



Please cite this paper as follows:

Mashhadi, A., Dehghani, H., & Bárcena, E. (2025). A systematic review of empirical LMOOC publications (2014-2021). *Journal of Research in Applied Linguistics*, 16(1), 90-104. <https://doi.org/10.22055/RALS.2023.44703.3138>

Research Paper

A Systematic Review of Empirical LMOOC Publications (2014-2021)

Amir Mashhadi¹, Hoorá Dehghani², & Elena Bárcena³

¹Corresponding Author, Department of English Language and Literature, Shahid Chamran University of Ahvaz, Ahvaz, Iran; mashhadi.scu@gmail.com

²Islamic Azad University, North Tehran Branch, Tehran, Iran; houra.dehghani@gmail.com

³Department of English Language, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain; mbarcelona@flog.uned.es

Received: 03/09/2023

Accepted: 22/05/2024

Abstract

This paper presents the results of a systemic review of empirical Language Massive Open Online Course (LMOOC) publications from 2014 to 2021, by identifying their research foci, types of technology platforms, target language skills/components, pros and cons, as well as future research trends through 4 databases, namely Google Scholar, ERIC, Taylor & Francis, and Scopus. A descriptive statistical approach was used to conduct a systemic analysis of 60 empirical LMOOC studies including peer-reviewed articles and book chapters. Thirty-two empirical LMOOC conference papers were also added to expand the qualitative data of the research using thematic data analysis. In the empirical LMOOCs analyzed, Spain and China were found to be the most prolific countries in terms of publications. Most publications were journal articles, mostly published in Computer Assisted Language Learning, among others. From institutions publishing empirical LMOOCs, Universidad Nacional de Educación a Distancia in Spain had the highest number of publications. With most publications having, at least, 2 authors, these studies were collaborative in nature. Considering their research foci and analysis methods, most studies were student and design-focused using mixed-methods approaches. As for the data collection methods, descriptive statistics were used the most frequently, followed by inferential statistics, content analysis, and social network. Audiovisual tools stood out among other technology platforms and almost half of the studies focused on the overall language proficiency. Implications of findings for language education are discussed in terms of the pros and cons and future trends of empirical LMOOC studies.

Keywords: Empirical LMOOCs; Systematic Review; Research Focus; Technology; Target Language Skills; Pros and Cons; Future Trends.

1. Introduction

Massive Open Online Courses (MOOCs) are the logical progression of Open Educational Resources (OERs), which are publicly available learning resources and media for use in teaching, assessment, and learning (Bárcena & Martín-Monje, 2014). Since their launch over a decade ago (Downes, 2008), MOOCs have become the fastest growing distance learning modality due to the accessibility and flexibility they provide, particularly for those who wish to maintain education outside formal education environments (Panagiotidis, 2019). The field of language teaching and learning, among others, is no exception, and a growing number of online courses with unrestricted access and potentially unlimited participation has been developed to meet the needs of language learners. Language Massive Open Online Courses (henceforth, LMOOCs [Bárcena & Martín-Monje, 2014]) contain audiovisual materials designed to be accessed online or downloaded for convenience, particularly in nonacademic environments (Beirne et al., 2017). These free online courses are also developed for a large number of consumers who want to learn, practice, or improve a language without the one-on-one help of a tutor (Ding & Shen, 2020). They are available on a variety of platforms, ranging from large multinationals to smaller local LMOOC providers. In response to the increasing attention to the potentials of this educational modality (Perifanou & Economides, 2014), the number of research studies on empirical MOOCs in the field of foreign language education is expanding (del Peral Pérez & Castrillo de Larreta-Azelain, 2023; Hashemifardnia et al., 2021).



Language learning through LMOOCs is reported to be demanding for students due to the overabundance of resources available, causing concentration loss (Zeng et al., 2020). In addition, LMOOCs have long been chastised for low completion rates (Liyanagunawardena et al., 2013; Li-Wei Hsu, 2021); inefficient student support (Alario-Hoyos et al., 2015); lack of a business model for the institutions; unsolved tension between structure and autonomy, and rudimentary interfaces due to lack of investment (Bárcena & Martín-Monje, 2014). These shortcomings still remain, and "more empirical evidence is needed in order to provide suggestions for the practice of using LMOOCs in language learning and teaching" (Li-Wei Hsu, 2021, p. 3).

In response to the growing need for empirical evidence in the general field of MOOCs, several systemic review studies have been carried out in recent years. For instance, Liyanagunawardena et al. (2013) characterized MOOC studies from 2008 to 2012 into introductory, concept, case studies, educational theory, technology, participant-focused, provider-focused, and others, while also offering a quantitative breakdown of MOOC publications by type, publication year, and authors. In another systematic review, Veletsianos and Shepherdson (2016) studied the research components, citations, and applied methods of MOOC studies from 2013 to 2015. Zhu et al. (2018) also explored the topics and methodologies of empirical MOOC studies from 2014 to 2016. Zhu et al. (2018) argued that although MOOC studies were mainly conceptual at the outset, more empirical studies have emerged recently.

There are only a few systematic reviews of MOOC literature in the field of foreign language teaching and learning (Rinatovna et al., 2017). For instance, Drakidou et al. (2019) looked into the latest discoveries in three Technologically-Enhanced Language Learning (TELL) areas, namely Mobile-Assisted Language Learning (MALL), LMOOCs, and Social and Open Language Learning (SOLL) in order to gather data on the target languages, geographical distribution of publications, and the occurrence of terms showing up in the keywords to discover patterns, trends, and gaps across all the three TELL domains. Furthermore, Palacios Hidalgo et al. (2020) investigated the origin and definition of MOOCs, types and platforms, advantages and limits, and the results of systematic educational reviews on MOOC application in foreign language learning.

The most recent systemic review that focused exclusively on LMOOCs was performed by Sallam et al. (2020), in which types, topics, and trends of the existing studies on LMOOCs were reviewed systematically. This research contributes to the existing literature by examining empirical LMOOCs, with a focus on various aspects, including research foci, technology platforms, types of target language skills/components, future research trends, and the pros and cons in language education. Given the increasing demand for LMOOCs around the world and a lack of comprehensive empirical evidence of LMOOC studies as a foundation for future applications (Li-Wei Hsu, 2021; Sallam et al., 2020), the present study aimed at conducting a systematic review of 60 empirical LMOOC publications (i.e., peer-reviewed articles and book chapters) from 2014 to 2021 in order to shed light on the geographical distribution, authors, publishers, participants, research foci, technology platforms, and target language skills in these studies. This study also considered qualitative data from 92 publications including research papers, book chapters, and conference papers to investigate the pros and cons of LMOOCs and their future trends and common themes in response to the periodic need to undertake such a type of study in the community of LMOOC researchers and practitioners. Thus, in pursuit of these goals, the following research questions were formulated:

1. What is the outline of the geographical distribution, authors, publications, and participants in the empirical LMOOC studies from 2014 to 2021?
2. What are the research foci in these studies?
3. What types of technology platforms are used therein?
4. What data collection instruments and analysis methods are used in these studies?
5. What types of target language skills/components are mostly considered?
6. What are the pros and cons of LMOOCs for language education?
7. What are the future research trends in empirical LMOOC studies?

2. Methodology

2.1. Procedure

This study aimed to analyze empirical LMOOC studies from the inception of empirical LMOOCs in 2014 up to the year 2021, guided by the research questions and variables. Initially, the search framework and keywords were specified as follows:

- The research team searched through various reliable sources to find LMOOC-related empirical references. They persisted in doing this until they had elicited all known empirical LMOOCs.
- Whereas the first LMOOC was introduced in 2012 (Read & Barcena, 2014), the research studies that matched the criteria for this study were published between 2014 and 2021.
- Except for chaining from known sources, the Google Scholar, ERIC, Taylor & Francis, and Scopus databases were searched for data collection to provide the most thorough and comprehensive review possible.
- The search terms for screening the titles and abstracts of empirical research on LMOOCs included language MOOCs, LMOOCs, MOOCs and language learning, language MOOCs and English learning.

Given the aims of the study, certain criteria were considered for empirical research on LMOOCs: Firstly, the quantitative data were collected from 60 peer-reviewed empirical journal articles, articles published by official institutions, and book chapters. In order to obtain further qualitative data, the authors supplemented the existing sources with 32 conference papers on empirical LMOOCs. In order to target high-quality innovative empirical research, the quantitative and qualitative data analysis sections of this study excluded nonempirical journal and conference papers, review papers, book reviews, and reports. Secondly, the search was continued for English-language articles that (a) investigated MOOCs in language classrooms and (b) provided information about the participants, research methods, learning activities, tools, and outcomes. Subsequently, studies that were not centered on language education but rather on different subjects such as programming, physics, mathematics, and so on were excluded.

To improve efficiency, the article collection and analysis workload were distributed among the authors of this article. Each of them retrieved articles by screening the topics and abstracts of LMOOC studies using the abovementioned search keywords within the given time span and considering the two research criteria. The research team initially identified 1,152 research studies, which were reduced to 863 items following the elimination of duplicates. After excluding nonempirical studies, review papers, book reviews, and reports based on the first criterion, 239 studies remained. In the next step, the researchers of this study excluded 91 studies that were not primarily about LMOOCs, adhering to the second criterion. As a result, 56 papers were also omitted as they did not provide sufficient empirical data, leaving 60 and 92 studies available for quantitative and qualitative analyses, respectively.

For the quantitative analysis, the authors analyzed the studies (i.e., 60 peer-reviewed empirical journal articles and book chapters) individually by outlining their research foci, research methods, and target language skills/components. They also provided an analysis of the publications on LMOOCs in terms of their publication year, citation number, geographical distribution, type of publication (i.e., book chapters, journal articles, etc.), and target languages. The coded information was, then, exchanged, double-checked, and recoded and the items that were not in consensus were discussed until an agreement was reached. The same process was replicated for the coding and thematic analysis of the qualitative data.

2.2 Data Analysis

The following data analysis procedures were applied to answer the research questions: In order to answer the first research question 1 (RQ1), the countries of origin or delivery of the LMOOCs in the 60 empirical studies (i.e., peer-reviewed journal articles and book chapters) that were published in various journals between 2014 and 2021 were determined. Moreover, the author affiliations of all the studies along with the locations of only the first authors were looked up to see where they were from. Having investigated the first authors' affiliations, the authors identified the five most frequented institutions or universities. The studies were also examined in terms of the number of authors to find out

who the single authors of the studies were, as well as how much collaboration there was among the authors in the domain of empirical LMOOC studies (see Table 1).

In this study, the first five researchers who published the most papers and also the first five researchers whose publications were cited the most on Google Scholar are listed. The participants in the empirical LMOOC studies are also classified as university students, language learners with various backgrounds, language learners from the same educational context, teachers, school students, and refugees and migrants. Furthermore, the distribution of the selected papers within the aforementioned databases, the proportion of book chapters vs. journal articles, the journals with the most empirical LMOOCs published, and the years with the most empirical LMOOCs published are all provided.

Given their relevance to the research objectives in this study, the authors relied on the research themes used in Zhu et al. (2018) to respond to the research question 2 (RQ2) about the research foci of the empirical LMOOC studies. In effect, one of the authors classified the themes into the five groups as follows: (1) student-focused, (2) teacher-focused, (3) design-focused, (4) context and impact, and (5) other. Next, the analysis was evaluated by the other authors. If the second and the third authors disagreed with the decisions made by the first one, they would discuss the discrepancies until an agreement was reached. To respond to the research question 3 (RQ3), the authors classified the types of technology platforms used, namely audiovisual tools, discussion tools, writing tools, online reading tools, material recording, revision activities, social networks, and others.

As for research question 4 (RQ4) on the research methodologies often used, the empirical LMOOCs were grouped into three broad categories identified by Creswell (2003) as quantitative, qualitative, and mixed-methods studies. The authors also took into account the data collection instruments and the data analysis methods in order to have a better grasp of the methodologies adopted in the studies. According to Zhu et al. (2018), the data collection instruments were categorized as interviews, surveys, focus groups, tests, observations, discussion forums, platform databases, and learning analytics. The data analysis methods were, in turn, sorted as descriptive statistics, inferential statistics, content analysis, social network analysis, constant comparative method, grounded approach analysis, collaborative autoethnography, phenomenological analysis, and SWOT analysis, and so on. To assess the most commonly used data collection instruments and data analysis methods, the authors gathered the related information from the studies and classified them based on their occurrences. The same cross-check approach as indicated in the data analysis technique for RQ2 was likewise employed.

In response to (RQ5), the authors examined the related information regarding the pros and cons of LMOOCs for language education in the results and conclusion sections of the studies and classified the prevalent themes based on their frequency of occurrence. Tables 7 and 8 represent the summary of the advantages and disadvantages by frequency. Responding to RQ6, the authors categorized the prospective paths suggested in the peer-reviewed journal articles, book chapters, and conference papers published in the predefined period and identified those categories with the most publications in order to suggest future trends in which empirical LMOOC research may advance.

3. Results

3.1 Outline of the Geographical Distribution, Authors, Publications, and Participants in the Empirical LMOOC Studies From 2014 to 2021

The first countries of origin or delivery of the LMOOCs in the 60 studies were Spain, China, the UK, and the USA, followed by Ireland, Taiwan, and Turkey as represented in Figure 1. According to LMOOC authors' affiliations, the most common country locations in the 60 studies were China ($n = 14$), Spain ($n = 13$), Ireland ($n = 8$), the United States ($n = 7$), the United Kingdom, and Turkey ($n = 4$), whereas the first authors were most commonly located in Spain ($n = 13$), China ($n = 10$), the United Kingdom ($n = 8$), and the United States ($n = 5$), followed by Taiwan, Ireland, and Turkey ($n = 4$):

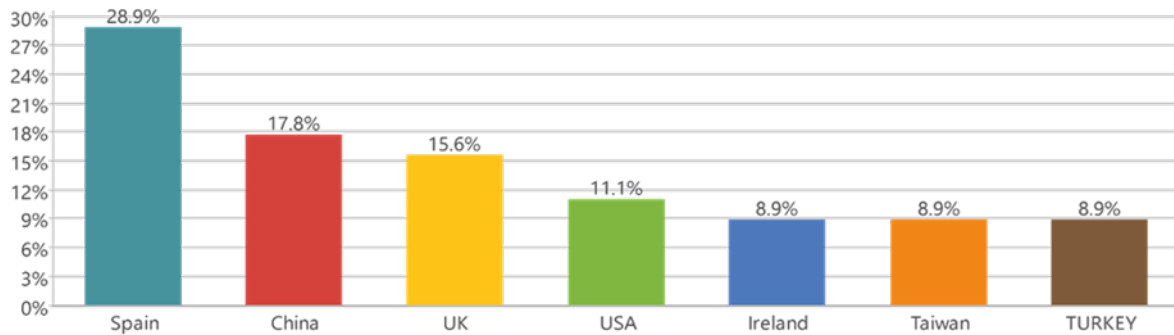


Figure 1. Geographical Distribution of the LMOOCs

In the case of the first authors' institutions publishing empirical LMOOCs, Universidad Nacional de Educación a Distancia ($n = 10$), The Open University ($n = 4$), Coventry University ($n = 4$), Dublin City University ($n = 4$), and the University of Iceland ($n = 3$), followed by the Human-Computer Interaction Institute, University of Science and Technology, Eskişehir Osmangazi University, and the University of Hospitality and Tourism, with two publications each, released the highest number of publications.

When it comes to the empirical LMOOC researchers (the first authors and coauthors), Conchur Mac Lochlainn, Elaine Beirne, and Mairead Nic Giolla Mhichil were the researchers with the highest number of publications, each with four publications, followed by Elena Martin-Monje, Kolbrun Fridriksdottir, Ruben Chacon-Beltran, and Timothy Read, with three publications and Amy Ogan, Beatriz Sedano, Clare Wright, Elena Barcena, Evelyn Yarzebinski, Jessica Hammer, Judith Odili Uchidiuno, Maria Dolores Castrillo, Maria Orisini-Jones, Muhammad Ozgur Yasar, and Zsuzanna Barkanyi with two publications came next. These 18 authors indeed made up 17% of all authors and were involved in 34% of all publications. The remaining researchers had only one publication. In the 36.67% of studies, more than two authors were involved. Following that, 33.33% of studies had two authors, while only 30% had one, indicating the prominence of collaboration and teamwork in this burgeoning discipline. In terms of citations, Alix Creuze, Tatiana Codreanu, and Tita Beaven ($n = 134$), Elena Martin-Monje ($n = 75$), Clare Wright, Jun Wang, Maria Dolores Castrillo ($n = 53$), and Na An ($n = 52$), whose total citations accounted for 25.2% of the rest of the examined researchers, were the most prolific ones.

As regards participants, university students made up 37.5% of those who took part in the empirical LMOOCs, followed by language learners from various backgrounds, teachers, refugees and migrants, language learners from a same context, and school students (see Figure 2):

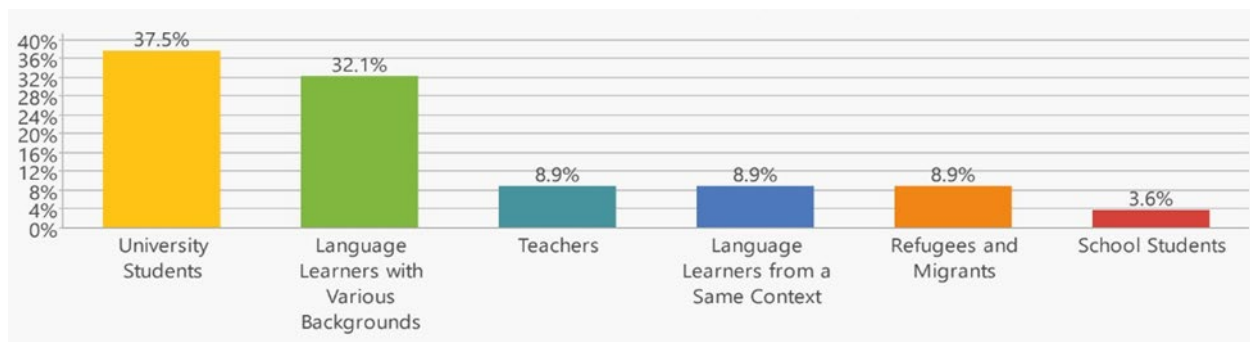


Figure 2. Education Levels (Status) of Participants

From the databases publishing the 60 empirical LMOOCs, the first place belonged to Google Scholar and Eric, with an equally distributed number of publications. Scopus and Taylor & Francis (8.3%) came in the second and third place, respectively (see Figure 3). The identified articles were published in a variety of journals, and the number of studies published in journals ($n = 50$) was bigger than the number of book chapters ($n = 10$). Among the journals, Computer

Assisted Language Learning published the highest number of empirical LMOOC studies, followed by ReCALL ($n = 4$). The remaining journals, including International Artificial Intelligence in Education Society, International Journal of Distance Education Technologies, and Interactive Learning Environments, all had the same number of publications ($n = 3$) during the given time span (see Table 1). In terms of publishing year distribution, according to the data, no study was published from 2009 to 2013. However, the number of research articles had a slight increase from 2014 to 2016, indicating that some researchers began to place more emphasis on studying empirical LMOOCs. In 2016, there were, in turn, three studies, which increased to nine in 2017. This increase in the number of publications suggested that more researchers were getting interested in this area, and that empirical LMOOCs generally followed a similar trend to LMOOCs, as based on Fang et al. (2019), the number of LMOOCs rose abruptly from 2016 to 2018 after a relatively constant number of LMOOCs from 2014 to 2016 (see Table 2):

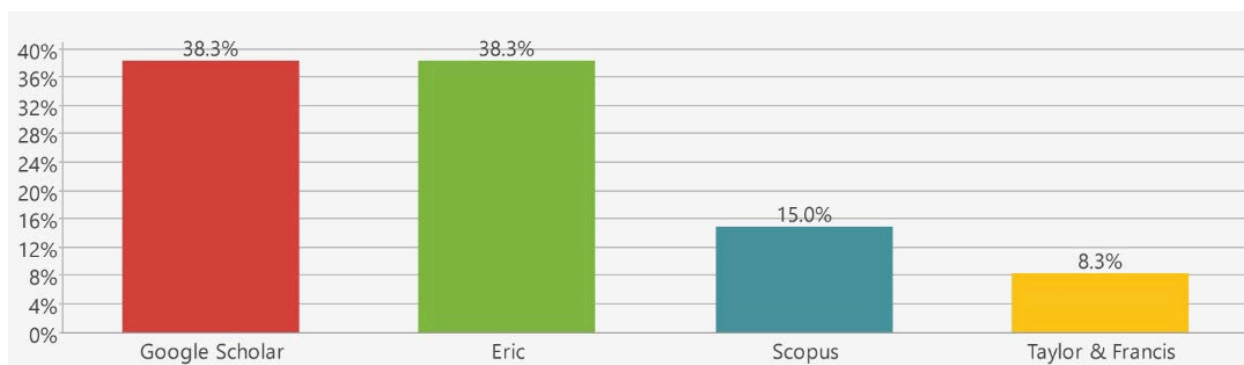


Figure 3. Databases of the Selected Publications

Table 1. *The Five Most Frequent Empirical LMOOC Publications*

No.	Publications	Frequency	Percentage
1	Book Section	10	16.67%
2	Computer Assisted Language Learning	8	13.3%
3	ReCALL	4	6.67%
4	International Artificial Intelligence in Education Society	3	5%
5	Journal of Interactive Media in Education	3	5%
6	International Journal of Distance Education Technologies	3	5%
7	Interactive Learning Environments	3	5%

Table 2. *Distribution of Empirical LMOOCs by Year (2014-2021)*

Year	Number of Publications	Percentage
2014	2	3.33%
2015	1	1.67%
2016	3	5%
2017	9	15%
2018	10	16.67%
2019	10	16.67%
2020	13	21.67%
2021	12	20%

3.2. The Research Foci in the Empirical LMOOC Studies

The empirical studies on LMOOCs were organized into five groups: student-focused ($n = 25$), design-focused ($n = 20$), context and impact ($n = 11$), instructor-focused ($n = 4$). Based on the findings of the systematic analysis of empirical LMOOCs, students were the focus of 41.6% of studies covering areas such as student behaviors, perceptions, motivation, satisfaction, emotions, performance, interaction, engagement, and retention. The second most prevalent type of empirical LMOOC research was focused on course design (33.3%), with topics including structural aspects of

LMOOCs, task design, prediction of knowledge scales, different learning models (flipped, blended, etc.), development of adaptive self-assessment activities, automatic text analysis, as well as design, creation, and implementation of courses for different cultural groups.

In the empirical LMOOC investigations, studies focusing on the context and impact of LMOOCs came in third (18.3%). Topics such as the flexibility or effectiveness of LMOOCs, LMOOC implementation in higher education, and the influence of integrating MOOCs into language learning were employed more frequently in such investigations. Unlike the three abovementioned research areas, only a small percentage of studies (6.6%) were instructor-focused, considering the topics related to instructors, such as instructors' experience and achievements in developing and running LMOOCs, their beliefs and perceptions regarding LMOOCs, and their roles in LMOOCs. The aforementioned categorization of results in terms of research foci, is in response to the urgent need to address the impediments to participants' LMOOC completion, and focus more on the implementation of such courses as examined through various teaching and learning theories.

3.3. Types of Technologies in the Empirical LMOOC Studies

Among the technologies, audiovisual tools were the most used ones, followed by discussion tools, revision activities and writing tools. Other categories, online reading tools, social networks, and material recording tools were considered less frequently in the studies examined (see Table 3). It seems that as the field evolved, MOOCs distanced themselves more and more from standard online courses and became multimodal, and nowadays, based on the results of this study, it can be said that LMOOCs have their own methodologies and modalities:

Table 3. *Types of Technologies in the Empirical LMOOC Studies*

Types of Technologies	Frequency	Percentent
Audiovisual Tools	30	29.1%
Discussion Tools	22	21.4%
Revision Activities	13	12.6%
Writing Tool	10	9.7%
Not Mentioned	9	8.7%
Online Reading Tool	8	7.8%
Social Networks	7	6.8%
Material Recording	4	3.9%

This study divided the research methods of LMOOC studies into three categories: quantitative, qualitative, and mixed-methods. In terms of the research methods utilized in the 60 studies, mixed-methods were the most prevalent method ($n = 32$); 17 studies were quantitative, and the remaining 11 studies were qualitative. Different research foci and research methods were, likewise, reflected differently in LMOOCs. Student-focused studies accounted for 43.57% of mixed-methods studies, while design-focused studies (21.88%), context-focused studies (18.75%), and other categories (12.5%) came next. In turn, teacher-focused studies accounted for the smallest percentage of mixed-methods studies. Qualitative studies, on the other hand, were more design-focused (35.29%), while student-focused and context-focused qualitative LMOOCs were represented equally (23.53%). The remaining categories were other (11.76%) and teacher-focused (5.88%). In quantitative studies, student-focused studies accounted for 36.36%, followed by design-focused (27.27%), teacher-focused (18.18%), context-focused (0.09%) and other studies (9.09%), respectively. The results imply that the domain of empirical LMOOCs is in need of more qualitative evidence within either qualitative or mixed-methods studies with more focus on teachers.

3.4. Data Collection Instruments and Analysis Methods in the Empirical LMOOC Studies

Considering the data collection instruments, survey approaches were the most commonly used data gathering instrument, followed by platform databases, interviews, quiz or test results, discussion forums, observation, focus group interviews, learning analytics, and other techniques, respectively (see Figure 4). The surveys are frequently used in conjunction with one (20.51%), two (46.15%), and more than two (33.33%) data collection instruments and are only used as the sole data collection instrument in 26.53% of the studies. The surveys are, in turn, used more frequently with interviews (24.49%) and platform databases (20.41%) than other modes of data collection. Moreover, 26.32% of the empirical LMOOCs used platform databases with one and just over half of them (52.63%) used platform databases with

two other modes of data collection. Being used alone in 36% of the studies, platforms are most often used with surveys in most cases (40%). Whereas a considerable percentage of interviews are used with more than two other means of data collection (47.06%), a small percentage is used with one (23.53%) or two other instruments (29.41%). Interviews are often accompanied by surveys (50%), and only 12.5% of the empirical LMOOCs used interviews as the sole data collection instrument.

Having analyzed the data collection methods, the authors probed the data analysis methods in the 60 studies. Among the data analysis methods, descriptive statistics were used the most frequently, followed by inferential statistics, content analysis, and social network analysis, respectively (see Figure 5):

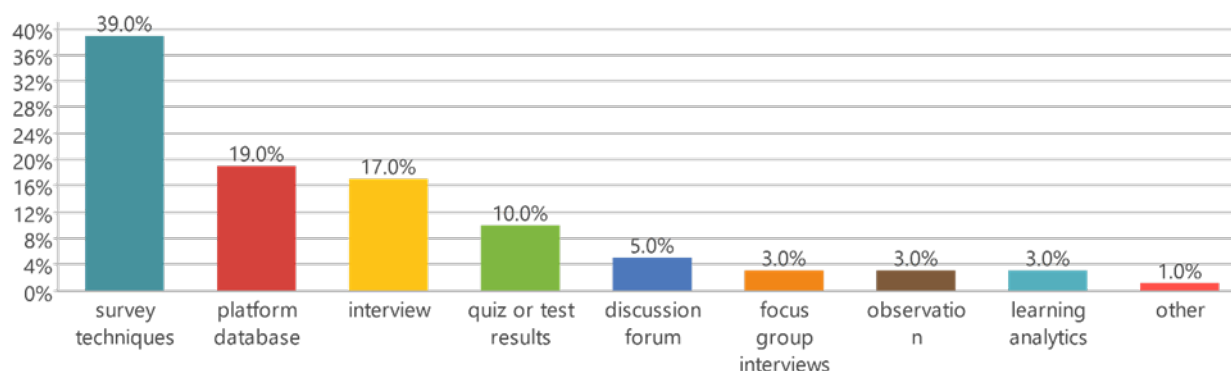


Figure 4. Data Collection Instruments of the Empirical LMOOC Studies

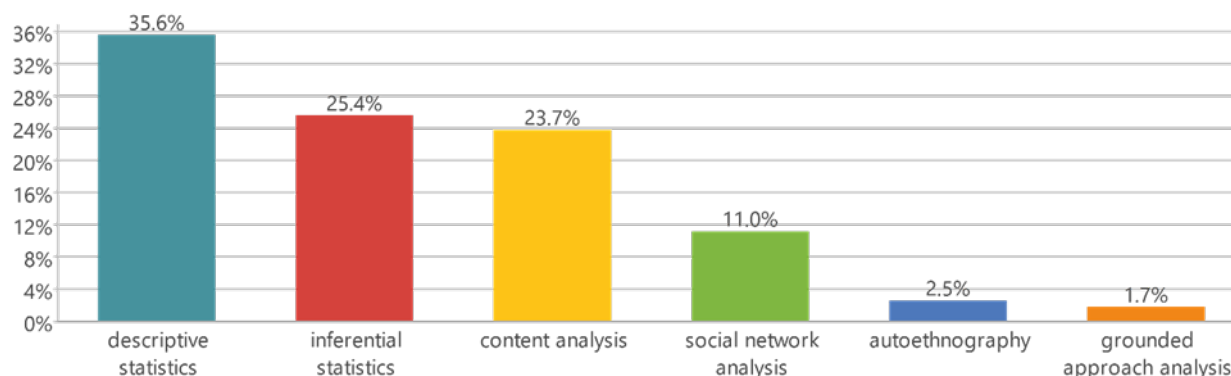


Figure 5. Data Analysis Methods

3.5. Target Language Skills/Components in the Empirical LMOOC Studies

As shown in Table 4, among the target language skills/components, nearly half of the empirical LMOOCs focused on the overall language proficiency (multiple skills), with writing coming in the second place. The remaining target language skills/components were considered to be speaking, reading, vocabulary, and listening based on their frequency. The results indicate that speakers need to use a repertoire or range of skills both for occupational and social purposes. Within the individual skills, writing is an easier goal than orality in technology-supported environments. Conversely, speaking, which is a productive skill, appears before reading and listening, which are receptive skills, and therefore are usually more prominent in a learner's interlanguage. On the other hand, the productive skills, such as writing and speaking, may present more of a challenge as they require more feedback and are particularly difficult to practice through LMOOCs, as it happens with other forms of online language learning:

Table 4. *Frequency of the Target Language Learning Skills/Components in the Empirical LMOOC Studies*

No.	Language Learning Skills/Components	Frequency	Percentage
1	Multiple Skills	40	47.6
2	Writing	13	15.5
3	Speaking	11	13.1
4	Reading	7	8.3
5	Vocabulary	7	8.3
6	Listening	6	7.1

3.6. Pros and Cons of LMOOCs for Language Education

In order to answer RQ5, the authors investigated the advantages and disadvantages of LMOOCs for language education based on their frequency of occurrence in the results and conclusion sections of the studies analyzed. They are, in turn, categorized into more general themes such as sample-based, design-based, instruction-based, and others (see Tables 5 and 6):

Table 5. *The Most Cited Advantages for LMOOCs*

No.	Advantages	General Themes	Frequency	Percentage of Analyzed Publications
1	Improving learners' engagement and interaction with others (face-to-face or Computer Mediated Communication)	(Sample-based)	42	15.97
1	Improving learners' knowledge and skills	(Sample-based)	42	15.97
2	Using convenient and easy-to-use platforms (to include more functionalities, multi-modal materials, convenient evaluation/testing, and flexibility)	(Design-based)	38	14.45
3	Reflecting on students' satisfaction with the applicable technology or approach	(Sample-based)	29	11.03
4	Developing autonomy (self-regulated learning)	(Sample-based)	23	8.75
5	Increasing motivation (sample-based)	(Sample-based)	20	7.6
6	Learning from others (peers, instructors, etc.)	(Instruction-based)	16	6.08
6	Exhibiting openness (allowing everyone to attend and have easy access to extensive data)	(Design-based)	16	6.08
7	Increasing the rate of course completion	(Sample-based)	14	5.32
8	Enhancing future prospects (job certificate, etc.)	(other)	8	3.04
9	Reducing affective filters (i.e., anxiety, stress, lack of confidence, etc.)	(Sample-based)	7	2.66
10	Improving technological literacy (ICT/digital literacy)	(Design-based)	5	1.9
11	Being free of charge	(Sample-based)	3	1.14

Table 6. *The Most Cited Disadvantages for LMOOCs*

No.	Disadvantages	General Themes	Frequency	Percentage of Analyzed Publications
1	Lack of participation, interaction, connection, and feedback from peers and teachers	(Sample-based)	21	23.89
2	Inefficient design	(Design-based)	14	15.91
3	Technical problems (log in issues, low speed, etc.)	(Design-based)	13	14.77
4	High dropouts	(Sample-based)	9	10.23
5	Motivation loss	(Sample-based)	7	7.95

5	Technology illiteracy	(Instruction-based)	7	7.95
6	Lack of sense of intimacy with peers or instructors (cultural mismatch, etc.)	(Sample-based)	5	5.68
7	Lack of time	(Design-based)	4	4.55
7	Negative feelings; loss of self-confidence, and anxiety, etc.	(Sample-based)	4	4.55
8	Unaffordable costs (certificates, laptop, etc.)	(Design-based)	2	2.27
9	Inefficient method of testing/evaluation	(Instruction-based)	1	1.14
9	Inadequate accessibility for individuals with impairments	(Sample-based)	1	1.14

3.7. The Future Research Trends in the Empirical LMOOC Studies

The suggested future research directions in the empirical LMOOCs are grouped into 11 themes to address RQ6. Each theme was, in turn, assigned a general category as sample-based, instruction-based, reliability-based, and result-based. The frequency of each category was, then, counted to infer the significance of each theme. The themes presented in Table 7 emerged as recurrent in the studies examined:

Table 7. *The Most Frequent Future Research Trends in the Empirical LMOOCs*

No.	Future Research Trends	General Themes	Frequency	Percentage of Analyzed Publications
1	Focusing on the learners, their motivation, and experience throughout the LMOOCs to reshape the design of the courses (Sallam et al., 2018)	(Sample-based)	21	25.3
2	Improving instructional design of the course	(Instruction based)	12	14.16
3	Conducting studies with larger data and complementing them with interviews, logs, course analytics, self-reports, etc.	(Reliability-based)	10	12.05
4	Applying the same method or technology to other skills, sub-skills, and contexts.	(Reliability-based)	9	10.84
5	Arguing or comparing different approaches (e.g., blended vs. flipped) and their impact on language learning and retention	(Result-based)	8	9.64
6	Replicating the same study or doing a follow-up investigation with the same design to make sure of the stability of the LMOOCs	(Reliability-based)	6	7.23
7	Researching new ways to facilitate large-scale education in further contexts	(Reliability-based)	5	6.02
8	Applying more advanced approaches, platforms, or technologies to improve language teaching or learning	(Result-focused)	4	4.82
9	Investigating teachers' beliefs, styles, behavior, strategies, and performance, etc.	(Sample-based)	4	4.82
10	Stating the significance of social learning in LMOOCs and how it should be further investigated to establish efficient communication among participants, which will open up a new research area for language MOOC design.	(Instruction-based)	3	3.61
11	Attempting to overcome the xMOOCs/cMOOCs dichotomy and find a model that is most suitable for language teaching and learning (Sallam et al., 2018)	(Instruction-based).	1	1.2

4. Discussion and Conclusion

This article aimed to provide a better understanding of a number of key aspects of empirical LMOOCs, namely their geographical distribution, authors, publishers, and participants, as well as their foci, types of technology, and target language skills/components. Furthermore, several studies composed of journal articles, book chapters, and conference papers were qualitatively analyzed to identify their common pros and cons, and help the authors infer future trends of empirical LMOOCs in nonformal education. The results show that Spain and China are the most productive countries in terms of empirical LMOOC publications and affiliated first authors, which conforms to the findings of Sallam et al. (2020). Given that the majority of LMOOC researchers are from Spain and China, more contribution from the international community seems necessary so the research findings about the applicability and design issues of LMOOCs can be examined in wider geographical contexts.

Universities are also found to be most active in offering empirical LMOOCs in comparison to other educational institutions. Among the former, Universidad Nacional de Educación a Distancia, trailed by the Open University, Coventry University and Dublin City University had the most publications in that order. Similarly, according to the findings of the research work undertaken, the participant profile of LMOOCs (and the corresponding empirical studies) is highly heterogeneous, although it is led by university students. This can probably be attributed to the digital learning culture of learners born close to or in the 21st century. The younger generations have a variety of choices with regard to study courses and show a distinct inclination towards technology-based learning objects (courses, materials, resources).

Multiple interpretations are plausible for these findings. Firstly, when compared to other educational institutions, universities predominantly exhibit a dual emphasis on both research and teaching. LMOOCs, in general, constitute a relatively novel educational modality in comparison to other formal and nonformal courses, which attracted the early attention of researchers because they conform an extremely challenging educational scenario (the enormous imbalance in the teacher-student ratio; the equally remarkable heterogeneity of the participants in terms of personal and academic profile, strengths and limitations in their capabilities; and also their needs and expectations for the course; the lack of academic and administrative requirements; the basic technology of the platforms and the highly limited resources that institutions are prepared to provide to their LMOOC programs with no apparent immediate return; to name a few).

Secondly, unlike primary and secondary centers that provide compulsory education, tertiary institutions provide noncompulsory education, which means that there is a noticeable interest in attracting students and displaying an attractive academic offer because there is an additional competition between the different universities in a given area. Universities are aware of the latest methodological and technological developments (and, to a large extent, have a leading role in them) and make an effort to be updated and innovative. In fact, LMOOCs have proven to exert a positive contribution towards the social visibility of the excellence of the scholars and courses of the best universities. Thirdly, universities have the implicit role of being academic levelers for students coming from many different secondary schools. LMOOCs can play a fundamental role offering precourse training for those who might not have the necessary background to start a given course comfortably and, at the same time, may offer specialization contents for high-level students on a particular topic which there is no time for during term time or may raise interest due to their advanced profile. Fourthly, LMOOCs can be an affordable and convenient solution to the 21st century demands for lifelong learning, this being a scenario which typically includes second language skills. Fifthly and finally, universities have a social responsibility towards those outside the educational system. Hence, LMOOCs in this sense can be seen as real bridges towards formal education for many human groups who, for various reasons, have been unable to enroll onto university degrees straight after secondary school, cannot have their documentation officially recognized or do not possess it any more (e.g., displaced people, such as refugees and migrants). The very inherent qualities that only LMOOCs offer, namely, the integration of online, open, and free-of-charge training, attract people who do not have access to formal degrees delivered in the standard classroom-based modality. Additionally, according to empirical data, motivation is improved when learners with highly individualistic learning profiles or social anxiety issues are placed on the same platform to learn (Miller & Lu, 2003), although no research has been undertaken to date on the potential behavioral effects that LMOOCs may have on these participants.

Due to these circumstances, it is not surprising that universities boast the largest selection of LMOOCs among the institutions that offer this type of online courses worldwide. Because some of the contributions of LMOOCs are transportable to other educational levels, it follows that, in line with Zhu et al. (2018), it is highly advisable that

practitioners from organizations other than universities become more engaged in the creation of LMOOCs to broaden perspectives on their applicability and design issues for the research community. The authors claim that such diversification will provide valuable insights into LMOOC functionality in varying settings in future applications.

A correlation can be also drawn between the universities to which LMOOC researchers are ascribed (online, distance-learning, distributed across campuses) and the intensity of their dedication. In fact, it can be said that, for various reasons including limited institutional support, LMOOCs are not a major field of research in applied linguistics and a reduced subset of researchers account for a remarkable percentage of the total publication production (although the research presented shows evidence of a clearly rising trend that has regularly gained momentum since 2017). A further feature of LMOOC research is that it is collaborative, as reflected both in the coauthorship of the articles (many of them analyzing the result of previously created courses) and in their acknowledgement sections, which often make reference to team effort. Moreover, although contrary to the findings of Sallam et al. (2020), journal articles are proven to play a more significant role than book chapters in publishing about empirical LMOOCs. This result is probably rooted in the fact that the present study was mostly based on western LMOOC research.

Concerning research foci, most empirical LMOOCs studies focused on the issues related to participants, particularly of a psychopedagogic nature (e.g., student behaviors, perceptions, motivation, satisfaction, interaction, retention, expectancies, self-regulation, profiling, etc.), and course design (e.g., organization and structuring, task design, prediction of knowledge scales, learning models, etc.). Conversely, little attention was paid to contextual considerations (e.g., the flexibility or effectiveness of LMOOCs, LMOOC implementation in higher education, the influence of integrating MOOCs into language learning, etc.), which is coherent with the few data that can be obtained about such issues by most MOOC learning analytic tools, and the scarce information made public by university MOOC programs.

Furthermore, instructor-related issues (e.g., expertise, perceptions, didactic roles, etc.) are minor research topics. This trend can be generalized to research on other educational modalities, that in the last couple of decades has been influenced by a strong transversal learner-centeredness, and it is only in last few years that researchers' attention has turned towards teacher needs. It should be mentioned, however, that the negative results of LMOOCs in terms of academic accomplishment and success in the publications were largely linked to insufficient instruction, which is related to the shifting role of teachers as facilitators who provide motivational support and methodological scaffolding at the expense of academic tutoring and guidance. Research projects where interviewed participants had the opportunity to express their views on the LMOOC undertaken, revealed their frustration when no expert was available to clear up their questions. Therefore, although being self-educated and autonomous is a valuable asset in today's world and, more especially in using LMOOCs for self-study, the research presented revealed that not all participants possessed the necessary aptitudes and skills to acquire knowledge in a self-directed manner (Palacios Hidalgo et al., 2020).

Whereas audiovisual and discussion tools were the most used technologies in empirical LMOOCs, online reading support tools and social networks were found to be the least exploited technologies. Following Vygotsky's (1978) socioconstructivist theory that highlights the significance of interpersonal interaction in learning, insufficient opportunities for teacher-student and student-student interaction were deemed to be a key design feature that hindered knowledge acquisition/construction (Bartalesi-Graf, 2017; Gilliland et al., 2018). Accordingly, LMOOC designers were seen to need to incorporate various social networking and reading features into the creation of such courses to examine their impact on language acquisition. Although there is consensus on the benefits of tools for both reading support and social learning, respectively due to the intellectual impact of literacy skills and to the inherently social nature of disciplines such as a second language, as mentioned above, institutions have not come to terms with the fuzzy business model underlying MOOCs and hence, do not dedicate sufficient funding to the technological and methodological evolution of their MOOC platforms.

Considering the methods underlying empirical LMOOC research, almost one-third were mixed, followed by quantitative and qualitative methods. This needs to be addressed in further related studies (Fang et al., 2019; Zhu et al., 2018), but it is noteworthy that mixed and qualitative approaches to LMOOC research have recently witnessed a sharper increase in number than quantitative ones. This is coherent with researchers' perspective that more in-depth and fine-grained results about the functionality and efficiency of massive courses can be obtained with qualitative data collection instruments, like selected (semi)structured interviews, and forum dynamics observation, particularly for an educational modality with striking dropout rates. However, external pre and postquestionnaires are increasingly frequent, and there is

a recent claim on the part of the empirical LMOOC research community for more sophisticated data analysis approaches like social network analysis and thematic analysis to supplement descriptive statistics and content analysis (Bárkányi, 2021; Jitpaisarnwattana et al., 2021; Read & Martín-Monje, 2021).

Finally, future trends of LMOOCs are generally envisaged in terms of four general categories: sample-based, instructional-based, result-based, and reliability-based studies. However, the most recent studies place more emphasis on replicating the same research parameters in different LMOOC topics and delivery contexts, rather than introducing innovative design features or parameters (Ding & Shen, 2019; Hashemifardnia et al., 2021; Qin & Tan, 2020), building new theoretical frameworks to facilitate online large-scale language education (Hashemifardnia et al., 2021; Hsu, 2021; Meri-Yilan, 2020); and also undertaking contrastive analyses of different types of LMOOCs or of the impact of specific design elements therein in terms of their impact on language learning and long-term assimilation (Hsu, 2021; Yaşar & Polat, 2021).

This article was an attempt to provide an overview of the latest research on empirical LMOOCs. It analyzed the published research work of a reduced but committed community on what has been accomplished in terms of the implementation of Sir John Daniel's old dream of MOOCs as an opportunity for the democratization of knowledge. Regarding second languages, this educational modality reveals a compromise between innovative research and transfer of special social value on the one hand, and real-world institutional and professional practicalities on the other hand. Bibliometric studies like this will be regularly offered as a service to the research community not to lose sight of the overall impact of the plethora of efforts that conform it.

Funding

No funding was received for conducting this study.

Data Availability

The authors confirm that all data collected or analyzed during this study are included in this published article.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- Alario-Hoyos, C., Estévez Ayres, I. M., San Agustín, M. P., Leony Arreaga, D. A., & Delgado Kloos, C. (2015). MyLearningMentor: A mobile app to support learners participating in MOOCs. *Journal of Universal Computer Science*, 21(5), 735-753. <https://doi.org/10.3217/jucs-021-05-0735>
- Bartalesi-Graf, D. (2017). From a small liberal arts college to the world: Our blended courses, SPOC, and MOOCs in Italian studies. In *Proceedings of European Conference on Massive Open Online Courses* (pp. 3-9). Springer. https://doi.org/10.1007/978-3-319-59044-8_1
- Bárcena, E., & Martín-Monje, E. (2014). Introduction. Language MOOCs: An emerging field. In E. Martín-Monje & E. Bárcena (Eds.), *Language MOOCs: Providing learning, transcending boundaries* (pp.1-10). Berlin: De Gruyter Open. <https://doi.org/10.2478/9783110420067.1>
- Bárkányi, Z. (2021). Motivation, self-efficacy beliefs, and speaking anxiety in language MOOCs. *ReCALL*, 33(2), 143-160. <https://doi.org/10.1017/S0958344021000033>
- Beirne, E., Mhichíl, M. N. G., & Cleircín, G. Ó. (2017). LMOOCs, classifying design: Survey findings from LMOOC providers. In K. Borthwick, L. Bradley, & S. Thouésny (Eds.), *CALL in a climate of change: adapting to turbulent global conditions—Short papers from EUROCALL 2017* (pp. 30-34). Research-publishing.net. <https://doi.org/10.14705/rpnet.2017.eurocall2017.686>
- Creswell, J. W. (2003). A framework for design. In J. W. Creswell (Ed.), *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed., pp. 208-227). Thousand Oaks, CA: SAGE Publications.

- del Peral Pérez, J. J., & Castrillo de Larreta-Azelain, M. D. (2023). Empowering teachers in LMOOC design by using a taxonomy of participants' temporal patterns. *Journal of Research in Applied Linguistics*, 14(2), 116-134. <https://doi.org/10.22055/RALS.2023.43579.3048>
- Ding, Y., & Shen, H. Z. (2020). English language MOOCs in China: Learners' perspective. *The EuroCALL Review*, 28(2), 13-22. <https://doi.org/10.4995/eurocall.2020.13090>
- Downes, S. (2008). Places to go: Connectivism & connective knowledge. *Innovate: Journal of Online Education*, 5(1), 1-6.
- Drakidou, C., Pareja-Lora, A., & Read, T. (2019). A bibliometric approach to the analysis of the Technologically-Enhanced Language Learning (TELL) literature. *Argentinian Journal of Applied Linguistics*, 7(2), 8-33.
- Fang, J. W., Hwang, G. J., & Chang, C. Y. (2019). Advancement and the foci of investigation of MOOCs and open online courses for language learning: A review of journal publications from 2009 to 2018. *Interactive Learning Environments*, 28, 1-19. <https://doi.org/10.1080/10494820.2019.1703011>
- Gilliland, B., Oyama, A., & Stacey, P. (2018). Second language writing in a MOOC: Affordances and Missed Opportunities. *TESL-EJ*, 22(1), 1-25.
- Hashemifardnia, A., Shafiee, S., Esfahani, F. R., & Sepehri, M. (2021). Effects of massive open online course (MOOC) on Iranian EFL learners' speaking complexity, accuracy, and fluency. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ)*, 22(1), 56-79.
- Hsu, L. (2021). What makes good LMOOCs for EFL learners? Learners' personal characteristics and information system success model. *Computer Assisted Language Learning*, 1-25. <https://doi.org/10.1080/09588221.2021.1899243>
- Hsu, R. L. W. (2021). A grounded theory exploration of language massive open online courses (LMOOCs): Understanding students' viewpoints. *Sustainability*, 13(5), 2577. <https://doi.org/10.3390/su13052577>
- Jitpaisarnwattana, N., Reinders, H., & Darasawang, P. (2021). Learners' perspectives on interaction in a language MOOC. *JALT CALL Journal*, 17(2), 158-182. <https://doi.org/10.29140/jaltcall.v17n2.472>
- Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *International Review of Research in Open and Distributed Learning*, 14(3), 202-227. <https://doi.org/10.19173/irrodl.v14i3.1455>
- Meri-Yilan, S. (2020). The online interface and social inclusion: A MOOC study in Turkey. *Journal of Interactive Media in Education*, 2020(1), 1-10. <https://doi.org/10.5334/jime.558>
- Miller, M., & Lu, M. Y. (2003). Serving nontraditional students in e-learning environments: Building successful communities in the virtual campus. *Educational Media International*, 40(1-2), 163-169. <https://doi.org/10.1080/0952398032000092206>
- Palacios Hidalgo, F. J., Huertas Abril, C. A., & Gómez Parra, M. (2020). MOOCs: Origins, concept, and didactic applications: A systematic review of the literature (2012-2019). *Technology, Knowledge and Learning*, 25(4), 853-879. <https://doi.org/10.1007/s10758-019-09433-6>
- Palacios-Hidalgo, F. J., Huertas-Abril, C. A., & Gómez-Parra, M. E. (2020). EFL teachers' perceptions on the potential of MOOCs for lifelong learning. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, 15(4), 1-17. <https://doi.org/10.4018/IJWLTT.2020100101>
- Panagiotidis, P. (2019). MOOCs for language learning: Reality and prospects. In Graziano, K. (Ed.), *Proceedings of the Society for Information Technology and Teacher Education International Conference*. Las Vegas: Association for the Advancement of Computing in Education (AACE), 540-546.
- Perifanou, M. A., & Economides, A. A. (2014). MOOCs for foreign language learning: An effort to explore and evaluate the first practices. In L. Gómez, A. López, & I. Candel (Eds.), *Proceedings of INTED2014 conference 10th, Valencia, Spain* (pp. 3561-3570). IATED Academy.

- Qin, M. X., & Tan, X. (2020). Examining a SPOC experiment in a foundational course: Design, creation, and implementation. *Interactive Learning Environments*, 13, 1-18. <https://doi.org/10.1080/10494820.2020.1722710>
- Read, T., & Barcena, E. (2021). The role of activeness for potentiating learning in LMOOCs for vulnerable groups. *Journal of Interactive Media in Education*, 1, 1-13. <https://doi.org/10.5334/jime.628>
- Read, T., & Martín-Monje, E. (2021). Mobile and blended, please! Migrants and refugees' learning choices in a language MOOC. *The JALT CALL Journal*, 17(3), 256-276. <https://doi.org/10.29140/jaltcall.v17n3.500>
- Rinatovna, G., Vladimirovna, E., Bizyanova, F., & Haidar, I. (2017). Moosle-ee: Massive open online social learning environment for English e-learning system. *Revista San Gregorio*, 20, 6-13.
- Sallam, M. H., Martín-Monje, E., & Li, Y. (2020). Research trends in language MOOC studies: A systematic review of the published literature (2012-2018). *Computer Assisted Language Learning*, 33(3), 1-28. <https://doi.org/10.1080/09588221.2020.1744668>
- Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013-2015. *International Review of Research in Open and Distributed Learning*, 17(2), 198-221. <https://doi.org/10.19173/irrodl.v17i2.2448>
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, 23(3), 34-41.
- Yaşar, M. Ö., & Polat, M. (2021). A MOOC-based flipped classroom model: Reflecting on preservice English language teachers' experience and perceptions. *Participatory Educational Research*, 8(4), 103-123. <http://dx.doi.org/10.17275/per.21.81.8.4>
- Zeng, S., Zhang, J., Gao, M., Xu, K. M., & Zhang, J. (2020). Using learning analytics to understand collective attention in language MOOCs. *Computer Assisted Language Learning*, 1-26. <https://doi.org/10.1080/09588221.2020.1825094>
- Zhu, M., Sari, A., & Lee, M. M. (2018). A systematic review of research methods and topics of the empirical MOOC literature (2014-2016). *The Internet and Higher Education*, 37, 31-39. <https://doi.org/10.1016/j.iheduc.2018.01.002>



© 2025 by the authors. Licensee Shahid Chamran University of Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution–NonCommercial 4.0 International (CC BY-NC 4.0 license). (<http://creativecommons.org/licenses/by-nc/4.0/>).

