Who Wants to Learn English Online for Free?

Rubén Chacón-Beltrán, Raymond Echitchi

Universidad Nacional de Educación a Distancia

Correspondence concerning this article should be addressed to Rubén Chacón-Beltrán, Universidad Nacional de Educación a Distancia, C. de Bravo Murillo, 38, 28015 Madrid, Spian.
E-mail: rchacon@flog.uned.es

This study of demographics is aimed to help LMOOC designers develop courses that are more appealing to prospective learners, and thus fight low completion rate which remains one of the main drawbacks of MOOCs. In addition, as the world battles against the COVID-19 pandemic, looking for alternative learning approaches is unavoidable. The data presented in this paper were collected between 2016 and 2020 by means of a questionnaire that over 29,000 participants completed upon registration. The questionnaire, which included three multiple-choice questions aimed at obtaining responses regarding age, level of education and gender, revealed that most learners were middle-aged adults who held a university degree. In addition, our findings seemed to indicate that female learners are more likely to take the courses than their male counterparts. The aforementioned findings, which provide an insight into the demographics of EFL MOOCs in Spanish-speaking contexts, are a good starting point for further research which could ultimately help educational authorities know the impact of EFL MOOCs and enable the latter to reach a wider audience.

Keywords: LMOOCs, English Language, Distance Learning, Demographics

Introduction

This study seeks to provide an insight into the profile of learners enrolled in Language Massive Open Online Courses (LMOOCs) through the analysis of learner demographics, namely age, level of education and gender. It comes in a context where traditional education has been hit hard by the global COVID-19 pandemic thus making educational authorities push for a transition towards distance learning. The contribution of English as a Foreign Language (EFL) MOOCs to fostering distance education is possible, so long as extensive research is carried out in order to find out more about the people who turn to those courses to learn English. On the surface, one of the major challenges the design of LMOOCs poses is the heterogeneity of participants (Bárcena -Madera & Martin-Monje, 2014; Martín-Monje et al., 2018) and the resulting difficulty in designing contents that suit all users. Though a wide range of factors might affect learners' motivation to complete the courses (Jordan, 2014; 2015; Martín-Monje et al., 2018; Li, 2019), we believe course designers' inability to come up with materials that appeal to a diversity of students is one of the main reasons why learners end up dropping out.

This article analyses two LMOOCs aimed specifically at learners of English as a Foreign Language. Our research was motivated by the belief that the analysis of learner demographics in both MOOCs could provide valuable information which could eventually help educational authorities develop MOOCs that are appealing to a wider audience. Further motivation for this study was the scarcity of publications addressing specifically MOOC demographics in Spanish-speaking contexts, apart from a few ones (Alario-Hayos et al., 2014; Martín-Monje at al. 2018) which emphasized engagement rather than personal background. Therefore, to get more information regarding LMOOC learners' personal background, three variables were taken into account, namely, age, level of education and gender. Then, results were compared to similar publications, and this led to conclusions that we believe could guide course designers and educational authorities. Therefore, after reviewing the literature on MOOCs and presenting the method we used, we shall present our findings and discuss them in an attempt to answer the following research questions:

RQ1.   Are learner demographics in our courses different from other MOOCs'?
RQ2.   Do learner demographics vary depending on course level?
RQ3.   What was the impact of COVID-19 on learner demographics?

Literature review

cMOOCs and xMOOCs

The development of the first MOOCs was motivated by the need to make education available to everyone through the creation online platforms that broke away from traditional learning settings where students were seldom actively involved in the learning process. Cormier (2008) coined the acronym MOOCs, which he used to describe Stephen Downes and George Siemens's "Connectivism and Connective Knowledge" online course. Since then, MOOCs have evolved and most of them now seem to be different from Downes and Siemens's course.

MOOCs are often divided into two broad types, namely cMOOCs and xMOOCs (Veletsianos & Shepherdson, 2016). CMOOCs, with "c" standing for "connectivist", were designed with a view to breaking away from traditional pedagogy. On cMOOCs platforms there are no teachers or specific curricula per se, but all members of the online community can serve as facilitators by eliciting discussions and sharing knowledge. Early MOOC developers such as Bousquet[1] (2012, cited in Beaven et al. 2014) believed that "good MOOC's" should rather be student-centred and based on a connectivist approach to learning. Therefore, the courses they ran were meant mainly to be platforms where learners would connect and build knowledge through online exchanges and networking with no need to rely on a teacher or specific syllabus.

As concerns xMOOCs,Sokolik (2014, p. 18) explains that their development was inspired by "the open course model originally formed as MITx, which was then joined by other universities, and has evolved into edX.org". Bárcena-Madera and Martín-Monje (2014, p. 8) go further to indicate that the use of 'x' in xMOOCs reflected the desire to make MOOCs "eXtensions", or as Read (2014, p. 99) explains, "continuation[s] of other types of eLearning courses that institutions have undertaken...". While cMOOCs were developed as a reaction against traditional learning approaches, xMOOCs tend to mirror formal learning environments and therefore distinguish between learners and instructors. Furthermore, xMOOCs are often centred around a specific topic and characterised by clear assessment methods which may entitle course participants to a certificate (See Reeves & Helberg, 2014, and Read, 2014 for a contrastive analysis of cMOOCs and xMOOCs).

Nowadays, most MOOCs seem to follow a content-based approach, and this is exactly why the acronym "MOOCs" may very often refer to xMOOCs, as they are more common and have been the subject of most publications in the field. Nevertheless, some researchers still believe that cMOOCs are inherently superior, thus making Sokolik regret that "there has developed a kind of hierarchy, where many consider cMOOCs to be superior in form and function" (2014, p.18). Sokolik further indicates that both xMOOCs and cMOOCs have positive features that can be used to foster learning. CMOOCs' emphasis on building a sense of community and promoting interaction are great advantages in any learning environment, while the traditional learning approaches in xMOOCs may help learners feel comfortable and at ease, when trying to acquire knowledge.

The covert conflict between cMOOCs and xMOOCs may be the reason behind Barcena-Madera and Martín-Monje's (2014) desire to emphasize the use of LMOOCs, which stands for Language MOOCs. LMOOCs, as described by the aforementioned authors, "are dedicated Web-based online courses for second languages with unrestricted access and potentially unlimited participation" (2014, p. 2).Though LMOOCs can be either cMOOCs or xMOOCs, Read (2014) suggests that developing a MOOC that only mirrors a traditional learning setting while completely avoiding the need to connect will prove inefficient. He then concludes that "a middle ground is required that enables a hybrid-xMOOC to be designed, including cMOOC features..." (p. 102). It is therefore not suprising to see that most LMOOCs tend to draw on the strengths of both xMOOCs and cMOOCs to result in what Sokolik refers to as "an eclectic mix of practices and tools aiming to engage students in the use of the target language in meaningful and authentic ways" (2014, p. 20).

---

[1] Bousquet, M. (2012, July 25). Good MOOC's, bad MOOC's. The Chronicle of Higher Education. URL: https://www.chronicle.com/blogs/brainstorm/good-moocs-bad-moocs

Profiles of MOOC Participants

Research on the profiles of MOOC participants has targeted various issues such as user demographics, engagement and achievement. As concerns demographics, the variables that have attracted more scholarly interest include gender, level of education and age. With regard to gender, there have been conflicting reports regarding which gender group would most likely make use of MOOCs. For instance, while Christensen et al. (2013) and Despujol et al. (2014) revealed that 56.9% and 56% (respectively) of their respondents were male, Morris et al. (2015) and Bayeck (2016) on the other hand found women to represent respectively 59.83% and 60% of learners in their courses. Therefore, as MacLeod et al. (2016) indicate, it may not be a good idea to make general claims about gender distribution, as the latter may vary depending on the contents of the course. Courses whose contents are appealing to the male population (due to social factors that go beyond the scope of this paper) would likely attract more males, and the other way around. As concerns level of education, research points to the fact that MOOCs mostly attract university graduates. For instance, Christiansen et al. (2013) studied 32 MOOCs offered on the Coursera platform and found that 83% of their research population had a post-secondary degree, with 79.4% holding a bachelor's degree. Similar findings were reported by Despujol et al. (2014), Dillahunt et al. (2014) and Bayeck (2016). Finally, most researchers agree that MOOC learners are more likely to be adults. For instance, Morris et al. (2015) obtained a mean age of 36.17 years old for participants in their study, while Martín-Monje and her colleagues (2018) found that overwhelming majority of learners in their MOOC were aged between 31 and 50. Finally, most MOOC users in the study carried out by Despujol et al. (2014) were adults whose median age stood at 33 years old. Though most researchers indicate that MOOCs tend to attract learners in their 30's or 40's, some studies have revealed that on some occasions the majority of users could be slightly younger. For instance, in their study of MOOCs in the Republic of Korea, Yong Kim et al. (2019) found that learners were aged between 20 and 29 years old. Furthermore, Christensen et al. (2013) revealed that most students enrolled in the University of Pennsylvania's Coursera MOOCs were younger than 30.

Other publications on MOOC user profiles have focused on classifying them according to engagement. For instance, Grünewald et al. (2013) grouped MOOC participants into *inactive*, *passive*, *reacting*, *acting* and *supervising* learners. After registration, inactive learners would never actually log into the course. Passive learners might go through the course to gain knowledge without partaking in any activities, as opposed to more active learners, who would only react to questions asked by their peers (reacting), initiate discussions through various means (acting) or moderate discussions and even summarise what has been learned from such discussions (supervising). The aforementioned classification is not very different from that of Hill (2013), who divided MOOC users into five self-explanatory categories, namely *no-shows*, *observers*, *drop-ins*, *passive* and *active* participants. Alario-Hoyos et al. (2014) slightly modified this categorization by talking about *non-engaged* and *engaged*, rather than passive and active participants. She went on to add two more categories, namely *latecomers* and *drop-in latecomers*. Finally, Martín-Monje et al. (2018), used Anderson et al.'s (2014) taxonomy to classify LMOOC students into five categories, namely *all-rounders* (who watch all videos and complete most assignments), *viewers* (who watch course videos but rarely complete assignments), *solvers* (who complete assignments but rarely interact with videos), *collectors* (who watch some videos and complete some assignments) and *bystanders* (who are completely inactive learners). No matter the taxonomy used, research has proved that most MOOC users tend to be made up of more passive learners than active ones. This definitely leads to low completion rate, which remains a real issue (Jordan, 2014, 2015; Martín-Monje et al., 2018).

Heterogeneity and Low Completion Rates

MOOCs' free and open access unavoidably leads to learner heterogeneity which not only applies to variables such as age, location, level of education, profession, but also affects motivation and engagement amongst learners. In fact, some users may enrol out of mere curiosity without really being interested in the course contents (Cross, 2013; Christensen et al. 2013). Furthermore, lack of time may be a great factor behind low completion rate (Conole, 2013), as taking a MOOC tends to be seen as a side activity, even amongst highly motivated learners. Finally, Onah et al. (2014) suggested that some learners may drop out of MOOCs due to other reasons such as "course difficulty and lack of support", "lack of digital and learning skills" and "bad experiences".

Whatever the case, a careful analysis of the profiles of MOOC participants is never a bad idea. Given the worldwide coverage of MOOCs, knowing more about the people who tend to enrol in specific courses could help instructors design materials that are appealing to their target population. Research on MOOC participants has the potential to make MOOCs more efficient, since having specific learners in mind can indeed contribute to the design of more engaging contents and activities. Finally, getting to know more about the profiles of MOOC users may contribute to the sustainability of the courses and lead to further research on how to incorporate MOOCs into traditional education.

Methodology

Background

This article is centred around two MOOCs that were designed by UNED's Department of Foreign Languages and Linguistics and were first run in 2013. The two courses build on the strengths of both cMOOCs and xMOOCs and are dedicated to the teaching of English to Spanish-speaking learners. All editions of the courses discussed in this paper were hosted on the OpenEdx platform and included video presentations, extra materials, assessment and discussion forums. In addition, course instructors would often intervene to elicit discussions, answer questions and guide learners throughout the month each edition lasted.

Empieza con el Inglés: Aprende las mil palabras más usadas y sus posibilidades comunicativas (beginner English: Learn the thousand most common words and their use) is one of the two courses discussed in this paper and will be referred hereafter as Empieza A1. It targets beginner-level learners of English and is supposed to be completed in 4 weeks. The course is made up of 6 modules where lexical items belonging to common domains of use are presented, translated into Spanish and discussed. The other course studied in this paper is *Starting to Write English with no Mistakes*, (hereafter Starting B1) which is a lower intermediate course whose goal is to introduce students to writing in English. This course consists of 12 modules in which users are taught how to plan, produce and self-correct their essays or other pieces of writing with the help of dictionaries and/or normative corpora such as the British National Corpus.

Apart from their distinct contents, it is also worth mentioning that Spanish is the main medium of instruction and discussion in Empieza A1 whereas English is the main language of teaching in Starting B2.

Participants

Both courses were meant to target specifically learners of English as a Foreign language. Furthermore, as the courses were designed by staff at a well-known university in Spain and advertised mainly on the university website and social media platforms, we assumed participants would likely be native speakers of Spanish or a language spoken in some parts of Spain, such as Catalan, Valencian, Basque or Galician. In fact, registered users' IP addresses, which we were able to obtain, indicated that most of them were based in Spain (72%), while others logged in from Latin American countries like Mexico (2.6%), Peru (1.3%), and a wide range of other locations in Europe or America.

Research instrument and procedure

The data that will be presented and analysed subsequently was obtained through a questionnaire made up of three questions that course participants were asked to complete while registering for the course. In other words, in addition to entering an email address and a password, they would provide information about their age, level of education and gender. While we asked an open-ended question to elicit responses regarding age, closed-ended questions were used for level of education and gender, so learners were provided with a limited number of choices. As concerns level of education, they had to choose between nine options, namely "no formal education", "primary", "middle", "secondary", "associate", "bachelor's", "master's", "doctorate" and "other", while for gender they were provided with three choices, i.e., "female", "male" and "other". The results presented in this paper cover fourteen editions of the courses that were run between April 2016 and April 2020. The 14 editions of the courses added up to a total of 32,133 registered learners and about 93% of them completed the

questionnaire. After downloading those results from the course analytics section on OpenEdx, all we had to do was to compare and constrast the frequencies obtained in Empieza A1 and Starting B1.

Results

Before going any further, we would like to point out that the very nature of MOOCs makes it impossible (or at least very difficult) to generalize the findings of a study such as ours. Nevertheless, we still believe that this research can be a point of reference for researchers interested in the demographics of EFL MOOCs in Spain as most registered users were indeed based in the country.

Age

As we mentioned earlier, studying the age of learners is a suitable way to obtain more information about the demand for a specific type of knowledge. In addition, it can foster the design of materials that are more engaging, and therefore reduce learner dropout. As concerns this study, both Empieza A1 and Starting B1 seemed to attract an audience mostly made up of middle-aged adults as found in Table 1. Overall, the median age in both courses is about 40 years and doesn't really fluctuate over time. Nevertheless, users in Empieza A1 were found to be slightly older than the ones in Starting B1, and this is why the overall median age in Empieza A1 is about 44 years old, while registered learners in Starting B1 were about 40 years old on average.

Table 1

Median age over time

	EMPIEZA A1	**STARTING B1**
Editions		
April/May 2016	44	40
November/December 2017	47	42
May/June 2018	44	42
November 2018/ February 2019	43	38
April/June 2019	43	41
November/December 2019	44	40
April 2020	42	33
Overall median age	**43.85**	**39.42**

A close look at Table 1 reveals a possible outlier in the April 2020 edition of Starting B1, where the median age dropped to 33. This can be explained by the number of learners aged 25 and under, which was larger than usual. The higher turnout of learners under 26 is probably due to the spread of the coronavirus pandemic, which made governments across Europe and beyond impose national lockdowns. This led to the closure of all educational institutions in Spain, which could have made students look for alternative ways to learn English. Whatever the case, the number of learners under 26 in both courses increased gradually over the years even far before the pandemic (See Table 2). Though we are unclear as to what could have caused the increase, this finding seems to indicate that more and more youth may resort to EFL MOOCs to complement formal education. As regards Empieza A1, in April 2016, 4.3% of the users who enrolled in the course were younger than 26, 36.6% were aged between 26 and 40 (both inclusive), and 41% were 41 and older. Though more and more learners under 26 have been enrolling in the courses, they still fall behind older users. Hence, in April 2020, there were 7.1% of users under 26, 39.3% aged between 26 and 40, and 53.6% aged 41 and older. A similar trend was witnessed in Starting B1, where the frequency of learners under 26 increased from 3.1% in the first edition to 20.2% in the April 2020 edition.

Table 2

Percentages of users aged 25 and under

|	EMPIEZA A1	STARTING B1
Editions		
April/May 2016	4.3	3.1
November/December 2017	3.4	4.8
May/June 2018	5.7	6.3
November 2018/ February 2019	7.2	9.9
April/June 2019	7.2	7.5
November/December 2019	8.5	11.3
April 2020	7.2	20.2
Mean percentage	**6.2**	**9**

Another striking finding about age is that Empieza A1 seemed to attract a higher proportion of older participants (aged 41 and above) than Starting B1. This is very obvious when having a look at Figures 1 and Figure 2. While there is some disparity between the number of learners aged between 26 and 40 on the one hand and learners older than 40 in Figure 1, the difference is not that striking in Figure 2. Figures 1 and 2 also reveal that, on the whole, Starting B1 seemed to be appealing to a larger number of younger learners than Empieza A1.

Figure 1

Evolution of users' age in Empieza A1
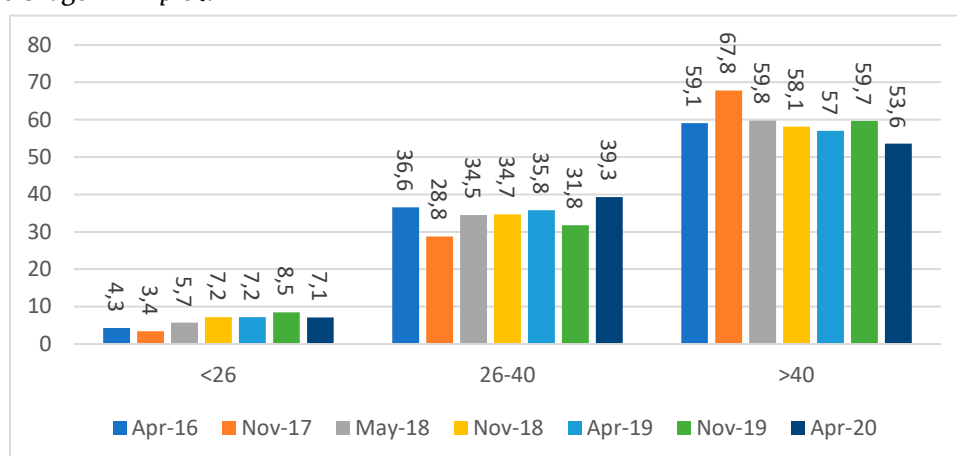
Figure 2

Evolution of users' age in Starting B1
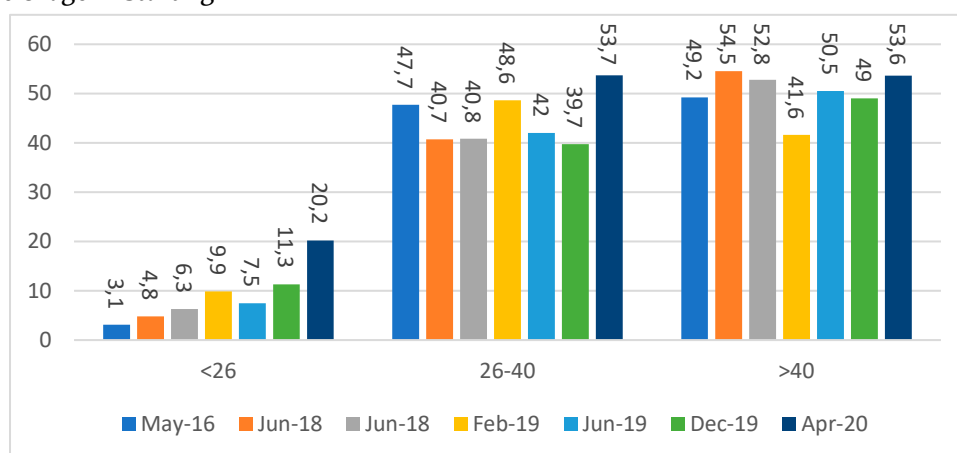
At this juncture, it is important to indicate that even though there were many more older learners in our MOOCs, this does in no way mean that EFL MOOCs could not appeal to larger numbers of users under 26. The fact that they flocked into the courses during lockdown is clear evidence that MOOCs can be an alternative way to teach English, not only to older learners but also to youth. Furthermore, the higher percentage of younger learners in Starting B1 seems to indicate that the growing importance of English in Spain and the fact that its teaching is more widespread than ever before may be contributing to higher proficiency amongst younger learners. Though further research should be carried out to confirm this, we believe EFL MOOCs in Spain may have to emphasize intermediate and advanced proficiency in order to be appealing to many more youths.

Level of Education

This research was also motivated by the desire to know users' levels of education and eventually relate it to the two proficiency levels that each respective course targeted. To achieve this, users were asked to provide information about their highest educational attainment, with options ranging from "primary school" to "doctorate". In addition, people with little or no formal schooling could indicate it by selecting "other" or "no formal education".

The most salient finding here is that across all editions of both Empieza A1 and Starting B2, the majority of participants were university graduates. On average, nearly 60 percent of registered users in both courses held a bachelor's or any other associate degree.

Table 3

Percentages of users holding an undergraduate degree

	EMPIEZA A1	STARTING B1
Editions		
April/May 2016	62.1	60.7
November/December 2017	61	62.7
May/June 2018	59.7	60.8
November 2018/ February 2019	55.8	57.8
April/June 2019	57.7	56.7
November/December 2019	57.3	58
April 2020	58.1	56.9
Mean percentage	**58.8**	**59**

Furthermore, there was a high percentage of registered users who said they hold a master's degree (See Table 4), which might be surprising since we are dealing here with beginner and lower intermediate courses. These findings could thus indicate that tertiary education may not always guarantee proficiency in English in Spanish contexts. When looking closely at the percentages of master's degree holders, one could notice a difference between Empieza A1 and Starting B1, with the latter attracting a higher percentage (27% versus 21.1%, overall). Nevertheless, this difference is not significant enough for one to assume that master's degree holders in Spanish EFL contexts are more likely to be intermediate learners of English than beginners.

Finally, it is important to point out that other researchers who did not target specifically EFL LMOOCs also found most MOOC learners to be university graduates (Christensen et al., 2013; Despujol, 2014, Dillahunt et al., 2014 & Bayeck, 2016). Therefore, it would be too far-fetched to even attempt to establish a clear causal relation between the high presence of bachelor's and master's degree holders in the courses and their English skills.

**Table 4**

Percentages of master's degree holders

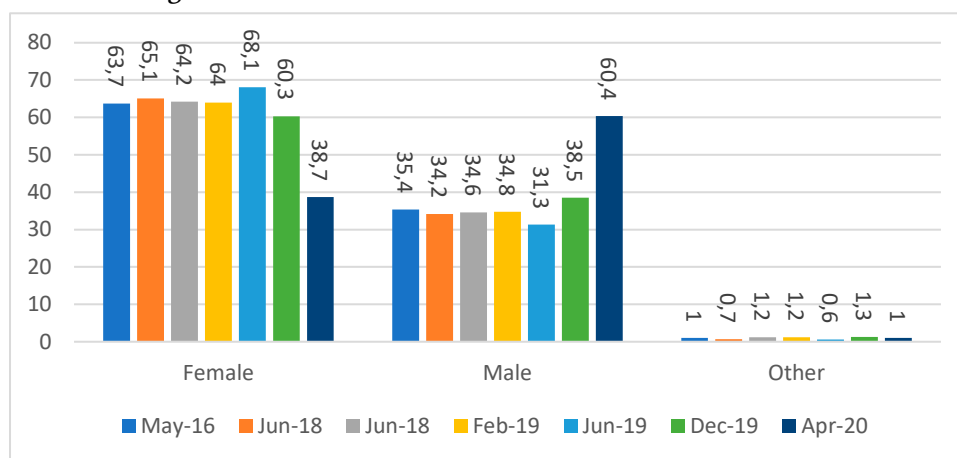|	*EMPIEZA A1*	*STARTING B1*
Editions		
April/May 2016	20.9	26.9
November/December 2017	20.3	26.5
May/June 2018	21.4	25.2
November 2018/ February 2019	22.3	27.8
April/June 2019	21.4	29.5
November/December 2019	21.5	27.5
April 2020	20.4	25.7
Mean percentage	**21.1**	**27**

In general, those participants who hold only a primary or secondary education certificate represent a very tiny percentage of respondents. Many reasons may account for this situation. In Spain, it is compulsory to complete 10 years of education and students are often encouraged to graduate from high school and go to university. Moreover, perhaps primary and secondary school leavers do not find the courses to be very interesting because they are most likely employed in low-skilled jobs that may not require the ability to speak a foreign language such as English. In addition, members of this group may be less aware of ICT educational tools and less skilled when it comes to accessing and making use of them.

To sum up, the findings of this paper indicate that our respondents would most likely be middle-aged individuals who hold a university degree. Their quite advanced age may indicate the need to upgrade their skills in a context where English language proficiency is in great demand.

Gender

The last variable that this research attempted to study was gender. The aim here was to find out which gender group was more likely to enroll in our courses, and compare our findings to what was obtained by other researchers. Therefore, learners in both Empieza A1 and Starting B1 were asked to provide information about their gender, with three options available, namely "female", "male" and "other", which would enable those learners who were not clear about their gender or did not abide by binary approaches to gender to express themselves.

Figure 3

Gender distribution in Empieza A1
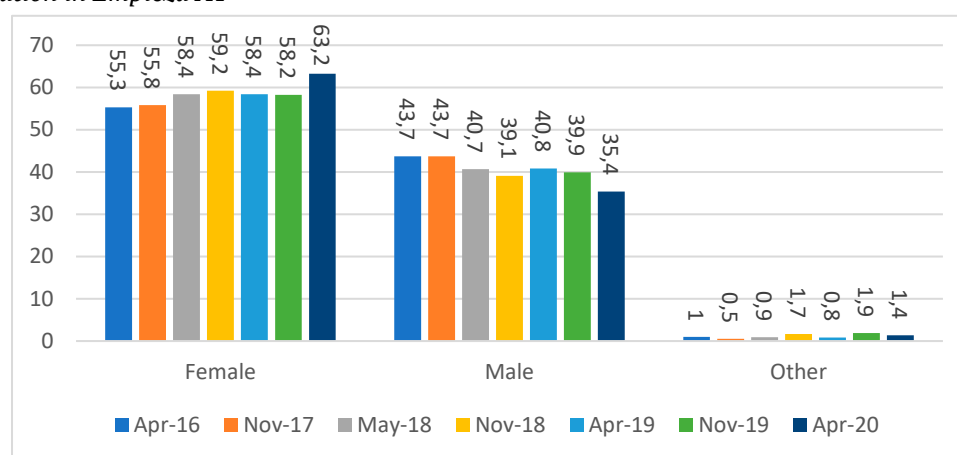
Figure 4

Gender distribution in Starting B1

Findings revealed that overall, women were more likely to enrol in the courses than men. Out of the fourteen editions of both Empieza A1 and Starting B1, men outnumbered women only in one edition of Starting B1 (See Figures 3 and 4). Though these findings concur with those of Morris et al. (2015) and Bayeck (2016), they are contrary to what Christensen et al. (2013) and Despujol et al. (2014) obtained.

It is very difficult to explain the imbalance between the number of men and women in our courses, and this is certainly not because of low demand for beginner and intermediate English amongst Spanish men. In fact, the few studies which correlate English proficiency with gender in Spanish-speaking contexts suggest that female learners actually do better than their male counterparts (Fernandez Fontecha, 2010; Jimenez Catalan, 2010). In addition, the percentage of male learners in the April 2020 edition of Starting B1 (run during the COVID-19 lockdown), which stood at 60.4%, further indicated how difficult making general claims about gender distribution could be. Therefore, we suspect that course content and other socioeconomic reasons beyond age, level of education, employment or proficiency might actually explain gender distribution in MOOCs. Therefore, it would be too risky to make assumptions regarding gender distribution in EFL MOOCs without considering those other variables.

Discussion

This section will attempt to provide answers to the research questions that were asked at the beginning of this paper. In order words, we shall indicate whether or not the findings of this paper are specific to EFL MOOCs, discuss the link between course proficiency level and demographics, and finally discuss the impact of the COVID-19 pandemic on LMOOC learner demographics.

EFL MOOC Vs MOOC Demographics (RQ1)

The findings of this research, which were related to age, level of education and gender did not really clash with what was obtained by researchers who studied MOOCs in other fields of study. The findings of this paper validate previous publications (Christiansen et al., 2013; Despujol et al., 2014, Dillahunt et al., 2014 and Bayeck, 2016) which indicated that MOOCs tend to attract university graduates. Furthermore, just like Morris et al. (2015) and Despujol al. (2014), we found that most MOOC participants are middle-aged adults. Finally, the fact that men inexplicably outnumbered women in one edition of the course is evidence that it is impossible to establish clear patterns relating to gender. All in all, this research has made us understand that EFL MOOCs are not very different from MOOCs in general with regard to their tendency to attract university graduates and adults. In addition, it has consolidated the belief that making general claims regarding gender might not be a good idea (MacLeod et al, 2016).

Course Proficiency Level and EFL MOOC Learner Demographics (RQ2)

This study has confirmed that course proficiency level somehow determined the demographics of the courses we studied. To begin with age, though the bulk of participants in both courses were middle-aged adults, Empieza A1 attracted a larger number of older learners than Starting B1. In addition, the proportion of learners under 26 in Starting B1 skyrocketed in the last two years while it remained quite steady in Starting A1.Though these findings cannot be generalized until similar studies are carried out, it goes without saying that the spread of English education in the past 20 years has indeed had a positive effect on language proficiency amongst younger learners in Spain, making them more likely to take intermediate and advanced proficiency courses.

As far as academic level is concerned, the only thing we can say for sure is that master's degree holders in Starting B1 were greater in number than the ones in Empieza A1, although the difference was not that significant. Therefore, even though it seems postgraduate degree holders would most likely enrol in an intermediate course, it is impossible to make a definite claim in a study like ours.

Finally, gender never really affected course proficiency level. In both the beginner and intermediate course, there were more female than male learners, overall.

COVID-19 and LMOOC Learner Demographics (RQ3)

We included two editions run during COVID-19 because we wanted to compare the demographics obtained before and after the spread of the novel coronavirus in order to find out whether or not the pandemic had affected learner demographics in our courses. While overall there were no significant changes in both Empieza A1 and Starting B1 in terms of level of education, gender distribution and age, the percentage of younger users in the 2020 edition of Starting B1 was quite different from previous editions. Over 20% of learners under 26 enrolled in this edition of the course, which is quite remarkable given that in previous editions the average percentage of learners within that age range was around 9%. It seems that due to lockdown, many students, who in other situations would be busy attending formal educational institutions, decided to enroll in the course in order to improve their English writing skills. As mentioned earlier in this paper, this change affected only the B1 course. This seems to indicate that most younger Spanish learners nowadays already have beginner knowledge of English thanks to greater emphasis on English language teaching in formal educational settings.

The 2020 edition of Starting B1 was also the only one in which men outnumbered women. Nevertheless, it is impossible to correlate this change with COVID-19 because both men and women were affected by the restrictions governments imposed. Therefore, a more logical thing to do is to consider this either as a random happening or the result of other factors that go beyond the scope of this paper.

Conclusion

This paper studied the profiles of learners enrolled in two MOOCs over a period of four years in order to gauge the demand for free online English language courses in Spanish-speaking contexts and draw conclusions that might help course designers develop materials appealing to a wider audience. After studying a population which totalled over 32,000 participants located mostly in Spain, we were able to obtain data pertaining to three main variables that have been used to study population groups, namely age, level of education and gender.

With regard to age, our findings revealed that most learners were individuals in their late thirties or forties. The overall median age of registered learners in both Empieza A1 and Starting B1 might be explained by various reasons. One of them is that until the beginning of the 21st century, the teaching of English as a Foreign Language in Spain was not as widespread as it is nowadays. Furthermore, those who received quality English education at school might have failed to practise the language on a daily basis. If they lived in a context where knowing English was not that necessary, they might have lost their skills over time. This might explain why our sample is made up of an overwhelming majority of university graduates, comprising a large number of undergraduate degree holders and a decent percentage of postgraduate degree holders too, including those who completed a PhD. Finally, this study has proved that both Empieza A1 and Starting B1 tend to be more appealing to female learners, even though it is impossible to explain why women outnumber men in the courses.

As concerns the questions that guided this study, our findings have revealed that EFL MOOCs parallel other types of MOOCs when it comes to age distribution and level of education amongst users. What's more, our analysis of gender also confirmed the unpredictability of this variable, something that was pointed out in our review of the literature on MOOC demographics. The findings of this research also seem to indicate that level of education could somehow influence users' decision to enroll in a beginner or intermediate course. Nevertheless, a more comprehensive study has to be carried out to generalise these results and provide a definite answer to our second research question. Finally, COVID-19 indeed had a clear impact on course demographics, especially as concerns the intermediate course we studied. As a matter of fact, Starting B1 attracted a much larger number of younger learners after lockdown, and we believe those younger learners enrolled in an attempt to find alternative ways to learn English, since schools were closed.

Limitations and bases for further research

It is important to end this article by indicating that our intention was never to make definite claims, which would be very difficult as learner dropout rates in MOOCs tend to be high. We are aware that not all our respondents eventually completed the courses, this is why our results are based on initial interest rather than course completion. In addition, the findings of this study, which are limited to the research sample and the courses we ran, need to be confirmed by further research. Until more studies on the profiles of EFL MOOC users in Spain are carried out, everything stated in this paper should be taken with a pinch of salt. Notwithstanding these limitations, this research might guide educational authorities working on the development of LMOOCs aimed at fostering the study of English as a Foreign Language, especially at beginner and intermediate levels. For instance, one thing that has been confirmed is that it is necessary to find ways to make MOOCs appealing to a wider target population, since university graduates still make up the bulk of registered users. Therefore, it might be a good idea to work towards the promotion and dissemination of such courses amongst people with lower qualifications and the development of contents that might be of interest to them.

Acknowledgements

Declaration of Competing Interest

None declared

References

Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Parada G. H. A., & Muñoz-Organero, M. (2014). Delving into participants' profiles and use of social tools in MOOCs. *IEEE Transactions on Learning Technologies*, *7*(3), 260-266. http://dx.doi.org/10.1109/TLT.2014.2311807

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses. In *Proceedings of the 23rd International Conference on World wide web* (pp. 687–698). ACM. http://dx.doi.org/10.1145/2566486.2568042

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

Bayeck, R.Y. (2016). Exploratory study of MOOC learners' demographics and motivation: The case of students involved in groups. *Open Praxis*, *8*(3), 223–233. http://dx.doi.org/10.5944/openpraxis.8.3.282

Beaven, T., Hauck, M., Comas-Quinn, A., Lewis, T., & de los Arcos, B. (2014). MOOCs: Strinking the right balance between facilitation and self-determination. *MERLOT Journal of Online Learning and Teaching*, *10*(1), 31-43.

Caraker, R. (2016). Spain and the context of English language education. *Research Bulletin*, *92*, 23-35.

Chacón-Beltrán, R. (2014). Massive online open courses and language learning: The Case for a beginner english

course. *Procedia. Social and Behavioural Sciences*, 141, 22-246. http://dx.doi.org/10.1016/j.sbspro.2014.05.042

Chacón-Beltrán, R. (2018). Vocabulary learning strategies outside the classroom context: What adults learn in a technology-based learner-centred environment. *The Language Learning Journal*, *46*(5), 583-593. https://doi.org/10.1080/09571736.2018.1503135

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. J. (2013). The MOOC phenomenon: Who takes massive open online courses and why? *Social Science Research Network Working Paper.* http://dx.doi.org/10.2139/ssrn.2350964

Conole, G. (2013). *Designing for learning in an open world.* Springer. https://doi.org/10.1007/978-1-4419-8517-0

Cross, S. (2013). *Evaluation of the OLDS MOOC curriculum design course: Participant perspectives, expectations, and experiences.* OLDS MOOC Project. http://oro.open.ac.uk/37836/

Despujol, I. M., Turro, C., Busqueis, J., & Canero, A. (2014). Analysis of demographics and results of student's opinion survey of a large scale MOOC deployment for the Spanish speaking community. In *Proceedings of Frontiers in Education Conference* (pp. 1–8). IEEE.

Dillahunt, T., Chen, B., & Teasley, S. (2014). Model thinking: Demographics and performance of MOOC students unable to afford a formal education. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 145–146). ACM. https://doi.org/10.1145/2556325.2567851

Fernandez Fontecha, A. (2010). Gender and motivation in EFL vocabulary production. In R.M. Jimenéz Catalán (Ed.), *Gender perspectives on vocabulary in foreign and second languages* (pp. 93-116). Palgrave Macmillan. https://doi.org/10.1057/9780230274938_5

Grünewald, F., Mazandarani, E., Meinel, C., Teusner, R., Totschnig, M., & Willems, C. (2013). OpenHPI - A case-study on the emergence of two learning communities. In *Proceedings of the Global Engineering Education Conference* (pp. 13–15). IEEE. http://dx.doi.org/10.1109/EduCon.2013.6530277

Jimenéz Catalán, R.M. (2010). Gender tendencies in lexical acquisition and use. In R.M. Jimenéz Catalán (Ed.), *Gender perspectives on vocabulary in foreign and second languages* (pp. 117-138). Palgrave Macmillan. https://doi.org/10.1057/9780230274938

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distance Learning*, *15*(1), 133-160. https://doi.org/10.19173/irrodl.v15i1.1651

Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *The International Review of Research In Open and Distributed Learning*, *16*(3), 341–358. http://dx.doi.org/10.19173/irrodl.v16i3.2112

Li, K. (2019). MOOC learners' demographics, self-regulated learning strategy, perceived learning and satisfaction: A structural equation modeling approach. *Computers and Education*, 132, 16-30. https://doi.org/10.1016/j.compedu.2019.01.003

MacLeod, H., Sinclair, C., Haywood, J., & Woodgate, A. (2016). Massive open online courses: Designing for the unknown learner. *Teaching in Higher Education*, *21*(1), 13-24.http://dx.doi.org/10.1080/13562517.2015.1101680

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. http://dx.doi.org/10.1080/09588221.2017.1378237

Morris, N.P., Hotchkiss, S., & Swinnerton, B. (2015). Can demographic information predict MOOC learner outcomes. In *Experience track : Proceedings of the EMOOC stakeholder summit* (pp. 199-207). Whiterose University Consortium. https://eprints.whiterose.ac.uk/86184/

Oliveira, M., Penedo, A., & Pereira, V. (2018). Distance education. Advantages and disadvantages of [sic] the point of view of education and and society. *Dialogia*, 29, 139-152. http://dx.doi.org/10.5585/Dialogia.n29.7661

Onah, D.F.O., Sinclair, J., & Boyatt, R. (2014). Dropout rates of Massive Open Online Courses: Behavioural patterns. *Proceedings of the 6th International Conference on Education and New Learning Technologies* (pp. 5825-5834). EDULEARN14. https://doi.org/10.13140/RG.2.1.2402.0009

Read, T. (2014). The architectonics of Language MOOCs. In E. Martín-Monje, & E. Bárcena-Madera (Eds.), *Language MOOCs: Providing learning, transcending boundaries* (pp. 91-102). De Gruyter Open.

Reeves, T., & Helberg, J. G. (2014). MOOCs: Let's get real. *Educational Technology*, *54*(1), 3-8.

Sokolik, M. (2014). What constitutes an effective language MOOC. In E. Martín-Monje, & E. Bárcena-Madera (Eds.), *Language MOOCs: Providing learning, transcending boundaries* (pp. 16-32). De Gruyter Open.

Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013–2015. *The International Review of Research in Open and Distributed Learning*, *17*(2), 198–221. https://dx.doi.org/10.19173/irrodl.v17i2.2448

Yong Kim, Ock Tae Kim, Jin Gon Shon. (2019). A Historical journey into K-MOOCs leading to possible

collaborations with North Korea. In K. Zhang, C.J. Bonk, T.C. Reeves, & T.H. Reynolds (Eds.) *MOOCs and open education in the global south: Challenges, successes and opportunities* (pp. 17-27). Routledge. http://dx.doi.org/10.4324/9780429398919-3