ReCALL (2023), 35: 3, 356–371
doi:10.1017/S0958344022000246

RESEARCH ARTICLE

LMOOC research 2014 to 2021: What have we done and where are we going next?

Yining Zhang
Tsinghua University, China (yiningzhang@tsinghua.edu.cn)

Ruoxi Sun
Tsinghua University, China (srx20@mails.tsinghua.edu.cn)

Abstract

This study reviews 71 high-quality studies of massive open online courses focused on languages (LMOOCs) that were published from the inception of LMOOCs to 2021. The purpose of this study is to gain a deeper understanding of the current state of research and identify fruitful directions for future LMOOC research. First, we reviewed three basic sets of characteristics of these studies: (1) research trends – for example, publication types and years; (2) research contexts – for example, countries in which the studies were conducted, the subjects' target languages, language-ability levels, skills, and whether the focal courses are for specific purposes; and (3) research design, including data collection, data analysis, and theoretical frameworks. We then utilized a text-mining approach called Latent Dirichlet Allocation that uses machine-learning techniques to identify research-topic commonalities underlying the collected studies. In this way, a total of nine topics were identified. They were: (1) core elements of LMOOCs; (2) interaction and communication in LMOOCs; (3) innovative LMOOC teaching practices; (4) LMOOC standards and quality assurance; (5) LMOOC implementation, participation, and completion; (6) LMOOC teaching plans; (7) LMOOC learning effectiveness and its drivers/obstacles; (8) learners and learning in LMOOCs; and (9) inclusiveness in LMOOCs. These were then diagrammed as a ThemeRiver, which showed the evolutionary trend of the nine identified topics. Specifically, scholarly interest in Topics 5, 7, and 9 increased over time, whereas for Topics 1 and 6, it decreased. Based on our results, we highlighted specific directions for future LMOOC research on each of the identified research topics.

Keywords: massive open online courses (MOOCs); language MOOCs (LMOOCs); online teaching; online learning; systematic review; topic modeling

1. Introduction

In recent years, massive open online courses (MOOCs) have been used to deliver and assess learning content in various domains at a global scale. They prominently include language MOOCs, or LMOOCs (Martín-Monje & Borthwick, 2021), which are "dedicated Web-based online courses for second languages with unrestricted access and potentially unlimited participation" (Bárcena & Martín-Monje, 2014: 1). According to Ding and Shen (2019), LMOOCs differ from other MOOCs in two ways: first, they tend to include more varied learning materials and activities; and second, their learning videos function not only as substitutes for live lectures but also as a way for students to engage with authentic target-language environments.

Cite this article: Zhang, Y. & Sun, R. (2023). LMOOC research 2014 to 2021: What have we done and where are we going next?. *ReCALL* 35(3): 356–371. https://doi.org/10.1017/S0958344022000246

© The Author(s), 2023. Published by Cambridge University Press on behalf of EUROCALL, the European Association for Computer-Assisted Language Learning.

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

In part due to these unique features, LMOOCs have been hailed as "the most attractive ... of all types of online language courses" but also the "most ... criticized" (Martín-Monje, Read & Bárcena, 2017: 11). In particular, such criticism focuses on low completion rates, lack of social interaction and low learner autonomy (e.g. Friðriksdóttir, 2019; Hsu, 2021a; Jitpaisarnwattana, Reinders & Darasawang, 2021; Read & Bárcena, 2021). Nevertheless, a recent study by Shah (2020) found that, in the wake of the pandemic, language learning had become one of the 10 most-followed MOOC subjects.

LMOOC research has closely followed the spread of LMOOCs themselves. The first peak in this work occurred in 2014, and included the seminal book *Language MOOCs: Providing Learning, Transcending Boundaries*, edited by Martín-Monje and Bárcena (2014), and Godwin-Jones's (2014) exploration of LMOOCs as an emerging mode of language-teaching delivery. Another milestone project, Sallam, Martín-Monje and Li's (2020) literature review of the LMOOC research published between 2012 and 2018, provided a thorough description of the LMOOC research status quo, together with a summary of recurring research trends and possible future directions for LMOOC research.

Given the relatively short period during which any LMOOC research has been conducted, and the recentness of the comprehensive literature review published by Sallam *et al.* (2020), it could easily be argued that another review of the LMOOC literature is premature. However, we consider one to be quite necessary for the following three reasons. The first is the rapid progress of ongoing endeavors to increase LMOOC quality. Although some post-2017 studies have concluded that LMOOCs are effective (e.g. Meri-Yilan, 2020; Read & Bárcena, 2021; Shalatska, 2018; Wang, An & Wright, 2018; Xue & Dunham, 2021; Zancanaro & Domingues, 2018), others continue to report low completion rates (e.g. Friðriksdóttir, 2019; Hsu, 2021a; Zeng, Zhang, Gao, Xu & Zhang, 2020), lack of intimate and human-like social interaction (Colibaba, Dinu, Gheorghiu & Colibaba, 2018; Hsu, 2021a; Jitpaisarnwattana *et al.*, 2021; Uchidiuno, Ogan, Yarzebinski & Hammer, 2018), and low levels of learner autonomy (Agonács, Matos, Bartalesi-Graf & O'Steen, 2019). Therefore, further inquiry across languages, proficiency levels and contexts is always welcome (Hsu, 2021b), and its outcomes can be expected to provide us with a more "contextual, grounded, and conceptually coherent" understanding of LMOOCs (Mac Lochlainn, Nic Giolla Mhichíl & Beirne, 2021: 111).

The second reason is the increasing volume of high-quality LMOOC research that has been published in the past three years. For the purposes of the present research, LMOOC studies are defined as high quality if they are peer-reviewed journal articles or chapters from edited books included in the Web of Science Core Collection. In 2021, for example, *ReCALL* published a special issue in which five articles focused on the design and delivery of LMOOCs. Meanwhile, LMOOC studies began to emerge in journals such as *Computer Assisted Language Learning* and *Language Learning & Technology*. In short, the emergence of these new high-quality studies, coupled with rapid evolution in LMOOC teaching practices, means that it is already necessary to reexamine trends in LMOOCs from their birth to the present time.

Our third reason for re-reviewing the LMOOC literature involves new departures in the data-analysis approaches used by such reviews themselves. A leading traditional method whereby literature review studies generate research topics is content analysis (Yu, Jannasch-Pennell & DiGangi, 2011). Today, however, there is increasing interest in adopting model-based approaches and machine-learning techniques to examine the intellectual structure of the studied domains. Latent Dirichlet Allocation (LDA), for instance, has the capability to group meaningful and interpretable research topics from numerous texts (Blei, Ng & Jordan, 2003), and has been successfully applied as a dimension-reduction technique in some prior educational research.

Thus, the purpose of the present study is to gain a thoroughly up-to-date understanding of LMOOC research, while also deepening our understanding of its major research topics through the use of LDA. Based on the results of that review, it also discusses potential research directions for future LMOOC research. Our systematic review is guided by the following research questions:

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

1. What were the publication trends in LMOOC research between 2014 to 2021?
2. What were the contexts of LMOOC research published during that period?
3. What research methods, data-collection techniques, data-analysis approaches and theoretical frameworks were used in such research?
4. What major research topics of the 2014–2021 LMOOC research can be discerned by LDA, and how did scholars' interest in these topics change over time?

2. Prior reviews of the LMOOC literature

One of the earliest reviews of LMOOC research was Godwin-Jones (2014), who concluded that LMOOCs should be treated as an emerging technology that has an impact on both internationalizing and localizing language learning. In the most comprehensive review of LMOOC studies to date, Sallam *et al.* (2020) identified eight recurring themes of research papers published between 2012 and 2018: (1) conceptualization of LMOOCs; (2) limitations of LMOOCs; (3) exploration of the most suitable models for language teaching; (4) potential of LMOOCs for languages for specific purposes courses; (5) learner-oriented LMOOC research; (6) teacher-oriented LMOOC research; (7) instructional design issues in LMOOCs; and (8) social learning in LMOOCs. It should be noted that the papers reviewed were mostly conference papers. On the one hand, this reflected the fact that the field is relatively new and developing; and on the other, it showed the importance of reexamining the overall trends and major themes in such research, given LMOOCs' ever-increasing popularity (Jitpaisarnwattana *et al.*, 2021; Lebedeva, 2021), amounting even to a kind of LMOOC fever in 2020 (Shah, 2020) and the related appearance of many more high-quality LMOOC studies in the field.

3. Methods

3.1 Literature-search approach and inclusion criteria

Appendix I. (all appendices appear in the supplementary material) presents our literature-search process using a PRISMA flow chart (UNC Health Sciences Library, 2021). To ensure the inclusion of high-quality studies only, we used Web of Science as the only database for our literature search. The keywords for that search were language MOOC, LMOOC, MOOC language learning, MOOC foreign language learning, MOOC second language learning, and MOOC EFL.

As of January 19, 2022, this approach yielded 1,325 studies. After removing 660 duplicated records, we screened the remaining 665, and excluded those not published in English or not published in the Web of Science Core Collection. Conference proceedings, book reviews, meeting abstracts, and editorial material were not considered in this research. This led to the exclusion of 454 further studies.

We then evaluated the eligibility of the remaining 211 studies by reading the abstract of each one, and checking whether it was accessible online. At this stage, we removed 136 of these studies because they did not actually focus on LMOOCs, and four additional ones because their full texts were not accessible online.

Thus, our literature review included 71 articles, the titles and authors of which are listed in Appendix II.

3.2 Data coding and analysis

To answer the first three of our four research questions, we developed a coding protocol (Appendix III) that included codes for country, language, language-ability level, language skills, research methods, specific populations, specific purposes, data-collection techniques, data-analysis approaches, and theoretical frameworks. To increase validity, two researchers

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

reviewed and coded the articles independently, and then the first author verified the codes. Consensus was sought through discussion, and eventually this resulted in an interrater agreement of .94.

3.3 Research topics and trends

To answer our fourth question, we applied LDA, which regards the document set as a bag of words and assumes that such words belong to certain topics that are already posited (Blei *et al.*, 2003). LDA implementation can be summarized as using the joint distribution to calculate the conditional distribution of hidden variables for a given observed variable (Blei *et al.*, 2003). Through such inference algorithms, LDA eventually presents those latent topics in the form of probability distribution over a vocabulary.

To perform LDA, we first combined the titles, keywords, and abstracts of the 71 reviewed studies into a single dataset. We then preprocessed that dataset by replacing contractions with their fully spelled out versions, and removing pronouns, numbers and all extra blanks. Our full list of stopwords can be found in Appendix IV.

To determine the optimal number of topics, we generated models of K topics from the preprocessed data. We limited the value of K to numbers from 6 to 10 inclusive, in keeping with the approach used by Sallam *et al.* (2020). From among the automated coherence measures originally designed to enable LDA to determine the optimal number of topics, we adopted the UMass measure. This measure is based on documents' co-occurrence counts, which are used as the basis for calculating the conditional likelihood of simultaneous occurrence of words, thus preventing inference of topics that could represent more than one concept (Mimno, Wallach, Talley, Leenders & McCallum, 2011). A higher UMass coherence value indicates higher topic coherence and a better solution for topic modeling (Asnani & Pawar, 2018). To present the UMass value more precisely, we added 20, 40, and 80 iterations into the algorithm.

Then, to represent trends in LMOOC research topics between 2014 and 2021, we drew a ThemeRiver diagram using Python.

4. Results

The following sections show (1) the research trends (i.e. types of publication, top LMOOC research publications, number of studies published per year); (2) research contexts (i.e. countries in which the studies were conducted, the subjects' target languages, course levels, language skills, specific population, specific purpose); (3) research methodology; and (4) major research topics and change in topic intensity over time.

4.1 Publication trends

4.1.1 Types of publication

Of the 71 high-quality LMOOC research studies published from 2014 to 2021, 86% ($n = 61$) were journal articles, and the remainder were 10 book chapters from three different edited volumes.

4.1.2 Top LMOOC research publications

The 61 journal articles were published in a total of 35 high-quality journals. Table 1 presents those journals that published at least two of these 61 studies, in descending order by number of studies published.

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Table 1. Journals that published at least two high-quality language massive open online course (LMOOC) studies, 2014–2021

Rank	Publication name	Studies (*n*)
1	*Computer Assisted Language Learning*	10
2	*International Journal of Emerging Technologies in Learning*	6
3	*ReCALL*	5
4	*International Journal of Computer-Assisted Language Learning and Teaching*	3
5	*Journal of Interactive Media in Education*	3
6	*Interactive Learning Environments*	2
7	*Interactive Technology and Smart Education*	2
8	*Sustainability*	2
9	*Turkish Online Journal of Distance Education*	2

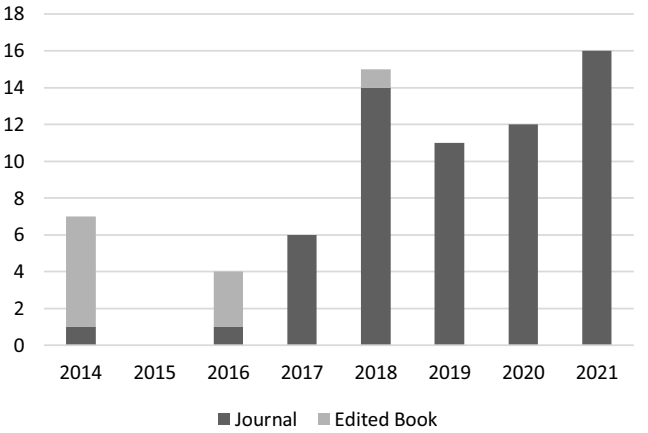


Figure 1. Number of collected language massive open online course (LMOOC) studies by year and publication venue

4.1.3 Number of studies published per year

As Figure 1 shows, the number of studies that were published in 2014 was seven (six of them as chapters in the same book). None followed in 2015. A possible explanation for this lack of publication in 2015 could be that the edited volume that appeared in 2014 could have absorbed all the available studies, thus leading to the following year being "dry." There were four in 2016 (three as chapters in another edited volume). In 2017, there were six, and in 2018, 15, all but one in journals. The final three years from which data were collected yielded 11, 12, and 16 journal articles, respectively. Despite year-on-year fluctuations, it is clear that LMOOC research received more and more high-level scholarly attention over the period in question.

4.2 Contexts of published LMOOC research

4.2.1 Countries in which LMOOC research was conducted

The top 10 countries in which LMOOC research was conducted are China ($n = 17$), Spain ($n = 16$), the UK ($n = 6$), the US ($n = 6$), Portugal ($n = 4$), Ireland ($n = 3$), Russia ($n = 3$), Turkey ($n = 3$), Iceland ($n = 2$), and Malaysia ($n = 2$).

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

4.2.2 Target language

A total of 13 languages were taught in the collected studies. In more than half of the studies (56.3%, $n = 40$), the target language was English, with the next most taught language being Spanish (14.3%, $n = 10$). The other languages were Italian, Irish, Chinese, French, Icelandic, Japanese, Portuguese, Russian, Dutch, Romanian, and Hindi.

4.2.3 Course levels

Just over half of the sampled studies ($n = 36$) did not specify their focal LMOOCs' course levels. Among the remainder, 18 studied LMOOCs for intermediate-level learners, 12 for elementary-level learners, and two for advanced placement learners.

4.2.4 Language skills

Around three quarters of the sampled studies ($n = 53$) focused on teaching integrated language skills. Among the other 18, speaking was the primary focus in most cases ($n = 7$), followed by writing ($n = 4$), cultural related ($n = 4$), vocabulary ($n = 3$), pronunciation ($n = 3$), reading ($n = 1$) and listening ($n = 1$).

4.2.5 Specific population

The great majority of the sampled LMOOC studies (88.7%, $n = 63$) did not focus on any specific population of learners. Of the remaining eight studies, four discussed the implications of LMOOCs for adult learners, especially elderly ones; three explored how LMOOCs could be used by refugees; and one was for learners with dyslexia.

4.2.6 Specific purpose

An even larger proportion of the collected LMOOC literature (90.1%, $n = 64$) was for general language learning rather than language for a specific purpose. Of the remaining seven studies, five examined LMOOCs that taught English for specific purposes (ESP), and the other two, English for academic purposes (EAP). Only two of those seven studies explicitly described the purpose of the LMOOC: one, that it was for English in medical education, and the other for hospitality English.

4.3 Research methodology

We classified the sampled studies into three types: quantitative, qualitative, and mixed methods. Uptake of each was more or less equal, with slightly more using quantitative designs (35%, $n = 25$) than mixed methods (32%, $n = 23$) or qualitative ones (31%, $n = 22$).

As for data-collection approaches, 43.7% of the studies ($n = 31$) used one data collection method and 40.8% ($n = 29$), more than one. The rest of the studies ($n = 11$) did not specify their data-collection approaches. Survey methods were adopted the most ($n = 43$), trailed distantly by interviews ($n = 11$) and collection of tracking data from learning platforms ($n = 11$).

As for data-analysis approaches, slightly under half of the studies ($n = 34$) used one data-analysis approach, and one third ($n = 24$) used more than one. The remaining studies ($n = 13$) did not mention their data-analysis methods. Descriptive statistics were adopted the most ($n = 39$), followed by thematic analysis ($n = 19$), inferential statistics ($n = 14$), content analysis ($n = 5$), learning analytics ($n = 4$), and grounded-theory analysis ($n = 1$).

As for the theoretical constructs used in the reviewed studies, the top three theoretical concepts were motivation ($n = 5$), heutagogy ($n = 3$) and learning autonomy ($n = 3$).

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Table 2. Topic categories and keywords of the sampled language massive open online course (LMOOC) literature that emerged from Latent Dirichlet Allocation

Topic no.	Label	Keywords	Topic intensity
T1	Core elements of LMOOCs	Mode, goal, feedback, structure, completion, retention, college, gain	.03
T2	Interaction and communication in LMOOCs	Interaction, communication, speak, international, characteristic, access, example, regard	.05
T3	Innovative LMOOC teaching practices	Innovative, content, participant, think, interest, teaching, profile, video, week	.08
T4	Standards and quality assurance in LMOOCs	Standard, Common European Framework of Reference for Languages (CEFR), paradigm, proposal, structure, didactic, argue, app	.02
T5	LMOOC course implementation, participation, and completion	Implement, participation, completion, college, micro, lesson, intrinsic, empower	.03
T6	LMOOC teaching planning	Plan, efficiency, deal, want, fit, label, typically, depend, space	.02
T7	LMOOC learning effectiveness and its factors	Learner, effectiveness, efficacy, quality, learn, characteristics, attribute, platform, model	.13
T8	Learners and learning in LMOOCs	Student, learner, learn, teacher, English, design, open, massive	.61
T9	Inclusiveness in LMOOCs	Refugee, migrant, inclusive, retention, collaboration, impact, help, demographic, share	.04

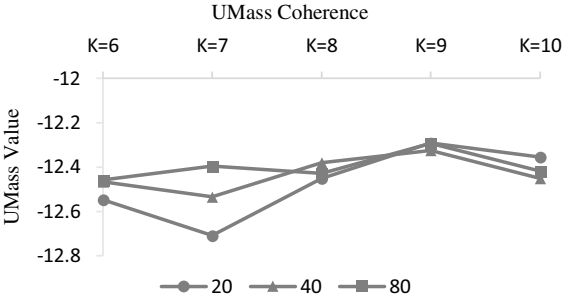


Figure 2. UMass coherence values at 20, 40 and 80 iterations

4.4 Major research topics and change in topic intensity over time

As shown in Figure 2, the model was of highest UMass coherence value when K = 9, and the optimal number of topics within that dataset was therefore determined to be nine.

Table 2 presents the topic-modeling results regarding these nine emerging topics, their associated keywords, and their intensity. Most of our labels for these topics are self-explanatory. However, it should be noted that the first one, "Core elements of LMOOCs," involves crucial LMOOC features such as teaching modes, teaching goals, feedback, course structure, completion rates, and retention rates; and that the second, "Interaction and communication in LMOOCs," includes among other matters, teachers' practices of speaking to LMOOC learners.

Figure 3 presents the evolution of topic intensity of the nine research topics from 2014 to 2021. Among them, T7 exhibited an obvious increase over time. On the other hand, T3 and T6 showed obvious decreases. The intensity of T8 was stronger than that of the other topics throughout the

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

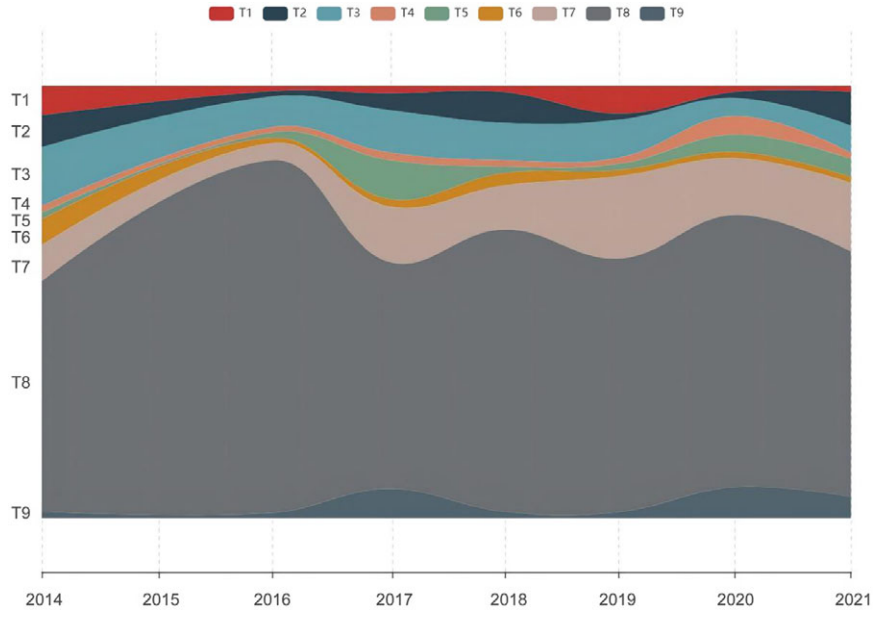


Figure 3. Changes in topic intensity, 2014–2021, of the nine research topics identified via Latent Dirichlet Allocation
Note. None of the 71 collected studies were published in 2015.

time period studied, but it fluctuated from one year to the next without exhibiting a clear tendency to increase or decrease. Similarly, the intensity of the remaining topics also fluctuated.

5. Future directions for LMOOC research

Based on the identified nine key LMOOC research topics and the changes in the popularity of those topics over time, this section provides detailed recommendations for future LMOOC research, including for subtopics within each topic, research designs, data-analysis approaches, and theoretical underpinnings.

Topic 1: Core elements of LMOOCs. Although our ThemeRiver diagram shows a decreasing trend in this research topic's popularity, understanding LMOOCs' core elements is still vital. Future researchers should obtain a more thorough understanding of each LMOOC course object in isolation. Usually, course elements in LMOOCs follow the xMOOC approach closely, and can be divided into two general types: video lectures and supporting materials. In the case of the former, the successful design and production of videos will directly impact LMOOC students' perceptions of their learning. Sokolik (2014) also suggested that videos in LMOOCs should reflect outside-the-box thinking, and that one of many options for achieving this is to make authentic story-based language learning videos. Nevertheless, the present systematic review of 71 studies found only one (Ge, Liang & Peng, 2021) that examined the videos used in LMOOCs. Filling this surprising gap in the literature is a clear and important avenue for future research.

Besides videos, supporting materials such as class slides, supplementary audio and video resources, text files, and external links are commonly supplied within LMOOCs. However, none of these specific elements were adopted as research objects by the studies we sampled, perhaps because many of those studies' authors seemed to regard them as a single package (e.g. Castrillo & Sedano, 2021; Zancanaro & Domingues, 2018). We therefore call for a more thorough understanding of each of the course objects in LMOOCs.

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Topic 2: Interaction and communication in LMOOCs. More research is needed to explore effective ways to enhance both learner–learner and learner–instructor interaction, especially amid the rise of innovative and intelligent communicative tools such as instant voice, text messages, and social media posts as aids to traditional LMOOC learning (Lebedeva, 2021). After all, LMOOCs are not only about the "goal" of learning but also about the "vehicle" (Bárcena, Read & Sedano, 2020: 39). It is therefore vital to set the promotion of interaction as "a specific e-literacy goal" (Wright & Furneaux, 2021: 33); indeed, such literacy should become an integral part of online-learning instructional design in general.

In addition, the impact of the promotion of interactivity within LMOOCs warrants further study. This is because, on the one hand, LMOOCs are designed to be learner centered and social (Bárcena, Read, Martín-Monje & Castrillo, 2014) and, on the other, because studies have found some learners treat such courses as resource banks, or even monologic learning spaces (Hsu, 2021a; Mac Lochlainn *et al.*, 2021). Scholars have also critiqued LMOOCs for their lack of face-to-face interaction (Mellati & Khademi, 2020; Uchidiuno *et al.*, 2018). This, as Mac Lochlainn *et al.* (2021) pointed out, suggests the need for scholars to debate whether it matters "if learners view LMOOCs as a means of furthering their offline goals rather than fostering virtual communities" (p. 123; see also Hsu, 2021a). In short, even though social interaction is vital to learning, there should be more empirical studies about the interrelationship of enrollment type, interaction, and learning outcomes, especially ones that take account of potential moderators and/or mediators of that interrelationship at the course, teacher, student, cultural and/or other contextual levels.

Topic 3: Innovative LMOOC teaching practices. Based on the insight that taking LMOOCs on mobiles is not identical to doing so on personal computers (Moreno & Traxler, 2016), one potential future direction for studying innovative teaching practices in LMOOCs involves the perspective of mobile-assisted language learning (MALL) on both instructional design and teaching methodologies. Yet only one study we reviewed adopted a MALL framework (Read & Bárcena, 2020).

A second direction involving this topic would be to explore more individualized LMOOCs with multiple learning paths. Consistent with Anderson, Huttenlocher, Kleinberg and Leskovec's (2014) classification of MOOC learners into five types – that is, viewers, all-rounders, solvers, bystanders, and collectors – individual LMOOC students exhibit a variety of both learning behaviors (e.g. Hsu, 2021a; Mac Lochlainn *et al.*, 2021; Martín-Monje, Castrillo & Mañana-Rodríguez, 2018) and learning sequences (Jitpaisarnwattana *et al.*, 2021). In this sense, compulsory tasks that require learners to submit uniform and rigid answers are arguably inappropriate to the LMOOC setting (Martín-Monje *et al.*, 2018), and should be replaced by more flexible options that take account of variation in learners' educational and cultural backgrounds (Bárcena & Read, 2015).

A third direction would be to reconceptualize assessment activities. Previous studies have reported inconsistent results regarding the effects of LMOOCs' adoption of self-assessment, peer assessment, and assessment by computers or artificial intelligence tools (e.g. Suen, 2014). Also, given that most students' interest and motivation to participate in a given course is likely to diminish over time, it is important for LMOOC providers to consider adopting more innovative and effective assessment activities (Hashim, Salam, Mohamad & Sazali, 2018; Sokolik, 2014, 2016; Uchidiuno *et al.*, 2018).

A fourth direction involves thinking outside the box. Specifically, although LMOOCs can be carried out in either xMOOC or cMOOC formats, and "many consider cMOOCs to be superior [to xMOOCs] in form and function" (Sokolik, 2014: 18), almost all LMOOCs so far have been xMOOC oriented, centering on a specific topic and mirroring formal learning (Martín-Monje *et al.*, 2018). It would be hard to argue against the idea that LMOOCs should provide something more than a lecture-based, instructor-centered approach – for example, encourage those taking them to set up on their own personalized learning networks (Godwin-Jones, 2014). Clearly, future

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

studies exploring how cMOOC-based or other new formats of LMOOCs can be effectively designed and utilized are warranted.

Topic 4: Standards and quality assurance in LMOOCs. Many studies have made important contributions to our understanding of what constitutes an effective LMOOC course (e.g. Chacón-Beltrán, 2017; Fang, 2018; Friðriksdóttir, 2021; Read & Bárcena, 2020; Sokolik, 2014). However, all of them have lacked clear, detailed, theoretically and empirically grounded rubrics that capture a range of dimensions of LMOOC design and teaching. Creating one could represent an important new research direction, especially if the resulting rubric is easy for LMOOC educators to use. In short, as Luo and Ye (2021) pointed out, "as millions of people learn from LMOOCs and millions of dollars are invested in LMOOCs every year, there is a pressing need to establish benchmarks for quality assurance" (p. 178). In this context, it is also worth noting that the data used in most previous studies of LMOOC quality and standards have been drawn from these course creators' experiences and reflections rather than from learners' perspectives – an imbalance that future studies should redress.

Topic 5: LMOOC implementation, participation, and completion. Further research on this topic should, in the first instance, continue exploring how best to develop specific language skills within LMOOCs. Past studies have suggested that LMOOCs can be effective for the development of receptive skills such as reading and listening, but much less so for productive skills such as speaking and writing (Sallam *et al*., 2020). Our finding that the two most popular language skills (apart from integrated skills) were both productive (i.e. speaking and writing) may suggest that LMOOC researchers are already making the necessary efforts in this direction. On the other hand, at just 15.5% of the total sample, the quantity of such studies is still limited, and more are clearly warranted – especially those that take account of obstacles to productive-skill acquisition that may be posed by course logistics.

Secondly, future LMOOC research on this topic needs to explore best practices for delivering culture-related content (Colibaba *et al*., 2018; Fuchs, 2020; Luo & Ye, 2021; Mac Lochlainn, Nic Giolla Mhichíl & Beirne, 2020a; Mac Lochlainn, Nic Giolla Mhichíl, Beirne & Brown, 2020b). Our review found that only four out of 71 studies studied culture-focused LMOOCs, and their treatment of such courses' instructional design was especially limited. Given the capability of LMOOCs to provide learners with authoritative materials, it is of great importance for future LMOOC research to discuss how teaching objectives, learning content, learning activities, and assessments that cater to different LMOOC study contexts can work synergistically in the delivery of cultural content.

A third promising avenue for research on this topic involves the teaching of LMOOCs for specific purposes, which, although relatively new (Sokolik, 2016), look likely to grow in popularity, perhaps even more quickly than LMOOCs as a whole (see also Uchidiuno *et al*., 2018). Nevertheless, only seven out of our 71 sampled studies examined the LMOOCs for EAP or ESP. Further studies involving the specific purposes of LMOOCs are therefore to be welcomed, and in particular, our results suggest that they should (1) seek more innovative pedagogical possibilities for accommodating language learners' needs rather than focusing on traditional course tools, such as dictionaries, translators, or interactive transcripts; (2) study how to facilitate students' acquisition of dual learning outcomes – for example, both medical knowledge and medical linguistic communication skills (Colibaba *et al*., 2018); and (3) give due attention to issues around the building, within LMOOCs for specific purposes, of learning communities that can facilitate both knowledge creation and collaboration (Sokolik, 2016).

Another fourth direction for the study of T5 is the localization and globalization of LMOOCs. Language learners can, of course, enroll directly in LMOOCs provided by countries where their target language is spoken (Doğan, Sunar, Duru & White, 2018). In practice, however, this does not always happen, and 41 of the 71 studies we reviewed focused on LMOOCs provided by local institutions, such as Chinese instructors teaching English. Many of these localized courses were designed around the goal of providing more indigenized learning to learners in the countries

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

where they were hosted. Therefore, it is important to conduct some direct comparative studies of LMOOC learning that situate course providers, teachers, teaching pedagogies, teaching beliefs, teaching contents, and students in different cultural contexts.

Topic 6: LMOOC teaching planning. We argue that this topic needs to be escalated into a broader one covering teachers and teaching in LMOOCs. The first new research direction suggested by such an expanded viewpoint involves teacher training. Future studies could usefully focus on (1) best practices in LMOOC teacher training (King, Luan & Lopes, 2018; Kormos & Nijakowska, 2017); (2) adaptations to diverse cultural groups and learners with specific needs (Kormos & Nijakowska, 2017); and (3) teacher training based on interventions in teachers' beliefs, autonomy, roles, and awareness, online teaching and learning theories, and specific online-activity designs (Castrillo, 2014; King *et al*., 2018; Koukis & Jimoyiannis, 2019; Orsini-Jones, Zou, Hu & Wei, 2017; Phi, 2017).

In addition, there has been little study to date of interaction among the multiple stakeholders in LMOOC design and teaching, despite the critical importance of extensive and close stakeholder collaboration to such courses' success (Castrillo & Sedano, 2021; Zancanaro & Domingues, 2018). What might be called the ecological sphere of these stakeholders includes, but is not limited to, learners, teaching assistants, universities, NGOs and other support associations.

Another direction for T6 research is the integration of LMOOCs into traditional classroom teaching. Although seven out of the 71 sampled studies focused on the use of LMOOCs as small private online courses, the conclusions of this research stand would be more reliable, valid, and nuanced if it examined additional languages, proficiency levels, language skills, and cultural backgrounds.

Topic 7: LMOOCs' learning effectiveness and its factors. Various factors such as motivation, demotivation, self-regulation, anxiety, the Big Five personality traits, and learning styles were identified in the collected literature as key differences (e.g. Agonács *et al*., 2019; Hsu, 2021a, 2021b). As Ding and Shen (2019) pointed out, much more research attention should be paid to the psychological and behavioral dimensions of LMOOC learners' autonomous behavior, and especially to how/why they adopted their autonomy, heutagogy, self-determination, self-efficacy, and use of metacognitive, motivation-control and emotional-control strategies.

Topic 8: Learners and learning in LMOOCs. Arguably, the most promising future direction in T8 research involves learner behavior. To date, many studies have been based on behaviors that can be recorded and tracked on learning platforms. However, behaviors other than those that such platforms can document certainly exist, and could be important reflections of the social dimension of learning. For example, Mac Lochlainn *et al*. (2021) found that as well as writing or recording their own target-language output as their LMOOC required, learners were listening to and even assessing the target-language output of their peers. In light of such complex behavioral dynamics, future LMOOC studies should actively seek out learning behaviors that are not recorded by learning platforms and/or that were not envisioned or encompassed by courses' initial designs.

Another potentially fruitful direction for future research on this topic is to attempt to understand LMOOC learners holistically as people, in line with Ushioda's (2011) argument that language learners should be understood holistically rather than simply as learner types (see also Martín-Monje & Borthwick, 2021). Specifically, this would require analysis not only of LMOOC learners' course-related data, such as their grades, but also of their relationships with their environmental contexts at both micro and macro levels.

Finally, future T8 researchers should study post-LMOOC learning. Language learning is a relatively long process, and language proficiency needs to be regularly or continuously topped up as part of a lifelong learning process. Chacón-Beltrán (2017) noted that the beginning-level learners of English that he studied received explicit instruction on how to continue their English learning after completing their LMOOC. We expect that many existing LMOOCs offer such instruction, either explicitly or implicitly, through videos or discussions. However, none of the studies we sampled focused on this phenomenon.

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Topic 9: Inclusiveness in LMOOCs. Recent studies point to the possibility of using LMOOCs to provide equitable, socially inclusive forms of language learning (e.g. Meri-Yilan, 2020; Rodrigo, 2014). Nearly 10% of our collected studies were of LMOOCs geared toward populations widely seen as marginalized by mainstream education, including refugees (e.g. Bárcena *et al.*, 2020; Castrillo & Sedano, 2021), people with dyslexia (Kormos & Nijakowska, 2017), low-proficiency adults (López, 2019), and senior citizens (Mac Lochlainn *et al.*, 2020a, 2021). LMOOC education for these groups should be carefully contextualized. Accordingly, research that takes account of LMOOC students' unique combinations of cultural background, educational level, and learning styles should be strongly encouraged.

However, inclusiveness in LMOOCs does not merely involve learners' characteristics but also the representation of minority languages. All the LMOOCs studied by our 71 sampled studies collectively taught only 13 target languages. More broadly, it is undeniable that most current LMOOCs reflect a Western-oriented epistemology (e.g. Bárcena *et al.*, 2020). But, as Mac Lochlainn *et al.* (2020b) pointed out, the true potential of MOOCs is "not only enablement or enhancement but also the possibility of empowerment" (p. 212). Lack of research on the social justice, inequality and diversity implications of LMOOCs could make it impossible to achieve the goal of empowerment (Meri-Yilan, 2020), or worse, deepen inequality because "MOOC producers are firmly moored in the global North and critics argue that MOOCs reproduce neo-colonial, hegemonic educational practices" (King *et al.*, 2018: 285). This would obviously deviate from the original conception of MOOCs as reaching those "underprivileged people who cannot access regular paid education and are often in precarious, unstable situations" (Bárcena *et al.*, 2020: 39). Researchers should therefore investigate the existing barriers to LMOOCs' inclusion of more languages, especially minority ones and non-Western epistemologies, and propose concrete ways of overcoming such barriers.

Research design. Design-focused research (Appel & Pujolà, 2021) was not identified as a distinct LMOOC research topic by LDA. Yet in spite of, or because of, this absence, more such research is clearly also needed. As Levy (2013) commented, a design-based approach could "provide one pathway forward to CALL, particularly because of its engagement, from the outset, with the concerns of the practitioner and their interaction with systems" (p. 38). It is through an iterative process of interchange between practitioners and researchers that the instructional designs of LMOOCs are best refined (Appel & Pujolà, 2021). In addition, grounded-theory research that engages in "granular analysis of activities" (Mac Lochlainn *et al.*, 2021: 113), and uses the results to generate concepts and theories particular to LMOOC learning, is strongly needed (Hsu, 2021a).

Data analysis. The two general foci of this future work should be big data and small data. With regard to the former, methods including data mining, learning analytics and machine learning are promising (e.g. Friðriksdóttir, 2021; Zeng *et al.*, 2020). As for small data, content-analysis, case-study, and/or autoethnographic methods may be effective ways to document and trace LMOOC educators' actions, reflections and changes. In particular, we believe that autoethnographic approaches (Ellis, 2004) could contribute vital new knowledge to the field of LMOOC teaching and learning.

Theoretical foundations. Our review identified 25 theoretical constructs that had been adopted by scholars examining LMOOC design and/or implementation. Among these, the most popular were motivation, heutagogy, and learning autonomy. This is not surprising as, by their nature, MOOCs require learners to be autonomous and self-motivated. Like Sallam *et al.* (2020), we are eager to see future studies that use more diverse and influential theories and concepts drawn from applied linguistics and general educational psychology. On the other hand, there should also be more self-regulated learning-related topics in LMOOC research. This has to do with both the high impact of self-regulated learning research in applied linguistics, and the nature of LMOOCs themselves, as enabling self-regulated language learning.

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

6. Limitations

This study has the following limitations. First, because it focused on published papers in peer-reviewed journals and book chapters, it may have missed important trends and topics in LMOOC research that appeared only in conference proceedings, dissertations and theses. Also, our review was limited to publications in English.

7. Conclusion

The present systematic review of 71 high-quality LMOOC research papers published between 2014 and 2021 has found that, in general, LMOOC research is promising. Our text mining of the selected studies identified nine major research topics in this body of literature. Based on these results, we have proposed future directions for LMOOC research. Perhaps a broader question that is worth pursuing is how the findings from LMOOC research can further augment, modify, or even redefine our existing knowledge of what constitutes a "good" language learner. The lessons that we learn from recent LMOOC research, together with findings in the future, will undoubtfully enlarge and reshape our vision of language learning based on large-scale data, more diverse participant populations, and a focus on complex learning dynamics that reflect real-world learning challenges. Eventually, this will help students to become more flexible and adaptive as they compete in the ever-changing, and sometimes chaotic, lifelong-learning environment.

Supplementary material. To view supplementary material referred to in this article, please visit https://doi.org/10.1017/S0958344022000246

Acknowledgements. This study is funded by the Ministry of Education in China Humanities and Social Science Project, Grant/Award Number: 20YJC880126.

Ethical statement and competing interests. This study did not involve human subjects, human material, human tissues, or human data. The authors declare no competing interests.

References

Agonács, N., Matos, J. F. Bartalesi-Graf, D. & O'Steen, D. N. (2019) On the path to self-determined learning: A mixed methods study of learners' attributes and strategies to learn in language MOOCs. *International Journal of Learning Technology*, 14(4): 304–330. https://doi.org/10.1504/IJLT.2019.106553

Anderson, A., Huttenlocher, D., Kleinberg, J. & Leskovec, J. (2014) Engaging with massive online courses. In Chung, C.-W. (ed.), *Proceedings of the 23rd International Conference on World Wide Web*. New York: Association for Computing Machinery, 687–698. https://doi.org/10.1145/2566486.2568042

Appel, C. & Pujolà, J.-T. (2021) Designing speaking interaction in LMOOCs: An eTandem approach. *ReCALL*, 33(2): 161–176. https://doi.org/10.1017/S0958344021000045

Asnani, K. & Pawar, J. D. (2018) Improving coherence of topic based aspect clusters using domain knowledge. *Computación y Sistemas*, 22(4): 1403–1414. https://doi.org/10.13053/CyS-22-4-2401

Bárcena, E. & Martín-Monje, E. (2014) Introduction. Language MOOCs: An emerging field. In Martín-Monje, E. & Bárcena, E. (eds.), *Language MOOCs: Providing learning, transcending boundaries*. Warsaw: De Gruyter Open, 1–15. https://doi.org/10.2478/9783110420067.1

Bárcena, E. & Read, T. (2015) The role of modularity and mobility in language MOOCs. *Verbeia*, 1(2015): 28–35.

Bárcena, E., Read, T., Martín-Monje, E. & Castrillo, M. D. (2014) Analysing student participation in foreign language MOOCs: A case study. In Cress, U. & Delgado, C. (eds.), *EMOOCs 2014: Proceedings of the European MOOC Stakeholders Summit 2014*. Lausanne: Ecole Polytechnique Federale de Lausanne, 11–17.

Bárcena, E., Read, T. & Sedano, B. (2020) An approximation to inclusive language in LMOOCs based on appraisal theory. *Open Linguistics*, 6(1): 38–67. https://doi.org/10.1515/opli-2020-0003

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3: 993–1022.

Castrillo, M. D. (2014) Language teaching in MOOCs: The integral role of the instructor. In Martín-Monje, E. M. & Bárcena Madera, E. (eds.), *Language MOOCs: Providing learning, transcending boundaries*. Warsaw: De Gruyter Open, 67–90. https://doi.org/10.2478/9783110420067.5

Castrillo, M. D. & Sedano, B. (2021) Joining forces toward social inclusion: Language MOOC design for refugees and migrants through the lens of maker culture. *CALICO Journal*, 38(1): 79–102. https://doi.org/10.1558/cj.40900

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Chacón-Beltrán, R. (2017) The role of MOOCs in the learning of languages: Lessons from a beginners' English course. *Porta Linguarum Revista Interuniversitaria de Didáctica de Las Lenguas Extranjeras*, 28: 23–35. https://doi.org/10.30827/Digibug.54001

Colibaba, A., Dinu, C., Gheorghiu, I. & Colibaba, S. (2018) MedLang MOOC–An innovative palliative care learning tool. Innovative approaches and teaching methodologies. In Grosu-Rădulescu, L.-M. (ed.), *Foreign language teaching in Romanian higher education* (Vol. 27). Cham: Springer International Publishing, 125–147. https://doi.org/10.1007/978-3-319-93329-0_7

Ding, Y. & Shen, H. (2019) Delving into learner autonomy in an EFL MOOC in China: A case study. *Computer Assisted Language Learning*, 35(3): 247–269. https://doi.org/10.1080/09588221.2019.1681464

Doğan, G., Sunar, A. S., Duru, İ. & White, S. (2018) Who is the English as a second language speaker in this MOOC? *International Journal of Information and Education Technology*, 8(3): 213–217. https://doi.org/10.18178/ijiet.2018.8.3.1035

Ellis, C. (2004) *The ethnographic I: A methodological novel about autoethnography*. Walnut Creek: AltaMira Press.

Fang, G. (2018) Japanese informatization teaching model based on MOOC. *International Journal of Emerging Technologies in Learning (iJET)*, 13(07): 124–136. https://doi.org/10.3991/ijet.v13i07.8800

Friðriksdóttir, K. (2019) The effect of tutor-specific and other motivational factors on student retention on Icelandic Online. *Computer Assisted Language Learning*, 34(5–6): 663–684. https://doi.org/10.1080/09588221.2019.1633357

Friðriksdóttir, K. (2021) The effect of content-related and external factors on student retention in LMOOCs. *ReCALL*, 33(2): 128–142. https://doi.org/10.1017/S0958344021000069

Fuchs, C. (2020) Cultural and contextual affordances in language MOOCs: Student perspectives. *International Journal of Online Pedagogy and Course Design*, 10(2): 48–60. https://doi.org/10.4018/IJOPCD.2020040104

Ge, J., Liang, X. & Peng, T. (2021) Investigating the design strategies of EFL learning videos from the perspective of social presence. *The Asia-Pacific Education Researcher*, 31(3): 243–252. https://doi.org/10.1007/s40299-021-00556-y

Godwin-Jones, R. (2014) Global reach and local practice: The promise of MOOCs. *Language Learning & Technology*, 18(3): 5–15.

Hashim, H., Salam, S., Mohamad, S. N. M. & Sazali, N. S. S. (2018) The designing of adaptive self-assessment activities in second language learning using massive open online courses (MOOCs). *International Journal of Advanced Computer Science and Applications*, 9(9): 276–282. https://doi.org/10.14569/IJACSA.2018.090937

Hsu, L. (2021b) What makes good LMOOCs for EFL learners? Learners' personal characteristics and Information System Success Model. *Computer Assisted Language Learning.* Advance online publication. https://doi.org/10.1080/09588221.2021.1899243

Hsu, R. L.-W. (2021a) A grounded theory exploration of language massive open online courses (LMOOCs): Understanding students' viewpoints. *Sustainability*, 13(5): Article 2577. https://doi.org/10.3390/su13052577

Jitpaisarnwattana, N., Reinders, H. & Darasawang, P. (2021) Understanding the roles of personalization and social learning in a language MOOC through learning analytics. *Online Learning*, 25(4): 324–343. https://doi.org/10.24059/olj.v25i4.2509

King, M., Luan, B. & Lopes, E. (2018) Experiences of Timorese language teachers in a blended massive open online course (MOOC) for continuing professional development (CPD). *Open Praxis*, 10(3): 279–287. https://doi.org/10.5944/openpraxis.10.3.840

Kormos, J. & Nijakowska, J. (2017) Inclusive practices in teaching students with dyslexia: Second language teachers' concerns, attitudes and self-efficacy beliefs on a massive open online learning course. *Teaching and Teacher Education*, 68: 30–41. https://doi.org/10.1016/j.tate.2017.08.005

Koukis, N. & Jimoyiannis, A. (2019) MOOCS for teacher professional development: Exploring teachers' perceptions and achievements. *Interactive Technology and Smart Education*, 16(1): 74–91. https://doi.org/10.1108/ITSE-10-2018-0081

Lebedeva, M. (2021) Instructional design of skill-balanced LMOOC: A case of the Russian language MOOC for beginners. *JUCS - Journal of Universal Computer Science*, 27(5): 485–497. https://doi.org/10.3897/jucs.67136

Levy, M. (2013) Design-based research and the quest for normalization in CALL. In Rodriguez, J. C. & Pardo-Ballester, C. (eds.), *Design-based research in CALL*. San Marcos: CALICO, 31–40.

López, M. L. C. (2019) When bilingualism goes beyond one's expectations: The learning of cognates amongst adult EFL students. *Estudios de Lingüística Inglesa Aplicada*, 1: 89–121. https://doi.org/10.12795/elia.mon.2019.i1.05

Luo, R. & Ye, Z. (2021) What makes a good-quality language MOOC? An empirical study of criteria to evaluate the quality of online language courses from learners' perspectives. *ReCALL*, 33(2): 177–192. https://doi.org/10.1017/S0958344021000082

Mac Lochlainn, C., Nic Giolla Mhichíl, M. & Beirne, E. (2020a) Diversity, exclusion and inclusion: A case study of welcome online & minority language representation in MOOCs. *Journal of Multilingual and Multicultural Development.* Advance online publication. https://doi.org/10.1080/01434632.2020.1823988

Mac Lochlainn, C., Nic Giolla Mhichíl, M. & Beirne, E. (2021) Clicking, but connecting? L2 learning engagement on an ab initio Irish language LMOOC. *ReCALL*, 33(2): 111–127. https://doi.org/10.1017/S0958344021000100

Mac Lochlainn, C., Nic Giolla Mhichíl, M., Beirne, E. & Brown, M. (2020b) The soul behind the screen: Understanding cultural enrichment as a motivation of informal MOOC learning. *Distance Education*, 41(2): 201–215. https://doi.org/10.1080/01587919.2020.1757408

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Martín-Monje, E. & Bárcena, E. (eds.) (2014) *Language MOOCs: Providing learning, transcending boundaries*. Warsaw: De Gruyter Open. https://doi.org/10.2478/9783110420067

Martín-Monje, E. & Borthwick, K. (2021) Researching massive open online courses for language teaching and learning [Editorial]. *ReCALL*, 33(2): 107–110. https://doi.org/10.1017/S0958344021000094

Martín-Monje, E., Castrillo, M. D. & Mañana-Rodríguez, J. (2018) Understanding online interaction in language MOOCs through learning analytics. *Computer Assisted Language Learning*, 31(3): 251–272. https://doi.org/10.1080/09588221.2017.1378237

Martín-Monje, E., Read, T. & Bárcena, E. (2017) The importance of forum interaction and success rates in language MOOCs. *EMOOCs 2017: European MOOCs Stakeholders Summit*. Madrid, Spain, 22–26 May.

Mellati, M. & Khademi, M. (2020) MOOC-based educational program and interaction in distance education: Long life mode of teaching. *Interactive Learning Environments*, 28(8): 1022–1035. https://doi.org/10.1080/10494820.2018.1553188

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

Mimno, D., Wallach, H., Talley, E., Leenders, M. & McCallum, A. (2011) Optimizing semantic coherence in topic models. In Barzilay, R. & Johnson, M. (eds.), *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Stroudsburg: Association of Computational Linguistics, 262–272.

Moreno, A. I. & Traxler, J. (2016) Mall-based MOOCS for language teachers: Challenges and opportunities. *Porta Linguarum Revista Interuniversitaria de Didáctica de Las Lenguas Extranjeras*, 1: 73–85. https://doi.org/10.30827/Digibug.54090

Orsini-Jones, M., Zou, B., Hu, Y. & Wei, L. (2017) Chinese English teachers' perspectives on "distributed flip MOOC blends": From BMELTT to BMELTE. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(4): 1–14. https://doi.org/10.4018/IJCALLT.2017100101

Phi, M. T. (2017) Becoming autonomous learners to become autonomous teachers: Investigation on a MOOC blend. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(4): 15–32. https://doi.org/10.4018/IJCALLT.2017100102

Read, T., & Bárcena, E. (2020) Toward a framework for language moocs and mobile assisted language learning. *Propósitos y Representaciones*, 8(1), e470. https://doi.org/10.20511/pyr2020.v8n1.470

Read, T. & Bárcena, E. (2021) The role of activeness for potentiating learning in LMOOCs for vulnerable groups. *Journal of Interactive Media in Education*, 2021(1): 1–13. https://doi.org/10.5334/jime.628

Rodrigo, C. (2014) Accessibility in language MOOCs. In Martín-Monje, E. M. & Bárcena Madera, E. (eds.), *Language MOOCs: Providing learning, transcending boundaries*. Warsaw: De Gruyter Open, 106–126. https://doi.org/10.2478/9783110420067.7

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*, 35(4): 764–791. https://doi.org/10.1080/09588221.2020.1744668

Shah, D. (2020) *The second year of the MOOC: A review of MOOC stats and trends in 2020*. https://www.classcentral.com/report/the-second-year-of-the-mooc

Shalatska, H. M. (2018) The efficiency of MOOCs implementation in teaching English for professional purposes. *Information Technologies and Learning Tools*, 66(4): 186–196. https://doi.org/10.33407/itlt.v66i4.2106

Sokolik, M. (2014) What constitutes an effective language MOOC? In Martín-Monje, E. M. & Bárcena Madera, E. (eds.), *Language MOOCs: Providing learning, transcending boundaries*. Warsaw: De Gruyter Open, 16–32. https://doi.org/10.2478/9783110420067.2

Sokolik, M. (2016) Academic writing in MOOC environments: Challenges and rewards. In Martín-Monje, E., Elorza, I. & Riaza, B. G. (eds.), *Technology-enhanced language learning for specialized domains: Practical applications and mobility*. Abingdon: Routledge, 165–176.

Suen, H. K. (2014) Peer assessment for massive open online courses (MOOCs). *The International Review of Research in Open and Distributed Learning*, 15(3): 312–327. https://doi.org/10.19173/irrodl.v15i3.1680

Uchidiuno, J. O., Ogan, A., Yarzebinski, E. & Hammer, J. (2018) Going global: Understanding English language learners' student motivation in English-language MOOCs. *International Journal of Artificial Intelligence in Education*, 28(4): 528–552. https://doi.org/10.1007/s40593-017-0159-7

UNC Health Sciences Library. (2021) *The LibGuides: Creating a PRISMA flow diagram: PRISMA 2020*. https://guides.lib.unc.edu/prisma

Ushioda, E. (2011) Language learning motivation, self and identity: Current theoretical perspectives. *Computer Assisted Language Learning*, 24(3): 199–210. https://doi.org/10.1080/09588221.2010.538701

Wang, J., An, N. & Wright, C. (2018) Enhancing beginner learners' oral proficiency in a flipped Chinese foreign language classroom. *Computer Assisted Language Learning*, 31(5–6): 490–521. https://doi.org/10.1080/09588221.2017.1417872

Wright, C. & Furneaux, C. (2021) 'I am proud of myself': Student satisfaction and achievement on an academic English writing MOOC. *International Journal of Computer-Assisted Language Learning and Teaching*, 11(1): 21–37. https://doi.org/10.4018/IJCALLT.2021010102

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press

Xue, X. & Dunham, R. E. (2021) Using a SPOC-based flipped classroom instructional mode to teach English pronunciation. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2021.1980404

Yu, C. H., Jannasch-Pennell, A. & DiGangi, S. (2011) Compatibility between text mining and qualitative research in the perspectives of grounded theory, content analysis, and reliability. *The Qualitative Report*, 16(3): 730–744.

Zancanaro, A. & Domingues, M. J. C. S. (2018) Massive open online courses (MOOC) for teaching Portuguese for foreigners: A case study. *Turkish Online Journal of Distance Education*, 19(2): 4–20. https://doi.org/10.17718/tojde.415602

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*, 35(7): 1594–1619. https://doi.org/10.1080/09588221.2020.1825094

About the authors

Yining Zhang is an associate professor in the Department of Foreign Languages and Literatures at Tsinghua University. Her research interests include online language teaching and learning, individual differences in language learning, and learning analytics in language teaching and learning.

Ruoxi Sun is an undergraduate student in the Department of Foreign Languages and Literatures at Tsinghua University. She is trying to apply mathematical and computational methods to the intersection field of AI and neuroscience.

Author ORCiD. Yining Zhang, https://orcid.org/0000-0001-5468-9702
Author ORCiD. Ruoxi Sun, https://orcid.org/0000-0002-9447-2185

https://doi.org/10.1017/S0958344022000246 Published online by Cambridge University Press