Language Learning & Technology
http://llt.msu.edu/issues/june2015/lin.pdf

June 2015, Volume 19, Number 2
pp. 85–117

A META-SYNTHESIS OF EMPIRICAL RESEARCH ON THE EFFECTIVENESS OF COMPUTER-MEDIATED COMMUNICATION (CMC) IN SLA

Huifen Lin, National Tsing Hua University

This meta-analysis reports the results of a systematic synthesis of primary studies on the effectiveness of computer-mediated communication (CMC) in second language acquisition (SLA) for the period 2000-2012. By extracting information on 21 features from each primary study, this meta-analysis intends to summarize the CMC research literature for the past decade by calculating an average effect size and performing a series of moderator analyses to factor out elements that might mediate the effect of such media in SLA. In total, 59 studies were identified as eligible after excluding three outlier studies, covering both published and unpublished studies. All studies were coded for learner characteristics (5 features), methodological characteristics (14 features) and publication characteristics (2 features), six of which were further analyzed as moderator variables. The results show that (a) there was a positive and medium overall effect for CMC used for instructional/learning purposes in SLA, (b) among the four language skills which CMC was intended to facilitate, writing skills produced the largest effect size, as did pragmatic competence, among the three language components, i.e. pragmatics, vocabulary and pronunciation explored in this meta-analysis; however this result should be interpreted as tentative since only one study measured pragmatic competence in the current meta-analysis, and (c) smaller group studies produced a larger effect size than those using larger groups or no grouping.

Keywords: Meta-Analysis; Computer-Mediated Communication (CMC); Second Language Acquisition; ESL; EFL

APA Citation: Lin, H. (2015). A meta-synthesis of empirical research on the effectiveness of computer-mediated communication (CMC) in SLA. *Language Learning & Technology, 19*(2) 85–117. Retrieved from http://llt.msu.edu/issues/june2015/lin.pdf

Received: August 8, 2013; **Accepted**: March 7, 2014; **Published**: June 1, 2015

INTRODUCTION

As the amount of research on computer-mediated communication (CMC) in the second language acquisition (SLA) field increases, researchers have tended to define it as a more comprehensive approach without losing its uniqueness. Thorne (2008) defined CMC as "multimodal, often (but not exclusively) Internet-mediated communication" (p. 325), while an earlier definition given by Herring (1996) described it as "communication that takes place between human beings via the instrumentality of computers" (p.1). Ostensibly the difference between these two definitions lies in the fact that in the former, communication is taking place via the Internet, which offers features that were not previously affordable or were only limited via computers.

As the functions of advanced communication technologies become more mature and easily accessible, the past two decades have witnessed an ever-increasing interest in using such communication technologies in the SLA field. According to Chun's paper (2007), which summarizes the hot topics in the top two U.S. computer-assisted language learning (CALL) journals, CALICO and Language Learning &Technology (LLT), (CMC) and Web-based instruction are the two most popular topics. This interest in CMC topics is not an isolated observation; CALICO, for example, dedicated an entire issue [2005, volume 22(3)] to the

topic of CMC and foreign language learning. The still growing and accumulating research interest in CMC in SLA necessitates a research synthesis that summarizes and more importantly systematically synthesizes the effects of such communication technologies for a very wide range of SLA purposes.

Meta-Analysis Related To CMC in SLA

In technology disciplines, the very early uses of meta-analyses were for computer-assisted instruction (CAI) in general during the 1980s, but meta-analysis research on CALL was very limited (Felix, 2005). Table 1 presents a summary of the previous meta-analyses on the overall effectiveness of technology in language learning. As indicated, there were approximately eight meta-analyses carried out during the period of 2003-2013, covering topics from electronic glosses specifically to more generally technology-supported SLA. The majority of these meta-analyses calculated effect sizes to represent the magnitude of the effectiveness of the technology-related treatment effects on SLA. The results, overall, favor the use of technology/computers as a means of facilitating SLA.

Thus far, only two meta-analyses have focused on CMC studies, i.e. Lin, Huang, and Liou (2013) and Sauro (2011). These two research syntheses present a comprehensive picture of what has been investigated in terms of synchronous CMC (SCMC) effectiveness over the period 1990-2010. However, the limitations of and approaches adopted in these two syntheses preclude an unequivocal answer to the question of the effectiveness of computer-mediated communication in SLA. The author of the current study identified some issues that were unexplored in these two syntheses. The first is related to the approaches adopted to synthesize the studies. Sauro's study employed a narrative review approach. According to Jackson (1980), narrative approach involves "…narrative discussions… [which] typify our attempts to make sense of the rapidly expanding research literature" (p.440). However his study does not deal with issues such as how treatment effects vary as a function of a single or a combination of other factors. Furthermore, the issue of "effectiveness" of SCMC in second language (L2) learning did not seem to be appropriately addressed in Sauro's synthesis. Most of the review related to the effectiveness of SCMC on the development of communicative competences described treatments/strategies adopted in primary studies but not the actual "effect" resulting from such differences in treatment design.

Whereas Lin et al. (2013) have examined the effectiveness of text-based synchronous CMC in SLA, their study suffers from some important limitations. First, the synthesis was only based on journal articles and dissertations, leaving the question open as to whether including other types of publications (e.g. conference papers, government reports, etc.) would produce different results. Additionally, out of the 10 studies, 8 were journal articles, which raises a concern of publication bias in that published studies are more likely to report significant results. Second, the focus of the synthesis is on the synchronous mode of CMC; studies exploring asynchronous communication were omitted and therefore the synthesis was not able to provide a comprehensive picture of CMC intervention in SLA. Third, their meta-analysis tried to identify potential moderators that might differentiate the effect of SCMC; however, due to the small sample size of each of the moderating variables, i.e. group size, treatment intensity, L2 proficiency, etc., the results were rendered uninterpretable. The current meta-analysis aimed to expand the previous meta-analyses and overcome their limitations. Specifically, it (a) included both published and unpublished studies related to CMC to present a comprehensive picture of the effectiveness of such interventions in the SLA field; (b) identified moderating variables that were not included in the previous meta-analyses, such as outcome measures; (c) analyzed the effects of the CMC interventions utilizing both fixed and random models of meta-analysis; (d) examined the effectiveness of CMC both at the language skill (i.e., reading, writing, speaking, and listening) and the language component (i.e., pronunciation, grammar and vocabulary) levels; and (e) made decisions and judgment calls at each step of the meta-analysis as transparent as possible following the transparency index established by Aytug et al. (2012).

Table 1. *Summary of Previous Meta-Analysis of CALL*

Study	Abraham	Grgurovic´ et al.	Lin et al.	Sauro	Taylor	Taylor	Yun	Zhao
Year	2008	2013	2013	2011	2006	2009	2011	2003
Publication Type	Journal article Dissertation Thesis Book chapter	Journal article, Dissertation Conference proceeding Report	Journal articles Dissertations	Journal articles	NA	Journal article Dissertation Thesis Report (ERIC)	Journal articles	Journal articles
Type of Comparisons	Computer-mediated glosses vs. paper-based (print) glosses	Computer technology supported vs. not-supported	Text-based SCMC vs. other means of communication	SCMC	CALL vs. traditional L1 Glosses	CALL-Based vs. paper-based Glosses	Multiple glosses vs. single gloss	Overall effect of technology applications in language learning
Targeted outcomes	Reading Vocabulary	NA	NA	Communicative competence	Reading	NA	Reading Vocabulary	NA
No. of studies	11	37	10	97	18	32	10	9
Sample size	542	NA	562	NA	875	1,845	1,518	419
Number of ES	NA	52	19	NA	18	NA	35	9
Strategy	Quantitative	Quantitative	Quantitative	Narrative	Quantitative	Quantitative	Quantitative	Mixed
Period	~2007	1970- 2006	1990-2012	1990-2010	~2002	~2006	1990-2009	1997-2001
Effect size	0.73 and 1.40	0.23 and 0.35	0.33	NA	1.09	0.51	0.46	1.12
Statistics	Cohen's d	Hedges' g	Cohen's d		Hedge's g	Hedge's g	Hedges' g	Cohen's d
No. of Moderators	3[a]	9[b]	4[c]	NA	NA	NA	8[d]	NA

Notes: [a]level of instruction, text type, outcome measure ; [b]type of technology, degree of integration, treatment length, proficiency level, L1, setting, L2, sample size; [c]L2 proficiency, setting, L2, treatment length; [d]sample size, L2, learner proficiency ,publication year, country, research design, outcome measure, vocabulary test type.

Research Questions

The following three research questions guided this meta-analysis:

1. What is the overall effect of CMC on L2 learning compared to face-to-face communication or no opportunities for communication at all?
2. Does the effectiveness of CMC differ depending on the language skills/components that it is intended to develop or reinforce?
3. To what extent do the following learner and methodology features affect the effectiveness of CMC: (a) learners' L1 and L2 (b) research context (c) grouping size (d) treatment length and (e) outcome measures, and is there a relationship between publication type and such effectiveness?

Variables

Several variables examined in this meta-analysis as moderators were mainly concerned with the methodological design of the primary studies. Factors such as the target language skill, group size, treatment length, learning context and learners' first language (L1) and L2 were suspected to make a difference in the effectiveness of CMC. Certain language skills might be more effectively developed via functions or affordances available in certain CMC tools or modalities; therefore it is necessary to explore the relationship between these two variables. Communications involving multiple participants determine the nature of interaction and the degree of overlapping utterances, which might, in the long run, be either productive or counterproductive in terms of achieving the intended communication aims. Additionally, the level of exposure and the context in which this exposure is provided might also differentiate the effects, given our common conception that most language skills need a significant amount of time to master, and a learning context in which distractions are avoided and guidance and immediate assistance can be asked for may differentiate the effects between long vs. short and in-class vs. out-of-class treatment conditions. Furthermore, certain learner L1s might be beneficial or detrimental to L2 learning due to inter-language transfer or interference.

Language Skills/Components

Four targeted language outcomes were pursued to represent the effect of CMC (speaking, reading, writing and listening) along with three language components (vocabulary, grammar and pronunciation). Each primary study was classified into one of the seven outcome categories after examining its description of the dependent variable. A study may be classified into more than one category if it measures more than one skill (e.g., Chenoweth, Ushida, & Murday, 2006; Liang, 2006). In this study, we firstly examine the overall effect of CMC on different SLA outcomes and then discuss possible moderators that might mediate such an effect.

Publication Bias

Regardless of the equal contribution of empirical studies that found non-significant results to our understanding of the line of research, it seems common for journal editors to have a preference for publishing studies that found significant differences favoring the treatment instruction, resulting in a so-called publication bias. This meta-analysis therefore explores if there is publication bias in the research on CMC, and its possible impact on the reported efficacy of CMC.

Treatment length

It is commonly held that the longer L2 learners are exposed to the target language or practice it, the more likely they are to acquire the language. Research on CMC has been conducted in studies ranging from a one-shot case study (Lee, 2009) lasting for 30 minutes to longitudinal studies lasting for more than two years (Chenoweth, Ushida, & Murday, 2006). Given that L2 skills take time to develop, it is worth examining if the length of treatment would have an impact on different levels of L2 learning.

Group size

CMC provides an electronic venue in which negotiation for meaning or negotiation work triggers students' language development, and consequently contributes to successful SLA (Smith, 2004). Group size determines the dynamics of the interaction and the possible nature of the negotiation and turn-taking patterns.

Research context

Research context refers to the context in which CMC activities are conducted. Students might be more motivated to participate in activities that are integrated into class than when asked to do so after class without support or monitoring.

Outcome measurement

The measurement adopted in evaluating the effect of instructional interventions has served as a moderator variable in previous meta-analyses (i.e., Li, 2010; Norris & Ortega, 2000). Given the wide variety of standardized and unstandardized tests employed to measure different language learning outcomes, it would be valuable to discover whether certain measurements affect or mediate the effectiveness of CMC interventions.

Learners' L1

SLA research has discussed the interference of students' L1 in their learning of the L2, and the negative or positive transfer of L1 on L2 acquisition (Bennui, 2008; Fagan & Cheong, 1987). Although it is beyond the scope of this meta-analysis to examine how students' L1 interferes with the L2 in a CMC environment, it is interesting to explore whether learners' L1 has a role to play in the effectiveness of CMC in SLA.

Learners' L2

Some languages are easier to learn than others, and learners might employ different learning strategies when learning different languages. It is worth investigating whether CMC technologies are more suitable for specific target languages than others.

METHODS

Identifying Primary Studies

The following searching strategies were performed to first identify and then retrieve the potential primary studies to be included in this meta-analysis with reference to previous meta-analysis searching steps (Li, 2010; Sauro, 2011). Firstly, electronic databases in the fields of (applied) linguistics, language learning/teaching and education overall were searched. The key words that were used for searching included computer-mediated communication, online interaction, computer-assisted-language learning, and computer-assisted instruction, in combination with tools commonly used for such purposes such as e-mail, chat-rooms, blogs, and classroom management systems (CMS). Searches were limited to those related to language learning/teaching in an SLA context. The electronic databases used included, but were not limited to, Linguistics and Language Behavior Abstracts (LLBA), ERIC, Springer online Journal Archives, JSTOR, and Linguistics Abstracts Online. Secondly, electronic and print journals in the field of CALL, CAI, SLA and education overall were searched for relevant articles dating back to 2000. The major journals searched included, but were not limited to, Language Learning & Technology, CALICO, System, Computer-Assisted Language Learning, The Modern Language Journal, the Journal of Computer-Assisted Instruction, Computers & Education, Foreign Language Annals, the British Journal of Educational Technology, Second Language Writing, Language Learning, and Second Language Research. Third, previous CMC (Lin, et al., 2013; Sauro, 2011) and CALL meta-analyses (Felix, 2005,

2008; Nguyen, 2008) were carefully read and references examined to identify CMC-related primary studies. Fourth, the reference sections of the retrieved CMC primary studies were scanned for potential studies. Fifth, Google Scholar was used with the same key words/combinations to identify lab (technical) reports, conference papers, and proceedings that were not officially published or may not have been identified through steps 1 to 4. Finally, ProQuest Digital Dissertation (PQDD) and Networked Digital Library of Theses and Dissertations (NDLTD) and the national library of the research team's country were searched using the same keywords/combinations to retrieve research reports, unpublished theses and dissertations. It has to be noted that, as a result, most research reports, unpublished theses and dissertations retrieved are from Taiwan, the country of the research team.

Inclusion Criteria

With reference to the inclusion/exclusion criteria followed in previous meta-analyses, and considering the purpose of the current meta-analysis, the following criteria were developed and followed when determining the eligibility of the retrieved studies. A study must have met the following criteria in order to be included in this meta-analysis:

1. The study was published between 2000 and 2012 (including online first studies). Technical reports and conference papers were also included in this meta-analysis to address the concern raised in previous meta-analyses that studies reporting significant findings were more likely to be published in journals (Norris & Ortega, 2006) and thus presented a potential bias to the results.
2. The study investigated some form of CMC (e.g., e-mail, chat, video/audio conferencing, discussion forums, CMS, Moodle, etc.) either exclusively or in conjunction with other instructional strategies/intervention as long as the effect of CMC can be teased out by making comparisons between treatment groups for which the only difference between them is the CMC intervention.
3. The study addressed the effect of CMC by examining or evaluating the change related to SLA during the process or by administering a posttest, both conditions requiring quantitative data.
4. The study employed an experimental or quasi-experimental design.
5. Studies recruited participants who were L2 or foreign language learners.
6. The studies included should report adequate quantitative information for effect sizes to be calculated.
7. For studies reporting across several sources, only one report was included in the meta-analysis. For example, if the same study was reported both in a dissertation and a journal, only one source of report was used.

The Coding Scheme and Inter-Coder Reliability

A team that included one meta-analyst and five research assistants who had all earned a Master's degree in TESOL, and had received substantial training in research methodology and quantitative data analysis conducted the meta-analysis. Each coder was assigned roles both as a primary and secondary coder in that he/she had to independently code the assigned ten studies and then acted as a second coder of another. As such, each study was coded at least twice. The inter-coder reliability was calculated by comparing the agreement between codes given by the two coders for each variable, and a ratio was calculated by dividing the number of agreed-upon codes by the total number of codes generated for all variables. Each coder was asked to note down the page number or write the author's description for high inference variables such as type of outcome measurements if the code demands subjective judgments. Disagreement between codes as well as high-inference codes went through another coding by another two coders and

final codes were assigned based on resolved discussions or the best estimation. The final inter-coder reliability was 100% for publication and learner variables, and 95% for methodological variables. The coding scheme consisted of three major aspects of the retrieved studies, i.e., publication (one feature), learners (three features) and methodology (nine features) as presented in Table 2. A brief description is provided for features and descriptors that are not immediately clear in the following. For a complete coding of the features please refer to Appendix A.

Table 2. *The Brief Coding Scheme*

Features	Descriptors
PUBLICATION TYPE	Journal article; Book or book chapter; Dissertation/; Thesis; Technical report; Conference paper (proceedings); Other
LEARNERS	
L1	English/German/Japanese/Chinese/Spanish/Russian/others
L2	English/German/Japanese/Chinese/Spanish/Russian/others
Age	Learners' age
METHODOLOGY	
Research context	In class; After class; Both
Group size	Pair (2 people); small (3–6 people); large (more than 6 people); Mixed (more than one grouping); Class
Treatment Length	Short: less than or equal to 10 weeks Medium: between 11 and 24 weeks Longitudinal: more than 24 weeks
Tutorial/Training	Mechanical; Content; Both
Outcome measures	Standardized test; institutional assessment; researcher-developed assessment; impressionable judgments; Other (specify)

Outcome Measures

Four categories were developed to code each study in terms of how outcomes were measured. Referring to Li (2010), Keck, Iberri-Shea, Tracy-Ventura, and Wa-Mbaleka (2006) and Thomas (1994), an outcome measure was coded as a standardized test if it was a well-established national or international standardized test such as the Test of English as a Foreign Language (TOEFL), GEPT (Taiwan's official General English Proficiency Test), or American Council on The Teaching of Foreign Languages(ACTFL) proficiency test; a measure was coded as an institutional assessment if it was developed and designed by an institution to screen or place students into different proficiency levels, such as the Tunghai English Placement Exam (Chiang, 2007), and BYU WebCAPE (Sanders 2005); a measure was coded as an in-house assessment if it was developed by the researcher for the purpose of the study, and was coded as mixed if more than one measurement was adopted, e.g. Kost (2004) and Hung (2007).

Group size

Four categories were identified for grouping size. If a group consisted of two people, it was coded as "pair"; if three to six people, it was coded as "small"; if more than six people, it was coded as "large"; "mixed" was coded for studies in which the number of students in each group was not equal; "no grouping" was coded for studies in which the entire class participated as a group.

Treatment Length

All primary studies were examined for how long a treatment lasted, and three categories were created to

identify studies with different treatment lengths. If the total period was no more than 10 weeks, it was coded as "short"; if it was 11-24 weeks, it was coded as "moderate"; if more than 24 weeks, it was coded as "long." It should be noted that the treatment length was defined as the time period in which the entire research was conducted and did not necessarily reflect the exact amount of time that participants received the intervention. Although the latter is more desirable for precisely pinning down the effect of CMC, very few studies provided this information. Additionally, for studies that adopted an asynchronous mode, exact intervention time is difficult or even impossible to measure.

Language Outcomes

Four targeted language skills outcomes, i.e. speaking, reading, writing and listening, and three components of the language skills, vocabulary, grammar (grammatical competence) and pronunciation, were pursued to represent the effects of CMC. Each primary study was classified into one of the seven categories after examining its description of the dependent variable. It should be noted that a study may be classified into more than one category of language outcome if it measures more than one skill (i.e., Chenoweth, Ushida, & Murday, 2006; Liang, 2006).

Effect Size Calculation for Individual Studies

In this meta-analysis, a standardized mean difference, Hedge's g, was computed for each study, and a summary/average effect size was computed to represent the overall effect of the CMC intervention. In practice, the standardized mean difference is calculated by dividing the mean difference in a study by its pooled standard deviation, which then becomes comparable across studies. Additionally, in order to minimize the variance among studies, a weighted mean is computed for each study to better estimate the summary mean. To do so, each study is weighted by the inverse of its variance (see Formula 1).

$$\hat{g} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)SD_1^2+(n_2-1)SD_2^2}{(N_{\text{total}}-2)}}} \times \left(1 - \frac{3}{4(n_1+n_2)-9}\right)$$

Formula 1

This would give greater weight to studies with large sample sizes because it is commonly agreed that larger samples provide better estimates of population parameters (Bangert-Drowns, Hurley, & Wilkinson, 2004; Hoffler & Leutner, 2007). A summary effect is then calculated to estimate the mean of a distribution of the weighted effect sizes calculated from each of the included studies. These procedures were followed by studies that report sufficient information based on which effect size calculation is possible. For studies that report incomplete information, the procedures suggested by Borenstein, Hedges, Higgins and Rothstein (2009) were followed.

Data Analysis

The independent set of weighted effect sizes was calculated following the aforementioned procedures. Furthermore, to test if the calculated effect sizes were significantly different from zero, confidence intervals were calculated. If the confidence intervals include zero, it means that the true mean effect size may include zero and thus the null hypothesis that the treatment effect is zero cannot be rejected at the $p < .05$ statistical significance level. Additionally, the width of the confidence level is also examined to represent the precision of the effects. To test the null hypothesis that all studies share a common effect size, a test of heterogeneity via Q-test was performed. A significant p-value < .05 would indicate that the true effects vary between studies and serve as evidence that there exist variables that might moderate the effectiveness of CMC intervention.

RESULTS

Phase I: The Research Synthesis

In total, 59 studies conducted between 2000 and 2012 were found as eligible studies for the meta-analysis. A list of the studies with substantive features related to publication and methodology is provided in Appendix B. Learner characteristics can be found in Appendix C. A first glimpse of both lists revealed a wide range of communication tools employed in the included studies. As indicated in Figure 1, approximately 22 studies carried out communication tasks utilizing different types of chat-enabling systems such as chat-rooms, Skype, IRC, and ICQ. The second most popular tools used were blogs followed by e-mails; both are prevalent tools for asynchronous communication. Classroom management systems (CMS) and self-developed platforms were also used in a couple of studies to support communication.

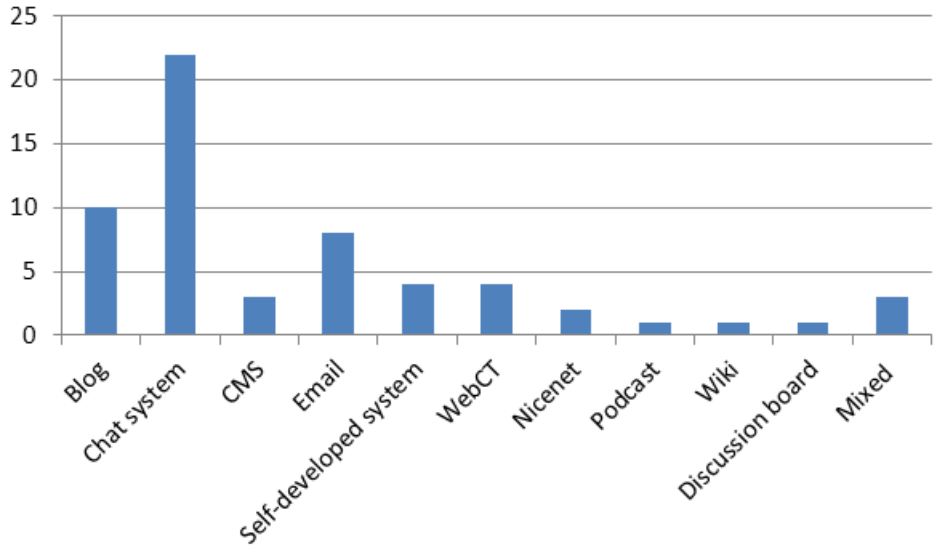

Figure 1. Frequency count of communication tools used in the primary studies.

To provide a profile of the study features, we summarized and tallied the number of frequency for each feature. Table 3 shows the learner characteristics of the included studies. Regarding participants' L1, it was found that three out of the six L1s across the included studies dealt with Asian language speakers, i.e. Chinese, Korean and Japanese. This pattern of L1 speakers indicated a major research interest in CMC with the Asian learner population. In fact, in Jeon and Kaya's (2006) meta-analysis on the effects of L2 instruction on inter-language pragmatic development, they also found a similar pattern of research interest in participants who are Asian language speakers. In this meta-analysis, Chinese L1 speakers (N=32), in particular, were the most dominant L1 group among the three Asian language groups. Furthermore, close to a quarter (N=14) of the studies recruited mixed L1 participants. Further examination of this group of studies revealed that the L2 of the mixed L1 speakers were English, German and Spanish and the studies were all conducted in the U.S. Only one study dealt with either Arabic or Turkish L1 speakers, suggesting a population of L1 speakers that might merit more CMC research. Also shown in the table is the overwhelming number of studies whose target language is English, signaling that English, as the most popular and major tool of international communication, is still the L2 that receives most attention in second/foreign language classrooms. Another interesting picture which emerged from this synthesis is that rather than Asian languages, two European languages, i.e. Spanish and German, emerged as the L2, second to English, that most CMC researchers were interested in. These patterns regarding study participants' L1 and L2 revealed that CMC studies have concentrated on Asian learners whose target languages are either an American or a European language and that there is very little research interest in

learning any Asian languages as a second or foreign language. Regarding participants' age, of the 42 studies in which data on participants' age is available, more than half recruited participants between 19 and 24 years old from colleges. Five studies involved young learners between 9 and 15 years old, and nine recruited secondary or high school learners. These patterns reveal that adult learners at college level are still the prevailing target population of most L2 studies, and little research attention has been paid to very young learners.

Table 3. *Learner Categories of the Included Studies*

L1 Background		Age		L2 Background	
	N[a]		N		N
Arabic	1	9–15	5	English	46
Chinese	32	15.5–18.5	9	German	2
English	6	19–4	27	Spanish	10
Japanese	1	>24	1	Mixed	1
Korean	1	NA	17		
Turkish	2				
Mixed[b]	14				
NA	2				

Notes: [a]The number of studies with such a characteristic; [b] The target language may be a second or foreign language for individual students from a multilingual class

The methodology characteristics of the included studies can be found in Table 4. As shown, the two target skills that received the most research attention in the CMC studies were writing (*N*=29) and speaking (*N*=23). Comparatively, very little research concentrated on the empirical effects of CMC on the receptive skills of listening and reading. This pattern of research interest regarding target skills might have to do with the communication tasks themselves that could be best mediated by certain communication devices. Widely used tools such as chat rooms, bulletin boards or e-mail are ideal venues for practicing speaking and writing rather than listening and reading. Treatment length was further classified into three categories at the cut-off points of less than 10 weeks, between 11 and 24 weeks and more than 24 weeks. It was found that the majority of studies featured medium length treatments (*N*=33) lasting between three and six months, and only two featured long-term treatments that lasted more than six weeks. This might indicate that the right amount of exposure of learners to treatments is no longer than half a year to produce the optimal effect of communication via technology on learners' target language development. A general consensus among SLA researchers is that treatments have to be offered for long enough for certain effects to be observed (Creswell, 2009); however, longer treatments involve the likelihood of participants being exposed to extraneous variables. Thus far, there has been no agreement on the most appropriate length of treatment. Our meta-analysis revealed that at least for SLA in the CMC context, most researchers agreed on treatment lengths of between three and six months.

Trends of grouping across the studies were also synthesized. Twenty studies grouped students into pairs to maximize the amount of interaction for each individual learner. Thirteen studies put students into small groups of three and six learners. It was also observed that five studies did not employ any specific kind of grouping and only one employed large grouping. The overall pattern that more than half of the included studies adopted small groups clearly indicates the importance of maximizing interaction opportunities and reducing turn-taking among interlocutors in CMC studies. Most tasks were carried out in class (*N*=29) rather than after class (*N*=20). As for the different types of outcome measures, the results found a variety employed in the included studies. In-house measures developed by the primary researcher for the purpose

of the study was the most prevalent type of outcome measure (*N*=45); standardized measurements were also used in eight studies. Institutional measurements, mostly used for screening purposes for beginners in individual institutions were scarcely used by CMC researchers, probably because the purposes of assessment differed significantly.

Table 4. *Methodology Categories of the Included Studies*

Skill		Outcome M		Treatment Length [a]		Grouping [b]	
	N		N		K		N
Overall	3	Mixed	2	Long	2	Class	5
Listening	3	In house	45	Moderate	33	Large	1
Speaking	23	Institutional	3	Short	23	Small	13
Reading	8	Standardized	8	NA	1	Pair	20
Writing	29	NA	1			Mixed	4
Pronunciation	2					NA	16
Vocabulary	3						
Pragmatics	1						
Tutorial		Research Context		Reliability			
	N		N		N		
Mechanical	19	In-class	29	Test R	12		
Content	28	After-class	20	Inter-rater R	27		
Both	10	Both	6	NA	20		
NA	2	NA	4				

Notes: [a]Treatment length: Long >24 weeks; medium: 10-24 weeks; short: <10 weeks; [b]Grouping: Pair (2 people);small (3-6 people); large (more than 6 people);Mixed (more than one grouping); Class (no grouping).

Phase II. The Quantitative Meta-Analysis

The earlier section presents the overall profile of the 59 eligible studies included in this meta-analysis with regard to their substantive learner and methodological features. In this section we provide the results of the second phase analysis, namely, the results of effect size calculation. We present the results pertaining to each research question we prescribed for the current meta-analysis.

Research question 1: What is the overall effect of CMC on L2 learning compared to face-to-face communication or no opportunities for communication at all?

Our first research question asked about the overall effectiveness of computer-mediated communication on second language learning compared to face to face or no form of communication at all. To address the concern raised in previous meta-analyses that statistically biased effect sizes may exist due to the inclusion of studies that did not include a control/comparison group, such as a one-group only pretest/posttest design (Norris & Ortega, 2006), this meta-analysis first went through three analyses to examine whether there are significant gaps in the effect sizes calculated when including or excluding these studies. The data shown in Table 5 include the mean effect size (Hedge's *g*), standard error, and the upper and lower limit under the 95% confidence level for the three separate analyses. As indicated, the mean effect sizes obtained when pooling all the studies is .461 under the random-effects model and 0.408 under the fixed-effects model (*K*=59), including the one-group studies being 0.323(*K*=15) under the random-effects model and 0.165 under the fixed-effects model, and including only two-group studies being 0.499(*K*=44) under the random-effects model and 0.587 under the fixed effects model. The

obtained effect sizes under the three conditions of pooling were very close. Furthermore, the pair-wise comparison between one group studies and two group studies in effect size also shows no significant difference, $Q(1)=1.214$, $p=.271$. Since no significant differences in effect size were found between pools of studies including or excluding studies that employed a one-group only design, we interpreted the results based on the pool of 59 studies. As indicated, the overall mean effect size (Hedge's *g*) of the 59 studies was 0.41 (fixed-effects model) and 0.44 (random-effects model). According to Cohen (1988), the effect at this magnitude was moderate with approximately 79% of variance not accounted for by chance. The confidence interval is between 0.29 to 0.63 in the random-effects model and 0.34 and 0.47 under the fixed-effects model, both not including zero, indicating that the effect size derived from the included studies is trustworthy and does not happen by chance. If we interpreted the calculated effect size in terms of standard deviation, the CMC experimental group students scored on average of 0.40 (random-effects) or 0.44 (fixed-effects) of a standard deviation above the scores of the control group with or without other means of communication.

Table 5. Over*all Effectiveness of CMC (Hedge's g) Comparing Studies of Different Research Designs*

Number of studies	K	Hedge's g^a	SE	95% Confidence Interval	
				Lower limit	Upper limit
All studies	59	0.441/0.408	0.074/0.032	0.291/0.344	0.632/0.471
One-group studies	15	0.323/0.165	0.114/0.050	0.099/0.068	0.547/0.262
Two or more groups	44	0.499/0.587	0.112/0.043	0.280/0.504	0.719/0.671

Note: [a] The value shown on the left is based on the random-effects model and on the right is the fixed-effects model value.

Research question 2. Does the effectiveness of CMC differ depending on the language skills/components that it is intended to develop or reinforce?

Our second research question explored if the effectiveness of CMC would differ depending on the specific language skill it is employed to facilitate. Table 6 displays the overall effect of CMC intervention (as compared to face to face or no such intervention). The relative effect of CMC in terms of different language skills (reading, writing, speaking, listening, overall) is also shown with the number of studies contributing to that effect size, the calculated mean effect size (Hedge's *g*), the standard error and both the lower and upper limit at the 95% confidence interval. CMC, in general, has a small to medium effect on language learning (Hedge's $g=0.461$). The 95% confidence intervals do not include zero meaning that such an effect does not happen by chance. For language skills, a wide range of effect sizes was found across the different language skills and their components. The magnitude of the effects ranges between 0.96 (pragmatics) and 0.32 (listening). Surprisingly, the use of CMC resulted in a negative effect on vocabulary learning (-0.616). Following Cohen's (1988) guidelines, the calculated effect sizes reflect a combination of small, medium and large magnitudes. That is, CMC has a small effect on listening, speaking and reading, a moderate effect on writing and pronunciation, and a large effect on pragmatics. To sum up, the results of Table 6 indicate that CMC is most effective in developing writing ($g=0.638$), followed by reading ($g=0.519$) and listening ($g=0.505$). It is least effective for the development of speaking ($g=0.307$) under the random-effects model. A similar pattern was also revealed under the fixed effects model. The overall effectiveness for the four different skills was found to be significant because the Q-text shows that there were significant differences in their mean effect sizes, $Q(4)=22.233$, $p < .001$. As for the language components, the random effects model found the largest effect size for pragmatic competence, followed by pronunciation, and a negative effect was found for vocabulary. An identical pattern was found under the fixed effects model. The above results have to be interpreted with caution, though, due to the small number of studies included in the categories of listening and reading, and the three components. The problem is particularly severe for pragmatics, which included only one study.

With some categories of studies underrepresented and others overrepresented, we cannot draw a complete and accurate picture with confidence from the current meta-analysis regarding the causal effect of CMC on different language skills development.

Table 6. *Effects of Computer-Mediated Communication by Independent Variable*

Independent variables	K[b]	Hedge's *g*[c]	SE	95% Confidence Interval	
				Lower limit	Upper limit
Overall Effect[a]	59	0.441/0.408	0.074/0.032	0.291/344	0.632/471
Listening	3	0.320/0.170	0.268/0.075	-0.206/0.022	0.845/0.317
Speaking	23	0.320/0.286	0.171/0.056	-0.014/0.176	0.654/0.395
Reading	8	0.461/0.311	0.159/0.061	0.149/0.192	0.773/0.430
Writing	29	0.620/0.614	0.124/0.049	0.376/0.518	0.863/0.710
Pronunciation	2	0.635/0.635	0.188/0.188	0.266/0.266	1.004/1.004
Vocabulary	3	-0.616/-0.616	0.254/0.254	-1.114/-1.114	-0.118/-0.118
Pragmatics	1	0.959/0.959	0.231/0.231	0.506/0.506	1.412/1.421

Notes: [a]The overall effect (Hedge's g) is computed based on the principle of "one study, one effect size" (Li, 2010);.[b]the number of studies that contributed to the effect size. Notice that one study may include more than one language skill, thus the total K might not add up to 59; [c]statistics on the left are computed based on Random-effects model(Q(7)=28.530, p < .001) while the statistics on the right are based on the fixed-effects model(Q(7)=63.690, p < .001).

Research question 3. To what extent do the following learner and methodology features affect the effectiveness of CMC: (a) learners' L1 and L2 (b) research context (c) grouping size (d) treatment length, and (e) outcome measures, and is there a relationship between publication type and such effectiveness?

Our last research question aimed to identify factors from the publication, learner and methodological levels that might moderate the effectiveness of CMC. To answer the research question, publication type and selected features of the learner characteristics and methodologies of the included studies were analyzed to examine whether they would moderate the overall effectiveness of computer-mediated communications. A series of Q-tests was performed, and the obtained Q statistics and associated p values (set at .05 level) were used to examine whether a certain variable was a significant moderator. The moderators examined included publication type, research context, group size, treatment length, learners' L1 and L2, and outcome measures. Table 7 presents the results of the moderator analysis followed by a more detailed description of the pair-wise comparison results for the significant moderators. For publication type, over half of the studies were journal articles that generated a small effect size (g=.322). Comparatively, the five conference papers generated the largest effect size (g=1.015), while the effect sizes calculated for the unpublished theses and dissertations were roughly moderate. In terms of research context, studies in which CMC tasks were conducted both in and after class produced the largest effect, which is almost twice the effect of those only conducted after class. The research context as revealed from the descriptors seems to favor a setting in which CMC activities were carried on into a non-class setting. Regarding group size, a clear picture has emerged that smaller groups did contribute to the beneficial effect of computer-mediated communication compared to larger groups. Twenty studies in this meta-analysis divided students into pairs and resulted in a close to large effect of language learning. The large group involving more than six interlocutors, as can be expected, generated the least effect. The analysis comparing different treatment lengths reveals that a medium length of treatment lasting between 10 and 24 weeks induced the best effect, followed by a treatment length of less than 10 weeks. The long-term treatment, surprisingly, resulted in the smallest effect. These seemingly unexpected results might again need to be interpreted with great caution because there were only two studies in this meta-analysis that

employed treatments that were longer than 24 weeks, which render the comparisons unreliable.

Our analysis of students' L1 as a moderator of CMC effectiveness revealed an astounding picture. Studies using Arabic, Japanese and Turkish L1 speakers generated the largest effect size, while the Korean and English L1 speakers had the smallest effect sizes. While there has been discussion on such issues as learners' L1 perhaps aiding or interfering with their L2 learning depending on their inter-language development, it is not easy to draw a clear cause-and-effect relationship between learners' L1 and the CMC effectiveness in the current meta-analysis. Given that variables other than learners' L1 might also play a significant role in such a relationship, the results are presented as suggestive rather than informative. Regarding outcome measures employed in the primary studies, the results indicated that studies using standardized tests yielded the largest effect size, while both in-house and mixed tests generated small but approximately equal effect size. Studies that employed institutional tests, on the contrary, generated a negative effect size.

Table 7. *Moderator Analysis Results*

Moderators	K^a	Hedge's g^a	SE	95% Confidence Interval		Q^b
				Lower	Upper	
Publication type						
Journal article	31	0.322/0.231	0.105/0.042	0.125/0.149	0.539/0.312	8.965*
Conference paper	5	1.015/1.041	0.203/0.118	0.618/0.809	1.412/1.273	60.611***
Thesis	11	0.533/0.702	0.271/0.083	0.002/0.539	1.063/0.866	
Dissertation	12	0.477/0.499	0.165/0.079	0.155/0.345	0.800/0.654	
Research Context						
In-class	29	0.512/0.611	0.132/0.050	0.253/0.512	0.770/0.709	1.458
After-class	20	0.350/0.224	0.142/0.056	0.073/0.115	0.628/0.334	30.31***
Both	6	0.650/0.283	0.224/0.070	0.211/0.147	1.090/0.420	
Group size						
Class	5	0.314/0.282	0.177/0.138	-0.033/0.012	0.660/0.552	4.922
Large	1	0.015/0.015	0.243/0.243	-0.461/-0.461	0.491/0.491	15.560**
Small	13	0.478/0.478	0.074/0.074	0.334/0.334	0.623/0.623	
Pair	20	0.624/0.696	0.176/0.059	0.279/0.580	0.968/0.812	
Treatment length						
Long	2	0.054/0.231	0.437/0.122	-0.803/-0.008	0.910/0.469	1.081
Medium	33	0.504/0.368	0.105/0.040	0.298/0.289	0.711/0.447	8.006*
Short	23	0.423/0.547	0.179/0.063	0.072/0.424	0.773/0.671	
L1						
Arabic	1	1.117/1.117	0.189/0.189	0.747/0.747	1.486/1.486	69.740***
Chinese	32	0.501/0.360	0.113/0.040	0.279/0.281	0.723/0.438	81.139***
English	6	0.197/0.317	0.253/0.121	-0.299/0.179	0.694/0.554	
Japanese	1	1.737/1.737	0.361/0.361	1.029/1.029	2.445/2.344	
Korean	1	-1.382/-1.382	0.423/0.423	-2.211/-2.211	-0.553/-0.553	
Turkish	2	1.653/1.6536	0.219/0.219	1.224/1.224	2.082/2.082	
Mixed	14	0.253/0.322	0.168/0.076	-0.076/0.174	0.581/0.471	

L2						
English	46	0.538/0.448	0.102/0.035	0.339/0.379	0.738/0.517	4.692
German	2	0.168/0.168	0.179/0.179	-0.183/-0.183	0.520/0.520	7.737*
Spanish	10	0.188/0.204	0.195/0.095	-0.195/0.018	0.570/0.389	
Outcome measures						
Standardized	8	0.673/0.350	0.222/0.067	0.238/-0.225	0.978/0.219	3.370
In-house	45	0.457/0.457	0.111/0.041	0.240/0.377	0.674/0.537	7.075
Institutional	3	-0.053/0.158	0.328/0.113	-0.696/-0.064	0.590/0.381	
Mixed	2	0.470/0.447	0.301/0.196	-0.120/0.062	1.059/0.832	

Notes: [a] The random-effects values are on the left of each column and the fixed-effects values are on the right; [b]This moderator analysis only included studies from which the data for the certain variable were available; the first value shown is based on the random-effects model followed by the value based on the fixed-effects model

To further examine whether the differences between the categories of moderators are significant, a series of pair-wise comparisons was conducted. Generally speaking, the random-effects model tends to give a more conservative estimate of the effect size than the fixed-effects model. As indicated in Table 8, two of the seven moderators were found to be significant, and might impact the effectiveness of the CMC intervention under the random-effects model, while six moderators were found to be significant under the fixed-effects model. The pairwise comparisons indicated that the effect sizes calculated from conference papers were significantly more pronounced than those of unpublished dissertations, theses and journal articles. Unpublished dissertations and theses also generated significantly larger effect sizes than journal articles.

As for research context, pair-wise comparisons revealed that studies which employed CMC activities both in and after class produced larger effect size than studies that used them only in class. The difference in effect size between after-class studies and both studies were not significant. In terms of group size, pair-grouping studies overwhelmingly outperformed studies that used either small, large or no grouping at all. Regarding treatment length, significant differences were found between short and long treatments and medium and short treatments, with short treatments generating larger effect size than long treatments, and medium-length treatments producing larger effect size than short treatments.

Our last significant moderator is learners' L2. Of the three pairwise comparisons, the only significant difference is between English and Spanish L2 learners, with studies of English as an L2 producing bigger effect sizes than Spanish L2 studies. Effect sizes contributed by each primary study based on skills can be found in Appendix B.

Table 8. *Follow-up Pairwise Comparisons of Significant Moderators Under Either the Fixed-Effects or the Random-Effects Model*

Pairwise comparisons	*Q* statistics[a]	Results
Publication Type		
Journal article vs. Thesis	0.476/25.652***	Thesis > Journal article
Journal article vs. Conference paper	8.927**/41.737***	Conference paper>Journal article
Journal article vs. Dissertation	0.553/9.098**	Dissertation > Journal article
Dissertation vs. Thesis	0.030/3.141	
Dissertation vs. Conference paper	4.237*/14.531***	Conference paper>Dissertation
Thesis vs. Conference paper	2.032/5.473*	Conference paper>Thesis

Research Context		
After class vs. In class	0.693/26.424***	In class > After class
After class vs. Both	1.278/0.432	
In class vs. Both	0.284/14.567***	Both > In class
Group Size		
Class vs. Large	0.989/0.914	
Class vs. Pair	1.547/7.605**	Pair > Class
Class vs. Small	0.738/1.578	
Large vs. Pair	4.122*/7.407**	Pair > Large
Large vs. Small	3.330/3.330	
Pair vs. Small	0.581/5.261*	Pair > Small
Treatment Length		
Long vs. Medium	1.005/1.154	
Long vs. Short	0.611/5.350*	Short > Long
Short vs. Medium	0.155/5.749*	Medium>Short
L2		
English vs. German	3.216/2.341	
English vs. Spanish	2.535/5.861*	English > Spanish
German vs. Spanish	0.005/0.030	

*Notes: [a]Values on the left are based on the random-effects model and on the right are from the fixed-effects model; * indicated significant at the .05 level; ** at the .01 level; *** at the .001 level.*

CONCLUSION AND DISCUSSION

The purpose of this meta-analysis was to scrutinize the effect of CMC intervention on SLA in terms of the four language skills and their components. It also identified the moderator variables that might impact its overall effectiveness. The analysis found that overall, CMC has a positive and medium effect on language learning as measured from the immediate posttests. Furthermore, CMC seems to be more effective in facilitating the acquisition of writing skills but less effective in developing oral skills. In the following we discuss the findings and provide plausible explanations.

First, with regard to difference in effect size as a result of grouping, the descriptive statistics found that putting students in pairs generated the largest effect size, followed by small groups of three to six students; interactions involving large groups generated the smallest effect size. Pair grouping significantly outperformed either small group, large group or no grouping at all (i.e., class-based). Paired grouping studies are found to generate almost twice the effect of the class-based studies. Jenks (2009) examined the interactional features of multi-participant voice-based chat rooms. In his review of previous studies on the same topic, but conducted in text-based chat rooms, he indicated that overlapping utterances are one unique feature of communication involving multiple participants, and that these utterances tend to be "disrupted, disjointed or delayed" (Herring, 1999) and "…counterproductive to achieving comprehensible communication" (Jenks, 2009, p. 28). In voice-based CMC involving multiple participants, incomprehensibility often accompanies overlapping talk because previous utterances are not available as they are in text-based CMC. In order to take control of the ongoing discussions, students employ different strategies such as the use of abbreviations and emoticons, especially when they cannot see each other. The above strategies to enhance comprehensibility both in text or voice-based CMC due to multiple participants would not be necessary when only two interlocutors are involved in the interaction, and so

rapid interaction is more likely in one-to-one than in multi-participant CMC. Definitely, the social dynamics with regard to turn-taking, equality, consensus and interruption are also different between one-to-one and multiple-participant CMC (Warschauer, 1997), and these characteristics unique to the process of each mode of communication would also have an impact on subsequent learning outcomes.

Second, studies incorporating CMC activities as in-class activities generated a significantly larger effect than those using after-class activities. This finding is not surprising at all given that learners tend to be more concentrated and focused on in-class activities rather than when engaged in the same activities after class where distractions might be abundant. Direct manipulation of in-class activities is also easier as any questions that arise may be more likely to be immediately resolved. This meta-analysis also found that incorporating CMC activities both in and after class generated a larger effect size than conducting CMC activities only in-class. This finding would encourage an expansion of instructional activities from a classroom setting to after-class settings in which autonomous and independent learning might take place.

Third**,** with regard to treatment length, the study showed that short treatments (≤ 10 weeks) generated the largest effect size, followed by moderate treatments (11–24 weeks). The longer treatments (> 24 weeks) generated the smallest effect size. Pairwise comparisons showed a significant difference between short and moderate and short and long treatments, with short treatments outperforming the others. This finding is contradictory to our expectation given the fact that language skills take time to develop, and any subtle changes might not be immediately measureable. The meta-analyst suspects that this finding might be attributable to outcome measures since all long treatments were measured using either in-house or institutional assessments, while the short treatments were measured using a mixture of assessments including standardized tests and, as will be discussed later, studies employing standardized tests yielded the largest effect sizes. The above discussion aside, treatment length in this meta-analysis was defined as the total period of time during which the primary study was conducted, which might not actually reflect the intensity of the treatments the participants were exposed to. Furthermore, task design, language skill and interlocutor types may all have to be considered before we can pin down the effect of treatment length on SLA.

Fourth, in terms of the effect of outcome measure, it was found that the majority of the included studies employed in-house measurements (K=45), which are defined as assessments developed by the researcher specifically for the purpose of a study. Among the 59 studies, only 8 employed standardized tests. The descriptive statistics show that standardized tests under the random-effects model produced the largest effect size. Institutional measurements generated the smallest effect size consistently under both models. Close examination found that one third of the in-house measurements failed to report any kind of reliability, while all standardized tests reported either test or inter-rater reliability. Cross-tabulation also found that in-house measurements were attributable to the majority of negative effect sizes, i.e. out of 14, 12 were contributed by in-house measurement studies. In the SLA area, in-house and tailor-made assessments seem to be the norm of measurement to accommodate the diversities of student level, instructional treatments and testing logistics. However, failure to report test validity and reliability might render the results debatable. The choice of measurement is critical in empirical studies. It is not likely to draw a causal relationship between the treatments and type of measurement. However, in this meta-analysis, standardized tests were associated with small, medium and large effect sizes evenly and therefore seem to be more trustworthy compared to in-house and other types of measurement for the majority of which the report of any form of reliability is missing.

Fifth, with regard to publication type, this meta-analysis found that conference papers yielded the largest effect size among the four different publication types. Out of the six conference papers included, four generated an effect size larger than 1, which might explain why this type of publication "outperforms" the other types which generated more evenly-distributed and moderate effect sizes. It might make little sense to argue that the effect of CMC would be contingent on the type of publication, where it would appear especially that publication bias examined previously in the former part of the paper showed no or little

bias in the included studies. Cross tabulation incorporating the research context showed that all of the conference papers were based on studies conducted in foreign language contexts, while the journal articles were based on studies in both foreign and second language contexts. And, as previously discussed, studies conducted in foreign language contexts yielded larger effect sizes than those conducted in second language contexts.

Sixth, to explore if learners' L2 has an impact on the effectiveness of CMC, effect sizes were calculated for the four major target languages examined in the primary studies. It was found that CMC integrated into the instruction of L2 English was significantly more effective than for German and Spanish, although a significant difference is only found between English and Spanish, with the former generating almost twice the effect of the latter under the fixed-effects model, and three times under the random-effects model. Cross-tabulation of the results related to learners' L1 indicated that out of the 46 L2 English studies, 34 were conducted with Chinese L1 learners, while out of the 10 L2 Spanish studies, five were conducted with English L1, and three with mixed L1 participants. The high motivation associated with the learning of English as an international language in the era of globalization is one plausible explanation for this finding. Furthermore, as far as the meta-analyst knows, English has always been the number one foreign language that Chinese learners endeavor to have a good command of. The enthusiasm for learning English is also revealed in government policies of making English required instruction in the curriculum from primary school level in all Chinese-speaking countries.

Finally, to ascertain whether learners' L1 moderates the effectiveness of CMC, effect sizes were calculated for the L1s in the primary studies. The majority of the included studies were conducted with Chinese L1 learners (K=32), while more than a quarter of the primary studies (K=14) have mixed L1 participants. Pair-wise comparisons revealed a complicated picture. Overall, studies conducted with Japanese, Korean and Arabic L1 learners generated a significantly larger effect size than the other L1 studies. The fact that only 1 study each was conducted with learners of these three L1s, and that there are 14 studies with students from mixed L1 backgrounds renders a reasonable interpretation of the results impossible. However, this finding also highlights the need for future research to conduct studies with participants from less-studied L1 backgrounds.

The results of this meta-analysis contributed to our understanding of SLA theory from the socio-cultural and interactionist perspectives. Interaction opportunities in a socio-cultural context provide likelihood for negotiations in that input becomes comprehensible, and learners' attention can be drawn to form and ways of developing discourses (Gass & Varonis, 1994). This meta-analysis lends empirical supports to the Interaction Hypothesis (Long, 1996) that online interactions/communications mediated by computers/technology can generate similar or even superior opportunities for L2 learning than are found in face-to-face settings. Systematic variables of the wide CMC context, though, should be considered. Based on the results of this meta-analysis, contextual variables such as research context, grouping, amount of exposure to the interaction, i.e., treatment length, and learners' L2 may exert different effects of such interactions. The CMC interactions may vary significantly when L2 learners carried out the tasks in class or after class. The number of interlocutors who interact simultaneously online (e.g., pairs versus small group versus large group versus whole-class) and learners' L2 are also variables that might systematically moderate the effect on amount and quality of online interactions.

LIMITATIONS AND FUTURE RESEARCH SUGGESTIONS

This meta-analysis is not without its limitations, which may influence the generalizability of the results. First, it included a higher ratio of certain types of publications, i.e. conference papers, theses and dissertations from the meta-analyst's country. The convenience of more immediate and easier access to these studies may have caused Chinese L1 participants to be overrepresented in the sample. Second, the division of certain variables, such as treatment length and task type into different levels for moderator analysis is somewhat arbitrary and involves a certain degree of subjectivity. Different division criteria

imposed on these variables may cause the results to be totally different. Third, certain studies which failed to provide sufficient quantitative data for effect size calculation but which met the inclusion criteria had to be excluded from the final pool. Thus this meta-analysis only reveals part of the picture of the CMC findings in our field. This limitation also extends to studies that have little visibility or that the meta-analyst has little or no access to.

The results revealed the following issues to be taken up in future research. First, English is still the dominant L2 that draws intense research interest in SLA. This phenomenon is probably due to the overwhelming population that speaks the language and hence the need for learners to sharpen their English skills. There is a need for future studies to look at less-studied languages, such as Chinese, Japanese and Arabic. The disproportionate representation of individual L2s in the dataset makes it impossible to pin down the differential effects of different L2s on language learning outcomes. The same observation could also be extended to the unequal number of studies with participants from different L1 backgrounds, educational levels and employing various interlocutor types. More research is needed that involves primary and secondary school students, learning Spanish, French, and German as a second language, and having native speakers and teachers as interlocutors. Second, the current meta-analysis also revealed that the outcome measures employed by primary researchers to assess certain treatment effects were quite diversified, ranging from e-mail messages, blog entries to summary writing to evaluate writing ability. Effects are likely to fluctuate depending on the ease of the measurement tasks, that is, effects might be greater in studies that employ indirect measurements to assess writing, such as the use of a cloze test, than in studies in which essay writing is used to evaluate students' writing ability. Third, it was found that only a few studies assessed the delayed effect of the CMC interventions. This raises the question of whether the linguistic knowledge acquired via CMC interactions would be sustained over time, and internalized as part of the learners' language system. As far as the meta-analyst knows, there were only a couple of meta-analyses that have specifically scrutinized the overall effectiveness of CMC tools in SLA. However, the findings of this meta-analysis can be described as tentative at best. As with the majority of meta-analyses across disciplines, this one attempts to incorporate as many studies as possible to provide a comprehensive picture of the question under study (Cooper, 2003; Ortega, 2010); however, the obtained overall effectiveness could be misleading without employing a "thick description of the context, content, people and procedures involved" (Kern, 2006, p.189). Future meta-analysts are therefore called to supplement quantitative measures with thick description whenever possible. Additionally, to tackle the long-lasting critique confronting meta-analyses which compare apples to oranges, prospective meta-analysis may be an alternative, where well-planned empirical studies, sound in their research design and theoretical framework, are conducted for the purpose of a future meta-analysis (Berlin & Ghersi, 2005; Oswald & Plonsky, 2010).

APPENDIX A. Codebook on the publication, learner and methodology variables of included studies

A. Publication characteristics

- A.1. Study ID: A number is assigned for each study
- A.2. Author(s): Last name(s), First name(s)
- A.3. Publication year: Last name(s), First name(s)
- A.4. Publication type
  - A.4.1. Journal article
  - A.4.2. Thesis
  - A.4.3. Dissertation
  - A.4.4. Conference proceedings
  - A.4.5. Conference paper

A.4.6. Report
A.4.7. Other (Please specify)
99. Not reported

A.5. Citation: Complete citation based on APA 6.0 style

B. Learner characteristics

B.1. Learners' L1: What is the learners' native language?
B.1.1. Chinese
B.1.2. Japanese
B.1.3. Spanish
B.1.4. English
B.1.5. Arabic
B.1.6. German
B.1.7. French
B.1.8. Mixed
B.1.9. Other (Please specify)
99. Not reported

B.2. Learners' L2: What is the learners' target language?
B.2.1. Chinese
B.2.2. Japanese
B.2.3. Spanish
B.2.4. English
B.2.5. Arabic
B.2.6. German
B.2.7. French
B.2.8. Mixed
B.2.9. Other (Please specify)
99. Not reported

B.3. Learners' L2 proficiency: Did the researcher measure learners' initial L2 language proficiency? If yes, what was the learners' L2 language proficiency?
B.3.1. Low
B.3.2. Low intermediate
B.3.3. Intermediate
B.3.4. Upper intermediate
B.3.5. Advanced
B.3.6. Mixed
B.3.7. Not sure (Please write in author's description)
99. Not reported

B.4. Measurement of Learners' L2 proficiency: What measure was used to measure learners' initial L2 proficiency?
B.4.1. In-house
B.4.2. Institutional
B.4.3. Standardized

B.4.4. Impressionistic
B.4.5. Other (Please specify)
B.4.6. Not sure (Please write in author's description)
99. Not reported

B.5. Educational level: What is the learners' educational level?
B.5.1. Primary school
B.5.2. Secondary school
B.5.3. High school
B.5.4. College or above
B.5.5. Other (Please specify)
B.5.6. Not sure (Please write in author's description)
99. Not reported

B.6. Age: What is the age range of the learners?
B.6.1. Report the age range
99. Not reported

C. Methodology characteristics

C.1. Experimental design: What is the research design of the study?
. C.1.1. Pretest posttest control group design
C1.2. Pretest posttest within group design
C.1.3. Posttest only control group design
C.1.4. Not sure (Please write in author's description)
99. Not reported

C.2. Randomization: Are learners assigned randomly to the treatment groups?
C.2.1. Yes
C.2.2. No
C.2.3. Not sure (Please write in author's description)
99. Not reported

C.3. Research setting: Is the target language a second or foreign language to the learners?
C.3.1. Second language
C.3.2. Foreign language
C.3.3. Not sure (Please write in author's description)
99. Not reported

C.4. Research context: When do L2 learners engage in communication tasks?
C.4.1. In class
C.4.2. After class
C.4.3. Mixed
C.4.4. Not sure (report exact descriptions from the study)
99. Not reported

C.5. Technique of CMC: What is the technology/platform used to carry out the communication?
C.5.1. Skype

C.5.2. Messenger
C.5.3. Bulletin board
C.5.4. E-mail
C.5.5. CMS (classroom-management system)
C.5.6. Blog
C.5.7. Researcher-developed platform
C.5.8. Other (Please specify)
C.5.9. Not sure (Please write in author's description)
99. Not reported

C.6. CMC modality: In what modality is the communication activity employed?
C.6.1. Text-based
C.6.2. Voice-based
C.6.3. Mixed
C.6.4. Not sure (report exact descriptions from the study)
99. Not reported

C.7. CMC temporality: Does the communication take place synchronously or asynchronously?
C.7.1. Synchronous
C.7.2. Asynchronous
C.7.3. Both
C.7.4. Not sure (Please write in author's description
99. Not reported

C.8. Interlocutors: Who is/are the interlocutor(s) of the L2 learners?
C.8.1. Peers
C.8.2. Instructors/researchers
C.8.3. Native speakers
C.8.4. Mixed
C.8.5. Not sure (Please write in author's
99. Not reported

C.9. Communication task: What is/are the communication task(s) employed in the study?
C.9.1. Opinion exchange
C.9.2. Information gap
C.9.3. Problem-solving
C.9.4. Puzzle
C.9.5. Mixed
C.9.6. Other (Please specify)
C.9.7. Not sure (Please write in author's

C.10. Targeted language skills: What is/are the language skill(s) being measured?
C.10.1. Listening
C.10.2. Speaking

C.10.3. Reading
C.10.4. Writing
C.10.5. Pronunciation
C.10.6. Grammar
C.10.7. Vocabulary
C.10.8. Mixed (specify)
C.10.9. Not sure (report exact descriptions from the study)
99. Not reported

C.11. Outcome measurement: What type of outcome measurement is employed in the study?
C.11.1. In-house: Please describe the test and give page number
C.11.2. Institutional
C.11.3. Standardized
C.11.4. Impressionistic
C.11.5. Mixed
C.11.6. Other (Please specify)
C.11.7. Not sure (Please write in author's description
99. Not reported

C.12. Reliability: What is the type of reliability reported in the study?
C.12.1. Test reliability
C.12.2. Inter-rater reliability
C.12.3. Not sure (Please write in author's description)
99. Not reported

C.13. Treatment length: How long does the study last?
C.13.1. Less than or equal to 10 weeks
C.13.2. Between 11 and 24 weeks
C.13.3. More than 24 weeks
99. Not reported

C.14. Tutorial: What tutorial is provided to learners to prepare for their communication tasks?
C.14.1. Technique-based
C.14.2. Content-based
C.14.3. Both
C.14.4. Not sure (Please write in author's description)
99.

C.15. Group size: How many students are there in a group?
C.15.1. 2 people
C.15.2. 3-6 people
C.15.3. > 6 people
C.15.4. Mixed (more than one grouping)
C.15.5. Class
99. Not reported

C.16. Control group: Is there a control/comparison group in the study?

C.16.1. Yes (describe the treatment condition and provide page number)

C.16.2. No

C.16.3. Not sure (Please write in author's description)

99. Not reported

APPENDIX B. List of the included studies and their substantive features

Author	Year	Platform	Publication[a]	T. L.[b]	Cont[c]	Group[d]	Measure[e]
Abrams	2003	WebCT	JA	M	IN	M	IH
AbuSeileek	2007	Discussion Forum	JA	M	IN	M	IH
Ahn	2006	Skype	DN	M	IN	S	IH
Alastuey	2010	Skype	JA	M	IN	P	IH
Alastuey	2011	Skype	JA	M	IN	S	STD
Arslan & Sahin-Kizil	2010	Blog	JA	M	BO	N/A	IH
Blake et al.	2008	Adobe Breeze	JA	S	AF	N/A	IN
Blake	2009	WebCT	JA	S	AF	N/A	IH
Camacho	2008	Chatroom	TH	S	IN	P	IH
Chang	2007	Chatroom	JA	S	IN	N/A	IH
Chang	2008	MSN	DN	M	AF	P	IH
Chang & Hsu	2008	MSN	CP	M	IN	P	IH
Chen	2008	Mixed	TH	S	IN	P	STD
Chiang	2007	Chatroom	JA	L	BO	S	IN
Chung	2004	E-mail	TH	S	IN	S	IH
Coniam & Wong	2004	ICQ	JA	S	AF	N/A	IH
Fellner & Apple	2006	Blog	JA	S	IN	N/A	IH
Fitze	2006	WebCT	JA	S	IN	C	IH
Fuente	2003	Mixed	JA	S	N/A	P	IH
González-Bueno & Pérez	2000	E-mail	JA	M	AF	C	IH
Huang & Chang	2009	MSN	CP	M	IN	P	IH
Huang & Hung	2008	MSN	CP	M	BO	S	STD
Huang & Hung	2010	Blog	JA	M	BO	C	IH
Hung	2007	E-mail	DN	M	BO	P	M
Jian	2005	RD platform[1]	TH	S	IN	S	IH
Jou	2008	E-mail	DN	S	AF	N/A	STD
Kost	2004	IRC	DN	M	IN	S	M
Lee	2009	MSN	TH	S	N/A	P	IH
Lee & Liou	2009	Wiki	CP	M	AF	P	IH
Li	2008	Chatroom	TH	M	AF	Small	IH
Li	2009	Blog	TH	M	IN	S	IH

Liang	2006	Blog	DN	M	AF	P	STD
Lin	2009	NiceNet	DN	S	IN	N/A	IH
Liu	2007	Mixed	DN	S	IN	N/A	STD
Loewen & Erlam	2006	RD platform	JA	S	IN	N/A	IH
Lord	2008	Podcast	JA	M	AF	S	IH
Lu & Liou	2004	E-mail	JA	M	AF	P	IH
Payne & Whitney	2002	Chatroom	JA	M	IN	S	IH
Peng & Hsu	2006	NiceNet	JA	S	IN	N/A	IH
Pérez	2000	E-mail	JA	M	AF	P	N/A
Pyun	2003	MSN	DN	S	AF	P	IH
Sanders	2005	WebCT	JA	L	AF	N/A	IN
Satar & Ozdener	2008	RD platform	JA	S	AF	P	IH
Sequeira	2009	Moodle	DN	M	IN	P	STD
Shang	2007	E-mail	JA	M	AF	P	IH
Simsek	2010	Blog	CP	N/A	IN	N/A	IH
Song & Usaha	2009	Moodle	JA	M	IN	S	IH
Sun	2010	Blog	JA	M	AF	C	IH
Sun	2012	Blog	JA	M	AF	N/A	IH
Thurston et al.	2009	RD platform	JA	S	N/A	P	IH
Tsai	2007	Blog	TH	S	IN	C	IH
Volle	2005	E-mail	JA	M	IN	M	IH
Wang	2010	Chatroom	TH	M	AF	S	IH
Xiao	2007	Skype	DN	S	AF	P	IH
Yang	2006	MSN	DN	M	IN	L	IH
Yang	2011	E-meeting	JA	M	BO	N/A	STD
Yanguas	2012	Skype	JA	S	IN	P	IH
Zheng	2010	Chatroom	TH	M	IN	M	IH
Zhou	2009	Blog	TH	M	N/A	N/A	IH

Notes: [1] Researcher-developed platform; [a]Publication type:JA=journal article; TH=thesis;DN=dissertation; CP=conference paper; [b] Treatment length: S=short (<= 10 weeks); M=medium(11 - 24 weeks); L= longitudinal(>24 weeks); [c]Context: IN= in class; AF=after class; BO=: in and after class; [d]Grouping: P= 2 people; S= 3-6 people; L=more than 6 people; M=more than one grouping; C= no grouping; [e]Measurement: IH=in house; STD=standardized test; IN=institutional assessment; M=mixed

APPENDIX C. Learner characteristics of the included studies

Primary Study	Year	L2	L1 [a]
Abrams	2003	German	Mixed
AbuSeileek	2007	English	Arabic
Ahn	2006	English	Mixed
Alastuey	2010	English	Mixed
Alastuey	2011	English	Mixed

Arslan & Sahin-Kizil	2010	English	Turkish
Blake et al.	2008	Spanish	Mixed
Blake	2009	English	Mixed
Camacho	2008	English	Mixed
Chang	2007	English	Chinese
Chang	2008	English	Chinese
Chang & Hsu	2008	English	Chinese
Chen	2008	English	Chinese
Chiang	2007	English	Chinese
Chung	2004	English	Chinese
Coniam& Wong	2004	English	Chinese
Fellner & Apple	2006	English	Japanese
Fitze	2006	English	Mixed
Fuente	2003	Spanish	English
González-Bueno & Pérez	2000	Spanish	English
Huang & Chang	2009	English	Chinese
Huang & Hung	2008	English	Chinese
Huang & Hung	2010	English	Chinese
Hung	2007	English	Chinese
Jian	2005	English	Chinese
Jou	2008	English	Chinese
Kost	2004	German	Mixed
Lee	2009	English	Chinese
Lee & Liou	2009	English	Chinese
Li	2008	English	Chinese
Li	2009	English	Chinese
Liang	2006	English	Chinese
Lin	2009	English	Mixed
Liu	2007	English	Chinese
Loewen & Erlam	2006	English	Mixed
Lord	2008	Spanish	N/A
Lu & Liou	2004	English	Chinese
Payne & Whitney	2002	Spanish	Mixed
Peng & Hsu	2006	English	Chinese
Perez	2000	Spanish	English
Pyun	2003	English	Korean
Sanders	2005	Spanish	Mixed
Satar & Ozdener	2008	English	Turkish
Sequeira	2009	Spanish	English

Shang	2010	English	Chinese
Simsek	2010	English	N/A
Song & Usaha	2009	English	Chinese
Sun	2010	English	Chinese
Sun	2012	English	Chinese
Thurston et al.	2009	Mixed	Mixed
Tsai	2007	English	Chinese
Volle	2005	Spanish	English
Wang	2010	English	Chinese
Xiao	2007	English	Chinese
Yang	2006	English	Chinese
Yang	2011	English	Chinese
Yanguas	2012	Spanish	English
Zheng	2010	English	Chinese
Zhou	2009	English	Chinese

Note: [a]Mixed L1 indicates students are from a multilingual class.

ACKNOWLEDGEMENTS

This study reported part of the data collected from a two-year grant research project supported by the National Science Council of Taiwan, grant number NSC99-2410-H-007-082-MY2. The remaining data was presented in an earlier study published in the same journal, titled, "Establishing an Empirical Link Between Computer-Mediated Communication (CMC) and SLA: A Meta-Analysis of the Research" in 2014. The author would also like to express her immense gratitude to the anonymous reviewers who offered many suggestions and comments on how to improve this paper.

REFERENCES

*Indicates study included in this meta-analysis

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, *21*(3), 199–226.

*Abrams, Z. I. (2003). The effect of synchronous and asynchronous CMC on oral performance in German. *Modern Language Journal, 87*(2), 157–167.

*AbuSeileek, A. F. (2007). Cooperative vs. individual learning of oral skills in a CALL environment. *Computer Assisted Language Learning*, *20*(5), 493–514.

*Ahn, H. (2006). The impact of individual learner characteristics and synchronous computer-mediated communication on language production in learners of English. Unpublished Doctoral Dissertation, University of Arizona, Arizona.

*Alastuey, M. C. B. (2010). Synchronous-voice computer-mediated communication: Effects on pronunciation. *CALICO Journal*, *28*(1), 1–20.

*Alastuey, M. C. B. (2011). Perceived benefits and drawbacks of synchronous voice-based computer-mediated communication in the foreign language classroom. *Computer Assisted Language Learning*, *24*(5), 419–432.

*Arslan, R. Ş., & Şahin-Kızıl, A. (2010): How can the use of blog software facilitate the writing process of English language learners? *Computer Assisted Language Learning*, *23*(3), 183–197.

Aytug, Z. G., Rothstein, H. R., Zhou, W., & Kern, M. C. (2012). Revealed or concealed? Transparency of procedures, decisions, and judgment calls in meta-analyses *Organizational Research Methods*, *15*(1) 103–133.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*, 29–58.

Bennui, P. (2008) A study of L1 interference in the writing of Thai EFL students. *Malaysian Journal of ELT Research*, *4*, 72–102.

Berlin, J. A., & Ghersi, D. G. (2005). Preventing publication bias: Registries and prospective meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 35–48). Hoboken, NJ: Wiley.

*Blake, C. (2009). Potential of text-based internet chats for improving oral fluency in a second language *The Modern Language Journal*, *93*(2), 227–240.

*Blake, R., Wilson, N., Pearson, N., Cetto, M., & Pardo-Ballester, C. (2008). Measuring oral proficiency in distance, face-to-face and blended classrooms. *Language Learning & Technology*, *12*(3), 114–127, Retrieved from http://llt.msu.edu/vol12num3/blakeetal.pdf.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). Complex data structures. In M. Borenstein, L. Hedges, J. Higgins, & H. Rothstein (Eds.), *Introduction to meta-analysis* (pp. 215–244). Chichester: Wiley.

*Camacho, R. (2008). *Developing writing fluency through synchronous computer-mediated communication.* Unpublished Master's Thesis, Brigham Young University, Utah.

*Chang, Y.-Y. (2007). The potential of synchronous text-based computer-mediated communication for second language acquisition. *Issues in Information Systems*, *8*(2), 355–361.

*Chang, Y.-Y. (2008). *A study of synchronous text-based computer-mediated communication and language learning outcomes.* Unpublished Doctoral Dissertation, Tamkang University, Taipei, Taiwan.

*Chang, Y. C., & Hsu, H.-H. S. (2008). The factors of synchronous computer-mediated communication on EFL writing. In Proceedings of The Second Tamkang International Conference on Second Language Writing (pp.26–36). Taipei, Taiwan: Tamkang University.

*Chen, F. (2008). *The effects of voiced-based synchronous and asynchronous CMC on EFL learners' oral proficiency.* Unpublished Master's Thesis, National Taiwan Normal University, Taipei, Taiwan.

Chenoweth, N. A., Ushida, E., & Murday, K. (2006). Student learning in Hybrid French and Spanish courses: An overview of language online. *CALICO*, *24*(1), 115–145.

*Chiang, M. H. (2007). A novel idea: English as foreign language reading via virtual literature circles. *English Teaching & Learning*, *31*(4), 1–37.

Chun, D. (2007). Come ride the wave: But where is it taking us? *CALICO Journal, 24*(2), 239-252.

*Chung, M.-C. (2004). *Effects of three theme-based projects on English learning for a CMC group and a non-CMC group in junior high school in Taiwan*. Unpublished Master's Thesis, National Kaohsiung

Normal University, Kaohsiung, Taiwan.

Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Coniam, D., & Wong, R. (2004). Internet relay chat as a tool in the autonomous development of ESL learners' English language ability: An exploratory study. *System*, *32*, 321–335.

Cooper, H. M. (2003). Editorial. *Psychological Bulletin*, *129* (1), 3–9.

Creswell, J. W. (2008). *Research design: Quantitative, qualitative and mixed methods approaches (3rd ed.).* Thousand Oaks, CA:Sage.

Fagan, E. R. & Cheong, P. (1987). Contrastive rhetoric: Pedagogical implications for the ESL teacher in Singapore. *RELC Journal*, *18*(1), 19–30.

Felix, U. (2005). What do meta-analyses tell us about CALL effectiveness? ReCALL, 17(2), 269–288.

Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learned in two decades of research? *ReCALL*, *20*(2), 141–161.

*Fellner, T., & Apple, M. (2006). Developing writing fluency and lexical complexity with blogs. *The JALT CALL Journal*, *2*(1), 15–26.

*Fitze, M. (2006). Discourse and participation in ESL face-to-face and written electronic conferences. *Language Learning & Technology*, *10* (1), 67–86, Retrieved from http://llt.msu.edu/vol10num1/pdf/fitze.pdf

* Fuente, M. J. (2003), Is SLA interactionist theory relevant to CALL? A study on the effects of computer-mediated interaction in L2 vocabulary acquisition, *Computer Assisted Language Learning, 16*(1), 47–81.

Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, *16*, 283–302.

Glass, G. V. (1976). *Primary, secondary, and meta-analysis of research*. Presidential address to the Annual Meeting of the American Educational Research Association, San Francisco, April 21, 1976.

*González-Bueno, M., & Pérez, L. C. (2000). Electronic mail in foreign language writing: A study of grammatical and lexical accuracy, and quantity of language. *Foreign Language Annals*, *33*(2), 189–198.

Grgurović , M., Chapelle, C. A., & Shelley, M. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, *25*, 165–198.

Herring, S. C. (1996). *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Amsterdam, The Netherlands: John Benjamins.

Herring, S. C. (1999), Interactional Coherence in CMC. *Journal of Computer-Mediated Communication*, 4: Retrieved from http://onlinelibrary.wiley.com.nthulib-oc.nthu.edu.tw/doi/10.1111/j.1083-6101.1999.tb00106.x/abstract

Hoffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, *17*, 722–738.

*Huang, S. Y. & Chang, Y. (2009). A mediation effect of computer-mediated communication on EFL writing. In G. Siemens & C. Fulford (Eds.), In The Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (pp. 2588–2597). Chesapeake, VA: AACE.

*Huang, H.-T. D. & Hung, S.-T. A. (2008). Chatting to write: Synchronous computer-mediated communication in an EFL writing classroom. In C. J. Bonk, M. M. Lee, & T. Reynolds (Eds.), The Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher

Education (pp. 1638–1645). Chesapeake, VA: AACE.

*Huang, H.-T. D., & Hung, S.-T. A. (2010). Effects of electronic portfolios on EFL oral performance. *Asian EFL Journal*, *12*(2), 192–212.

*Hung, P.-Y. (2007). *The effects of exchanging emails with American key pals on Taiwanese college students' writing in English*. Unpublished Doctoral Dissertation, Kent State University.

Jackson, G. B. (1980). Methods for integrative reviews. *Review of Educational Research*, *50*(3), 438–460.

Jenks, C. J. (2009). When is it appropriate to talk? Managing overlapping talk in multi-participant voice-based chat rooms. *Computer Assisted Language Learning*, *22*(1), 19–30.

Jeon, E. H., & Kaya, T. (2006). Effects of instruction on interlangauge pragmatic development. In J. Norris & L. Ortega, (Eds.), *Synthesizing research on language learning and teaching* (pp. 165–211). Amsterdam, The Netherlands: John Benjamins.

*Jian, Q. W. (2005). *Effects of online peer response on EFL college writing.* Unpublished Masters' Thesis, National Tsing Hua University, Hsinchu, Taiwan.

*Jou, Y-A. E. (2008). *The effect of email keypal project on the enhancement of reading and writing performance of elementary school-aged students in Taiwan*. Unpublished Doctoral Dissertation, Alliant International University, California.

Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition. In J. Norris & L.Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Amsterdam, The Netherlands: John Benjamins.

Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly*, *40*(1), 183–209.

*Kost, C. R. (2004). *An investigation of the effects of synchronous computer-mediated communication (CMC) on interlanguage development in beginning learners of German: Accuracy, proficiency and communication strategies*. Unpublished Doctoral Dissertation, University of Arizona, Arizona.

*Lee, J. Y. (2009). *The effect of computer-mediated communication (CMC) interaction on L2 vocabulary acquisition: A comparison study of CMC interaction and face-to-face interaction*. Unpublished dissertation, Iowa State University, Iowa.

*Lee, S. L., & Liou, H. C. (2009). A case study of Wiki-base pair versus individual writing in an EFL College Context. In the Proceedings of 27th International Conference on English Teaching and Learning in the R.O.C. (pp.28–44). Kaohsiung, Taiwan: Crane.

*Li, L. J. (2008). *The impact of different levels of instructional strategy on synchronous CMC in improving college students' oral English proficiency and learning motivation*. Unpublished Master's Thesis, National Cheng Kung University, Tainan, Taiwan.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*(2), 309–365.

*Li, W.-L. (2009). *The effects of integrating blogging into peer feedback revision on English writing performance and attitude of vocational high school students in Taiwan*. Unpublished Master's Thesis, National Taiwan Normal University, Taipei, Taiwan.

*Liang, M-Y. (2006). *Interaction in EFL online classes: How web-facilitated instruction influences FL university students' reading and learning*. Unpublished Doctoral Dissertation. Indiana University, Indiana.

*Lin, S. M. (2009). *How computer-mediated communication affects ELL students' writing processes and writing performance*. Unpublished Doctoral Dissertation. Oklahoma University, Oklahoma.

Lin, W.-C., Huang, H.-T., & Liou, H.-C. (2013). The effects of text-based SCMC on SLA: A meta-analysis. *Language Learning & Technology*, *17*(2), 123–142. Retrieved from http://llt.msu.edu/issues/june2013/linetal.pdf.

*Liu, C. N. (2007). *Pragmatics in foreign language instruction: The effects of pedagogical intervention and technology on the development of EFL learners' realization of "Request"*. Unpublished Doctoral Dissertation, Texas A & M University, Texas.

*Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, *19*(1), 1–14.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition (pp.413-468).* San Diego, CA: Academic Press.

*Lord, G. (2008). Podcasting communities and second language pronunciation. *Foreign Language Annals*, *41*(2), 374–389.

*Lu, K. Y., & Liou, H. C. (2004). Effects of E-mail exchanges on EFL senior high school students' English writing. *English Teaching & Learning*, *28*(4), 95–119.

Networked Digital Library of Theses and Dissertations (NDLTD)

Nguyen, L. V. (2008). Computer mediated communication and foreign language education: Pedagogical features. *International Journal of Instructional Technology and Distance Learning, 5*(12).

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*, 417–528.

Norris, J. M., & Ortega, L. (2006). The value and practice of research synthesis for language learning and teaching. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 3–50). Philadelphia, PA: John Benjamins.

Ortega, L. (2010). Research syntheses. In B. Paltridge & A. Phakiti (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 111–126). London, UK: Continuum.

Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, *30*, 85–110.

*Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: output, working memory, and interlanguage development. *CALICO Journal*, *20*(1), 7–32.

*Peng, C. Y., & Hsu, P. Y. (2006). Implementing online journal practice into EFL writing class. *Chaoyang Journal of Humanities and Social Sciences*, *4*(2), 1–20.

*Pérez, L. C. (2000). Electronic mail: An extension of my Spanish class boundaries? *Cincinnati Romance Review*, *19*, 138–145.

ProQuest Digital Dissertation (PQDD)

*Pyun, O. C. (2003). *Effects of networked language learning: A comparison between synchronous online discussions and face-to-face discussions*. Unpublished Master's Thesis, The Ohio State University, Ohio.

*Sanders, R. E. (2005). Redesigning introductory Spanish: Increased enrollment, online management, cost reduction and effects on student learning. *Foreign Language Annals*, *38*(4), 523–532.

*Satar, H. M., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and

anxiety: Text versus voice chat. *The Modern Language Journal*, *92*(4), 595–613.

Sauro, S. (2011). SCMC for SLA: A research synthesis. *CALICO Journal, 28*(2), 369–391.

*Sequeira, C. A. (2009). *Synchronous computer-mediated communication and second language proficiency*. Unpublished Doctoral Dissertation, University of Oregon, Oregon.

*Shang, H. F. (2007) An exploratory study of e-mail application on FL writing performance. *Computer Assisted Language Learning*, *20*(1), 79–96.

*Simsek, Ö. (2010). The effect of weblog (blog) students' writing performance. *Procedia Social and Behavioral Sciences*, *2*, 953–958.

Smith, Bryan. (2004). Computer mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition 26*, 365–398.

*Song, W., & Usaha, S. (2009). How EFL university students' use electronic peer response into revisions. *Suranaree Journal of Science & Technology 16*(3), 263–275.

*Sun, Y. C. (2010). Extensive writing in foreign language classrooms: A blogging approach. *Innovations in Education and Teaching International*, *47*(3), 327–339.

*Sun, Y. C. (2012). Examining the effectiveness of extensive speaking practice via voice blogs in a foreign language learning context. *CALICO Journal*, *29*(3), 494–506

Taylor, A. M. (2006). The effects of CALL versus traditional L1 glosses on L2 reading comprehension. *CALICO Journal, 23*(2), 309–318.

Taylor, A. M. (2009). CALL-based versus paper-based glosses: Is there a difference in reading comprehension? *CALICO Journal, 27*(1), 147–160.

Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning*, *44*, 307–336.

Thorne, S. (2008). Computer-mediated communication. In N. Van Deusen-Scholl & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 1415–1426). Philadelphia, PA:Springer.

*Thurston, A., Duran, D., Cunningham, E., Blanch, S., & Topping, K. (2009). International on-line reciprocal peer tutoring to promote modern language development in primary schools. *Computers & Education*, *53*, 462–472.

*Tsai, Z. Y. (2007). *Effects of a teacher-centered blog on English reading comprehension*. Unpublished Master's Thesis. National Kaohsiung Normal University, Kaohsiung, Taiwan.

*Volle, L. M. (2005). Analyzing oral skills in voice E-mail and online interviews. *Language Learning & Technology*, *9*(3), 146–163. Retrieved from http://llt.msu.edu/vol9num3/pdf/volle.pdf

*Wang, C. Y. (2010). *A study comparing the effects of synchronous CMC and FTF interaction on L2 oral proficiency development for students with various working memory capacities*, Unpublished Master's Thesis, National Tsing Hua University, Hsinchu, Taiwan.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The Modern Language Journal*, *81*(4), 470–481.

*Xiao, M. (2007). *An empirical study of using Internet-based desktop videoconferencing in an EFL setting*. Unpublished Doctoral Dissertation, Ohio University, Ohio.

*Yang, M. L. (2006). *Exploring Taiwanese EFL students' responses to synchronous CMC: Effects on language use, learning and transfer, and perceptions*. Unpublished Doctoral Dissertation. University of

Texas at Austin, Texas.

*Yang, Y. F. (2011). Engaging students in an online situated language learning environment. *Computer Assisted Language Learning*, *24*(2), 181–198.

*Yanguas, I. (2012). Task-based oral computer-mediated communication and L2 vocabulary acquisition. *CALICO Journal*, *29*(3), 507–531.

Yun, J. (2011).The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis, *Computer Assisted Language Learning, 24*(1), 39-58, DOI: 10.1080/09588221.2010.523285.

Zhao, Y. (2003). Recent developments in technology and language Learning: A literature review and meta-analysis, *CALICO Journal, 21*(1), 7–27.

*Zheng, S. N. (2010).*The transferability from synchronous computer-mediated communication to oral discussion*. Unpublished Master's Thesis, National Tsing Hua University, Hsinchu, Taiwan.

*Zhou, H. J. (2009). *Effects of weblog-based writing instruction on English writing performance of vocational high school students*. Unpublished Master's Thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.