A systematic review of empirical studies on computer-assisted language learning

Esin Saylan, Mehmet Kokoç and Zeynep Tatlı

Editor: Kerry Earl Rinehart


Link to this volume: https://doi.org/10.15663/wje.v28i1

Copyright of articles

Authors retain copyright of their publications.

Articles are subject to the Creative commons license: https://creativecommons.org/licenses/by-nc-sa/3.0/legalcode

Summary of the Creative Commons license.

Author and users are free to

Share—copy and redistribute the material in any medium or format

Adapt—remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms

Attribution—you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use

Non-Commercial—you may not use the material for commercial purposes

ShareAlike—if you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original

No additional restrictions—you may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Open Access Policy

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge.
A systematic review of empirical studies on computer-assisted language learning

Esin Saylan, Mehmet Kokoç, Zeynep Tatlı
Trabzon University
Turkey

Abstract

The aim of this study was to examine the current trends and main findings of the empirical computer-assisted foreign language learning studies. The database, Web of Science, was reviewed and a total of 36 studies between 2000 and 2020 were analysed. The systematic analysis was performed with the content analysis method. The content of the reviewed articles was analysed in the following categories: The technological trends, the technological context used, variables, duration of the studies, data collection tools, data analysis strategies, skills/factors analysed, underlying theories, and countries of the studies, target languages, participants, and years of the studies. Along with other insights this review found, the most common technological trend was educational software, while the technological contexts used varied greatly. Key variables from mostly small studies (less than 100 participants generally tertiary students), focused on technological contexts, language learning skills (typically vocabulary and speaking) and participants’ attitudes. Underlying theories of these studies varied greatly, Technology Acceptance Model and Cognitive Load Theory being common. This review study can serve as a guide for practitioners who plan designing computer-assisted language learning and teaching activities.

Keywords

Computer-assisted language learning; foreign language learning; learning technologies

Introduction

The rapid advances in digital technologies and their ever-increasing opportunities impact the way of gaining and utilizing information and create new opportunities for education, facilitating learning beyond traditional classroom environments. Given the rise of digital technologies in the field of language education in the last two decades due to the widespread use of Internet and the preferences of the new generation students, information and communication technologies have become integrated into foreign language syllabi and are being heavily relied on in language teaching and learning as tutors, communication facilitators, test or authentic data source or social interactors (Yang, 2010). Computer-assisted Language Learning (CALL) as an interdisciplinary field has used the advanced learning and
teaching technologies in the context of different perspectives. CALL technologies attempt to help language teachers improve the quality of their teaching with the use of computers as the main media in the form of computer screen-displayed text, visuals, sound, calculations, control, storage, and other functions (Liu et al., 2020). CALL technologies are not limited to the use of computers in class but all the digital technologies in and out of class with the aim of language learning.

Previous studies showed that CALL technologies have been employed for speaking, writing, and pronunciation skills such as Skype, E-teaching analytical rubrics, automatic speech and error recognition systems (Alfehaid, 2018; Karal, Kokoç, & Çakır, 2017; Neri et al., 2008; Poursalehi et al., 2014; Romaña Correa, 2015; Strik et al., 2009) and through podcasts, web-based activities, web-delivered videos among many others for the improvement of listening skills (Al Qasim & Al Fadda, 2013; Chen et al., 2014; Meihami et al., 2013). For writing skills, various digital technologies have been used, such as feedback alternatives, corrective programs, word processor, error analysis, and machine translation (Garcia & Pena, 2011; Zaini & Mazdayasna, 2014). Also, CALL technologies have been used for the development of reading skills through the use of images, sounds, graphs, and animations integrated into reading activities on computers (Bhatti, 2013; Meihami & Varmaghani, 2013); for vocabulary and grammar achievement through systematic vocabulary repetition and extra software exercises on computer applications (Alahmadi, 2019; Enayati & Gilakjani, 2020; Tozcu & Coady, 2004; Zapata & Sagarra, 2007); and, though limited in number, for student/teacher perspectives, motivation, and autonomy, (Ghufron & Nurdianingsih, 2020; Gómez-Parra & Espejo-Mohedano, 2020; Halvorsen, 2020; Kessler & Bikowski, 2010; Raby, 2007). Taken together, these studies support the notion that effective use of CALL technologies should be integrated into language teaching and learning context (Fathi & Ebadi, 2020).

There have been several studies focusing on CALL, reporting positive results of CALL applications. There have been various review studies on CALL (Golonka et al., 2014; Wang & Vásquez, 2012; Mohsen & Balakumar, 2011; Uzunboylu & Ozcinar, 2009; Felix, 2005; Stockwell, 2007; Hubbard, 2005) and meta-analysis studies (Abraham, 2008; Peterson, 2010; Chiu, 2013; Grgurović et al., 2013) concentrating on its effects within specific language skills but limited in number and not making differences between second and foreign language learning. Therefore, a study looking at the trends and main findings of CALL in all language skills in foreign language learning context over the last two decades is required. In addition, it is necessary to have a comprehensive study of empirical CALL in foreign language classrooms to pave the way for future studies and provide new insights for researchers by displaying the results of previous research. This study reports a comprehensive systematic review of empirical CALL studies in the 21st century between 2000 and the end of May 2020 on Web of Science, which is a flagship publisher-independent global citation database. The aim of this study is to explore the current state of empirical CALL studies by investigating the theoretical perspectives framing the studies, identifying the technologies and data collection tools along with the variables studied. This research combined the results of 36 independent studies dealing with the use of CALL in foreign language learning environment to identify the effect of CALL applications in the learning process of different language skills (reading, writing, listening, speaking, vocabulary, and grammar), and the primary aim of this study was to find the trends of empirical CALL studies in foreign language learning context and to display their results. Bearing this question in mind, the following additional research questions were asked:

- What are the 21st century technological trends in an empirical CALL foreign language learning environment?
- What technological learning contexts have been used in the available literature?
- Which variables have been used in the available literature?
- How long have the empirical CALL studies taken place in the available literature?
- What data collection tools and data analysis techniques have been used in the available literature?
• What skills/factors have been examined in CALL studies in recent years?
• Which underlying theories have been examined in the available literature?
• Which countries have studied empirical CALL in the available literature?
• What were the target languages studied in empirical CALL studies in the available literature?
• What level and number of participants joined in empirical CALL studies in the available literature?
• In which years have the empirical CALL studies been performed?

Method

This study followed a systematic review method. A systematic review is a literature review with specific methodology to get a synthesis of existent evidence for a focused question through a specific protocol by independent reviewers (Bearman et al., 2012). A systematic review study aims to provide a comprehensive, detailed summary of primary research for a specific research question after careful identification, selection, synthesis of the high-quality evidence pertinent to priori selected criteria via transparent, explicit, and systematic methodology (Harris et al., 2014).

Procedure and data collection

With the aim of getting a deeper understanding of the current state of empirical CALL studies in a foreign language learning environment, this systematic review was conducted based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009) guidelines. The first step was to conduct a comprehensive literature search for studies focusing on CALL in foreign language learning contexts. The literature search covered the articles in the period from 2000 to the end of May 2020, the first two decades of the 21st century, and included a computer search of Web of Science (WOS), the premier research platform for information in the sciences, social sciences, arts, and humanities. Two main search terms, “computer-assisted foreign language learning” and “computer-based foreign language learning” refined by the term “experimental studies”, were searched in the database within the time span of 2000–2020 in the category of Topic, which searches title, abstract, author keywords, and keywords plus indexed in SCI-EXPANDED, SSCI, A&HCI journals. The following criteria were considered in the selection of the articles:

• The study was published between 2000 and 2020 (the latest online search date was 27 May).
• The study investigated some form of CALL (e.g., e-mail, online discussions, videos, podcasts, electronic glosses, etc.).
• The study employed an experimental or quasi-experimental design.
• The studies included adequate information for the methodology used.
• Studies recruited participants who were foreign language learners (second language learners were excluded).
• The study should be reported in English, as the researchers’ comprehension languages were English and Turkish and translation of other languages would result in difficulties interfering with the analysis process.
• The publication type of the study should be a journal article (dissertations, unpublished manuscripts, proceeding papers, book reviews, reviews, editorial materials, and case studies were excluded).

The literature search initially identified 811 full text articles in all indexes but reduced to 329 when limited to SCI-EXPANDED, SSCI, A&HCI indexes. After refined by the term “experimental studies”, 45 studies were gathered. When the inclusion criteria was applied for languages and eligible data for
systematic analysis, the final number of eligible studies was 36 (see Appendix), with a total of 3,419 students (See Figure 1).

**Figure 1.** The flowchart of study identification process.

**Data analysis**

Technology used and technological learning context, variables, duration of the studies, data collection tools, data analysis methods, language skills and other factors studied, theories considered, country of the study, target language, and level and number of participants were extracted from each study (For a table of results, please contact the authors.) Two independent evaluators coded variables on all 36 papers, with 0.93 inter-evaluator agreement. Upon the existence of discrepancies between evaluators, articles were reviewed again and a consensus decision was reached by all three authors.

**Findings**

The 21st century trends of empirical CALL studies were analysed in terms of technological trends, technological learning contexts, variables, duration of the studies, data collection tools, data analysis strategies, analysed skills/factors, underlying theories, countries of the studies, number and education level of participants, and years of the published articles.

**Technological trends in empirical CALL studies**

In response to RQ1, the general technological trends in the available literature were categorised by the authors. The most commonly used technology was educational software (n=12) followed by Web 2.0 tools (n=6). The other technologies used were mobile technology (n=3), e-book (n=3), multimedia (n=3), recognition technology (n=1), artificial intelligence (n=1), podcasts (n=1), e-learning (n=1),
Microsoft Office (n=1), email (n=1), virtual reality (n=1), eye-tracking (n=1), and LMS (n=1) (Table 1).

Table 1. Distribution of the Technologies Used in the Available Literature

<table>
<thead>
<tr>
<th>Technological trends</th>
<th>Number (frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational software</td>
<td>12</td>
</tr>
<tr>
<td>Web 2.0 tools</td>
<td>6</td>
</tr>
<tr>
<td>Mobile Technology</td>
<td>3</td>
</tr>
<tr>
<td>E-book</td>
<td>3</td>
</tr>
<tr>
<td>Multimedia</td>
<td>3</td>
</tr>
<tr>
<td>Recognition Technology</td>
<td>1</td>
</tr>
<tr>
<td>Artificial Intelligence</td>
<td>1</td>
</tr>
<tr>
<td>Podcasts</td>
<td>1</td>
</tr>
<tr>
<td>E-learning</td>
<td>1</td>
</tr>
<tr>
<td>Microsoft Office</td>
<td>1</td>
</tr>
<tr>
<td>E-mail</td>
<td>1</td>
</tr>
<tr>
<td>Virtual Reality</td>
<td>1</td>
</tr>
<tr>
<td>Eye-tracking</td>
<td>1</td>
</tr>
<tr>
<td>LMS</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>36</strong></td>
</tr>
</tbody>
</table>

The educational software used in the studies were instructional streaming video (ISV), Chinese Listening and Speaking Diagnosis and Remedial Instruction (CLSDRI), a company-installed software, computer-supported collaborative learning (CSCL) system, instructional lexis software, dubbed LexisBOARD, Bayesian networks, a computer-assisted task-based language instruction (CATBI) tool, computer-assisted form-focused language instruction (CAFFI) tool, a CALL system on personal digital assistants (PDAs), verbal oral feedback system, foreign languages in the elementary school (FLES) software, and computer-generated animations/static visuals. Web 2.0 tools in the studies were online summary writing-pal (SW-PAL), web-based collaborative writing on Google Docs, CSCL English writing tool, Web-discussion board on the University Internet System (UIS), synchronous computer-mediated communication, and voice over instant messaging (VoIM) English teaching assistants (ETAs). The mobile technologies used in the studies were English File Pronunciation (EFP) application, outdoor u-learning (mobile devices/GPS) and Indoor CALL. The e-book used in the studies were electronic gloss, data-driven learning (DDL), e-vocabulary book, online-dictionary, and electronic computer glosses. The multimedia tools used in the studies were captions, digital video cases, captioned interactive videos and interactive videodiscs. The recognition technology used was KinectV2, a recognition device. The artificial intelligence was intelligent personal assistants (IPA), Alexa. The podcasts used were for pronunciation. The e-learning tool used in the study was e-learning platforms for online writing. The Microsoft Word Office program was used in one study. The virtual reality used in one study was overspecified references in a virtual world (designed using GIVE platform). The eye-tracking tool used in one study was captions/automatic caption-filtering system via an eye-tracking system. Finally, the LMS
used in one study was the official LMS of a Turkish state university. (For a table of results, please contact the authors.)

**Technological learning contexts**

In response to RQ2, the technological learning contexts used in the studies were examined and a wide range of contexts was realised. The technological learning contexts are any analogue or digital technologies, educational applications, software, or e-learning services developed for Technology Enhanced Learning. The technological tools used in the studies examined were as follows: Summary writing online tool (n=1) (Chew et al., 2020), Intelligent Personal Asistants (n=2) (Chang & Hsu, 2011; Dizon, 2020), English File Pronunciation Application (n=1) (Fouz-González, 2020), Recognition Device (n=1) (Hwang et al. 2019), Podcasts (n=1) (Fouz-González, 2019), E-learning Platforms (n=1) (Zibin & Altakhaineh, 2019), Outdoor U-learning Devices and Indoor Call Applications (n=2) (Chang, 2018; Chang et al., 2018), Electronic/Computer Glosses (n=2) (Bowles, 2004; Lee et al., 2016), Online Dictionary (n=1) (Karras, 2016), Instructional Streaming Videos (n=1) (Huang & Chuang, 2016), CLSDRI-A Remedial Instruction System (n=1) (Hsiao et al., 2016), Google Docs (n=1) (Bikowski & Vithanage, 2016), Laptop Computer with a Headset (n=1) (Alvarez-Marinelli et al., 2016), Online Feedback through CSCL System (n=2) (Lan et al., 2015; Yang, 2016), Instructional Lexis Software LEXISBOARD (n=1) (Mirzaei et al., 2016), Microsoft Word Office (n=1) (Zaini & Mazdayasna, 2015), Online Feedback through E-mail (n=1) (Alipanahi & Mahmoodi, 2015), Virtual World (n=1) (Luccioni et al., 2015), Captions/Eye Tracking System (n=1) (Hsu et al., 2014), Bayesian Network (n=1) (Aslan et al., 2014), LMS (n=1) (Başal & Gürol, 2014), Computer Assisted Tools (CATBI-CAFFI) (n=1) (Arslanyılmaz, 2013), Web Discussion System (n=1) (Chang et al., 2013), Captions (n=2) (Lwo & Lin, 2012; Shea, 2000), Voice over Instant Messaging System-VOIM (n=1) (Yang et al., 2012), Digital videos (n=1) (Zottmann et al., 2012), Verbal Oral Feedback systems (n=1) (Özdener & Satar, 2009), Synchronous Computer-mediated Communication (n=1) (Şahin, 2009), Computer Generated Visuals (n=2) (Lin & Chen, 2007; Lin et al., 2006), and Computer-based Instruction System (n=1) (Nutta et al., 2002). These results indicate that no specific technological learning context have been used commonly in empirical studies, but a wide range of contexts were utilised for better computer-based language learning opportunities. The technological learning contexts and tools in the available literature focused on different language skills and factors effective in language learning.

**Variables included**

In response to RQ3, variables studied in the available literature were categorised as independent and dependent variables, yielding quite a wide range of results. The dependent variables changed according to the technological learning context used as stated before and dependent variables focused on academic achievement in basic language skills, such as Writing, Listening, Speaking, Pronunciation, Grammar, Vocabulary, Reading, Perceptions of Learners on Independent Variables, Cognitive Load, Learning Styles, Language Production, Analytical Skills, and Language Proficiency in all Skills.

**Duration of the studies**

In response to RQ4, the duration of each study in the available literature was examined yielding results of minimum >610 seconds and maximum 13 months. The results for the duration of empirical studies are given in Figure 2. The studies completed within the same day were stated as <1 day in the chart and
there were two studies indicated as semesters and one study did not indicate the duration of the procedure.

Note: f meaning frequency (number)

Figure 2. Duration of the empirical CALL studies in the available literature.

Data collection tools

In response to RQ5, the data collection tools in the available literature were analysed. As the study focused on only empirical studies, 28 studies used pre-test post-test strategies, two studies used post-test without a pre-test, one study used scores on midterms and final exams, one study used assigned tasks, two studies used weekly tests/questions, one study used only two tests for examining on the spot measurement, and one study used t-units to gather the required data, and seven studies used delayed test or second post-test design. Along with the mentioned test tools, 18 studies used questionnaires/scales, four studies used interviews, two studies performed classroom observations, one study used eye-movement tracking device, one study used checklist for activity performance, and one study used chat scripts for the required data gathering (Figure 3).

Note: f meaning frequency (number)

Figure 3. Data collection tools used in the available literature.

Data analysis strategies

In response to RQ6, the data analysis strategies of the available literature were analysed, and some studies were found to have used more than one analysis strategy (n=10). The most often used statistical analysis method was t-test (n=18) followed by ANOVA (n= 11). The others were ANCOVA (n=5), MANCOVA (n=3), Mann-Whitney-U Test (n=2), SPACOVA (n=1), Wilcoxon-Signed-Rank Test (n=1), Pearson’s Correlation (n=1), Regression Analysis (n=1), Levene’s and Shapiro-Wilk Tests (n=1), Post-hoc Bonferroni (n=1), Tukey Post hoc test (n=1), Hierarchical Linear Modeling (n=1), P-density
MANOVA (n=1), Path Analysis (n=1), Kruskall-Wallis Test (n=1), and Analysis of Within-Group Comparison (n=1) as seen in Figure 4.

Note: \( f \) meaning frequency (number)

Figure 4. Data analysis strategies in the available literature.

Skills/factors analysed

In response to RQ7, the skills and other factors examined in the available literature were analysed and some studies were found to have examined more than one skill and other factors (n=22). Among seven basic language skills categorized (Writing, Speaking, Listening, Vocabulary, Pronunciation, Grammar, Reading), the most studied skill was Vocabulary Comprehension (n=10) followed by Speaking (n=8). The others were Writing (n=7), Reading (n=7), Pronunciation (n=5), Grammar (n=3), Sentence Formation with all dimensions (n=1), and Language Production with all dimensions (n=1). Other than basic skills, among other various factors examined in the available literature, Students’ Perceptions of the technological learning context used or the activities performed were the most commonly analysed factors (n=7) followed by Student Motivation (n=3) and Cognitive Load (n=3). The other factors examined along with basic skills were Students’ Attributes, Ability Levels, Students’ Satisfaction, Attention, Teacher’s Perception, Students’ Feelings, Cognitive Skills, Analytical Skills, and Learning Styles, each studied in one study (n=1) as shown in Figure 5.

Note: \( f \) meaning frequency (number)

Figure 5. Skills and factors analysed in the available literature.
Underlying theories

In response to RQ8, the underlying theories, principles, and hypothesis that the available literature was based on were examined but none could be identified in eight studies. Among the remaining 28 studies, TAM model (n=3) and Cognitive Load theory (n=2) were the most commonly used theories. CALL was the main argument in all the studies though not explicitly stated in some. The other theories considered were the Zone of Proximal Development (ZPD), Swain’s (1993) Output Hypothesis, Interactionist Theory, Cognitive Theory of Multimedia Learning (CTML), Total Physical Response (TPR) Method, ARCS Learning-Motivational Model, blended learning, Technology-Assisted Learning Theory, Cognitive Theory of Multimedia Learning, Data-Driven Learning Theory, Technology-Assisted Sheltered Model, Collaborative Autonomous Pedagogy, computer-supported collaborative learning, The Lexical Approach, computer-mediated communication, Information Processing Theory, Felder and Silverman’s (1988) Learning Styles Model (FSLSM), web-based language education, computer-mediated communication (CMC), communicative language teaching (CLT), Generative Theory of Multimedia Learning, social constructivism, Cognitive Flexibility Theory, Interaction Hypothesis, Cognitive Theory of Learning, Dual Coding Theory, and Schmidt’s (1995) Noticing Hypothesis, as seen in Figure 6.

Countries of the studies

In response to RQ9, the countries where the available studies were performed were analysed and Taiwan was the leading country in empirical CALL studies (n=14) followed by the USA (n=6). The other studies were from Iran (n=3), Turkey (n=3), Spain (n=2), Argentina, Costa Rica, Germany, Jordan, Japan, Korea, Malaysia, and Vietnam (each n=1) (Figure 7).
Target Languages

In response to RQ10, the target languages used for the application of the CALL studies were examined and the most used foreign language was found to be English (n=29) followed by Spanish (n=3), Russian (n=2), Chinese (n=1), Turkish (n=1), and French (n=1). One study used both French and Russian as target languages as seen in Figure 8.

Participants

In response to RQ11, the education level of the participants and number of the participants in each study were analysed. Among the most common education level of the participants was University (n=25). The others were High School (n=3), Junior High School (n=2), Secondary School (n=2), Elementary School (n=3), and Graduate (n=1) (Figure 9). The studies covered had at least 16 participants and at most 816 participants (Figure 10).

Figure 7. The frequency and places of the empirical CALL studies in the available literature.

Figure 8. The target languages in the available literature.

Figure 9. Education level of the participants.
A systematic review of empirical studies on computer-assisted language learning

Publication date

In response to RQ12, the publishment years of the available literature was analysed and 2016 was observed to be the peak year (n=8) in the new millennia as for the empirical CALL studies with fluctuations in the other years after a stable minimum number of studies (n=1) in the first seven years of the two decades (Figure 8).

Discussion

In this study, 36 articles focusing on empirical CALL studies retrieved from Web of Science were analysed in terms of common trends and main results. The analysis showed a gradual increase in the number of empirical CALL studies after 2000, reaching to the highest level in 2016 followed by a sudden decrease in the following two years and a stable condition in the last two years considered as the final point for this study.

In the examined studies, the most common technological trends were Educational Software and Web 2.0 tools. The technological learning contexts in the available data which describe the application of any technological or computer-assisted language learning tools to teaching and learning context, for the facilitation of different language skills here, varied greatly. Shadiev & Yang (2020) identified 24 technologies in their review studies and, similarly, we identified 30 different technological learning contexts under 14 different technology categories, indicating that researchers have been using all the available technological contexts for the development of foreign language skills. Hubbard (2005) claimed in his review study that CALL is still unexplored in the field of language teaching; therefore, any apparent innovation can be examined to learn more about it, which seems to be the case in our review study as well. Macaro et al. (2012) reviewed the CALL studies between 1991 and 2010 and found Multimedia (22%) and CMC Technologies as the most frequently studied technologies. However, in
our study Multimedia was 8% with Educational Software being the leading technology (33%). This difference may have resulted from the different time spans, as our study includes a more recent time span with newer technological contexts introduced.

The variables in the available literature focused on the technological contexts used and their effects on the language skills taught. The most studied language skill in the available data was Vocabulary Comprehension/Retention followed by Speaking skills. Macaro et al. (2012) identified Vocabulary and Writing as the most studied skills in their review study with 117 studies. However, when they limited their studies from 2001 to 2010, they found Reading and Writing skills to be the mostly studied skills. In his review study focusing on the technologies used in language skills teaching in four major English-language journals focusing on CALL in the time period 2001–2005, Stockwell (2007) discovered a great variety in the technologies used for different language skills and areas of interest with the reasons for using the determined technology as well. He also found Vocabulary to be the most studied language skill, followed by Grammar, but Speaking was the least studied skill, despite the increase in the-five-year period, which he attributed to the development of such SCMC technologies as chat and audio-conferencing, and to ASR technologies. That Speaking found to be the second most studied language skill in our reviewed articles may represent the recent further advancements in those technologies.

Despite the inexistence of an underlying theory in 22% of the studies, the rest were theoretically based on various theories for CALL and learning, and TAM was the most often used model for getting learners’ perceptions of the technology used. Liu et al (2002) reviewed 70 studies and reported only 33 providing theoretical underpinnings for their research. Our review study provided a detailed list of the theories considered in CALL studies providing insights to future researchers as for the theories to consider while structuring their studies, which is the original part of this study. The studies in the review were done in different timespans: 22% were completed within the same day and 22% in one month. The fact that only one study continued over one year may indicate the difficulty in long-term attainment of the technological context used.

The most common sample group in the available data was university students followed by fewer number of studies at K-12 level and only one study with graduates, the reason of which as indicated by Turan & Akdag-Cimen (2020) may be easier access opportunities to the participants with higher self-regulation abilities. The greater number of the studies in the data analysed (n=27, 75%) included less than 100 participants, which may be due to the technological contexts used in the studies. The most popular sample size in the review study of Macaro et al. (2012) was 21–30 participants (14%) which was 19% in this study, >50-100 being the highest number of participants (39%). The reason of this difference is not clear and needs to be addressed in future studies. As for the study places, although Taiwan was observed to be the country studying empirical CALL most frequently in the defined time period, the most commonly studied target language was English, which was confirmed by other review studies (Shadiev & Yang, 2020; Ghanizadeh et al., 2015; Shadiev et al., 2017), which is attributed to English being the most popular and widely spoken language in the world.

The available data were mostly gathered with pre-test post-test design (77%) assisted by questionnaires (50%), which is related to the nature of the empirical studies. The data obtained were mostly analysed with t-test and ANOVA, some using more than one analysis strategy for more detailed data.

The results obtained from the available articles showed that independent variables studied were effective on the dependent variables in almost all of the studies and yielded positive results in terms of learner perspectives and attitudes. Similar to the results in this systematic review study, Golonka et al. (2014) found in their review study that learners enjoyed using various technologies in FL learning and preferred them over conventional methods and materials by getting more engaged in the process of learning to form more positive attitudes towards learning and the technology used. Liu et al. (2002) reported two studies which focused entirely on students’ attitudes. However, in this review study no
study entirely focused on student attitudes because only empirical studies, which examined at least one skill, were considered in this study. Liu et al. (2002) also reported that there was not convincing evidence on the use of computer technologies to improve language skills in all areas and that the majority of their reviewed articles yielded positive attitudes towards the technological context used. The studies in this study mostly confirmed the impact of technological tools on the examined skills or factors. For example, under the category of Web 2.0 tools, online summary writing-pal (SW-PAL) by Chew et al. (2020) was used to improve participants’ writing skills and it was found particularly useful in enhancing the examined writing skill along with student perspectives about the tool both as motivating and challenging. Lwo and Lin (2012) used different captions in multimedia L2 learning with respect to vocabulary acquisition and reading comprehension and found that participants relied on graphics and animation as an important tool for understanding English sentences. Chang et al. (2018) used outdoor u-learning mobile technology for improving English listening and examining its cognitive load effect and concluded that the mobile technology used provided better English listening achievement and that there was a significantly negative relationship between English listening achievement and extraneous cognitive load. Chang and Hsu (2011) used educational software which is a CALL system on personal digital assistants (PDAs) for improving reading skills and concluded that grouped students had better reading comprehension than individual students and the participants perceived the system as easy to use and were satisfied with it. Overall, the studies reviewed here mostly indicated evidence on the efficacy of the technology used and all reported positive attitudes and perceptions by the participants. (For a summary of the reviewed studies, please contact the authors.).

**Conclusion**

Previous systematic review studies on CALL provided valuable insights for the field of language learning by preparing a base for the current CALL methodology. This study also aims to present a comprehensive analysis of empirical CALL studies in foreign language learning by revealing the common trends as for technologies used, educational context, variables, underlying theories, participants, target languages, study time/duration/countries, skills/factors analysed, data collection and analysis strategies, and displaying the results of the studies.

The results here revealed a wide range of technologies and technological tools used for foreign language learning, mostly exploring the development and achievement in vocabulary, and speaking skills of university students in particular. The diverse range of technologies, content delivery platforms, and methods of presentation varied greatly compatible with the technological advancements in the new millennia. There are no common theoretical bases or fixed statistical analysis methods identified in the available data. The results of the studies mostly indicate the potential benefit and efficacy of the technological context used on foreign language learning.

This study provides an important reference base in the field of CALL; however, it has some limitations, like all studies. First, the current review is limited to the articles in the category of Topic in Web of Science with defined search terms. Future studies may focus on a wide range of databases in wider categories with more detailed defined terms. Second, the focused timespan of this study is 2000 to May 2020, which could be broadened to a larger date range in future studies to pave way for the comparison of the trends from the beginning of CALL studies to date. Third, this review displays the results of the available data without making comparisons. Future studies may compare the results of the studies with each other based on the skills studied to find the most efficient technological learning context for foreign language development.
References


Learning and Teaching (IJCALLT), 10(2), 83–95.  
http://dx.doi.org/10.4018/IJCALLT.2020040105

https://doi.org/10.1017/S0958344013000013

https://doi.org/10.1177/0363546513497567

https://doi.org/10.1080/0363546513497567

https://doi.org/10.1080/09588220500442632

https://doi.org/10.1007/s10639-015-9428-y

https://doi.org/10.1080/09588220903467335

https://doi.org/10.1080/15391523.2002.10782348

https://doi.org/10.3390/sym12050701

https://doi.org/10.1017/S0261444811000395

https://doi.org/10.18052/www.scipress.com/ILSHS.11.57

https://doi.org/10.18052/www.scipress.com/ILSHS.9.49

https://doi.org/10.7326/0003-4819-151-4-200908180-00135

https://doi.org/10.1017/S095834401100005X

https://doi.org/10.1080/09588220802447651


Appendix

Studies included in the systematic review


