zhao (2012) found several limitations that may have biased their results. The first was with Jacobs et al’s (1981) *ESL Composition Profile*. Zhao (2012) felt that this particular rubric did not truly address the relationship between the elements of *voice* and quality of writing. Furthermore, she argued that this rubric was not representative of what was being used in ESL classrooms to assess students’ writing. She argued that using a rubric that was utilized by the program being studied would have yielded more accurate results.

Likewise, Zhao (2012) found limitations with *the Voice Intensity Scale* designed by Helms-Park and Stapleton (2003). She argued that the rubric was developed based on theories surrounding *voice* that scholars such as Matsuda and Tardy (2007, 2008) called a “skewed representation” of the elements found within the L2 writing literature. The study does not appear to have considered much of the social constructivist literature, which suggests that an authorial voice within academic discourse may be attained both by presenting an individualistic self and by interacting with others within a written text (Cumming et al., 2005; Hyland, 2008; Palacas,
Zhao further argued that the study was limited in its scope in that the only correlations made were solely between *voice* and overall writing quality. According to Zhao, further studies would be necessary not only to consider to what extent *voice* functions in the overall presumed quality of writing, but also to examine the correlation between each of the components of voice rubric to understand any variance in writing when controlling for other variables.

Based on these observations, Zhao and Llosa (2008) replicated Helms-Park and Stapleton (2003) using data from high-stakes exit writing assessments of L1 students at the high school level in order to determine whether or not *voice* played the same role in L1 argumentative writing as it might in L2 argumentative writing. They also aimed to test the validity and the generalizability of the findings of Helms-Park and Stapleton’s (2003) study. This new study included 42 anchor essays collected from one administration within the New York State Regents Examination in English Language Arts and the *Voice Intensity Scale* rubric (2003). However, rather than using a prescribed rubric to measure the quality of writing, the essays were pre-rated using the rubric provided by the examination agency. The results of the study indicated that in contradiction to Helms-Park and Stapleton’s (2003) study, there were significant correlations between assessments of *voice* and the overall quality of writing that accounted for approximately 61% of the variance in the scores of the 42 writing samples.

However, Zhao (2012) observed that the context in which the data were collected varied greatly between studies. For example, in Zhao’s test, the writing samples were taken from timed high-stakes exit writing exams giving to L1 high school students whereas the samples from Helms-Park and Stapleton were gathered from in-class compositions done in an L2 classroom. Furthermore, the nature of the data sources did not allow for any real in-depth analysis of the relationship between *voice* and overall writing quality.
Subsequently, Zhao readdressed these issues in her 2010 two-phase doctoral study. In the first phase of her study, a voice rubric was designed based on Hyland’s (2008) theoretical model of voice then was used to score the strength of authorial voice in 200 argumentative writing samples taken from a high-stakes L2 academic writing exam (the TOEFL, Test of English as a Foreign Language). The results were analyzed and measured in order to be used in the second phase of the study, which included a revision of this rubric that included the discourse elements associated with voice in terms of ideas and content, manner of presentation, and reader/writer presence. Specifically, the discourse elements assessed were: central point articulation, directives, hedges, boosters, attitude markers, authorial self-mention, and reader reference. This rubric was used to score another 200 writing samples, and the results were analyzed to reveal a significant correlation between overall voice strength and the overall quality of writing. The data were further analyzed to explore the correlations between each of the voice components from the rubric and overall quality of writing. The findings demonstrated that most of the *voice* elements did show a significant correlation to overall writing though to a varying degree. However, elements such as authorial self-mention seemed to have a negative correlation when scored.

Coupled with the findings of Zhao and Llosa (2008), these results suggest that the teaching of *voice* has relevance for L1 and L2 composition, assessment, and learning. However, the limitations with these studies are that they only analysed test data, and do not address pedagogical or methodological issues of supporting students’ use of voice. Although Zhao framed her work within social constructivist theories, neither of her studies provided any direct insights into learning to write through collaborative processes. Of particular relevance both for the uses of CMC in classrooms and corrective peer feedback, elements of voice need to be considered from the perspective of both social and individual development. L2 writers not only
assist each other in their writing but also build a discursive identity within the online domain. Investigations must be done to analyze how to approach teaching *voice* in the L2 classroom.

**Gap in Research and the Present Study**

To the best of my knowledge, at the time of writing, there were no available studies studying a collaborative CMC such as Etherpad© with elements of both ACMC and SCMS to examine its effects on learning. Nor have I encountered any studies of this kind which additionally quantified learner's performance using pre and post-test tasks to measure changes from the intervention. Thus, it was my intent in this thesis to investigate learner's written performance to understand if the experience of CMC affects students’ writing performance, and if so, what variables may have influenced this learning.

Three research questions aimed to address this gap in the current research concerning the connection between collaborative CMC and its effects on English L2 writing. Research Questions 1 and 2 ask:

RQ1: Is the regular iterative use of CMC associated with an overall improvement in the student’s English writing?

RQ2: Is the regular iterative use of CMC associated with the emergence of voice in their writing?

These two questions followed from the findings of Zhao (2013), which showed that writing with a high level of voice was also deemed to be high quality writing. Furthermore, to try to understand whether or not either of these elements of writing are positively influenced by the repeated use of collaborative CMC. Research Question 3 asks:

RQ3: What facets of CMC might be responsible for this association?
Research Question 3 also considers whether ACMC or SCMC may have been influential in improving the quality of students’ writing and eliciting their written voice. This dual focus on ACMC and SCMC is important because much of the current research has focused primarily on ACMC, and has not considered whether either of these alternative modes of communication have differing influences over on the quality of L2 writing or the salience of voice in that writing.
Chapter 3. Method

In this chapter, I outline the methods of research utilized in this study. After first defining the focus of data and analyses, I describe Design Based Research (DBR) and the purpose for its use in this study. I next outline the context of the research, including the participants (see Appendix A), the tool used (see Appendix B), the instruments and tasks used (see Appendices C, D, E, F, G, H, & I), the raters and their training sessions (see Appendices J, K, &L), the questionnaire, the interviews (see Appendices M &N), and the analysis performed.

Focus on Written Compositions

The main focus of analyses was on three separate writing samples collected from each of the participants (see Appendices C, D, E, F, & G). One sample was collected as an initial, baseline measure, one sample was collected after the participants collaborated in F2F writing sessions, and a final sample was collected after the participants collaborated online using the Etherpad©. The baseline sample was taken in the second class. Participants were given 40 minutes to compose their in-class essays based on the writing prompt (see Appendix C) that was given to them in class. The time of 40 minutes was chosen as each class session lasted for 50 minutes. After the baseline sample was taken, students were then divided into groups of four and given four weeks to collaborate on two different pieces of writing (see Appendix D). During the four weeks, each student was placed in two different groups to collaborate. In the F2F sessions, the participants were allowed to collaborate at least once during regular class time (if time permitted, they were allowed another peer editing session in class). The majority of their collaboration was done outside of class. Students would meet and edit their work. Students submitted a rough draft of each essay, and then a final draft of those essays. Minimal instructor feedback on issues of organization and topic was given on the rough draft in order to help guide
their collaborative editing sessions. However, I did not edit the drafts for them, nor did I comment on every aspect that required revision or edits. Once the second essay was completed, the students were given 40 minutes to write an in-class essay without the use of any dictionary or other outside help. Similarly, with the Etherpad© sessions, participants were divided into groups in order to collaborate on two different documents; and again, each student was in a different group for each essay. Groups were assigned a URL, which allowed them work on a single document with their group members (see Appendix B). The students used the synchronous chat sessions to plan their writing and for editing purposes. They used the collaborative document portion of the Etherpad© to write and edit their essay both synchronously and asynchronously. After the students submitted their second essay, they, like the face to face group, performed an in-class essay without the use of dictionaries or other outside help (see Appendices E &F).

**Design-Based Research**

This mixed methods study followed principles of DBR. Design based research can demonstrate what elements in practice are successful as well as make predictions about correlations and relationships that are evident but not completely comprehended (Gregor, 2006). I have used DBR to explore the relationship between collaborative writing and overall improvement in L2 writing, hoping to establish a basis for future predictions and propositions about the state of this relationship.

Anderson and Shattuck (2012) described how DBR was developed by educational researchers as a means for contending with the difficulties of performing classroom-based research. This approach allows for the substantiation of the research by situating the design in a real educational context. Brown (1992) argued that effective pedagogical interventions should
have the capacity to move from the experimental classroom setting to an average classroom setting whose make-up consists of both average students and average teachers. This extension is especially important with regards to L2 research, which has a long legacy of laboratory experimental studies with limited pedagogical relevance. Brown further suggested that within the scope of DBR, interventions could be comprised of learning activities, assessments, or technological interventions. The present study focuses on an intervention that combines these components. Mingfong et al. (2010) posited that through the use of DBR, researchers create a contextual domain that offers opportunities for other researchers to implement such interventions into their own work in hopes of recreating findings or expanding upon those findings. Ultimately, DBR allows teachers as well as researchers to design and refine pedagogies and learning tools (MacKinnon, 2012).

Design-based interventions (DBI) can be viewed as evolutionary processes predicated on the notion that the design will undergo multiple iterations in order to achieve desired outcomes (Anderson & Shattuck, 2012). MacKinnon (2012) illustrated the process of the iterations and shows that this type of methodology has a characteristically dynamic element. The researcher must not only consider the data and analyze the constraints of the data, but then begin a review process leading to a redesign, a new implementation, and new evaluation process. The researcher must go through at least two iterations to see the full evolutionary process of the design. It is through these iterations and evolutions that the research tools and conceptual models activate understanding and promote adjustments of both the context and the intervention in order to maximize opportunities for learning (MacKinnon, 2012). Managing the iterations can be a complicated task as it is never quite clear at what point (if at all) the research is complete (MacKinnon, 2012). Although this caveat does exist, DBR has become an increasingly popular method for research done in the areas of technology in education.
One major aim of DBR is to develop a better means for educators to achieve a particular outcome. For example, researchers may find through several iterations of a particular practice or tool that they are not achieving the expected outcomes due to the lack of efficacy of the practice or tool. The cyclical nature of DBR, and its ability to reinvent itself allows researchers to reconfigure and retool the design of a model or intention in order to achieve the desired results. The consecutive development of multiple iterations allows researchers to identify and learn the best mode for accomplishing their desired end result (Gregor, 2006). In this study, the cycles included two modes of collaborative writing. Students were first asked to collaborate on writing in F2F group sessions. Over the course of these sessions, learners would attend their writing class where they would receive instruction about writing. Learners would then meet outside of class to collaborate on a single piece of writing. A rough draft of this writing was submitted to me, the instructor, and minimal feedback was given. Group members would then edit and revise their work to create a final draft. After this, learners were placed in a new F2F group to repeat the process. Once this cycle had been completed, the students were asked to perform an in-class writing task, which was timed and did not allow them to use dictionaries or other outside tools. Upon completion of this cycle and the in-class writing, participants were then grouped in order to collaborate outside of class using the Etherpad©. The Etherpad© cycle followed the same pattern as described above for the F2F cycle.

The current study incorporated a mixed-methods approach, within which both quantitative and qualitative data were collected and analyzed. The data were collected either concurrently or sequentially, and given a hierarchical priority in addition to integrating the data at one or more stages during the research process (Creswell et al., 2003). For the purposes of this study, participant’s writing samples from the baseline and two iterations were rated and scored for voice intensity, overall writing quality, and for separate voice intensity markers.
These scores were statistically analyzed and results compared with those from the Technology Satisfaction Questionnaire (TSQ) and further informed by individual interviews. The TSQ is a satisfaction questionnaire based on a Likert scale used to gauge learners’ satisfaction with written collaboration on the Etherpad© and also written collaboration in general, whether F2F or through the use of collaborative CMC. Likewise, the data gathered from the TSQ were coded and statistically analyzed. Finally, the interviews were recorded and later transcribed and coded then analyzed through qualitative analysis. The data were collected with the writing iterations taking place first, then the participants completed the TSQ and subsequently participated in interviews.

**Context of the Research**

This study was conducted within an English Language Bridging Program at a major university in Ontario. Dooey (2010) noted that such bridging or pathway programs are usually designed as alternative pathways to university admissions for students who are academically ready yet lack the language proficiency to enter directly into the university. The program in which the participants were enrolled is a bridging diploma program designed to increase the English language proficiency of students who have received a conditional acceptance into the university. After successful completion of the program, students are allowed to enroll in a full-time course load in their respective major. As well, the students are given a diploma showing the successful completion of their coursework. The courses within the program are English for Academic Purposes courses, which run for one academic year. The students are required to study full time and complete and pass all of their courses within an academic year. At the time of my research, there were eight classes in the program in total (2 writing, 2 reading, 2 presentation, 1 discourse, and 1 academic skills). As well, the students were required to take
first-year course of the program’s choosing. This was to give the students an opportunity to acculturate to an English-dominant university environment while having full support of the bridging program. The students would receive either a pass or fail mark in the eight bridging classes, and would be given a traditional letter grade for the first-year course. The research was conducted in the Advanced Academic Writing Course of the bridging program.

The writing course in which the study took place was a thirteen-week course. The course focused on teaching students how to write within an academic context, and comprised twelve modules, which incorporated the fundamentals of academic writing. During the thirteen weeks, students learned about how to properly structure their writing. They learned how to organize their essays in a logical manner. Students were also taught strategies for grammar use and word choice. The course dictated that students must begin writing in a formalized manner, and that they must use precise language. As well, students were taught proper citation methods for their field of study and were expected to use these citation methods in all of their written assignments.

Participants

This study involved 35 ESL students (who responded to the invitation letter shown in Appendix A). A frequency analysis of the demographic questions of the TSQ (see Appendix M) revealed that 50% of those participating in the survey were female. All of the participants identified themselves as being between the ages of 18 to 25. With regards to ethnicity, 93.3% identified themselves as Asian/Pacific where 6.7% identified themselves as “other”. The L1 of the majority of the students was Mandarin with the rest speaking Cantonese. Two of the students were raised outside of China; therefore, each spoke a second L1 (Japanese and Vietnamese), yet both were fluent in one of the two Chinese dialects. All of the students had conditional
acceptance into an undergraduate program in the university, and had accordingly been placed into the bridging diploma program. They had been accepted either into engineering, health sciences, business, social sciences, or humanities programs. All students had been admitted into the program with minimum scores on the following language proficiency tests: 70 on the Test of English as a Foreign Language (TOEFL), 5.0 on the International English Language Testing System (IELTS), 55 on the Canadian Academic English Assessment (CAEL), or 65 on the Michigan Language Assessment Battery (MLAB). In order to successfully graduate from the program and remove the English proficiency condition from their acceptance to the university, the students would need a minimum final score equivalent to 6.5 on the IELTS in all four of their terminal courses and an overall score of 6.5 on their final exams.

**Instructor**

I acted as both the instructor and researcher over the course of the study. I have been teaching English for Academic Purposes (EAP) at the university level since the fall of 2008. Prior to this, I received two Bachelor of Arts degrees, one in Spanish/Foreign Language and Literature and the second in English Literature. As well, I studied in the Teacher Preparation program at Rutgers University. In the summer of 2008, I received a TESOL certificate from the University of Toronto, and in the fall of 2009 I began my studies as a part-time student working toward an MA in Language and Literacies Education (LLE) at the University of Toronto.

Since 2008, I have taught EAP courses that cross the four language skills of writing, reading, speaking, and listening and across all levels of language proficiency. Recently, I began teaching in bridging programs aimed at international students who do have a high level of English leaving high school, but would struggle in a first-year university classroom setting. These courses are designed to give students instruction that focuses on the use of the English
language within a university context. Students learn how to write academically researched papers using an academic tone in their writing. They learn how to take part in seminars and discussions; as well as, give both group and individual presentations. Students are given strategies for listening and taking notes, and are given strategies for reading textbooks and acquiring new vocabulary. These programs typically offer opportunities for the students to become a part of the regular student body, and offer them mentorship on campus to set them up for future success.

Along with teaching both EAP courses and bridging programs, I also have over two years of experience teaching writing at the graduate level. The courses that I have taught have focused on the fundamentals of academic writing, (tone, register, etc.) and also on the development of a research question into a graduate-level research paper.

**Instructor as Researcher**

Since this study was to be executed within a language classroom, I felt it beneficial to act as both the researcher and the instructor. In doing this, I would be able to observe participating students directly and to control for any variance that might occur among different instructors in the two sections of the course. In order to assume this dual role, I had to obtain permission from both the Ethical Review Board at the University of Toronto and the institution in which the study took place. Potential participants were ensured that their participation or their decision not to participate in this study would in no way affect their mark in this course (see Appendix A). Therefore, I had a proxy present the study to the students, and the proxy gathered all signed consent forms, so that I would not be aware of which students I was teaching might, or might not, be participating in the research. The proxy sealed these in an envelope and sent them to my thesis supervisor at the University of Toronto. He then locked them safely away in his office.
until the course was over and their marks were submitted to the registrar. Students were also
told that no identifiable information would be used in the write-up of this study in order to
protect their identities. Students were offered no compensation for their participation in the
study. They were offered a copy of a short write-up about the research upon their request, and
told that participation in this study could benefit their learning experiences. All 36 students in
the program were active in all activities in order to ensure anonymity among the participants.
However, only 35 of those enrolled agreed to participate in the study.

Research Instruments and Procedures

The CMC Tool

Etherpad© (http://openetherpad.org/, July 2015) (see Appendix B) was used during the
data collection phase of this research. This is a collaborative CMC tool that utilizes both
synchronized chat and writing sessions as well as allowing for asynchronous writing sessions.
Learners collaborated on a piece of writing (see Appendices D and E) while using a chat session
on one side of the screen and either a synchronous or asynchronous writing document on the
other side of the screen. The document function allows for learners to edit their writing in a
collaborative fashion by recording any edits that are made to the document, and showing who
made each edit. Learners can then discuss these edits on the chat portion of the screen (see
Appendix B). Each user is assigned a color and they are able to create their own usernames,
allowing the users to see who is working on a document at any given time. I was also able to
have access to the tool to be able to monitor revisions and discussions. I created a specific
Etherpad© URL link for each group that was distributed by me via the class’ learning platform.
All parties involved were able to view this link from any remote location at any time. As well, I
was able to enter the session and give any necessary feedback to the learners or answer any questions that they might have.

**Writing Tasks and Cycles of Data Collection**

Data were collected from two sections of the same course of L2 learners. The purpose of having one instructor for both sections and having the researcher act as the instructor was to ensure that all of the participants received the same level of instruction throughout the data collection process. In order to achieve these ends, the instruction of the classes ran concurrently and on pace. The same materials and instructional media were used in both classes. All assignments and assessments were considered equivalent in each section of the course.

At the start of the research, participants were asked to perform a writing task to assess the features of voice in their writing at the onset of the study as well as to gauge the overall quality of each participant’s writing. Scores from this task were used as a baseline measure with which to compare later writing samples from the participants. A sample of the diagnostic prompt appears in Appendix C.

**Assessment Instruments**

The Voice Rubric designed by Zhao (2012) was used to measure the presence of voice in the writing (see Appendix F), and the REEP rubric measured the overall quality of writing (see Appendix G).

**Voice intensity rubric.** The *voice rubric* is an adoption of Zhao’s (2012) rubric, which focuses on seven discourse features that commonly indicate authorial voice in writing (as described in Chapter 2): content point, directives, hedges, boosters, attitude markers, authorial self-mention, and reader reference. When evaluating the students’ compositions for voice, raters
were not asked to judge the overall quality of voice; rather, they were to only record the frequency in the use of those voice elements in the writing samples. Once the voice ratings were completed, the same raters were asked to rate the same essays again but this second time for their overall quality of writing.

**REEP rubric.** This second rating was done using the REEP writing rubric adapted from Mansoor and Grant (2002). REEP stands for (A)Rlington Education and Employment Program, an adult ESL education program in Arlington, Virginia to help students become employable.

This rubric is scored out of a scale of 0 to 6 with 6 being the highest. The rubric contains four categories that have been deemed crucial for high quality writing: content vocabulary, organization and development, structure, and mechanics. Voice was included in the original REEP rubric as one of five categories; however, for the purposes of this study it was not used because voice was rated separately in the Voice Intensity rubric.

**Cycles of instruction.** Students were given forty minutes to write each task during classes of fifty minutes. After the completion of the baseline writing sample, several instructional sessions took place without using the CMC tool. These sessions were designed to incorporate traditional collaborative writing methods. The participants were randomly assigned to task groups using an online group making tool (http://www.aschool.us/random/random-pair.php; last viewed July 12, 2015). The size of these groups was determined based on the total number of students in the class. The groups were typically between 4-5 students. Participants collaborated on two different writing tasks within their groups (see Appendix D). At the end of these two sessions with traditional collaborative writing, another writing sample was taken (see Appendix B) for a future comparison of writing with and without the use of the tool. This writing sample prompt was identical to the baseline sample prompt in order to account for possible topic effects, and the same was true for the next writing sample described immediately
below. However, the participants were not advised ahead of time that the writing tasks would be based on the same topic.

Upon the completion of the first cycle of the study, the next phase began, which incorporated instructional sessions that allowed the participants to collaborate using Etherpad©. Two instructional iterations using the tool took place. During each iteration, the participants were assigned to groups using the same method noted above and asked to collaborate on a writing assignment (see Appendix E). Once the iterations were completed, another writing sample was taken from the individual participants in the same manner as previously noted for the baseline and the F2F iterations. This final sample was taken to assess voice and overall writing quality (see Appendix B). These assessments were later compared to the initial baseline sample as well as the sample taken without instruction using the tool in order to determine whether any emergence of voice or improvement in writing quality could be attributed to the multiple iterations using the tool.

**Questionnaire**

To further determine whether or how the tool played a role in the learning of voice in L2 writing, a satisfaction questionnaire (TSQ) (see Appendix M) was distributed to the participants for their completion. The questionnaire asked students to rank their answers to items on a Likert scale of 1 to 5 with 1 being strongly disagree and 5 being strongly agree. The questionnaire included a section that gathered demographic information from the participants. As well, there were eight questions designed to specifically ascertain participant’s level of satisfaction with the use of the Etherpad© as an instructional tool. The next eight questions were designed to ascertain student’s level of satisfaction with regards to the collaborative writing process as a whole.
Interviews

Finally, a group of participants was interviewed individually to discuss their experiences with the tool in order to determine what, if any, features of the tool may have led to the observed outcomes (see Appendix N). All participants were asked to volunteer, and out of 35 participants, 11 chose to volunteer for the interviews. A simulated recall session took place as part of the interview process in which the researcher and the participants looked at writing samples to discuss word choice and phrasing that characteristically are determiners of voice. Interview questions (see Appendix L) were also designed to gain an understanding of how the participants felt about the two different collaboration conditions (CMC and F2F), their level of participation, and whether or not they used any other CMC during their collaboration sessions. The interviews were done in person. I conducted the interviews with the assistance of a colleague for the purposes of note taking. The interviews were recording using a digital voice recorder, and written notes were taken of each interview. Participants were given a sheet with the interview questions (see appendix O) and two samples of their writing to use throughout the interview. The participants were allowed to freely answer the questions to the best of their memory and knowledge. I only intervened while they answered for clarification’s sake. Each interview took approximately 20 minutes. These data were audio recorded, transcribed, and coded for qualitative analysis.

Raters and Training

Two external raters were solicited to rate the level of authorial voice as well as the overall writing quality of the participants’ compositions (see Appendix J). These raters had both taught Academic Writing within an L2 English context, but were not currently the instructors of
the participants. Both raters hold four-year university degrees as well as Bachelor of Education degrees, post-graduate TESOL certificates, and relevant Masters Degrees. Rater 1 holds an M.Ed., had extensive training in L2 writing assessment, and had assessor certifications in both the Canadian Language Benchmark Placement Test (CLBPT) and the International English Language Test Systems (IELTS). Rater 2 holds an M.Sc. and had no specific training in L2 writing or assessment; however, she had taught writing in EAP programs at the secondary and post-secondary levels.

A rater training session (see Appendices J, K, & L) was organized to ensure that raters understood all aspects of the rating process as well as to ensure that there was a high degree of reliability among the two raters. The dates and times of the samples of students’ writing were not disclosed to the raters, nor was the location and name of the institution in which the samples were taken disclosed to avoid biasing the ratings. Furthermore, any personally identifying materials related to the participants was removed from the writing samples. The rater training session took place over a period of three hours in a single day. Both raters were present for the session, signed the consent form, and completed the Rater Background Questionnaire (see Appendix J). During the training session, the raters were trained in how to use the Voice Intensity Rubric (see Appendices H & J) and the REEP rubric (see Appendices I & K). In the training of the Voice Intensity rubric, raters were instructed to count the number of each of the voice intensity makers (i.e. content point, directives, hedges, boosters, attitude markers, authorial self-mention, and reader reference). An overall rating of the writing samples was reserved for the overall voice score. Raters were to use the numbers of each marker to help them determine an overall score.

After the completion of the rating using the Voice Intensity rubric, raters were then to look at the same piece of writing and rate it for overall writing quality. Before the initial ratings
were done during the training session, the raters were given a sample of writing that was deemed to have a high level of voice (see Robinson & Modrey, 1986 p.160). After reviewing that exemplar, the rating session began. Each rater was given one copy of the same sample of writing. They were asked to first rate it for voice and then rate the overall writing quality. They then presented their scores for overall voice intensity and overall writing quality as well as the numbers for each voice intensity marker. Each rater presented identical scores for each component being marked, so there was a perfect degree of inter-rater reliability during this session. The raters were then asked to discuss their rationales for the scores, and then the same steps were repeated with a second sample of writing. Again, the raters presented the same score for each of the components being rated at a high level of agreement. The rater training session and the rating took place after all of the data collection had occurred in order to avoid any bias as to when the students’ compositions had been written or by which student.

The raters rated each sample independently after the rater training session. Once the raters submitted their scores, these were coded for each rater and a correlational analysis was performed to determine the degree of inter-rater reliability, described below.

Analyses

Analyses consisted of two sets of quantitative analysis performed on the scores for the writing samples and the scores of the TSQ. As well, a qualitative analysis was performed on the interviews given by the participants. It is important to note that while 35 students participated in the study, most of the results refer to only 33 data points. This is a result of missing data due to not all students submitting a composition.

Analyses for Voice and Overall Writing Quality
Once the writing samples were rated by each of the two raters, I discovered discrepancies among their ratings when calculating correlations between the raters’ scores for each variable measured. There were 27 variables in total: for three of these, a correlation score between raters could not be calculated as one rater gave a constant score for the variable.

**Correlational analysis for the baseline scores.** The correlational analysis for the baseline voice intensity score showed that there was a low level of reliability between the two raters for this variable \((r=.24, n=33, p=.17)\). A similar study was run for the baseline overall writing quality scores. Again, the analysis showed that there was a low level of inter-rater reliability \((r=.03, n=33, p=.87)\). A correlational analysis was performed to determine if there was a significant correlation between the baseline scores for content points, and the results revealed that there was a low level of reliability between the two raters \((r=.33, n=33, p=.06)\). An analysis of the scores for directives showed that there was a moderate level of inter-rater reliability between the two raters \((r=.63, n=33, p=.00)\). Likewise, an analysis was performed to assess the correlations between the scores of the two raters for hedges. However, due to the fact that at least one of the variables was deemed to be a constant, an analysis could not be fully performed as these scores were considered to include a constant score. An analysis of the scores for boosters was then performed, and the analysis showed that there was a low level of reliability among the two raters \((r=.16, n=33, p=.38)\). Next, an analysis of the scores for attitude markers was performed, and the scores revealed that there was a moderate level of reliability among the raters \((r=.53, n=33, p<.00)\). An analysis of the scores for authorial self-mentions showed a high level of reliability among the two raters \((r=.90, n=33, p=.00)\). Finally, an analysis of the scores for reader references was performed and showed that there was a high level of inter-rater reliability \((r=.62, n=33, p=.00)\).
Correlational analysis for the scores after the F2F iterations. Once the correlational analysis had been performed on the baseline scores, the scores of the two raters for the F2F iterations were compared and analyzed for each variable. The first variable to be analyzed for a correlation were the scores for voice intensity. The analysis for the scores revealed that there was a moderate level of reliability between the two raters ($r=.58$, $n=33$, $p=.00$). Next, an analysis for the scores of overall writing quality revealed that there was a low level of inter-rater reliability ($r=.40$, $n=33$, $p=.02$). An analysis of the scores for content points was performed, and the analysis showed that there was a low level of inter-rater reliability ($r=.29$, $n=33$, $p=.10$). After this, an analysis of the scores for directives revealed that there was a moderate level of inter-rater reliability ($r=.39$, $n=33$, $p=.03$). A correlational analysis of the scores for hedges was performed: however, the analysis revealed that there was a low level of reliability between the two raters ($r=.17$, $n=33$, $p=.35$). Once the analysis for the hedges had been performed, an analysis for the scores of boosters was performed. This analysis showed that there was a very low level of inter-rater reliability ($r=.06$, $n=33$, $p=.73$). An analysis for the scores of attitude markers revealed low inter-rater reliability ($r=.41$, $n=33$, $p=.02$). Next, an analysis of the scores for authorial self-mentions showed that there was a moderate level of reliability between the two raters ($r=.68$, $n=33$, $p=.00$). Finally, an analysis of the scores for reader references was performed. The results showed low inter-rater reliability ($r=.21$, $n=33$, $p=.24$).

Correlational analysis for the scores after the Etherpad© iterations. Once the correlational analysis had been performed on the F2F iteration scores, the scores of the two raters for the Etherpad © iterations were compared and analyzed for each variable. The first variable analyzed for a correlation was voice intensity, which revealed a low level of reliability between the two raters ($r=.32$, $n=31$, $p=.08$). Next, an analysis for the scores of overall writing quality revealed that there was a moderate level of inter-rater reliability ($r=.49$, $n=.31$, $p=.01$).
An analysis of the scores for content points was performed, and the analysis showed a low level of interrater reliability ($r=0.11, n=31, p=0.56$). After this, an analysis of the scores for directives revealed that there was a moderate level of inter-rater reliability ($r=0.40, n=31, p=0.03$). A correlational analysis of the scores for hedges was performed: however, the analysis was not able to be completed as at least one of the variables was a constant. An analysis for the scores of boosters showed no inter-rater reliability ($r=0.00, n=31, p=0.99$). An analysis for the scores of attitude markers revealed low inter-rater reliability ($r=0.23, n=31, p=0.22$). An analysis of the scores for authorial self-mentions was not able to be completed because at least one of the variables was a constant. Finally, an analysis of the scores for reader references showed low interrater reliability ($r=0.01, n=31, p=0.96$).

**Creation of new variables.** After the correlational analyses were performed, it was noted that out of the original 27 variables scored by the raters, 13 were deemed to have low inter-rater reliability, three were not able to be analyzed for correlations, and 11 had moderate or high levels of reliability among the raters. Therefore, it was decided that the average of the two raters’ scores for each of the variables was to be used for subsequent analysis of the data. Hence, 27 new scores for variables were created using the averages of the scores of the two raters.

Upon the creation of the new variables as described, descriptive analyses were run to calculate the mean scores for the 33 students and their standard deviations.

**Analyses of Questionnaires**

Out of the 35 participants, 30 submitted a completed Technology Satisfaction Questionnaire. Lack of attendance did not permit that a questionnaire be collected from all of those participating in the study. After the questionnaires were completed and submitted, a descriptive analysis, a frequency analysis, and a Chi-square test were performed.
Analyses of Interviews

The interview consisted of five questions that were designed to gather information with regards to the research questions outlined in this thesis. The questions aimed to gauge whether participants had a preference for F2F collaborative sessions as opposed to sessions using the collaborative CMC tool. The questions also sought to understand whether participants thought that the CMC tool prompted them to contribute more in the collaborative sessions than during F2F collaborative sessions.

I discovered during the progress of the study that participants were using other forms of collaborative CMC other than Etherpad© (i.e. WeChat and QQ) to chat during both the F2F sessions and the Etherpad© sessions. WeChat is a messaging service that provides a platform for text messaging, voice messaging, and video conferencing with broadcasts being between one to many. WeChat is popularly used among Chinese populations, equivalent to that of Facebook within the English-dominant societies. WeChat allows participants to communicate in Chinese. Similarly, QQ is a social media platform created by a Chinese company that, among other things, offers users the opportunity to participate in chat-based text sessions as well as video-based chat sessions. Therefore, an interview question was developed to gain a better understanding as to why the participants chose to communicate via these media. Finally, a question was designed in which the interviewees were asked to look over some of their writing samples and discuss the decision making process and whether one form of collaborative writing helped them in this process over another form (i.e. F2F v. Etherpad©).

The interviews were analyzed based on themes that arose from the participant’s responses. Each interview was transcribed and coded according to thematic information.
Chapter 4. Results

This chapter reports the results of the data analyses described in Chapter 3. Results of the analyses of the students’ writing are reported first as the primary basis for the study. Next, the results of the TSQ are reported to support and help interpret the findings related to the analyses of writing. Finally, the results from the interview sessions lend further qualitative support to the findings of the study.

Results of Analyses of Writing

Below are the analyses of the scores given for each writing sample. Each iteration produced one sample of writing per participant (baseline, F2F, Etherpad©). Each sample was rated using the Voice Intensity Rubric and the REEP rubric. Scores were for Voice Intensity, Overall Quality of Writing, Content Point, Directives, Attitude Markers, Authorial Self-mention, and Reader Reference. One descriptive analysis was performed for each iteration. Following these descriptive analyses, results of Wilcoxon-Signed Ranks Test are reported for each variable across the baseline and two iterations.

Baseline Descriptive Statistics

The descriptive statistics (see Table 1) for the initial baseline writing were as follows according to each of their variables: Voice Intensity \((n=33, \text{Min}=2.00, \text{Max}=4.00, \text{M}=3.29, \text{SD}= .63)\); Overall Writing Quality \((n=33, \text{Min}=2.5, \text{Max}=5.00, \text{M}=3.60, \text{SD}= .60)\); Content Points \((n=33, \text{Min}=2.00, \text{Max}=4.50, \text{M}=3.39, \text{SD}= .68)\); Directives \((n=33, \text{Min}=1.00, \text{Max}=10.00, \text{M}=4.27, \text{SD}= 2.17)\); Hedges \((n=33, \text{Min}=.00, \text{Max}=5.00, \text{M}=1.05, \text{SD}=1.00)\); Boosters \((n=33, \text{Min}=.00, \text{Max}=3.50, \text{M}=1.29, \text{SD}= .91)\); Attitude Markers \((n=33, \text{Min}=.50, \text{Max}=4.50, \text{M}=2.93, \text{SD}=1.1)\).
$Max=6.00, M=2.97, SD=1.20$; Authorial Self-mention ($n=33$, $Min=.00, Max=6.00, M=1.77, SD=1.74$); and Reader References ($n=33$, $Min=.00, Max=12.00, M=3.32, SD=3.53$).

**Table 1. Descriptive Statistics for the Baseline Writing Samples**

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$Min$</th>
<th>$Max$</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice Intensity</td>
<td>33</td>
<td>2.00</td>
<td>4.00</td>
<td>3.29</td>
<td>.63</td>
</tr>
<tr>
<td>Quality Overall</td>
<td>33</td>
<td>2.50</td>
<td>5.00</td>
<td>3.61</td>
<td>.60</td>
</tr>
<tr>
<td>Content Point</td>
<td>33</td>
<td>2.00</td>
<td>4.50</td>
<td>3.39</td>
<td>.68</td>
</tr>
<tr>
<td>Directives</td>
<td>33</td>
<td>1.00</td>
<td>10.00</td>
<td>4.27</td>
<td>2.17</td>
</tr>
<tr>
<td>Hedges</td>
<td>33</td>
<td>.00</td>
<td>5.00</td>
<td>1.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Boosters</td>
<td>33</td>
<td>.00</td>
<td>3.50</td>
<td>1.29</td>
<td>.91</td>
</tr>
<tr>
<td>Attitude Markers</td>
<td>33</td>
<td>.50</td>
<td>6.00</td>
<td>2.97</td>
<td>1.20</td>
</tr>
<tr>
<td>Authorial Self-mention</td>
<td>33</td>
<td>.00</td>
<td>6.00</td>
<td>1.77</td>
<td>1.74</td>
</tr>
<tr>
<td>Reader References</td>
<td>33</td>
<td>.00</td>
<td>12.00</td>
<td>3.32</td>
<td>3.53</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**F2F Descriptive Statistics**

The descriptive statistics (see Table 2) for the second iteration, based on a sample of writing taken after students were involved in two F2F collaborative writing sessions, were for each scored variable: Voice Intensity ($n=33$, $Min=1.50, Max=4.00, M=3.30, SD=.67$); Overall Writing Quality ($n=3, Min=2.00, Max=5.00, M=3.77, SD=.80$); Content Points ($n=33$, $Min=1.50, Max=4.00, M=3.55, SD=.64$); Directives ($n=33$, $Min=.50, Max=9.00, M=3.67, SD=1.93$); Hedges ($n=33$, $Min=.00, Max=3.00, M=1.32, SD=.86$); Boosters ($n=33$, $Min=.00, Max=2.50, M=.67, SD=.63$); Attitude Markers ($n=33$, $Min=.50, Max=7.00, M=3.23, SD=1.45$); Authorial Self-mentions ($n=33$, $Min=.00, Max=1.00, M=.09, SD=.26$); and Reader References ($n=33$, $Min=.00, Max=7.00, M=.77, SD=1.60$).
Table 2. Descriptive Statistics of Samples Taken after the F2F Sessions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice Intensity</td>
<td>33</td>
<td>1.50</td>
<td>4.00</td>
<td>3.39</td>
<td>.67</td>
</tr>
<tr>
<td>Quality Overall</td>
<td>33</td>
<td>2.00</td>
<td>5.00</td>
<td>3.77</td>
<td>.80</td>
</tr>
<tr>
<td>Content Point</td>
<td>33</td>
<td>1.50</td>
<td>4.00</td>
<td>3.55</td>
<td>.64</td>
</tr>
<tr>
<td>Directives</td>
<td>33</td>
<td>.50</td>
<td>9.00</td>
<td>3.67</td>
<td>1.93</td>
</tr>
<tr>
<td>Hedges</td>
<td>33</td>
<td>.00</td>
<td>3.00</td>
<td>1.32</td>
<td>.86</td>
</tr>
<tr>
<td>Boosters</td>
<td>33</td>
<td>.00</td>
<td>2.50</td>
<td>.67</td>
<td>.63</td>
</tr>
<tr>
<td>Attitude Markers</td>
<td>33</td>
<td>.50</td>
<td>7.00</td>
<td>3.23</td>
<td>1.45</td>
</tr>
<tr>
<td>Authorial Self-mention</td>
<td>33</td>
<td>.00</td>
<td>1.00</td>
<td>.09</td>
<td>.26</td>
</tr>
<tr>
<td>Reader References</td>
<td>33</td>
<td>.00</td>
<td>7.00</td>
<td>.77</td>
<td>1.61</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Etherpad© Descriptive Statistics

The final iteration consisted of a writing sample taken after participants collaborated on two pieces of writing via the use of the Etherpad© Collaborative Writing Tool. The descriptive statistics (see Table 3) for each variable assessed were: Voice Intensity ($n=31$, $Min=2.50$, $Max=4.50$, $M=3.79$, $SD=0.54$); Overall Writing Quality ($n=31$, $Min=2.50$, $Max=5.50$, $M=4.01$, $SD=0.68$); Content Point ($n=31$, $Min=2.50$, $Max=4.50$, $M=3.74$, $SD=0.50$); Directives ($n=31$, $Min=0.00$, $Max=12.00$, $M=5.18$, $SD=2.71$); Hedges ($n=31$, $Min=0.00$, $Max=3.00$, $M=1.32$, $SD=0.97$); Boosters ($n=31$, $Min=0.00$, $Max=3.50$, $M=0.94$, $SD=0.79$); Attitude Markers ($n=31$, $Min=2.00$, $Max=7.50$, $M=4.10$, $SD=1.36$); Authorial Self-mention ($n=31$, $Min=0.00$, $Max=1.50$, $M=0.13$, $SD=0.34$); and Reader References ($n=31$, $Min=0.00$, $Max=3.50$, $M=0.57$, $SD=0.97$).

Table 3. Descriptive Statistics of Writing Samples after the Use of Etherpad©

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice Intensity</td>
<td>31</td>
<td>2.50</td>
<td>4.50</td>
<td>3.79</td>
<td>.54</td>
</tr>
</tbody>
</table>
A Wilcoxon Signed Ranks Test was performed for each of the seven major variables to see if the scores on any of them for the three writing samples changed significantly over time.

**Change in voice over iterations.** The first of the tests analyzed the Voice Intensity Rating to determine if there was a significant change in the scores between the baseline and the sample taken after the F2F sessions (Voice Intensity F2F and Voice Intensity Baseline), the baseline and the samples taken after the Etherpad© sessions (Voice Intensity Etherpad© and Voice Intensity Baseline), and between the samples taken after the F2F sessions as compared to those taken after the Etherpad© sessions (Voice Intensity Etherpad© and Voice Intensity F2F) (see Table 4). The Wilcoxon Signed Ranks Test show that there was no significant change between Voice Intensity F2F and Voice Intensity Baseline ($Z = .77, p = .44$). There was a statistically significant change between Voice Intensity Etherpad© and Voice Intensity Baseline ($Z = 3.00, p = .00$). As well, there was a statistically significant change between Voice Intensity Etherpad© and Voice Intensity F2F ($Z = 2.39, p = .02$).
Overall writing quality over iterations. A second Wilcoxon Signed Ranks Test (see Table 5) was performed to determine whether a change had taken place in the overall writing quality among the three samples. The results revealed that there was no statistically significant change between the baseline and the samples taken after the F2F sessions ($Z=1.22, p=.22$). There was, however, a significant change between the baseline and samples taken after the Etherpad© sessions ($Z=2.39, p=.02$), but there was no significant change between the Etherpad© sessions and the F2F sessions ($Z=1.10, p=.27$).

Table 5. Change in Overall Writing Quality among the Three Iterations$^a$

<table>
<thead>
<tr>
<th></th>
<th>F2F- Baseline</th>
<th>Etherpad©- Baseline</th>
<th>Etherpad©- F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z$</td>
<td>-1.22$^b$</td>
<td>-2.39$^b$</td>
<td>-1.11$^b$</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.22</td>
<td>.02</td>
<td>.27</td>
</tr>
</tbody>
</table>

a. Wilcoxon Signed Ranks Test  
b. Based on negative ranks

Analysis of Voice Intensity Markers

The next series of tests focused on the Voice Intensity Markers outlined in the Voice Rubric. According to Zhao (2010, 2012), these markers are critical for writing to have a high level of voice intensity. Thus, these markers were looked at individually to ascertain whether the CMC tool had any influence in the number of these markers in the participants' writing.
**Change in number of content points over iterations.** The first of these tests was performed on the Content Point Markers (see Table 6). The results show that there was no significant change between Content Point F2F and Content Point Baseline ($Z=1.13, p=.26$). However, there was a statistically significant change between Content Point Etherpad© and Content Point Baseline ($Z=2.35, p=.02$), but there was no statistically significant change between Content Point Etherpad© and Content Point F2F ($Z=1.54, p=.13$).

| Table 6. Change in the Number of Content Points among the Three Iterations$^a$ |
|---------------------------------|------------------|------------------|------------------|
|                                  | F2F- Baseline    | Etherpad©- Baseline | Etherpad©- F2F    |
| $Z$                              | -1.13$^b$        | -2.35$^b$         | -1.54$^b$        |
| Asymp. Sig. (2-tailed)           | .26              | .02              | .13              |

$a$. Wilcoxon Signed Ranks Test  
$b$. Based on negative ranks

**Change in number of directives over iterations.** The next test focused on the change in the number of Directives that appeared in the writing among the three samples taken (see Table 7). The results show that there was no significant change between Directives F2F and Directives Baseline ($Z=1.62, p=.11$). As well, there was no significant change between Directives Etherpad© and Directives Baseline ($Z=1.68, p=.09$). However, there was a statistically significant change between Directives Etherpad© and Directives F2F ($Z=3.18, p=.00$).

| Table 7. Change in the Number of Directives among the Three Writing Iterations$^a$ |
|---------------------------------|------------------|------------------|------------------|
|                                  | F2F- Baseline    | Etherpad©- Baseline | Etherpad©- F2F    |
| $Z$                              | -1.62$^b$        | -1.68$^c$         | -3.18$^c$        |
| Asymp. Sig. (2-tailed)           | .11              | .09              | .00              |

$a$. Wilcoxon Signed Ranks Test  
$b$. Based on negative ranks
a. Wilcoxon Signed Ranks Test
b. Based on positive ranks
c. Based on negative ranks

**Change in number of hedges over iterations.** A signed ranks test was then performed to determine if there was a change in the number of hedges that appeared in the writing (see Table 8). There was no significant difference between Hedges F2F and Hedges Baseline ($Z=1.31, p=.19$). Likewise, there were no significant results for Hedges Etherpad© and Hedges Baseline ($Z=1.08, p=.28$). Finally, there was no statistically significant difference between Hedges Etherpad© and Hedges Baseline ($Z=.12, p=.90$).

<table>
<thead>
<tr>
<th></th>
<th>F2F- Baseline</th>
<th>Etherpad©- Baseline</th>
<th>Etherpad©- F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Z</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.19</td>
<td>.28</td>
<td>.90</td>
</tr>
</tbody>
</table>

Table 8. Change in the Number of Hedges among the Three Writing Samples*

**Change in number of boosters over iterations.** A signed ranks test was performed to determine whether a change had occurred in the number of boosters that appeared in the writing across the three iterations (see Table 9). The results show that there was a statistically significant change between Boosters F2F and Boosters Baseline ($Z=2.90, p=.00$). However, there was no significant change between Boosters Etherpad© and Boosters Baseline ($Z=1.58, p=.12$). As well, there was no significant change between Boosters Etherpad© and Boosters F2F ($Z=1.27, p=.20$).

<table>
<thead>
<tr>
<th></th>
<th>F2F- Baseline</th>
<th>Etherpad©- Baseline</th>
<th>Etherpad©- F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Z</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.19</td>
<td>.28</td>
<td>.90</td>
</tr>
</tbody>
</table>

Table 9. Change in the Number of Boosters among the Three Writing Samples*
Change in number of attitude markers over iterations. Attitude markers were the next focus of the signed ranks tests (see Table 10). The results show that there was no significant change between Attitude Markers F2F and Attitude Markers Baseline ($Z=0.47, p=0.64$). However, there was a statistically significant change between Attitude Markers Etherpad© and Attitude Markers Baseline ($Z=3.08, p=0.00$). Likewise, there was a statistically significant change between Attitude Markers Etherpad© and Attitude Markers F2F ($Z=2.51, p=0.01$).

### Table 10. Change in the Number of Attitude Markers among the Three Writing Samples

<table>
<thead>
<tr>
<th></th>
<th>F2F- Baseline</th>
<th>Etherpad©- Baseline</th>
<th>Etherpad©- F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z$</td>
<td>-2.91&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-1.58&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-1.27&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.00</td>
<td>.12</td>
<td>.20</td>
</tr>
</tbody>
</table>

*a. Wilcoxon Signed Ranks Test  
b. Based on positive ranks  
c. Based on negative ranks*

Change in number of authorial self-mentions over iterations. A change in the number of Authorial Self-mentions was next assessed through the use of a signed ranks test (see Table 11). The results of show that there was a statistically significant change between Authorial Self-mention F2F and Authorial Self-mention Baseline ($Z=3.84, p=0.00$) as well as between Authorial Self-mention Etherpad© and Authorial Self-mention Baseline ($Z=3.79,$
However, there was no significant change between Authorial Self-mention Etherpad© and Authorial Self-mention F2F ($Z=.34$, $p=.73$).

Table 11. Change in the Number of Authorial Self-Mentions among the Three Writing Samples

<table>
<thead>
<tr>
<th></th>
<th>F2F-Baseline</th>
<th>Etherpad©-Baseline</th>
<th>Etherpad©-F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z$</td>
<td>$-3.84^b$</td>
<td>$-3.78^b$</td>
<td>$-0.34^c$</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.00</td>
<td>.00</td>
<td>.73</td>
</tr>
</tbody>
</table>

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Change in number of reader references over iterations. The final analyses performed was a signed ranks test to determine if there was change among the three writing samples with regards to reader references (see Table 12). The results show that there was a statistically significant change between Reader References F2F and Reader References Baseline ($Z=3.36$, $p=.00$) as well as between Reader References Etherpad© and Reader References Baseline ($Z=3.68$, $p=.00$). However, there was no significant change between Reader References Etherpad© and Reader References F2F ($Z=0.77$, $p=.44$).

Table 12. Change to the Number of Reader References among the Three Writing Samples

<table>
<thead>
<tr>
<th></th>
<th>F2F-Baseline</th>
<th>Etherpad©-Baseline</th>
<th>Etherpad©-F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z$</td>
<td>$-3.36^b$</td>
<td>$-3.68^b$</td>
<td>$-0.77^b$</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.00</td>
<td>.00</td>
<td>.44</td>
</tr>
</tbody>
</table>

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks
Correlational Analysis of Voice Intensity and Overall Writing Quality

**Correlational analysis in baseline samples.** A correlational analysis of the baseline samples comparing voice intensity and overall writing quality revealed that there was a statistically significant correlation ($r=0.52$, $n=33$, $p=0.00$) between these variables.

**Correlational analysis in F2F iterations.** A correlational analysis of the scores from the F2F iterations showed that there was a statistically significant relation between the scores for voice intensity and the scores for overall writing quality ($r=0.89$, $n=33$, $p=0.00$).

**Correlational analysis of Etherpad® iterations.** A correlational analysis of the Etherpad® revealed a statistically significant correlation between voice intensity and overall writing quality ($r=0.78$, $n=31$, $p=0.00$).

**Results from the Technology Satisfaction Questionnaire**

Participants were asked to identify their level of computer expertise as well as the number of language courses taken in which CMC was used as part of the teaching process. With regards to level of expertise, the majority of the participants (63.3%) estimated their level of expertise as intermediate, 16.7% claimed that they were novices, 10% claimed that they had no experience, 6.7% chose to select the not applicable choice, and 3.3% claimed that they were expert users. In response to the number of language courses taken that included the use of CMC as part of the learning process, 31% indicated that they had 3 years of experience, 20.7% noted that they had 2 years of experience, 17.2% noted that they had no years of experience, 6.9% noted that they had 1 and 4 years respectively, and 3.4% claimed that they had 7 and 10 years respectively. (Note that only 29 of the 30 participants responded to this question.)
Descriptive analysis focused on the frequencies of the scores for the 16 satisfaction questions (see Table 13), eight of which considered the participant’s satisfaction with the use of Etherpad© and eight of which considered their satisfaction with the collaborative writing process as a whole. As well, Chi-square tests (see Tables 14 & 15) were performed to assess whether the findings were significantly different from a random (uniform) distribution.

Table 13. Descriptive Analysis of Responses to Technology Satisfaction Questions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key</td>
<td>35</td>
<td>1</td>
<td>35</td>
<td>18.00</td>
<td>10.25</td>
</tr>
<tr>
<td>Satisfaction 1</td>
<td>30</td>
<td>1</td>
<td>5</td>
<td>3.50</td>
<td>1.01</td>
</tr>
<tr>
<td>Satisfaction 2</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>3.33</td>
<td>.84</td>
</tr>
<tr>
<td>Satisfaction 3</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>3.63</td>
<td>.77</td>
</tr>
<tr>
<td>Satisfaction 4</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>3.50</td>
<td>.86</td>
</tr>
<tr>
<td>Satisfaction 5</td>
<td>30</td>
<td>1</td>
<td>5</td>
<td>3.47</td>
<td>1.01</td>
</tr>
<tr>
<td>Satisfaction 6</td>
<td>30</td>
<td>1</td>
<td>5</td>
<td>3.27</td>
<td>.94</td>
</tr>
<tr>
<td>Satisfaction 7</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>3.30</td>
<td>.79</td>
</tr>
<tr>
<td>Satisfaction 8</td>
<td>29</td>
<td>2</td>
<td>5</td>
<td>3.59</td>
<td>.68</td>
</tr>
<tr>
<td>Collaboration 1</td>
<td>30</td>
<td>1</td>
<td>5</td>
<td>3.23</td>
<td>1.01</td>
</tr>
<tr>
<td>Collaboration 2</td>
<td>29</td>
<td>2</td>
<td>5</td>
<td>3.83</td>
<td>.85</td>
</tr>
<tr>
<td>Collaboration 3</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>3.80</td>
<td>.85</td>
</tr>
<tr>
<td>Collaboration 4</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>3.93</td>
<td>.74</td>
</tr>
<tr>
<td>Collaboration 5</td>
<td>29</td>
<td>2</td>
<td>5</td>
<td>3.79</td>
<td>.90</td>
</tr>
<tr>
<td>Collaboration 6</td>
<td>29</td>
<td>2</td>
<td>5</td>
<td>3.86</td>
<td>.92</td>
</tr>
<tr>
<td>Collaboration 7</td>
<td>29</td>
<td>1</td>
<td>5</td>
<td>3.10</td>
<td>1.15</td>
</tr>
<tr>
<td>Collaboration 8</td>
<td>29</td>
<td>1</td>
<td>5</td>
<td>3.83</td>
<td>.89</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 14. Chi-Squared Test of Etherpad© Satisfaction Questions 1 to 8

<table>
<thead>
<tr>
<th></th>
<th>Key</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td></td>
<td>.00</td>
<td>10.80</td>
<td>9.20</td>
<td>14.00</td>
<td>9.20</td>
<td>13.33</td>
<td>15.33</td>
<td>13.47</td>
</tr>
<tr>
<td>df</td>
<td>34</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 15. Chi-Squared Test of Collaboration Questions 1 to 8

<table>
<thead>
<tr>
<th>Key</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>0.00^a</td>
<td>11.33^b</td>
<td>10.31^c</td>
<td>10.00^d</td>
<td>18.27^d</td>
<td>17.76^c</td>
<td>5.90^c</td>
<td>8.76^c</td>
</tr>
<tr>
<td>df</td>
<td>34</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Asymp. Sig.</td>
<td>1.00</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
<td>.00</td>
<td>.00</td>
<td>.12</td>
<td>.07</td>
</tr>
</tbody>
</table>

Satisfaction Question 1

Question one of the first eight questions, which considered the participant’s satisfaction with the use of the Etherpad© CMC, asked participants whether they were able to learn from the Etherpad© sessions. The descriptive statistics for this question were $n=30$, $Min=1$, $Max=5$, $M=3.50$, $SD=1.01$ (see Table 13). The analysis of frequencies revealed that 46.7% of the participants gave this question a score of 3, 30% scored this question with an answer of 4, 16.7% of the participants answered with a score of 5, and 6.7% gave the question a score of 2. The Chi-square test showed that these results were significantly different from a random uniform distribution: $X^2 (3, n=30) =10.80$, $p=.01$ (see Table 14).

Satisfaction Question 2
Question two asked participants whether they were stimulated to become more active in the Etherpad© sessions. The descriptive statistics for this question were $n=30$, $Min=2$, $Max=5$, $M=3.33$, $SD=.84$ (see Table 13). The frequency analysis shows that 40% of the participants selected 3 as their ranking, 36.7% selected 4, 16.7% selected 2, and 6.7% selected 5. The Chi-square test shows that these results were significantly different from a random uniform distribution: $X^2 (3, n=30) = 9.20, p=.03$ (see Table 14).

**Satisfaction Question 3**

In question three, participants were asked whether they thought online chat and editing sessions assisted in the development of their writing. The descriptive statistics for this question were $n=30$, $Min=2$, $Max=5$, $M=3.63$, $SD=.77$ (see Table 13). The frequencies associated with question three revealed that 43.3% of the participants gave a score of 3, 40% of the participants gave a score of 4, 13.3% gave a score of 5, and 3.3% gave a score of 2. The Chi-square test shows that the results were significantly different from a random uniform distribution: $X^2 (3, n=30) = 14.00, p=.00$ (see Table 14).

**Satisfaction Question 4**

Question four asked participants whether as a result of this course, they would take another course which utilized a similar form of computer technology. The descriptive statistics for this question were $n=30$, $Min=2$, $Max=5$, $M=3.50$, $SD=.86$ (see Table 13). The frequency analysis of this question shows that 43.3% of the participants scored the question with 4, 33.3% gave it a score of 3, 13.3% gave it a score of 2, and 10% gave a score of 5. The results were significantly different from a random uniform distribution: $X^2 (3, n=30) = 9.20, p=.03$ (see Table 14).
Satisfaction Question 5

The fifth question asked participants whether the use of Etherpad© prompted them to participate more in discussions. The descriptive statistics for this question were $n=30$, $Min=1$, $Max=5$, $M=3.47$, $SD=1.01$ (see Table 13). The frequency analysis revealed that 40% of respondents gave this question a rating of 3, 30% a rating of 4, and 16.7% a rating of 5, 10% a rating of 2, and 3.3% a rating of 1. As well, the Chi-square test shows the results to be significantly different from a uniform distribution: $X^2 (4, n=34) =13.33, p=.01$ (see Table 14).

Satisfaction Question 6

Question six asked participants if they had put a great deal of effort into learning how to use the Etherpad© throughout the course. The descriptive statistics were $n=30$, $Min=1$, $Max=5$, $M=3.27$, $SD=.94$ (see Table 13). The frequency analysis shows that 36.7% of the respondents gave a score of 4 as well as 36.7% gave a score of 3. 16.7% scored this question with a 2, 6.7% with a 5, and 3.3% with a 1. The Chi-square test shows the results to be significantly different from a uniform distribution: $X^2 (4, n=30) =15.33, p=.00$ (see Table 14).

Satisfaction Question 7

Question seven asked participants if the use of Etherpad© helped to ensure that their level of learning in the course was of the highest quality. The descriptive statistics of this question were $n=30$, $Min=2$, $Max=5$, $M=3.30$, $SD=.79$ (see Table 13). The frequency analysis reveals that 50% of the respondents scored this question with a 3, 30% with a 4, 13.3% with a 2, and 6.7% with a 5. The Chi-square test was significantly different from a uniform distribution: $X^2 (3, n=30) =13.47, p=.00$ (see Table 14).
Satisfaction Question 8

The eighth and final question pertaining to the satisfaction of the use of the Etherpad© asked whether participants felt that overall the instruction that they received through the use of the Etherpad © met their learning expectations. For this question, only 28 respondents gave an answer. The descriptive statistics for this question were $n=29$, $Min=2$, $Max=5$, $M=3.59$, $SD=.68$ (see Table 13). The frequency analysis reveals that 48.3% of the respondents gave the question a score of 4, 41.4% a score of 3, 6.9% a score of 5, and 3.4% a score of 2. The Chi-square test shows that these results were statistically different in their distribution compared to a random uniform distribution: $X^2 (3, n=29) =18.59, p=.00$ (see Table 14).

Collaboration Questions

The final eight questions of the survey focused on the participants’ satisfaction with the collaborative writing process. As with the previous eight questions, they were asked to rate their satisfaction on a scale of 1 to 5 with 1 being the lowest and 5 being the highest. A descriptive analysis was performed on the results of these questions (see Table 13), and a summary of those results is given below.

Collaboration Question 1

The first of the eight questions asked whether participants felt that the collaborative learning experience was better when done in the collaborative communication platform (Etherpad©) as opposed to face-to-face collaborative sessions. The descriptive statistics for this question were $n=30$, $Min=1$, $Max=5$, $M=3.23$, $SD=1.01$ see Table 13). The frequency analysis showed that 36.7% gave a score of 3, 30% gave a score of 4, 20% gave a score of 2, 10% gave a
score of 5, and 3.3% gave a score of 1. The Chi-square test shows that the results were significantly different from a uniform distribution: \( X^2(4, n=30) = 11.33, p = .02 \) (see Table 15).

**Collaboration Question 2**

The second question asked participants whether they felt part of a learning community within any of the groups in which they worked. The descriptive statistics for this question were \( n=29, \text{Min}=2, \text{Max}=5, M=3.83, SD=.85 \) (see Table 13). The frequency analysis reveals that 48.3% of the respondents rated this question with a score of 4, 24.1% with a score of 3, and 20.7% with a score of 5, and 6.9% with a score of 2. The Chi-square test was significantly different from a uniform distribution: \( X^2 (3, n=29) = 10.31, p = .02 \) (see Table 15).

**Collaboration Question 3**

Participants were then asked if they actively shared their ideas with other group members. The minimum score given was 2 with a maximum score of 5. The descriptive statistics for this question were \( n=30, \text{Min}=2, \text{Max}=5, M=3.80, SD=.85 \) (see Table 13). The frequency analysis revealed that 46.7% of the respondents scored the question with a 4, 26.7% with a 3, 20% with a 5, and 6.7% with a 2. The Chi-square score shows that the results were significantly different from a uniform distribution: \( X^2 (3, n=30) = 10.00, p = .02 \) (see Table 15).

**Collaboration Question 4**

Question four asked participants if they were able to develop new knowledge and skills by working with other group members. The descriptive statistics for this question were \( n=30, \text{Min}=2, \text{Max}=5, M=3.93, SD=.74 \) (see Table 13). The frequency analysis shows that 56.7% of the respondents answered with a score of 4, 20% with a score of 5, 20% with a score of 3, and
3.3% with a score of 2. The Chi-square test shows that the results were significantly different from a random uniform distribution: \(X^2 (3, n=30) = 18.25, p=.00\) (see Table 15).

**Collaboration Question 5**

In question five, participants were asked if they were able to develop their editing skills via collaboration sessions. The descriptive statistics for this question were \(n=29, Min=2, Max=5, M=3.79, SD=.90\) (see Table 13). The frequency analysis revealed that 58.6% of the respondents responded to this question with a score of 4, 17.2% with a score of 5, 10% with a score of 3, and 3.8% with a score of 2. The Chi-square test showed a high level of statistical significance compared to a random uniform distribution: \(X^2 (3, n=29) = 17.76, p=.00\) (see Table 15).

**Collaboration Question 6**

Question six asked participants whether collaborative learning within their group was effective. Note that only 29 of the 30 respondents answered this question. The descriptive statistics for this question were \(n=29, Min=2, Max=5, M=3.86, SD=.92\) (see Table 13). The frequency analysis showed that 37.9% of the respondents answered with a score of 4, 27.6% with a score of 5, 27.6% with a score of 3, and 6.9% with a score of 2. The Chi-square test shows that the results for this question were not significantly different from a uniform distribution: \(X^2 (3, n=29) = 5.90, p=.12\) (see Table 15).

**Collaboration Question 7**

The seventh question asked participants if collaborative learning within the group was time consuming. Note that only 29 respondents answered this question. The descriptive statistics for this question were \(n=29, Min=1, Max=5, M=3.10, SD=1.15\) (see Table 13). The frequency
analysis reveals that 37.9% of the respondents rated this with a score of 4, 24.1% with 3, 20.7% with 2, 10.3% with 1, and 6.9% with 5. The Chi-square test shows that the results were not significantly different from a random distribution: $X^2 (4, n=29) = 8.76, p=.07$ (see Table 15).

**Collaboration Question 8**

In the eighth and final question, participants were asked whether they felt overall satisfaction with the collaborative learning experience within this course. Note that only 29 of the 30 respondents answered this question. The descriptive statistics for this question were $n=29$, $Min=1$, $Max=5$, $M=3.83$, $SD=.89$ (see Table 13). The frequency analysis showed that 58.6% of the respondents rated this question with a 4, 17.2% with a 5, 17.2% with a 3, 3.4% with a 2, and 3.4% with a 1. The Chi-square test shows that these results were statistically significantly different from a random distribution: $X^2 (4, n=29) = 29.79, p=.00$ (see Table 15).

**Findings from Interviews**

Participants were asked to take part in an individual interview session that consisted of questions related to the collaborative writing process and the use of CMC for the purposes of written collaboration (shown in Appendix N). Out of the 35 participants in the study, 11 chose to take part in the interviews.

**Interview Question 1**

Questions one and two of the five interview questions asked participants how they felt when they wrote collaboratively in both the F2F sessions and the Etherpad© sessions. They were also asked whether these experiences were positive or negative, and to give examples. With regards to how the participants viewed their experience collaborating F2F, there were
more negative responses than positive. The participants related that they felt that collaborating F2F eliminated misunderstandings as they were able to understand group members better (than in the CMC) due to the ability to see their faces and understand their body language. This awareness led to more clarity of ideas when discussing their work. One participant felt that it was beneficial for relationship building. Two participants spoke about the benefits of F2F collaboration with regards to division of labor. They said that it made it easier to divide the work thus helping them develop their “group leadership” skills (see Excerpt 1 in Appendix O). As for the negatives, one major negative among most interviewees was that in the F2F sessions, one person typically took charge identifying themselves as the “group leader”. In doing so, that person often dominated the discussion and divided the duties as they saw fit. There was less collaborative writing. Rather, group members were often assigned sections to write (e.g. introduction, single body paragraph, and/or conclusion). They would then email this to the group leader who would often edit the writing on his/her own. Many pointed to the fact that with F2F sessions it was difficult for group members to find a mutual time to meet outside of class. They suggested this as one reason for the need for a group leader and a single editor. One student said that a major difficulty he faced was that he typically did all of his composing on the computer, but in F2F sessions, it was difficult for everyone to sit with individual laptops and take notes as well as collaboratively edit (see Excerpt 2 in Appendix O). Another participant felt that since his language level was lower, the F2F sessions were quite difficult as he often felt rushed due to time restraints (see Excerpt 3 in Appendix O). This time pressure made it a challenge for him to participate in the editing process, and to understand the edits that were made on his behalf.

Interview Question 2
The responses to the same question about the Etherpad© were quite different. Only three of the participants felt that there were any negatives associated with this form of collaboration. The majority felt that it was an overall positive experience. One major positive was that participants could see the work of others while working on their part. This visibility helped them as it served as a model for their own writing. As well, the interviewees claimed that being able to constantly see the entire text and write directly on the Etherpad© helped them with their edits and revisions as there was always a physical record of changes. In fact, one participant said that he often used the video replay application of the tool to help him see changes that had been made by his peers in order to better understand how to improve his writing. Many of the participants said this access allowed for more connected writing that would appear to a reader as coming from one author. All of the participants acknowledged that one major advantage of this CMC tool was the ability to collaborate and work at any time in any place. They also felt the fact that the instructor was able to check their progress, prompting them to be more motivated to work hard on these papers (see Excerpt 4 in Appendix O). With regards to the negative aspects of the Etherpad©, the major complaint of the three students who did not find it a positive experience was the amount of time involved in using the tool. They felt that it took more time to type out the chat sessions whereas in a F2F session they could speak directly to the other party. As well, they felt the context lacked emotion, which often led to misunderstandings. They also felt that it was difficult to arrange for their group members to be online at the same time to discuss the writing (note that this was also a complaint of the F2F sessions). The final point of negativity relates to their perceived role of the group leader. They felt that they had less control over the whole process, but mostly with regards to editing and revising. They felt that group members could go onto the Etherpad© at any time and make changes to the document without each other’s consent. In fact, one of the three interviewees said that a group member had revised
her edits without her knowing, and that this made her feel really “sad” (see Excerpt 5 in Appendix O). She said this was the biggest reason that she found the sessions using the tool to be a negative experience.

**Interview Question 3**

The third question in the interview asked the interviewees what their contribution level to the group was, and whether they contributed more during the Etherpad© sessions versus the F2F sessions. They were also asked if they had a defined role within the group, and if so, what that role was. The majority said that they participated more in the Etherpad© sessions compared to the F2F context, primarily due to the reasons noted in the above paragraph. One interviewee said that he did not really contribute much to either session type as he prefers to work alone. Most cited that their contribution levels varied depending on the group’s needs and the group dynamics. Most of the interviewees saw themselves as the “group leaders”. One participant referred to herself not as the leader, but as an organizer or guide. She said that she would give feedback and try to guide her group members rather than taking control of their writing. She said that this was more easily accomplished on the Etherpad©. Four of the students saw themselves more as participants. They said they would often work at sharing ideas, and trying to work more democratically within the group. Two of these said that they contributed more in the Etherpad© sessions compared to the F2F condition, because they felt that their language level was lower: therefore, they found it more difficult to contribute in the F2F sessions. One student said that although she was the group leader in all of her four sessions, she felt that she contributed most in the F2F sessions because she was able to speak Chinese to her group members (see Excerpt 6 in Appendix O).
Interview Question 4

During the course of the study, it was discovered that the participants were collaborating outside of the F2F sessions and the Etherpad© sessions using WeChat. Therefore, an interview question was created to get a better understanding of why they chose to use this third form of communication. They were asked whether they used any other type of CMC during their collaboration sessions, and if so, to name them, and to explain why they chose these forms of communication. They all said that they exclusively used WeChat, and one person said he tried to get his group members to use Skype. They explained the format of WeChat, which is very similar to Skype. Users can use it for video conferencing, voice-only calls, and to send text messages. WeChat is highly popular among the Chinese speaking community, and the participants said their reasoning for choosing WeChat over other CMC was their familiarity with the platform (see Excerpt 7 in Appendix O). Their rationale for using this medium during both types of collaborative sessions was that unlike in F2F meetings, they could communicate at any time and in any place. Furthermore, they said that while on the WeChat they could speak in Chinese without the instructor’s knowledge (see Excerpt 8 in Appendix O). The participants also said that this form of CMC allowed them to see the other person, which eliminated much of the misunderstandings because they could see facial expressions and read their body language, which was more like face-to-face communication. Based on these observations, the interviewees felt that WeChat enriched all of the collaborative sessions.

Interview Question 5

The final question in the interview asked the interviewees to look at two samples of their writing, and to discuss the decision-making processes that they went through when writing. All of the participants said that they could not remember what they were thinking at the time that
they wrote the compositions. They spoke primarily about how they made decisions within the group context. They either had a group leader who divided the work, and then put the sections together and edited the paper to make a cohesive essay, or they collaborated as a group by making line-by-line decisions about vocabulary, grammar, and connectivity of the writing. All of the participants said that the latter process helped with the decision making on their individual essays, but that they could not point to exact examples due to a lack of memory with regards to those pieces of writing.
Chapter 5. Discussion

Chapter 5 summarizes then discusses the findings of this study. I consider theoretical implications as well as implications for both teachers and learners. Implications for researchers are also considered along with the limitations of this study. All of the above are discussed within the context of the three guiding research questions:

RQ1  Is the regular iterative use of CMC associated with an overall improvement in the quality of students’ English writing?

RQ2  Is the regular iterative use of CMC associated with the emergence of voice in their writing?

RQ3  What facets of CMC might be responsible for this association?

Summary of Findings

Results presented in Chapter 4 indicate that students participating in this research improved their English writing over the period of the course, and they developed certain aspects of authorial voice in that writing. But the design of the research makes it impossible to attribute these improvements and developments directly to their uses of collaborative writing or CMC as distinct from other associated factors such as the students’ practicing writing, using and studying English, taking courses at an English-medium university, and simply maturing. Nonetheless, findings from the research point to features of the collaborative CMC that were associated with, and that may have facilitated, the students’ English writing improvement and development of authorial voice. These indications suggest areas to explore and evaluate in future studies that might take a more controlled approach to inquiry than the present preliminary, exploratory investigation into the possible effects of collaborative CMC on English L2 writing.

Overall Writing Quality
Results of Wilcoxon Signed Rank tests shown in Table 5 in Chapter 4 indicate that the 35 participating students improved their English writing between the start and the end of the EAP bridging program. But there was no significant improvement in writing quality after uses of Etherpad©. A slight increase in the mean scores between the writing samples taken after the F2F sessions ($M=3.77$) and those taken after the use of the tool ($M=4.02$) (see Tables 1 to 3) suggests an upward trend, nonetheless. Results from the TSQ and interview sessions could indicate a relation between the use of collaborative CMC and an improvement in writing quality. For example, this type of CMC allows for learners to have a physical record of writing of their peers so that they may model vocabulary and structure. As well, the organic nature of this type of CMC allows for the learners to manipulate their writing and allow it to evolve over time and more collaboratively rather than in isolated F2F sessions. In other words, learners are not restricted by place and time when collaborating on a document. Furthermore, being able to examine an entire document as it is being composed was said to help the learners keep track of changes that they and their group members made thereby making the writing process more salient for them. These associations remain speculative, nonetheless, in view of the empirical data from the present research.

**Voice in Writing**

Wilcoxon Signed Ranks Tests showed no significant changes between the baseline writing samples and the F2F writing samples. However, there were some significant changes between the baseline writing samples and the samples taken after the use of the CMC tool. As well, there were some significant changes between the F2F writing samples and the samples taken after the collaborative sessions on e Etherpad©. The TSQ and the interviews indicate that uses of the Etherpad© may have prompted some aspects of the emergence of voice in the
students’ English writing. Participants rated their experience with the tool and in collaborative settings high. As well, the participant’s anecdotal commentary about the CMC tool’s ability to allow them to see models of writing and language point toward possible associations between the use of the Etherpad© and an increase in the intensity of voice in their writing. Nonetheless, as observed below, certain foci of teaching in the course went in opposing directions, and may actually have curtailed students’ development of voice in their writing.

**Voice Intensity Markers**

While there was an increase in some of the features of voice in the students’ English writing after collaboration on the Etherpad©, other markers showed no significant increase. Consideration of each voice intensity marker and the context surrounding the classroom instruction including the use of the tool may give more insight.

**Content points.** Raters found significant changes in main content points in students’ writing between the baseline sample and the sample taken after the use of the Etherpad©. The mean scores for each iteration showed a steady increase between each iteration (see Tables 1 to 3). These increases may be attributable to uses of the CMC or simply to time, practice, and instruction.

**Directives.** There was a significant increase in the number of Directives used between the writing samples taken after the F2F sessions as compared to those taken after the collaborative sessions using Etherpad©. However, there was no significant change between the number of Directives seen between the baseline sample and both the samples taken after the F2F collaboration sessions and those taken after the Etherpad© collaboration sessions. Furthermore, there was a steep dip in the mean score between the scores given after the F2F sessions and the original baseline and then a distinct increase in scores between the samples taken after the
Etherpad© sessions and those taken after the F2F sessions. Likewise, there was a sharp increase between the samples taken after the Etherpad© sessions and the baseline sessions. These trends suggest that uses of the Etherpad© may be associated with an increase in the appearance of Directives in the writing of the participants.

**Hedges and boosters.** Hedges and boosters describe opposing features of writing in that both act as communicative strategies to weaken or strengthen a statement. Boosters represent strong claims through making assertions of conviction and confidence whereas hedges weaken a claim through the use of qualifiers that can show doubt and a lower level of certainty in a statement or claim (Hyland, 1998). There were no significant changes in the number of hedges or boosters written by students through the iterations of this study. Neither rater found many incidents of hedges or boosters in the writing across the three samples. This trend is surprising because a major unit in this particular writing class urged students to use hedges in their writing as the students often made claims that were too strong. Therefore, there may have been an initial drop in the number of boosters in their writing, but as the students became more confident in using a variety of features in their writing, the numbers began to increase. A slight increase in the number of boosters between the F2F iteration and the Etherpad© iteration may indicate that while students were taught to make less strong statements and claims, when able to see their writing as a whole on the Etherpad© and to see the qualifying language that other classmates used, students felt more inclined to reintroduce boosters into their writing.

Perhaps, as Gilbert (1991) noted, second language learners often express an opinion or belief as if it were a statement of fact. Gilbert argued that this tendency can seem inappropriately direct to a reader. Wishnoff (2000) argued that instruction and textbooks can lead to the lack of hedges in writing. Wishnoff claimed that much instruction in L2 academic writing in English teaches students that there should be a direct linear argument and therefore a
lack of hedges. Wishnoff further argued that through training, advanced L2 learners could increase the number of hedges in their planned writing sessions. Therefore, the trends seen here may be due to the learners having had only 40 minutes to write their essays, which gave little time for planning and may have prompted a direct approach to their discourse.

**Attitude markers.** Uses of the Etherpad© may have been associated with the appearance of attitude markers in participants’ writing. There was a slight increase in the mean score for attitude markers between the baseline sample ($M=2.90$) and the sample taken after the F2F iteration ($M=3.22$) and then again in the Etherpad© iteration ($M=4.10$).

**Authorial self-mention.** Over time, the mean frequencies of authorial self-mentions decreased, particularly between the baseline ($M=1.77$) and the samples after the F2F iteration ($M=.09$), and then this trend leveled off between the F2F iteration and the Etherpad© iteration ($M=.13$). It should be noted that students were taught in the course that authorial self-mentions are not deemed to be appropriately academic in tone for the purposes of English for Academic Purposes (EAP) writing. This influence of instruction could be a reason for the lack of significant changes in this aspect of voice in students’ writing. It may simply be that due to instruction, the participants ceased using many authorial self-mentions in their writing.

**Reader references.** There was a sharp decrease in the mean frequencies of reader references between the baseline ($M=3.32$) and the F2F iteration ($M=.77$) and then again in the Etherpad© iteration ($M=.56$). As with the instruction on authorial self-mention, students were taught in class to avoid directly addressing readers in their writing as part of the EAP Writing curriculum. Therefore, this change may be due more to instruction rather than any other factors.

**Voice Intensity and Overall Writing Quality**
Following Zhao (2012), the research questions developed for this study were designed to investigate correlations between voice intensity and overall writing quality among English L2 students. Correlational analyses showed significant associations between voice and overall writing quality across the scores of all three sets of writing samples in the present study. These findings confirm those presented by Zhao (2012) and contradict those presented by Helms-Park and Stapelton (2003).

**Satisfaction with CMC**

The majority of participants who answered the TSQ questionnaire stated that they were satisfied with the Etherpad© as a means for collaborative writing. Most participants felt that their learning experiences with Etherpad© were beneficial in helping them improve their writing. Although most students agreed that they spent a great deal of time learning how to use the Etherpad©, they also noted that they would be willing to take a course that used a similar CMC format for instruction. Furthermore, the participants felt that the collaborative writing process in general was a positive experience. Overall, the participants were satisfied with the collaborative learning and writing processes, and that they felt that they benefited from them.

**Interviews**

The interviews helped to make a few issues apparent. The first is that the collaborative process, no matter what form, influence learning experiences in a positive way, supporting ideas advanced by Storch (2013). The collaborative sessions using the Etherpad© were perceived to allow for more involvement among the participants, and so true collaboration rather than just a division of labor, as had occurred in the F2F sessions. Finally, it is apparent that L2 learners need opportunities to communicate in their L1 in order to express their ideas clearly as well as
understand what they are learning. The Etherpad© sessions coupled with the WeChat showed how the present predominantly Chinese learners were able to spend the majority of their time communicating in English while finding a common ground in which they could check their understanding in Chinese. The Etherpad© sessions, as opposed to the F2F sessions, offered them an environment that was rich in the target language, and WeChat was used as a viable support. Many of the F2F sessions were either done outside of class with discussions in Chinese or on the WeChat, where Chinese was the only language used. In these sessions, students typically divided the work, and the target language was only used in the act of writing.

**Implications**

Although exploratory and descriptive, results of the present research have implications for theory, teachers, learners, and researchers.

**Implications for Socio-cultural Theory**

Findings from the TSQ and interviews indicate that the students in this study felt that the Etherpad© offered opportunities for collaboration as a social endeavor in which they took an active role in writing a whole document and were integrally involved in editing processes. Conversely, when the students worked in groups during the F2F sessions, the work was often divvied up by a group leader, and students worked individually on a single paragraph that was later pieced together with other paragraphs by the group leader to make a complete document. As well, students noted that during the F2F sessions, the editing and revising was often the work of the group leader or one single member of the group rather than group revising and editing as a whole. Furthermore, the interviews indicated that students were able to model the writing that they observed their group members using while writing a document during the Etherpad©
sessions, and they were able to see what their classmates had written as they wrote their parts. They stated that this helped them to discover new vocabulary and improve upon their grammar and writing style. Alternatively, as noted previously, the F2F sessions did not allow for this type of modeling. Since the students chose to use the F2F sessions as a means for decision making rather than collaborative writing, the modeling of proper written forms was not achieved.

Vygotsky (1981) contended, and Warschauer and colleagues have demonstrated for CMC (1999, 2002, 2005), that to develop mediational means tools must be included within human actions, manipulating and transforming the movement and formation of cognitive functions. Therefore, it can be assumed that when it comes to a written document, both the asynchronous and synchronous functions were valuable to the participants in aiding them with learning how to write academically and in developing their authorial voices. As the interviews suggest, many of the students felt that the computer was crucial to their writing process. Students noted that they found it difficult to collaborate in F2F settings because they could not see what their classmates were writing or vice versa. The Etherpad© and other such collaborative CMC gave these learners a platform in which they are able to manipulate in more social and collaborative terms. The learners could all simultaneously view and manipulate a single document without being restricted by time or space.

Furthermore, in the interview sessions, students noted that while collaborating on their essays, they noticed aspects of their group members writing (i.e., vocabulary, grammar, and stylistic features) that they had not been including in their own writing. They were then able to use these elements in a portion they were writing for the essay, and then reflect on whether or not these forms worked for them in the same manner that their peers had been able to use such elements. The process of being able to revise and edit a living document through their online collaborative sessions further enabled students to engage in meaningful output that would, in
principle at least, help to improve their L2 writing skills. All of these elements demonstrating metalinguistic functions in which learners thought about the language, internalized aspects that they felt to be proper, and modified future output (Swain 2001, 2005). In sum, the Etherpad© would appear to be a rich tool that encourages the output of meaningful forms in order to facilitate language learning.

In addition, the work that the students did via the use of the collaborative CMC (Etherpad©) and the F2F sessions allowed them to construct new types of knowledge and learn in a way that was student centered. As the instructor, I acted only as a guide. They were encouraged to make their own decisions about content, organization, grammar, and vocabulary. According to the information gathered from the interviews, the primary difference between the Etherpad© sessions and the F2F sessions was that the Etherpad© sessions were truly collaborative. The F2F sessions still had a high level of individual work. The students used those sessions more for planning rather than writing whereas in the Etherpad© sessions, students engaged collaboratively in both aspects of the writing process. In these respects, with an appropriate task, the Etherpad© and other similar collaborative CMC offer learners opportunities that align with the ideals of Constructivist Theory (Hyland, 2008; Palacas, 1989; Pegrum, 2009; Senior, 2010).

According to Swain et al. (2002) within the context of collaborative writing and peer revisions, collaborative dialoguing emerges which mediates L2 learning. Swain et al (2002) referred to these moments of collaborative dialoguing as LREs (language related episodes). In this study, I did not focus specifically on LREs; however, when reviewing the notes from the interviews and the results of the TSQ, it was evident that through the use of the Etherpad© the students were able to engage in a great deal of collaborative dialoguing. Again, this is compared to the F2F sessions, which appeared to have been more cooperative in nature. Likewise, both
Parks, Hamers, and Lemmonier (2005) and Hyland and Hyland (2006) argued that the use of CMC in peer-feedback and peer collaboration include dialoguing that allows learners to have a permanent record of the discourse between peers. This was a major assertion of the participants in the interview sessions. They felt that being able to see the conversations and the language features that were visible in the collaborative document housed on the Etherpad© helped them feel more confident to attempt new forms in their writing. Storch (2013) and Tan, Wigglesworth, and Storch (2010) argued that CMC is a feasible domain for encouraging learners to participate in collaborative writing activities, giving learners opportunities to improve language and proper grammatical structures by permitting them to consider word choices and breakdowns in communication. The present student interviews support this idea in that the participants precisely cite that the time to evaluate what others had written and the dialogue surrounding the writing helped them to improve upon features of voice in their writing and their overall writing skills.

Finally, within her research on the emergence of voice in writing, Zhao (2012) found that there was a correlation between writing that was deemed high in voice and that which was of high overall quality. My correlational analysis showed a similar, significant correlation between voice intensity and writing quality.

**Implications for Teachers**

The findings of this study have several implications for teachers and English writing classrooms. The research shows that through the collaborative writing process coupled with collaborative CMC learners can improve the intensity of students’ voice in their writing. Uses of CMC allow for modeling of writing that gives weaker students time to observe accurate forms from their peers while working on their own writing. As well, collaborative CMC permits
students to truly collaborate on a single document rather than simply working in a cooperative manner. However, teachers need to be aware of the time that it takes for students to learn how to use collaborative CMC environments effectively. Furthermore, it may be wise to find a collaborative CMC that allows students to work on a single document while using video chat rather than text-based chat sessions. As the students indicated in their interviews, the use of video chat sessions were easier and less time consuming; as well, these sessions eliminated many of the misunderstandings that may occur through text-based chat sessions. Finally, the correlation between voice and overall writing quality should encourage teachers to find opportunities in their L2 writing classes to incorporate collaborative writing, and specifically through the use of collaborative CMC.

**Implications for Learners**

The implications for the learners is that the results of this study show that collaborative writing sessions and uses of collaborative CMC students are able to increase certain aspects of the intensity of voice in their writing as well as potently improve their overall writing quality. Collaborative CMC such as the Etherpad© allows students to actively participate in the writing process. Taking on the role of both writer and editor in context may help them to improve their writing skills. Collaborative CMC is an organic platform that encourages students to not just view this type of writing as a cooperative task, but rather to truly collaborate and take notice of different language styles and forms that they may ultimately use in their own writing.

**Implications for Researchers**

The implications for researchers are many. Foremost, research is needed to identify and confirm the sources of increases in the intensity of voice in L2 writing as well as overall
improvement of writing quality. Particularly, more research needs to be done to understand exactly what elements of collaborative CMC elicit such improvements. It was not completely clear from the present data whether ACMC or SCMC played a greater role in the improvements in the students’ writing. Additionally, it would be interesting to investigate classroom applications of a tool that allowed for both simulated F2F communication and an opportunity to collaborate both synchronously and asynchronously on a single document. According to Deft and Lengel (1996), the richest media are those which most closely resemble F2F communication within a given context. Therefore, research that considers the use of a tool that combines both types of media should be considered. A comparison could be done to determine whether this type of CMC proves to be more influential in the improvement of writing than common forms of CMC available at present. Furthermore, researchers need to consider how students’ uses of English outside of class determine their improvements on tasks within classrooms in English. An important lesson from the present study is the importance for researchers to set up a control group to be able to evaluate the effects of unique classroom interventions such as the present use of CMC in order to be able to provide insights and evidence of the causality of pedagogical treatments.

Limitations

Although the study did help to answer some aspects of my research questions, there were limitations in the research being able to do so fully. One major limitation is that causality cannot be determined. Although students’ writing did improve after the use of the CMC and collaborative writing sessions, many other variables could have contributed to this improvement. For instance, the use of the collaborative CMC came at the end of the course. It is therefore difficult to ascertain whether any noticed improvement in students’ writing resulted
from the CMC or whether this was due to the students’ practice writing sessions. As well, their instruction was solely in English in a target-language rich environment.

The lack of a control group in the research design is another factor that limited understanding of causality. Unfortunately, the nature of the EAP program made it difficult to think of how a true control group could have been established. For this reason, multiple iterations of data collection and writing collaborations in differing media, including a baseline, were performed following principles of Design-based Research. This approach made the investigation more exploratory than explanatory.

Another limitation of the study surrounds the sample size. It would be beneficial to perform the study at a later date using a larger pool of participants. Further analysis could be performed using repeated-measures methods, which was not possible in this study due to the small sample size.

An unforeseen limitation of the study was that the participants were using WeChat in both F2F sessions and the Etherpad© sessions. Unfortunately, this element was difficult to foresee or control because it was done outside of the classroom and without my knowledge. However, this circumstance points to the need for future studies to be done into the types of CMC that are the most suitable for certain contexts and populations. WeChat did simulate F2F communication, which was important for the present predominantly Chinese participants. Nonetheless, while one major benefit of the Etherpad© was that the participants could collaborate on a single document at any time in any location, WeChat does not have the capacity for students to collaborate in this way. Skype is a possible option as users are able to view each other’s computers, yet they are not actually able to manipulate any document on that computer. For these reasons, the Etherpad© seems to be one of the best options for written collaboration at
this time. Perhaps future studies could look at multiple forms of CMC to see which would be
best suited for such a context.

A final limitation of the study was inter-rater reliability. Although there was a high
degree of certainty between the raters at the outset of the rater training session, once the raters
were apart and rating, there was a discrepancy between their scores. Even a statistical average of
each score was created, and a new variable for each set was created, there were still
discrepancies between the raters, which does not make the results fully reliable. In future
studies, it would be beneficial to have raters meet frequently to benchmark their scores in order
to maintain inter-rater reliability. As well, more rater training should be offered to the raters to
ensure consistency.
References


Appendix A: Participant Consent Letter

Dear potential research participant,

My name is Jessica Turetken. I am an M.A. student at the Ontario Institute for Studies in Education (OISE), at the University of Toronto. I am writing to ask for your participation in research I am doing for my thesis research. After reading the detailed information below, if you wish to participate in this study please complete and return to my colleague who is distributing this letter the consent form attached to the bottom of this letter at the end of today’s information session. I have asked my colleague to distribute and collect these letters so that I do not know who is, or is not, participating in this research until after the present course is completed, in order to protect your privacy and not to have your decision to participate in the research affect in any way your passing the present course. I will not know which students in this class participate or not in the research until after the course is completed and grades for the course are submitted. Thank you.

**Title of Research Project:** *The impact of collaborative CMC on the emergence of voice and the improvement of overall writing quality of L2 learners.*

**Principal Investigator:** Jessica Turetken, M.A. Candidate, Department of Curriculum, Teaching & Learning, OISE/University of Toronto

**Purpose of the Study:** This study aims to examine how the use of collaborative Computer Mediated Communication (CMC) encourages the emergence of authorial voice in Second Language writing, and how this leads to a higher overall quality of writing.

**Participants:** I am looking for volunteer participants to participate in this study between October and December 2014. In order to be a participant, you must meet the following conditions. You must:

1) be a non-native speaker of English;
2) be 18 to 26 years old;
3) have a personal interest in studying writing as part of an English for Academic Purposes course (either to continue your studies or to further your employment opportunities);
4) have at least intermediate or advanced proficiency in English, approximately equivalent to a score of 68 on the TOEFL iBT or a score of 4.5 on the IELTS, or better; and
**Benefits:** The elements of the study will help you to improve your overall writing quality by increasing the presence of authorial voice in your writing. As well, you will benefit from both the face to face collaborative writing sessions as well as the online sessions. The face-to-face sessions should improve your spoken English along with your written English. Online chat sessions have also been shown to lead to an improvement in spoken English as well as written English. Furthermore, you will be gaining a full understanding of how to use such online collaborative writing tools.

**What Participants will do:** The main part of the research involves collaborative writing sessions in both a face to face context as well as an online context. The sessions have been designed as part of the regular coursework, so that all students and instructor will benefit from the outcomes of the study. The study will begin with a baseline diagnostic writing. Participants will be given 45 minutes to complete this task. Phase one will then involve regular classroom instruction along with students working in groups of 3-4 in face to face sessions to complete a single co-authored document. Students will then perform a second writing diagnostic that will be used to gauge any changes in their writing. The final phase will have students collaborate on a single co-authored document by working in groups of 3-4 using an online collaborative writing tool. Once that piece of writing has been submitted by the group, students will perform a final writing diagnostic to again denote any changes in their writing. Participants will then be given a Technology Satisfaction Questionnaire (TSQ), which will be used to assess the participant’s feelings towards the use of the collaborative writing tool as well as their comfort level with technology before and after the use of the tool. Finally, about 5 participants will be asked to be interviewed in order to gain a better understanding of lexical choices made during the collaborative writing sessions as well as get a better sense of what participants felt were the most and least useful aspects of the tool.

**Participants’ Rights**

- **To Confidentiality:** As I will be both instructor and researcher, confidentiality is critically important. The first step to keeping the participant’s identities confidential is to have a colleague present this consent letter for signing. My colleague will then seal the consent letters, and they will be kept with my supervisor Dr. Alister Cumming until the course has ended. After each phase and each diagnostic sample has been taken, I (as the instructor) will mark the writings with feedback and return a copy to the participants. Another copy will be kept until the completion of the course to which time the data will be analyzed. All of these data will be kept confidential. If you choose to participate in the research I will assign a pseudonym (false name) to you, and all of the data I collect will use this pseudonym instead of your name to conceal your identity. Please recognize, however, that the research will involve a small group of students, so all participating students may know the identities of those participating in the research, so I must ask you to not reveal the other people’s identities. All of these data will be kept in a double-locked filing cabinet in my office and/or in encrypted files on a password-protected computer in my locked office. I, and my supervisor (Professor Alister Cumming), are the only people who will have access to this information. All of these data will be systematically destroyed seven years after my thesis is completed. Results from the research will be presented in my M.A thesis and may also be published in academic journals or presented at scholarly conferences. In all these reports, no information will appear that could identify you personally. Again, all references to your data will use a pseudonym.
• **To Ask Questions about the Research:** If you wish to ask questions about this research project, you may do so at any point. Please contact me (Jessica Turetken) at 416-579-0426 or jessica.turetken@mail.utoronto.ca or you may speak to me directly. You may also contact my supervisor, Professor Alister Cumming, regarding questions at 416-978-0276 or alister.cumming@utoronto.ca. The University of Toronto also has an office regarding ethics if you want more information about your rights as a research participant, or to verify the authenticity of this research. You may contact the Office of Research Ethics at 416-946-3273 or ethics.review@utoronto.ca.

• **To Withdraw at Any Time:** You may withdraw from this study at any time. If you quit the study, you can indicate to Dr. Cumming in writing by email or by voice mail at his office phone (both indicated above) whether you wish me to use or immediately to destroy the data from you. There will be absolutely no consequences if you wish to quit the study.

**Risks:** The only potential risk in this study is any personal discomfort you may feel in class discussions or similar activities and during the interview, though these activities will be similar to what occurs in any English writing course.

Please read and complete the attached consent form to indicate your willingness, or not, to participate in this study.

Kindly,

Jessica Turetken
OISE / University of Toronto
Phone: 416-579-0426
Email: jessica.turetken@utoronto.ca

PLEASE KEEP A COPY OF THIS LETTER FOR YOUR RECORDS

******************************************************************************
Consent Form for Participants

I have read the letter from Jessica Turetken hat describes the project, *The impact of collaborative CMC on the emergence of voice and the improvement of overall writing quality of L2 learners.* I understand that all of the data collected from me will be kept confidential and secure, and the data will be destroyed after a period of seven years. I also understand that I may withdraw from the study at any time without any negative consequences, and that if I do; all information collected from me will be destroyed and/or deleted immediately. I also understand that the data collected from me may be used in academic journals or academic presentations. In such cases, I understand that no identifying information will be presented about me. I also acknowledge that my questions regarding this study have been answered satisfactorily. As such,
☐ I agree
☐ I do not agree

to participate in this study.

Name (print): ___________________  Email or Phone Number: ___________________

Signature: _____________________  Date: _________________
Appendix B: Screenshot from Etherpad© Session
Appendix C: Writing Diagnostic: Baseline Sample, F2F Iteration Sample, and Etherpad© Iteration Sample

Instructions:

Read the prompt below and write a well-crafted essay using your best English. You will have 40 minutes to write, so please use your time well. There is space in your packet to outline and plan your essay. There is also a check-list included to help you proofread your essay.

Writing Prompt:

In modern society we are constantly bombarded with new technologies. Sometimes the use of these technologies can lead to an invasion of our privacy. Is it ever necessary for the government or other organizations to use technologies such as social networks to gain information about you? Why or why not?
Appendix D: Collaborative Writing Tasks for F2F Iterations

Collaborative Writing Task:

In your groups, you will need to agree to write on one current event topic. Your group will research the topic and write a short article. The aim of this article is to inform your reader as much as possible about the topic you have chosen. You will have time in class to collaborate on the research and writing of this article. Your article should consist of 3-7 pages of well thought out content. Below is a list of possible topics. However, you may choose your own topic from current events that you are seeing in the news.

Topics:

2014 World Cup
Censorship
Internet Privacy
Traffic/Public Transit
Globalization
Climate Change
Appendix E: Collaborative Writing Tasks for Etherpad© Iterations

Collaborative Writing Task:

In your groups, you will need to agree to write on one current event topic. Your group will research the topic and write a short article. The aim of this article is to inform your reader as much as possible about the topic you have chosen. You will collaborate with your group via your assigned Etherpad© link. Your article should consist of 3-7 pages of well thought out content. Below is a list of possible topics. However, you may choose your own topic from current events that you are seeing in the news.

Topics:

2014 World Cup
Censorship
Internet Privacy
Traffic/Public Transit
Globalization
Climate Change
Appendix F: Sample of Student Writing Taken after the Etherpad© Iteration that Was Deemed to Have High Voice Intensity and to Have High Overall Quality

In the society now, with the fast speed of the social networks development, the people’s privacy become a serious problem. According to the speed of the new technologies, it is easy for government or organizations to gain information from networks. If it is necessary for government or some organization to get others profile in detail through the use of the social networks confused the majority of people. To be objective, if it is necessary for them to do this should depend on the situation.

First of all, to know people’s well is a good way to help people who in need on time and increase the country’s society level at the same time. In the whole world, there are minority people who are weak or sick, some of them are homeless, some of them are injured, some of them are starving, and some of them may be need money to take education. For this situation, it is better for the government or other legal organizations to your people’s information by using technologies such as social networks. Thus, they will know who need what kinds of help and they can know the information that the person give them is correct. Based on this, the government or the legal organizations will not give the money to people who are lieing, so the government and the legal organizations will spend the money on the people who really need help. Besides, it’s also about a country’s security. Gain much more people’s profiles is a way to know how safe the citizens’ lifes are. Because of the government got everyone’s information by use the technologies, it is easy for them to distinguish if the person is worthy, dangerous or not. They will know if the person will threatened to the country, province, city or any citizens through analyze the people’s information. From the points above, people could say it is necessary to use new technologies to gain everyone’s information.

By contrast, there is another group of people still insist that it is unnecessary for governments to collect people’s information through the technologies. There are two kinds of
people in the group. The first kind of people are the one who thinks that there are no useful and truthful information online. Even though the government or organizations gain everybody’s profile on the social network, it might be the wrong data because day by day, people smarter than before. It is hard for them to believe the networks completely now. They become more careful when they write the information about themselves online. The second kind of people in the group is the one who keep saying that when the governments or organizations gain their information from networks, it also shows that there are higher possibilities for others to take people’s information on the social network. When the alleged organization get the datas, people could be threatened and lose their money or get hurt by this. The losing of personal datas in other people's life become dangerous. Thus, this kind of people thinks it is not necessary for governments or organizations to use technologies to gain people’s information.

In conclusion, there are advantages and disadvantages whether governments or organizations take the people’s information from the networks or not. To discuss if it is necessary for the get the information should depend on the situation.
Appendix G: Sample of Student Writing Taken after the Etherpad© Iteration that Was Deemed to Have Low Voice Intensity and to Have High Overall Quality

Now new technologies become a new part of people’s life it has more and more important influence for people’s life. As far as new technologies play a very important role in our life, it is necessary for government or other organization to use technology such as social network to gain information.

New technologies will become the easiest way for government or other organizations to gain information. First of all, such new technologies like social media contain a large amount of information. Facebook have billions uses and it will be very easy for government to collect informations with the help of Facebook company. Furthermore, people like google everyday, so google know all the information about “who” search “what” and “when”. So it will be very easy for government to absorb information related to the person they interested. In conclude new technologies will be the easier way for government or other organization to get information.

Use new technologies to gain information can take government save money. Compare to the tradition way, collect information on the internet is the cheaper way to gain information. For example, few person who search an internet can do the double work than who search information in library. Furthermore, government can create with the social media company, then they are absorb information easier because of the company’s help. In conclude, use new technologies to get information can save government money.

New technologies is necessary be used by government or other organization to gain information to cause it can save the governments money and it is also the easier way to gain information.
Appendix H: Voice Rubric (Format adopted from Zhao, 2012)

**Dimension 1: Presence and clarity of ideas in the content**

<table>
<thead>
<tr>
<th><strong>A. Content point</strong></th>
<th><strong>B. Directives</strong></th>
<th><strong>Overall voice evoked by the presence and clarity of ideas in the content</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>A clear central point is articulated</td>
<td>Directives are used _____ times in the essay</td>
<td>- The reader feels a clear presence of central ideas (point of view) throughout the text.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The writing shows a strong commitment to the topic through full development of the central idea (point of view) with adequate use of effective examples and details.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The reader feels that he or she is being invited to participate in the discussion of the topic and the construction of an argument through the author’s use of directives phrases when presenting ideas.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The idea (point of view) and the use of examples and details in the writing are unique, interesting, and engaging, indicating sophisticated thinking behind the writing.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The reader feels that there is a central idea (point of view) in the text, but it is not fully developed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The writing shows some commitment to the topic with proper use of some supporting examples and details. But the examples are not always appropriate or effective.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The reader occasionally feels that he or she is being invited to participate in the discussion of the topic; but more often, the reader feels a lack of interaction with the writer.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The idea (point of view) and the use of examples and details in the writing are safe and general, lacking uniqueness, sophistication, or thoughtfulness.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The reader cannot find a consistent central idea (point of view) in the text.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The writing does not show any commitment to the topic; rather, it is only an attempt (or a failed attempt) to answer a question. No examples or details are used to develop the topic.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The reader feels that the writer is not concerned with the reader, and the writing is a confusing monologue instead of a clear dialogue between the writer and the reader.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- The writing is generic and lifeless.</td>
</tr>
</tbody>
</table>
### Dimension 2: Manner of ideas presented

<table>
<thead>
<tr>
<th>C. Hedges</th>
<th>D. Boosters</th>
<th>E. Attitude Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedges are used ______ of times in the essay.</td>
<td>Boosters are used ______ of times in the essay.</td>
<td>Attitude markers are used ______ of times in the essay.</td>
</tr>
</tbody>
</table>

#### Overall voice evoked by the presence and clarity of ideas in the content

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>The reader feels a clear presence of a central idea (point of view) throughout the text. The writing shows a strong commitment to the topic through full development of the central idea (point of view) with adequate use of effective examples and details. The tone of the writing shows personality, adds life to the writing, and is engaging and appropriate for the intended reader. Word choice, and language use by extension, is varied, often interesting, sophisticated, and eye-catching to the reader.</td>
</tr>
<tr>
<td>4</td>
<td>The writer presents ideas and claims somewhat mildly with frequent use of unnecessary hedges; only occasionally does the writing show some degree of authority and confidence. The writer seems to have a stance on the topic under discussion, but no strong attitude is revealed in the writing. The tone of the writing is appropriate for the intended reader and the purpose of the writing, but lacks personality and liveliness. Occasional interesting word choice and language use may catch the reader’s attention, but the effect is inconsistent.</td>
</tr>
<tr>
<td>3</td>
<td>The writer presents ideas and claims very mildly, showing a lack of authority and confidence in what he/she is writing. The writer seems indifferent and does not have a clear stance on or attitude toward the topic under discussion. The writer writes in a monotone that does not engage the reader at all; oftentimes the reader finds him- or herself drifting off while reading the text. Word choice or language use is flat, general, and dull, and thus unable to catch the reader’s attention.</td>
</tr>
</tbody>
</table>
## Dimension 3: Writer and Reader Presence

<table>
<thead>
<tr>
<th>F. Authorial self-mention</th>
<th>G. Reader reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>First person pronouns are used ____ times in the essay.</td>
<td>Reader pronouns are used ____ times in the essay.</td>
</tr>
</tbody>
</table>

### Overall voice evoked by writer and reader presence

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>The writer reveals him- or herself in the writing either directly or indirectly, giving the reader a clear sense of who the writer is as a unique individual. The reader feels that the writer is aware of and able to engage the reader effectively in a direct or subtle way. The sharing of personal backgrounds and experiences, if any, is effective, genuine, and engaging to the reader.</td>
</tr>
<tr>
<td>4</td>
<td>The writer reveals him- or herself in the writing to some extent, leaving the reader with some sense of who he/she is. The reader feels that the writer is aware of and trying to engage the reader in a way, but with limited success. The sharing of personal backgrounds and experiences, if any, is genuine but not so engaging or effective to the reader.</td>
</tr>
<tr>
<td>3</td>
<td>The reader has little or no sense of who the writer is as a unique individual instead of a generic, faceless person. The reader feels that the writer is not concerned with the reader or completely fails to engage the reader in any way. The sharing of personal backgrounds and experiences, if any, is generic, ineffective, and even inappropriate, making the reader feel annoyed.</td>
</tr>
</tbody>
</table>

### Overall voice strength

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>The reader feels a very strong authorial voice in the writing.</td>
</tr>
<tr>
<td>4</td>
<td>The reader feels a fairly strong authorial voice in the writing.</td>
</tr>
<tr>
<td>3</td>
<td>The reader feels a somewhat weak authorial voice in the writing.</td>
</tr>
<tr>
<td>2</td>
<td>The reader feels a very weak authorial voice in the writing.</td>
</tr>
<tr>
<td>1</td>
<td>The reader cannot really feel the presence of authorial voice in the writing.</td>
</tr>
</tbody>
</table>
## Appendix I: REEP Writing Rubric (adapted from Monsoor & Grant, 2002)

<table>
<thead>
<tr>
<th>Score</th>
<th>Content/Vocabulary</th>
<th>Organization and Development</th>
<th>Structure</th>
<th>Mechanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>- Effectively addresses the task</td>
<td>- Multi-paragraph with clear introduction,</td>
<td>- Syntactic variety</td>
<td>- Appropriate spelling and mechanical</td>
</tr>
<tr>
<td></td>
<td>- Substantive amount of information</td>
<td>- development of ideas, and conclusions</td>
<td>- Well-formed sentences</td>
<td>conventions</td>
</tr>
<tr>
<td></td>
<td>- Varied and effective vocabulary usage</td>
<td>- Ideas are connected (sequentially and logically)</td>
<td>- Few or no grammatical errors (verb tense</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Appropriate supporting details</td>
<td>mistakes, comparative and/or superlative)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>- Effectively addresses the task</td>
<td>- Can write a paragraph with main idea and</td>
<td>- Attempts a variety of structural patterns</td>
<td>- Uses periods, commas, and capitals</td>
</tr>
<tr>
<td></td>
<td>- Extensive amount of information</td>
<td>- supporting details</td>
<td>- Some errors</td>
<td>- Most conventional spelling</td>
</tr>
<tr>
<td></td>
<td>- Varied vocabulary choice and usage,</td>
<td>- Attempts more than one paragraph and may exhibit</td>
<td>- Use correct verb tense</td>
<td></td>
</tr>
<tr>
<td></td>
<td>although may have some errors</td>
<td>- rudimentary paragraph structure (intro, body,</td>
<td>- Makes errors in complex structures</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- conclusion)</td>
<td>(passive, conditional, past perfect)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>- Addresses the task at some length</td>
<td>- Use details for support or illustration (reasons</td>
<td>- Has some control of basic structures</td>
<td>- Uses periods and capitals with some errors</td>
</tr>
<tr>
<td></td>
<td>- Begins to vary vocabulary choice</td>
<td>contrast), but development of ideas is inconsistent.</td>
<td>(simple present/simple past)</td>
<td>- May use commas with compound sentences</td>
</tr>
<tr>
<td></td>
<td>- Occasional vocabulary errors, but</td>
<td>- Some ideas may be well developed while others are</td>
<td>- Some complex sentences (e.g. with when,</td>
<td>- Mostly conventional spelling</td>
</tr>
<tr>
<td></td>
<td>meaning not obscured</td>
<td>- weak</td>
<td>after...)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>- Addresses at least part of the task with</td>
<td>- Limited in appropriate details-insufficient</td>
<td>- Restricted to basic structural patterns</td>
<td>- Some punctuation and capitalization though</td>
</tr>
<tr>
<td></td>
<td>some substance</td>
<td>- amount of detail or irrelevant information</td>
<td>(simple present, subject-verb), has some</td>
<td>frequent errors that distract from meaning</td>
</tr>
<tr>
<td></td>
<td>- Limited vocabulary choice</td>
<td>- Trouble sequencing</td>
<td>errors</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Occasional vocabulary errors, but</td>
<td>- May indicate paragraphing</td>
<td>- Correct usage of adverbials (because clause)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>meaning not obscured</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>- Addresses part of the task (some but</td>
<td>- Thought pattern can be difficult to follow, ideas</td>
<td>- Serious and frequent grammatical errors</td>
<td>- Frequent errors</td>
</tr>
<tr>
<td></td>
<td>little substance) or copies from the model</td>
<td>- not connected, not logical</td>
<td>- Meaning obscured</td>
<td>- Inconsistent use of punctuation</td>
</tr>
<tr>
<td></td>
<td>- Irrelevant information</td>
<td></td>
<td>- Sentence structure repetitive</td>
<td>- Spelling may distract from meaning</td>
</tr>
<tr>
<td></td>
<td>- Frequent vocabulary errors of function,</td>
<td></td>
<td></td>
<td>- Invented spelling</td>
</tr>
<tr>
<td></td>
<td>choice and usage with meaning obscured</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>- Little comprehensible information</td>
<td>- Weak, incoherent</td>
<td>- Frequent grammatical errors</td>
<td>- Lack of mechanics</td>
</tr>
<tr>
<td></td>
<td>- May not address question</td>
<td></td>
<td>- Mostly fragments</td>
<td>- Handwriting</td>
</tr>
<tr>
<td></td>
<td>- Limited word choice, repetitions</td>
<td></td>
<td>- 2/3 phrases/simple patterned sentences</td>
<td>and/or spelling obscure meaning</td>
</tr>
<tr>
<td>0</td>
<td>- No writing</td>
<td>- No comprehensible information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix J: Rater Background Questionnaire (Format adapted from Zhao, 2010)

Name: ____________________ Date: ____________________

Directions: Please provide complete answers to all of the following questions. The information gathered will be used only for research purposes and will be kept confidential. Your name and any identifying details will not appear in the final dissertation or any of the subsequent publications.

EDUCATIONAL BACKGROUND

Please describe your current and previous educational background

<table>
<thead>
<tr>
<th>Degree/Certificate</th>
<th>Field of Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduate</td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td></td>
</tr>
<tr>
<td>Professional training in writing instruction and assessment</td>
<td></td>
</tr>
</tbody>
</table>

Appendix K: Manual for Voice Rating (Format adapted from Zhao, 2010)

Before you start to rate:

1. The nature of the writing samples:
The student essays were handwritten in class during a 90 minute timed assessment period. Some essays may be incomplete as participants may not have finished within the 90 minutes. There may be spelling and grammatical mistakes as the participants were not allowed to use neither a dictionary nor any other aid. Everything will be exactly as the participants wrote it during the assessment session. The essays may also include a section in which the participant planned their essay. These items may cause the essays to appear awkward. However, you are not to rate the essays based on their overall quality (i.e. spelling, grammar, etc..). As well, you are not to rate anything from the planning phase of their writing. Keep in mind that you are ONLY rating the voice intensity as outlined in the voice rubric. DO NOT let the quality of the writing influence how you rate the essays.

When you are rating:

2. The construct being evaluated:
You are NOT SUPPOSED to judge the quality of the voice in your rating of the 7 voice categories. You are only recording the level of salience in the use of those voice elements in these writing samples, which were seen as an index of the intensity of an author’s voice in written discourse. After you have finished rating those 7 categories for salience, you will then rate the essay for its overall voice strength.

3. Voice versus accent:
The study is not interested in defining the accent of the author in the written texts. So, if you see all the broken sentences or unidiomatic expressions in the writing samples, do not take those as evidence of voice and say, “wow, there’s such a clear voice of an ESL writer in here!” Instead, just focus on rating the salience of the use of those individual voice elements identified in the preliminary voice rubric.

4. Hedging: use of can, could would
Need to be differentiated from subjunctive mood and past tense; e.g.,
    If I were…I could/would have…
    “If I hoped to be an English teacher, I would have taken many English classes.”
    “When I chose the subject that I didn’t like, I could not concentrate.”
    And also from “can” in the sense of being able to; e.g.,
    “Birds can fly.”
    “Without X, how can we live in the world where X is indispensable?”

5. Authorial self-mention:
In these writing samples, “we” can be used either as authorial self-identification or as reader references. Raters should be more careful when looking at this word. See the examples below:
Example 1: “When I was little, my brother and I always went to…., we
loved….” [authorial self-mention]
Example 2: “At different stages of life, we need to make choices; we need to make decisions to keep moving forward.” [Reader pronoun use]

6. Articulation of Central Point:
Reasons used to support the central point are not considered as articulation or restatement of the central point, unless the central point is re-stated when the reasons were given. For example, if the central point is “I definitely think to do X is important”, and the following paragraphs expand on that point by giving readers supporting reasons. So, paragraph 2 concludes with a statement as “This is why I thought it is important to do X” and paragraph 3 begins with “It is important to do X also because ….”, and the last paragraph concludes by restating the point “Therefore, I think to do X is really important”. In other words, count these (with explicit reference to the central point) as reiteration of central point.

7. Directives:
Cohesive devices/transitional words such as first, second, finally, in conclusion are not considered as directives. Only those phrases that refer readers to certain information either stated in the text or elsewhere outside of the text and those explicit signals to readers with regard to what is being stated or how to interpret what has been stated are counted toward this category.

After your training session, when rating actual essays:

8. Plan your rating sessions:
If possible, find a time to rate these essays right after this training session. Your mind is most familiar with the rubric and the application of the rubric right after the training; so it will be most efficient to do the ratings right after. Otherwise, you will need to re-familiarize yourself with the rubric and the whole rating process, which costs you extra time and may also undermine the purpose of having the training session today—to help maximize reliability in the ratings and validity of the study results.

Try to do 20 writing samples at a time—approximately 2 hours this takes. Do take a break before you go back to rate another 20 samples. Doing too many at a time may lead to fatigue, which could cause potential reliability issues. Doing too few at a time is not efficient considering the fact that you will need to revisit the rubric each time you start a new rating session; plus, it may also threat the reliability (particularly intra-rater reliability) in the ratings you gave.
Appendix L: Manual for Overall Quality Rating

Before you start to rate:

1. Nature of the samples:
The student essays were handwritten in class during a 90 minute timed assessment period. Some essays may be incomplete as participants may not have finished within the 90 minutes. There may be spelling and grammatical mistakes as the participants were not allowed to use neither a dictionary nor any other aid. Everything will be exactly as the participants wrote it during the assessment session. The essays may also include a section in which the participant planned their essay. Some of these items may cause the essays to appear awkward. Although you are rating the essays for the overall quality, do not rate any notes that the students have written in relation to the planning of their essay. You will only rate the essays based on the criterion outlined within the REEP Writing Quality rubric.

What you are rating:

2. The construct being rated:
You will be rating the participant’s essays for the overall quality of the writing. You will use the REEP Writing Quality rubric as part of this process. The rubric includes four categories related to writing quality: Content/Vocabulary, Organization and Development, Structure, and Mechanics. Essays should be scored for each category on a scale of 0-6 (with 0 being the lowest and 6 being the highest) depending on the presence criteria listed within the category
Appendix M: Technology Satisfaction Questionnaire

Instructions: This questionnaire is designed to measure your perceptions on the level of collaborative learning, social presence, and satisfaction. There is no right or wrong answer for each question. However, it is important for you to respond as accurately as possible by checking the most appropriate response.

SECTION 1. GENERAL INFORMATION

1. What is your gender?
   __ Female
   __ Male
   __ Not applicable

2. What is your age?
   __ under 18
   __ 18–25
   __ 26–35
   __ 36–45
   __ Above 45
   __ Not applicable

3. What is your predominant ethnic background?
   __ Caucasian
   __ African–American
   __ Latino
   __ Asian/Pacific Islander
   __ Other
   __ Not applicable

4. Please estimate your level of computer expertise.
   __ No experience
   __ Novice
   __ Intermediate
   __ Expert
   __ Not applicable

5. How many language courses have you taken so far that have used CMC as part of the learning process? Please circle the number.
   0 1 2 3 4 5 6 7 8 9 10 more than 10
### SECTION 2. SATISFACTION

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I was able to learn from the Etherpad© sessions</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>I was stimulated to become a more active participant in the writing process through the Etherpad© sessions</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Online chats and editing sessions assisted me in developing my writing</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>As a result of my experience in this course, I would like to take another course which uses similar technologies.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>The use of Etherpad© prompted me to participate more in discussions.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>I put a great deal of effort in learning how to use Etherpad© throughout this course.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>As a result of using Etherpad©, my level of learning that took place in this course was of the highest quality.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Overall, the instruction that I received through the use of Etherpad© met my learning experience.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

### SECTION 3: Collaboration

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Collaborative learning experience in the computer-mediated communication (Etherpad©) environment is better than in the face-to-face environment</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>I felt part of the learning community in my group</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>I actively exchanged my ideas with my group members</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>I was able to develop new knowledge and skills from other members in my group</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>I was able to develop editing skills through peer collaboration</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Collaborative learning in my group was effective</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Collaborative learning in my group was time-consuming</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Overall, I am satisfied with my collaborative learning experience in this course</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix N: Individual Interview Questions

Interview Questions:

1. Think back to when you were placed in groups to do collaborative writing both in face-to-face writing sessions as well as when using the Etherpad. How did you feel about working on one piece of writing as a group? Was the experience positive or negative? Can you describe a situation where both positive and negative things happened?

2. Did you feel that you contributed a great deal to the group writing process? If so, how? If not, why not? Was there a medium (face-to-face or through the Etherpad) that allowed you to become more involved in the collaborative process?

3. Did you or any of your group members collaborate via any other medium during the collaborative writing process (i.e. WeChat, QQ, etc…)? If so, why did you choose to communicate this way?

4. I will now show you a sample of your writing and ask you a few questions about word choices and other decisions that you might have made during that process. Take a moment to look back over the document and try your best to recall the decision making process that you went through when writing.
Appendix O. Excerpts from Interviews

Excerpt 1: Question 1 Interviewee 1

_Interviewer:_ Describe how you felt when you wrote collaboratively in face-to-face sessions? Was the experience positive or negative? Please give some examples.

_Angie:_ I think that when I write collaboratively in a group in face to face sessions, Like, I think there are two, umm both ways, like both negative and positives. Because the negative way..uh..some..like a few students don’t like to do their own work they just like follow the group. Like the group leader just did things and they didn’t think about the topic, they just like listen to others...But it is also like apostive point, because like the group leader can increase their skills like quickly, and the other positive point..uhm...some group members can set up a good relationship like between who don’t know each other before like they can become a good friends.

Excerpt.2: Question 1 Interviewee 2

(same question asked as above)

_Jack:_ Uhm..is this asking what I’m feeling when I write in a group?...First of all, I think when I’m writing I need software support, so I cannot write without a computer. But, in a group writing, especially in a face to face I cannot write with my own PC, because I work on my own desktop. So, I think that is one of the negative things. Another is I think when we have a group writing, especially a face to face, we cannot really take our time. I think I am the kind of person that needs to take time on writing. But, the positive thing is I think the group members can have more connection with others....the essay we wrote we just write separetely. Each person just have one paragraph. The connection was weak.
Excerpt 3: Question 1 Interviewee 3

(same question asked as above)

**Ethan:** Uhm, I think, I think face to face is very common. Is not always good, because sometimes we need to come to one location and somebody live far away, so they cannot come on time. Environment of the location. If we need to find a library is the best place to discuss the essay, but if there are not any library near by the group members, how can they meet up?...Uhm, I think, uh, in my opinion, when I edited face to face, make me feel rushed and uhm, sometimes, because you know, our level is different, and when you want to do something you want to do fast and finish my job with them.

Excerpt 4: Question 2 Interviewee 1

**Interviewer:** Describe how you felt when you wrote collaboratively in groups using Etherpad©? Was the experience positive or negative? Please give some examples.

**Leyla:** Uhm, I think it’s a little bit between positive and negative, because on the Etherpad our instructor could see who’s working and who’s not. Some people they will choose to add in things, like maybe some students start really fast and quickly, they start typing and writing a paragraph. If our teacher is to check the Etherpad© they will see who is working and who is not, so other people will start editing things or writing a little bit or even outlines, and we can find their which part is not...not really good or which part is really good that we never thought about. We could talk about on the Etherpad or talk about using phone, but we edit on the Etherpad©...Sometimes we just sit our work and do it at home, so we could not see each other’s work, how they doing and how they express their ideas, but on the Etherpad© we can see “yes” he understand our outline or he have got good ideas and we can share....Sometimes, when I
work with someone and I hand in the essay I find that other person, he have different ideas we have different topics...I think this always happen in the face to face...The Etherpad® made writing more connected.

Excerpt 5: Question 2 Interviewee 2

(same question as above)

Consuelo: I don’t know...we just separate the work. For example, one person just writes the introduction and conclusion and the other people write the body paragraph. And, someone wants to change your paragraph and you feel so sad and want to change it back... So sad, and you think I disagree with you...

Interviewer: Did you ever help anyone with their paragraph?

Consuelo: Sometimes, but before I change their paragraph, I would call them before.

Interviewer: So you would literally call them? You wouldn’t chat with them?

Consuelo: Just call them.

Interviewer: When it came to changing the paragraphs then...So, the difference between the Etherpad and the Face to face.

Consuelo: No disputes with the face to face. No disputes.

Excerpt 6: Question 3

Interviewer: Describe your contributions to group writing processes. Were you more involved in face-to-face or Etherpad sessions? How? What role did you take in your group? Please give examples.

Naomi: Uh, I will, uh, I will more like the Etherpad sessions, because you can keep the record and see what’s up. But, after the face to face sessions you will forgot something if you not take
the note and the other people will forget. Sometimes the plan in the face to face session is not working and we will not use after we come back home and start our part and we think “okay this is better so I use the new one. The old one we just talk, but not use. So, the rule taking in our group, we basically just talk about what we just write. Sometimes we write the thesis statement, but when we get home we use the topic but with different word. I don’t think that face to face is working….I just part of, because I am not a good writer. I cannot become the leader because sometimes if I choose not good topic, or give not good thesis statement, or cannot help people with their part.

Excerpt 7: Question 4 Interviewee 1

**Interviewer:** Did you or any group member communicate in any other medium while writing collaboratively, such as WeChat, QQ, etc.? If so, why did you choose to communicate this way? What did you do? Please give examples.

**Alexander:** We used WeChat…and someone of my group set up a group, establish a group and we all discuss in the group and can share our idea or to…tell jokes. A good way to relax, right?

Excerpt 8: Question 4 Interviewee 2

(question same as above)

**Jennifer:** Mmm..I think we always use like the WeChat. Communicate it means, communicate information is about the writing information or before the writing?

**Interviewer:** Anything.

**Jennifer:** For example, we always use the WeChat to like organize our time, like which time the group member is are free for F2F discussion, and because this Etherpad sometimes is very hard to discuss our opinion. So, sometimes we use the WeChat to use Chinese to translate our ideas…because WeChat is convenient to us.