

Accessibility of Open Educational Resources: how well are they suited for English learners?

Irina Rets¹, Tim Coughlan¹, Ursula Stickler², Lluisa Astruc²

¹ – Faculty of Wellbeing, Education & Language Studies, Institute of Educational Technology, Open University UK, Milton Keynes

² – Faculty of Wellbeing, Education & Language Studies, School of Languages and Applied Linguistics, Open University UK, Milton Keynes

Abstract

Open Educational Resources aim to offer learning to all, yet the language level used in resources could be a barrier to many potential learners. This paper examines the readability of 200 OER courses in English from two major OER course platforms. We compared the means of readability metrics between these OER courses at different educational levels and subject categories that the platforms offer using inferential statistics as well as cluster analyses. Results prove that there is a progression of difficulty between lower and higher educational levels with introductory courses being easier to read. However, the analysis also highlighted that more than 86% of the courses require an advanced level of English language proficiency. On the other hand, subject matter does not appear to be linked with the readability of the courses.

This study contributes further to the current discussion of the inclusiveness of OER and the factors that hinder its universal use. The study addresses a gap in the literature as, to our knowledge, no other studies have analysed the linguistic accessibility of OER English learners and consideration of the meaning of the educational levels assigned to OER courses has been limited.

Keywords (6): Open Educational Resources; accessibility; readability; English as a Foreign Language; inclusiveness; online learning

Introduction

Open Educational Resources (OER) are *learning, teaching and research materials in any format and medium*, including online courses, that are freely available in the public domain and that are licenced for use and adaptation without cost (William and Flora Hewlett Foundation, 2018, UNESCO, 2019). Thus, this OER definition suggests that OER content is licenced to allow its reuse, redistribution and revision. OER should enable learners to exercise their human right of equitable access to education by decreasing educational costs (UNESCO, 2012). While some other definitions of OER have suggested the commercial sale and use of OER (Downes, 2019), this research specifically builds on the OER definitions of William and Flora Hewlett Foundation (2018) and UNESCO (2019), as “no-cost” provision is an essential component in these definitions of OER, asserting that the purpose of OER is access *for all*. OER can be accessed any time and in any location by learners both in non-formal educational environments and for the purpose of finding support for their formal education.

OER present possibilities for learning at scale, with the major OER platforms receiving millions of visitors each year. Some examples include OpenLearn (2020), which recorded over 4.3 million unique visitors to its courses in 2017-2018. Another example is Saylor Academy (Saylor) (2020) which registers more than 800 thousand total visits on average monthly.

While it has been emphasised that OER can increase educational benefits, particularly in developing countries (UNESCO, 2012), a number of research studies on the use of OER in such contexts are critical of these claims. The predominant use of English as the language of instruction is among the frequently mentioned obstacles to the wider use of OER in developing regions (Hatakka, 2009; Kanwar, Kodhandaraman & Umar, 2010). The reasons for such commonplace use of English in OER courses is due to the fact that a high proportion of OER is created in English from top English-medium universities, as well as due to the increasing global spread of English (Cobo, 2013).

While English is widespread as a first or second language in most countries, counting hundreds of millions of speakers (McGreal, 2017), it is still not understood by the vast majority of the world's people well enough to successfully learn the course content in English (Chapple, 2015; Uchihara & Harada, 2018). This situation does not suit the universal use of OER and instead creates divides. The English language barrier limitation has already been documented among native speakers of Chinese (Huang, Lin & Shen, 2012), Russian (Knyazeva, 2010) and Italian (Banzato, 2012). Thus, for the global English-speaking audience, the linguistic complexity of the English language used in OER is an important concern.

Although greater production of OER in other languages, or translation of English content, would offer a means to increase access, there are currently 7117 languages spoken globally (Ethnologue, 2020). Even with just 23 languages currently accounting for more than half the world's population (Ethnologue, 2020), translating each OER course into 23 languages would require much additional work on the part of OER platforms.

An alternative solution to making English OER more accessible to the global OER audience is reduction of linguistic complexity of OER reading materials to improve their understandability. While the process of making OER linguistically accessible also requires some additional effort from the OER publishers, understanding which linguistic features differentiate between OER at different educational levels and subjects can help further improve automatic text simplification tools, that can potentially be applied to increase OER linguistic accessibility in the future. Linguistic accessibility is already included in a number of quality guidelines and is portrayed as an issue that determines the quality of learning resources. For example, Quality Guidelines for OER in Higher education (UNESCO, 2015) reflect on the importance for the OER to be comprehensible to enable others to successfully use it.

However, notwithstanding calls for more accessible OER as part of quality guidelines, and research findings that suggest the need for a new understanding of access to content that

addresses linguistic barriers, very few studies have explored accessibility of existing OER to English learners. Studies of linguistic accessibility have been conducted in areas such as freely available online patient education materials (e.g. Kher, Johnson & Griffith, 2017; Xie, Wang & Chinnadurai, 2018), however these resources are not OER and reflect concerns of medical, rather than educational outcomes.

Having identified this gap in understanding and the potential of linguistic accessibility to impact on the inclusivity of OER, the purpose of this study is to investigate the linguistic accessibility and difficulty level of OER reading materials. In order to achieve a wide view of current resources, we conducted a readability analysis of OER courses from two major platforms.

Literature review

Reading is one of the major channels of information intake during learning. Some readers may have a higher tolerance of uncertainty when dealing with a text they don't understand but generally "we all tend to lose heart if what we are reading seems to be too difficult" (Harrison, 1980). This quote exemplifies the notion of linguistic accessibility, which is the focus of this research.

In many academic contexts, it could be argued that the language used is intended to be of a higher complexity and require advanced reading skills (Lei & Yan, 2016). However, the intent of OER is to increase access to education, both for learners in developed countries, who are not ready or able to access college education, and also learners in developing countries (Literat, 2015). Thus, there should be a distinction between the expectations of linguistic accessibility in higher level academic texts, and the language used to communicate with learners who should potentially benefit the most from OER. As there are no formal entry

requirements for OER, and OER can be intended to provide a bridge into HE study, the expectations for literacy should not ideally be as high as those of higher education study.

Increasing linguistic accessibility can be one potential solution to overcome the challenges associated with inclusiveness of OER. Linguistic accessibility is closely connected with the notion of readability. The assessment of the readability or the relative ease of comprehension of a given written text can guide material designers and educators when preparing their materials in a way that would facilitate learners' comprehension of the text. Lower levels of reading skills in new students have been found to result in lower course completion rates in formal degree courses (Macdonald & Scott, 1997).

The ability to determine readability of a written text requires an evaluation of linguistic factors (Berndt & Wayland, 2013). Among those features most often used to assess text readability are semantic and syntactic attributes of the text. Research studies consistently find vocabulary a predictor of text complexity, in particular the measures of word length and frequency, with longer and less frequent words increasing text complexity (Harrison, 1980). Word length is measured in syllables per word, and the word frequency by how often the word tends to appear in ordinary usage. A further predictor of text complexity is sentence length, with longer sentences putting a greater load on information processing capacity (Crossley, Greenfield & McNamara, 2008).

Word and sentence length serve as the basis of a range of traditional readability formulas. Among the three formulas that are reported to be suitable for all kinds of texts, that are most commonly used to assess suitability for foreign language learners and for benchmarking to the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001) are the Flesch-Kincaid Grade, Flesch Reading Ease and Gunning Fog index (Textinspector, 2018). Each of these formulas is calculated according to a ratio of total words, sentences and syllables in a given written text.

Although readability formulas offer a quantitative and relatively fast prediction of text complexity, their usage as sole measurements of text complexity has been questioned. The claim has been put forward that readability formulas do not account for all factors related to difficulty and particularly qualitative factors and comprehension factors such as text cohesion (Crossley, Greenfield & McNamara, 2008). Among the recent trends in the study of text readability have been the development of more advanced computer readability tools (e.g. Coh-Metrix Second Language (L2) Reading Index, Textinspector) that combine traditional readability formulas with the analysis of other text features such as lexical diversity of the text, proportion of advanced lexis, level of nominalisation or cohesion (for the list of metrics and their interpretation, see Table 1). Research has shown that such sophisticated automated readability assessment yields better and more accurate results than the traditional readability formulas (Crossley, Greenfield & McNamara, 2008; Xia, Kochmar & Briscoe, 2016).

Readability analysis serves an important practical need as it helps to assess accessibility of reading materials to readers. While a number of studies have been conducted on the accessibility of printed textbook materials (Berendes et al., 2018; Maslin, 2007), the accessibility of OER and specifically accessibility of OER to English learners is underrepresented in the research literature, with most studies on the topic of online material accessibility concentrating in the field of healthcare, as mentioned in the previous section. While health information materials and Wikipedia articles are not necessarily produced by educational institutions, these are similar to OER in that they are online materials hosted on public platforms that aim to inform and educate a wide audience. Therefore, the focus on the linguistic accessibility of these types of materials is also relevant to OER.

These studies (e.g. Betschart et al., 2017; Kher, Johnson & Griffith, 2017; Sanghvi et al., 2012; Xie, Wang & Chinnadurai, 2018) looked at readability assessment of online health materials by triangulating the analysis of the metrics produced by a range of readability

formulas. All of these studies found the language of online health materials too complex for the average native English reader. Recommendations were made in these studies to update the websites hosting these materials to gear them towards improving the accessibility of medical education “as part of a systematic approach to increase reader comprehension” (Xie, Wang & Chinnadurai, 2018, p. 117).

Beyond this, the investigation of the relationship between the educational levels of materials, their subject matter, and their readability, has not received much attention in research. The study of Maslin (2007) compared readability across five top-selling U.S. first grade reading programs using the measures of overall word counts, average words per page, sentence length, the number of unique words, as well as using readability formulas for each passage sampled. The study found that there was a progression of difficulty with passages being easier in the beginning of the year and more difficult at the end. The evidence of grade-level-based complexification was further supported by the study of Berendes et al. (2018) who found that there were significant differences between textbook reading materials of Grades 5/6 and 9/10 for seven of the 10 linguistic features with 9/10 grade materials being more demanding. Three features (word length, ratio of genitive nouns to all nouns, and ratio of derived nouns to all nouns) showed significant differences for all grade comparisons. However, the accuracy of the classification model built in that study for grade-level-based comparison was only 75%.

Some work has directly explored whether subject matter has any influence on readability and this could be of interest to the inclusivity of OER, for example, by identifying subject matter in which particular effort is needed to make texts accessible. The study of Jatowt & Tanaka (2012) compared readability of Wikipedia subject categories, namely Wikipedia articles on biology, chemistry, computing, economics, history, literature, mathematics, and philosophy using the results of applied readability formulas. While different categories produced varying results in terms of their readability levels, the study found that articles in the

computing category were the most readable. However, more research based on wider readability measures and types of material would provide a clearer picture of the relationship between subject matter and readability.

The present study attempts to fill in the current gap in OER accessibility in relation to English learners and investigate the effect educational level and subject matter have on OER readability. As such, the research questions of this study are as follows:

- 1) To what extent are OER courses offered on major platforms accessible to English learners?
- 2) Is there a difference in the readability level of OER according to the stated educational level of the courses?
- 3) Is there a difference in the readability level of OER according to the subject matter?

Materials and methods

The two OER course platforms selected for this study are OpenLearn (2020) and Saylor Academy (Saylor) (2020). Both of these platforms were established in 1999 and currently offer a large variety of open courses through the medium of English. OpenLearn is a UK-based platform and Saylor is based in the U.S.A. OpenLearn offers OER courses across three educational levels – Introductory courses for the learners new to a subject; Intermediate – for the learners who have some familiarity with a subject area, this level corresponds to undergraduate level courses; and Advanced – for learners who want to gain a more critical understanding of a subject, this level corresponds to postgraduate courses (<https://www.open.edu/openlearn/about-openlearn/try>). Saylor assigns courses to five educational levels which generally reflect progression in a particular subject of study. For example, ‘101’ is an introductory course and leads to more advanced courses, such as ‘401’. Courses at level 0 generally indicate ‘remedial courses’ that prepare students for regular

university study. However, personal communication with the Saylor development team (personal communication, December 27, 2018) confirmed that this progression sequence might be rather loose as a topic that is covered in the lower level in one school may be covered in the upper level at another.

To select the courses from these platforms whose reading materials we would analyse as part of readability assessment, we contacted the OpenLearn team and obtained data on the 150 most popular courses on the platform in 2017-2018, in terms of unique visitor numbers to the introductory page of the course. As popularity data was not available for Saylor, we selected 50 courses at random from Saylor, ensuring that there were 10 courses for each of the five educational levels that the platform offers, and a diverse range of subject matters.

Having made a list of 200 courses from the two platforms, we downloaded the reading materials from each course into word documents which we then uploaded into Textinspector (2020) online readability tool. The tool can analyse a maximum of 10000 words from a text, so only the first 10000 words of the reading materials from each course were analysed for the purposes of this study. All the non-text elements (e.g. illustrations, tables, bibliography) were removed from the analysed materials, as the tool does not process such information. Readability assessment metrics produced by Textinspector and used for the analyses are described in Table 1 below.

Table 1. Description and interpretation of applied readability metrics automatically calculated by Textinspector

Readability assessment metric	Characteristics	Interpretation
1. Average sentence length	Average number of words per sentence	Having fewer words per sentence makes the text easier.
2. Type/token ratio (TTR); 3. VOCD 4. MTL	Measures of lexical diversity, proportion of unique vs repeated words in the text	The lower the measures are, the easier the text is.

5. Words with more than 2 syllables – Percentage; 6. Average syllables per sentence; 7. Average syllables per word; 8. Syllables per 100 words	Measures of word length	The fewer syllables the words have on average in the text, the easier the text is.
9. Flesch Reading Ease 10. Flesch-Kincaid Grade 11. Gunning Fog index	Formulas' calculations are based on the analysis of word length vs sentence length	Higher scores of Flesch Reading Ease indicate material that is easier to read. The results of the other two formulas correspond to years of formal education a person needs to understand the text on the first reading. The lower the resulting grade is, the easier the text is.
12. Noun elements per sentence	Average number of noun elements per sentence	Less nominalisation contributes to the text being easier.
13. A1; 14. A2; 15. B1; 16. B2; 17. C1; 18. C2	Proportion of lexis in the text that belongs to each language proficiency level in terms of CEFR	The more A1 lexis and the less C1 and C2 lexis the text has, the easier the text is.
19. 0-6K (combined) 20. 6K-100K (combined)	Measures of word frequency E.g. '1K' means they are the first 1000 most frequently used/used words in English.	The higher the percentage of the numbers before 6, the more frequently used vocabulary the text includes and, thus, the easier the text is.
21. Logical connectives	Measure of text organisation	More cohesive connectives between sentences in a text contribute to the text being easier as they make the links between sentences more explicit.
22. Scorecard	An instant score that refers to the level of the text in terms of CEFR using all readability factors mentioned above.	The Scorecard of above B2 level indicates a difficult text accessible to language learners of the highest level of proficiency.

To answer the first research question concerning accessibility of OER courses to EFL learners, we looked at the Scorecards (as described in Table 1) and counted the number of courses that require more than intermediate level of language proficiency as identified by Textinspector. Descriptions of the CEFR levels used in the analysis are provided in Table 2 below.

Table 2. Structured overview of all CEFR levels related to overall reading comprehension

Advanced	C2	Can understand with ease virtually everything heard or read.
	C1	Can understand in detail a wide range of lengthy, complex texts, identifying finer points of detail including attitudes and implied as well as stated opinions
Intermediate	B2	Can obtain information, ideas and opinions from highly specialised sources within his/her field. Can understand specialised articles outside his/her field, provided he/she can use a dictionary occasionally to confirm his/her interpretation of terminology
	B1	Can identify the main conclusions in clearly signalled argumentative texts. Can recognise the line of argument in the treatment of the issue presented, though not necessarily in detail.
Beginner	A2	Can identify specific information in simpler written material he/she encounters.
	A1	Can get an idea of the content of simpler informational material and short simple descriptions, especially if there is visual support.

Based on: structured overview of CEFR levels for reading comprehension, Council of Europe (2001)

To answer the second research question on the differences in readability levels between the courses at different educational levels, for OpenLearn courses we conducted independent sample T-tests for normally distributed variables and non-parametric Mann-Whitney U tests for non-normally distributed variables with educational levels as independent variables and the readability metrics produced by Textinspector and described in Table 1 as dependent variables. We used 21 variables in total excluding the scorecards as these are non-numeric data. We conducted analysis of variance ANOVA for Saylor courses as there are five educational levels, using the same variables.

We also identified the variables that show significant differences consistently between all educational levels and used these metrics as dependent variables for further analysis when answering the third research question.

The third research question was concerned with the effect of subject matter on the readability of the courses. To answer this question, we first conducted One-Way ANOVA using the fewer readability metrics identified earlier as dependent variables, and subject labels predefined by the selected OER platforms as independent variables. However, as the subject labelling differs between the two OER platforms and these labels contain a rather broad selection of courses (e.g. courses ‘Emotions and emotional disorders’ and ‘The ancient Olympics: bridging past and present’ belong to the same subject label), we also conducted Hierarchical and K-means cluster analysis. All statistical analyses were conducted using SPSS 24.

Given that the platforms differ in the way that they structure subject matter and level, it is worth clarifying that our primary aim in this study was not to compare the platforms. The use of multiple platforms is a means to assess whether the patterns in the findings were consistent and could be considered to have a level of generalisability.

Results

This section provides an overview of our analyses to answer the three research questions stated at the end of the literature review section.

RQ1. Accessibility of OER to English learners

The scorecards automatically produced by Textinspector for each course are presented in Table 3.

Table 3. Scorecards and corresponding percentage of courses for each card from OpenLearn and Saylor OER platforms

Proficiency Level	Proportion of OpenLearn courses	Proportion of Saylor courses
B1+	2%	0%
B2 / B2+	7%	14%
C1/C1+	35%	24%
C2/C2+	56%	62%

As can be seen from Table 3, the minimum level of English proficiency required to be able to follow current OER courses is upper-intermediate (B2 and B2+) level in terms of CEFR. Some courses were identified as intermediate (B1+), however, the percentage of such courses is very small (only 2%) and such courses were found only on one platform. Most courses (91% and 86% respectively on the two platforms) require advanced levels of language proficiency (between C1 and C2+).

RQ2. Difference in readability between courses at different educational levels

T-tests and Mann-Whitney U tests of the readability metrics of the OpenLearn courses at different levels showed that there is significant statistical difference in some metrics between all three levels. Comparison of the lowest (level 1) to the most advanced (level 3) educational level showed statistically significant difference for 14 readability metrics out of 21, intermediate (level 2) to the advanced (level 3) – for eight readability metrics out of 21 and for the lowest to intermediate level – 15 readability metrics out of 21.

The biggest difference was observed between introductory courses (level 1) and advanced courses (level 3) as the effect sizes for the variables that showed significant difference between the levels were bigger – e.g. the effect size for ‘words with more than two syllables, percentage’ is 0.78 – for level 1 vs 2, 0.67 – for level 2 vs 3 and 1.19 for levels 1 vs 3. The smallest difference in readability is observed between intermediate courses (level 2) and

advanced courses (level 3) as there are fewer variables that showed significant difference and the effect sizes for those variables are mostly medium (no effect sizes larger than 0.75).

Comparison of the means of readability metrics showed that level 3 (advanced) courses are the most difficult to read among the three levels and the introductory courses are the easiest. Advanced courses have the highest average sentence length (M=22.94 words per sentence as compared to level 1, M=20.96 and level 2, M=22.55) and word length (M=18.60 syllables per word as compared to level 1, M=1.58 and level 2, M=5.22). Readability formulas show that they are suitable for college level and estimated as ‘difficult to read’ as compared to level 1 courses which were estimated as suitable for 10th-12th grade and ‘fairly difficult to read’. Level 3 courses contain more words of lower frequency (M=21.36 words as compared to level 1, M=19.68 and level 2, M=21.17) and more advanced lexis (M=9.92 words as compared to level 1, M=7.68 and level 2, M=9.08).

However, ANOVA analysis of the readability metrics of Saylor courses showed much less pronounced difference between readability of the courses at different levels as compared to OpenLearn data, where there were statistically significant differences in some variables between all levels. The only statistically significant difference based on Saylor data was observed between courses at ‘remedial’ level 0 vs. courses at levels 1, 3 and 4 (with large effect sizes, all effect sizes were bigger than 1.14).

Comparison of the means of readability metrics of Saylor courses showed that level 0 courses are the easiest to read among the five levels according to the most metrics – they employ shorter sentences (M=19.54 words per sentence as compared to level 3, M=22.47 or level 4, M=21.21), shorter words (M=1.5 syllables per word as compared to level 1, M=1.74 or level 3, M=1.73) and easier words (M=23.22 A1 lexis as compared to level 4, M=16.76 or level 2, M=18.63), more word repetition (M=61.51 of diverse words as compared to level 3, M=80.01 or level 1, M=68.50) and they are suitable for 10th-12th grade students as compared

to the rest of the levels which were estimated as college level and ‘difficult to read’. While there were no variables that showed statistically significant difference between all five levels among Saylor courses, such variables for OpenLearn courses were measures of word length: ‘words with more than 2 syllables, percentage’, readability formulas: ‘Flesh Reading Ease’, ‘Flesh-Kincaid Grade’, ‘Gunning Fog index’ and proportions of advanced lexis: ‘A1’, ‘B2’ and ‘C1’.

RQ3. Difference in readability between courses belonging to different subjects

Having identified variables that show significant statistical differences between educational levels using one data set as described above, we reduced the amount of these variables as some of them measure similar properties of the text. Thus, we used three dependant variables – Flesh Reading Ease which subsumes the measures of word and sentence length, A1 and C1 which subsume the measures of easy and difficult lexis, to analyse the relationship between subject matter and text readability.

We first conducted ANOVA analyses to investigate if there are significant statistical differences in readability between the subject labels used in OpenLearn and Saylor. The analysis showed no statistically significant difference between any of the subject labels on both platforms in the three dependent variables. As the subject classification is not uniform across the two OER platforms and each classification is rather broad, we conducted cluster analysis to gain further evidence on the presence or absence of influence of subject matter on the readability of the courses.

Before starting the cluster analysis we eliminated the courses which were shown as outliers in the boxplot visualisation for the three selected dependent variables. Thus, we continued the analysis with 142 courses (8 outliers removed) from OpenLearn platform and 40

courses (10 outliers removed) from Saylor platform. The readability metrics were normalized in the interval [0, 1].

To decide on the number of clusters, we conducted hierarchical cluster analysis and examined the dendrogram (see excerpts in Figures 3 and 4 in Appendix 1). The hierarchical clustering algorithms indicated between 2 and 7 clusters as the interval to be tested for OpenLearn data and 2 to 6 clusters for Saylor data.

To understand what the best cluster solution is in the identified intervals of 2-7 clusters in OpenLearn data and 2-6 clusters in Saylor data, we conducted K-means analysis for each cluster solution and then identified the intra-cluster similarity and inter-cluster dissimilarities amongst clusters for each solution. The results are shown in Tables 4 and 5.

Table 4. Sum of squares between and within groups for 7 cluster solutions for OpenLearn courses which were identified from ANOVA analysis.

Cluster solution	Between groups	Within groups
2	0.852436	26.33321
3	0.130007	19.811275
4	0.563348	15.27024
5	0.151552	14.844481
6	0.433165	12.623916
7	0.401687	9.991971

Table 5. Sum of squares between and within groups for 6 cluster solutions for Saylor courses which were identified from ANOVA analysis.

Cluster solution	Between groups	Within groups
2	0.000002	12.292566
3	0.109072	5.877895
4	0.259627	6.607204
5	0.402800	6.182531
6	0.872593	3.548964

Having created a line graph for the inter- intra-cluster similarities, we identified at which point the clusters have maximum similarities within the cluster solutions and maximum dissimilarities between the clusters for the two OER platforms.

Figure 1. Intra-cluster similarity and inter-cluster dis-similarities amongst OpenLearn clusters

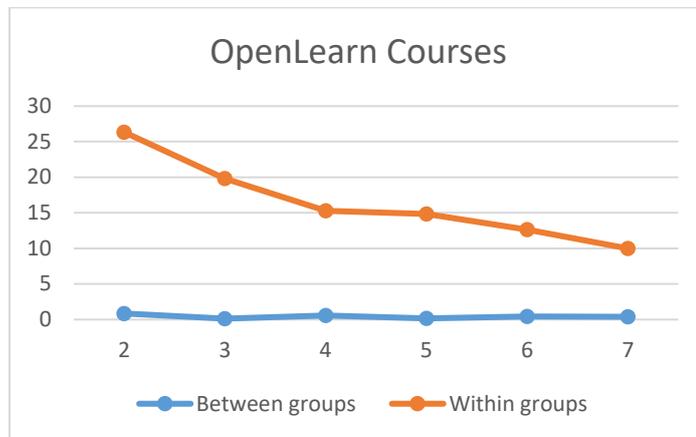
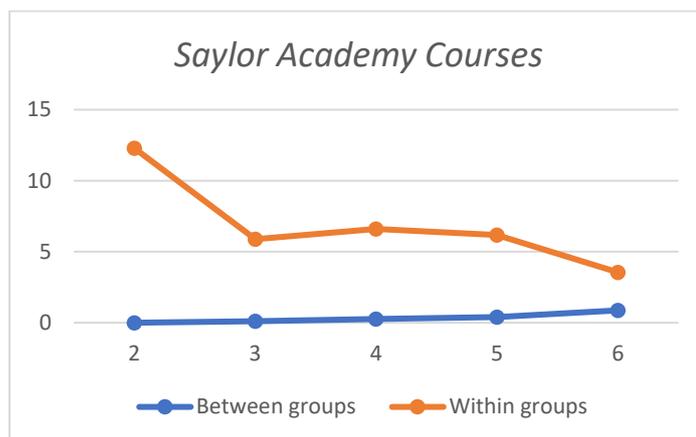


Figure 2. Intra-cluster similarity and inter-cluster dis-similarities amongst Saylor clusters



However, as the difference between the clusters is much smaller than within the clusters on both platforms with this difference being more pronounced with OpenLearn courses – we focused on within cluster similarity, which reduces after the fifth cluster solution both for OpenLearn and Saylor courses. The difference between the clusters also slightly increases after the five clusters which makes the five-cluster solution a good case for the analysis.

Having decided on the five-cluster-solution, we looked into the subjects of the OpenLearn and Saylor courses that were assigned to each cluster. To exemplify, such Open Learn courses as ‘the autistic spectrum: from theory to practise’, ‘organisations and

management accounting’ and ‘exploring the English language’, were assigned to cluster five. As for the fifth cluster with Saylor courses, it has such courses as, for example, ‘microbiology’, ‘introduction to businesses’ or ‘public relations’.

We also examined the descriptive statistics of the ANOVA analysis to identify which clusters are the easiest and the most difficult in terms of their readability. Table 6 below presents the results of the comparison of the means for the three dependent variables between the five clusters with OpenLearn courses.

Table 6. ANOVA descriptive statistics for the five cluster solutions for OpenLearn courses

		N	Mean	Std. Deviation
Flesch Reading Ease	1	42	51.89	6.00203
	2	53	39.64	4.89882
	3	15	57.94	5.37603
	4	7	42.83	7.01798
	5	25	30.59	6.67336
	Total	142	43.76	10.49247
A1	1	42	21.26	1.80811
	2	53	18.94	2.13985
	3	15	26.42	2.19972
	4	7	28.47	2.03405
	5	25	16.61	1.72837
	Total	142	20.48	3.80176
C1	1	42	5.05	.65835
	2	53	6.31	.57244
	3	15	4.34	.69573
	4	7	3.80	.49786
	5	25	7.14	.47367
	Total	142	5.75	1.16284

The table shows that the fifth cluster with OpenLearn courses is the cluster with the most difficult texts in all three measures, while clusters three and four are the clusters with the

easiest texts. Clusters two and one have medium difficulty and are positioned between clusters five and three.

The same type of ANOVA analysis was applied to Saylor course clusters to identify the clusters with the easiest and most difficult texts. Table 7 below presents the results of the comparison of the means for the three dependent variables between the five clusters for Saylor courses.

Table 7. ANOVA descriptive statistics for the five cluster solutions for Saylor Courses

		N	Mean	Std. Deviation
Flesch Reading Ease	1	4	35.38250	6.063065
	2	17	41.92647	3.789909
	3	3	58.41333	1.475477
	4	4	48.49000	2.106166
	5	12	34.93917	5.705148
	Total	40	41.06875	7.895633
A1	1	4	17.32750	.561508
	2	17	18.97706	1.777392
	3	3	23.18000	1.127431
	4	4	19.60000	1.079135
	5	12	15.50250	.960465
	Total	40	18.14725	2.534040
C1	1	4	7.86750	.325103
	2	17	5.72588	.453611
	3	3	4.97667	.351046
	4	4	6.94500	.571577
	5	12	6.49000	.290517
	Total	40	6.23500	.853016

The table shows that the third cluster with Saylor courses is the easiest in all three measures. The fifth cluster is the most difficult in two measures (Flesh Reading Ease and the amount of A1 lexis), and the first cluster is the most difficult in the third measure, amount of C1 lexis.

However, while there are statistically significant differences in the readability measures between the clusters, the clusters are widely mixed in terms of their subject areas. This can be seen from the excerpt from the dendrograms (See Figures 3 and 4 in Appendix 1) where very

different courses subject-wise are positioned close together, and this can further be seen from the courses assigned to the fifth cluster with both OpenLearn and Saylor courses as exemplified previously in this section.

Discussion

This study aimed to contribute to our understanding of accessibility of English language OER to English learners and whether there is any effect of educational level of OER and subject matter on OER accessibility. To that end, we focused on measuring readability of the selected OER materials quantitatively using an advanced online readability tool (Textinspector) that combines traditional readability formulas with the analysis of other semantic and syntactic features of the text as recommended by research literature (Crossley, Greenfield & McNamara, 2008; Xia, Kochmar & Briscoe, 2016).

The first research question of the study was concerned with the extent to which OER courses are accessible to English learners on two popular OER platforms. The analysis of the scorecards automatically produced by Textinspector showed that more than 86% of courses on both OER platforms were only considered suitable for the learners at the highest or advanced level of English proficiency. Thus, OER might not be accessible to English learners who do not read English fluently. This finding supports the results of other research studies on readability of online materials conducted in the field of online healthcare education (Betschart et al., 2017; Kher, Johnson & Griffith, 2017; Sanghvi et al., 2012; Xie, Wang & Chinnadurai, 2018). While those studies were aimed at English native speakers, they also found that the language used was too difficult for an average patient/reader. The findings of this study provide a new form of evidence that current English language in OER is creating a barrier, and prevents those readers who cannot read English to an advanced proficiency to learn through these OER, building on previous research (Banzato, 2012; Cobo, 2013; Hatakka, 2009; Huang, Lin & Shen,

2012; Kanwar, Kodhandaraman & Umar, 2010; Knyazeva, 2010;). As Wiley & Gurrell (2009) pointed out, “if the learner speaks English but only reads at a high-school level and the resource is written with a university-level vocabulary and in an academic style, this resource is not a high-quality resource for that person” (p. 19). A related issue is the question of how accessible is accessible enough? Previous research suggests that a certain threshold of linguistic competence is needed to be able to benefit from linguistic accessibility, and beginner level English learners demonstrate a “language competence ceiling” which prevents them from performing well, even with the texts at increased linguistic accessibility levels (Oh, 2001, p. 87). However, this research found that existing OER require the highest, or advanced, levels of English proficiency. As such, the recommendation from this research is for OER material writers to have not only advanced learners, but also intermediate proficiency-level English learners in mind when designing OER, in order to support this group of learners to access and benefit from their courses and materials.

The second research question was concerned with the effect educational level of the materials has on their readability as both OER platforms under investigation offer materials at levels that require different amount of background knowledge of the subject. On the one hand, independent sample T-tests and Mann-Whitney U tests conducted with the readability metrics on OpenLearn materials showed that there are statistically significant differences between OER at all three given educational levels. These differences concerned the measures of word and sentence length and amount of advanced lexis. This result supports the evidence on the contributors to text difficulty which have also been reported to be word and sentence structure (number of syllables per word and number of words per sentence) as well as word meaning (word rareness and corresponding level of proficiency) (Berendes et al., 2018; Harrison, 1980; Maslin, 2007). On the other hand, ANOVA analysis conducted with the Saylor courses showed that such differences occur only between ‘remedial’ level 0 courses and more senior courses.

The differences in the results between the two platforms might be due to the differences in learning design as OpenLearn website explicitly explains the differences between educational levels contrary to Saylor platform, and the development team of the latter confirmed that the sequence in difficulty progression between the levels might be loose (personal communication, December 27, 2018)

The evidence concerning the effect of educational level on readability that came from the analysis of the reading materials from both platforms showed that courses at the lowest educational level ('remedial' level 0 and introductory level 1 courses) are the easiest to read. This result supports the evidence described in studies of Maslin (2007) and Berendes et al. (2018) which showed there was a progression of difficulty depending on how much exposure the learners had had to the topic.

The third research question of the study was concerned with the effect subject matter has on the readability of OER courses. ANOVA analysis between the subject labels predefined by each platform showed no statistically significant difference between any of the subject categories on either of the platform. This result was further supported by cluster analysis – visual inspection of the dendrograms and the inspection of the course membership of each cluster showed that very diverse courses subject-wise are positioned close together and assigned to one cluster. Thus, both methods of analysis applied to the two OER platforms suggested no effect of subject matter on readability of the courses. This result is not fully in line with Jatowt & Tanaka (2012), who also observed varying results in readability levels of different Wikipedia subject categories but found articles in the computing category to be the easiest to read. As research is scarce on this topic, and the categorisation of courses is a complex activity, more evidence would be needed to draw firm conclusions. However, this study has not identified any significant links between subject matter and OER readability.

Conclusion

This study offers insights into the accessibility of English language OER across two popular OER platforms. While the findings of this study showed that the reading materials at introductory levels on both OER platforms are easier to read, the study demonstrated that the majority of English OER texts at different educational levels and subject categories are only suitable for native speakers or English learners with advanced language proficiency. Taking into consideration various OER guidelines and suggestions from other studies on online material readability, this study makes a case for raising the awareness of educators working in OER context about the current difficulty level of English language OER, and about the gap between many potential OER learners' abilities and the learning materials that purportedly enable inclusive education.

Some limitations arise from the fact that there is no standard combination of readability tests or consensus on the readability metrics that should be used to evaluate the difficulty of the text. The studies with a deep linguistic focus also include such metrics as type of nouns (e.g. genitive nouns, derived nouns), lexical co-referentiality, spatiality or temporal cohesion (Crossley, Greenfield & McNamara, 2008). Multimodal researchers also investigate the effect of some additional factors on readability such as the role of the layout or font types. In this study we used metrics that are best established in readability assessment, automatically provided by an online readability tool and that can be accessed by an audience with no specialised background in linguistics or multimodality.

Among the pedagogical implications drawn from the study is the recommendation for OER material writers to check the text difficulty level of their materials with advanced online readability tools (e.g. Textinspector) prior to publication and to be aware of the linguistic features of the text that contribute to the increase in its difficulty as shown in this and in earlier studies. Such features include primarily the usage of long words and sentences as well as

advanced lexis of low frequency of usage. As we have found evidence that OER with higher assigned educational levels are often more difficult to understand, authors may consider whether or not these really require advanced language use, or if they could be taught using simpler forms of English.

Given the lack of empirical research on the issue of linguistic accessibility of OER to English learners, we hope this study can begin a debate about these issues, and prompt those involved with creating OER to pay greater attention to language use. In order to further understand whether and how increasing linguistic accessibility ‘works’, it will be important to conduct behavioural studies with international OER learners. Further empirical investigation on whether increased linguistic accessibility contributes to increased completion rates of OER courses, and learner satisfaction, will help to increase the global benefits of producing and sharing OER.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported and funded by the Leverhulme Trust, Open World Learning.

References

- Banzato, M. (2012). Barriers to teacher educators seeking, creating and sharing open educational resources: An empirical study of the use of OER in education in Italy. *2012 15th International Conference on Interactive Collaborative Learning (ICL)*, (pp. 1–6).
- Berendes, K., Vajjala, S., Meurers, D., Bryant, D., Wagner, W., Chinkina, M., . . . Graham, S. (2018). Reading Demands in Secondary School: Does the Linguistic Complexity of

- Textbooks Increase With Grade Level and the Academic Orientation of the School Track? *Journal of Educational Psychology*, 110(4), 518-543.
- Berndt, A., & Wayland, J. (2014). Evaluating the readability of marketing research textbooks: An international comparison. *Journal of International Education in Business*, 7(1), 47-59.
- Betschart, P., Zumstein, V., Bentivoglio, M., Engeler, D., Schmid, H., & Abt, D. (2017). Readability assessment of online patient education materials provided by the European Association of Urology. *International Urology and Nephrology*, 49(12), 2111-2117.
- Chapple, J. (2015). Teaching in English Is Not Necessarily the Teaching of English. *International Education Studies*, 8(3), 1-13.
- Cobo, C. (2013). Exploration of Open Educational Resources in Non-English Speaking Communities. *International Review of Research in Open and Distance Learning*, 14(2), 106-128.
- Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge University Press.
- Crossley, S., Greenfield, J., & McNamara, D. S. (2008). Accessing Text Readability Using Cognitively Based Indices. *TESOL Quarterly*, 42(3), 475-493.
- Downes, S. (2019, December 5). *Evaluation of CL/4291 "Draft text of the Recommendation concerning Open Educational Resources"*, UNESCO. [Blog post]. Retrieved 2 March 2020 from <https://www.downes.ca/cgi-bin/page.cgi?post=70211/>
- Ethnologue (2020). Online database of the language of the world. Retrieved 2 March 2020 from <https://www.ethnologue.com/guides/how-many-languages>
- Harrison, C. J. O. (1980). *Readability in the classroom*. Cambridge University Press

- Hatakka, M. (2009). Build it and they will come? – Inhibiting factors for reuse of open content in developing countries. *The Electronic Journal of Information Systems in Developing Countries*, 37(0). Retrieved 10 June 2018 from <http://www.ejisdc.org/ojs2.../index.php/ejisdc/article/view/545>.
- Huang, W.-H. D., Lin, M.-F. G., & Shen, W. (2012). Understanding Chinese-speaking open courseware users: a case study on user engagement in an open courseware portal in Taiwan. *Open Learning: The Journal of Open, Distance and e-Learning*, 27(2), 169–182.
- Jatowt, A., & Tanaka, K. (2012). Is wikipedia too difficult?: Comparative analysis of readability of wikipedia, simple wikipedia and britannica. *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, (pp. 2607-2610).
- Kanwar, A., Kodhandaraman, B., & Umar, A. (2010). Toward sustainable open education resources: A perspective from the global south. *The American Journal of Distance Education*, 24(2), 65-80.
- Kher, A., Johnson, S., & Griffith, R. (2017). Readability assessment of online patient education material on congestive heart failure. *Advances in preventive medicine*, 2017.
- Knyazeva, S. (2010). Open educational resources in CIS: Preliminary study of the state-of-the-art and prospects for development. BSU. Retrieved 13 October 2018 from <http://elib.bsu.by/handle/123456789/13100>
- Lei, L., & Yan, S. (2016). Readability and citations in information science: evidence from abstracts and articles of four journals (2003–2012). *Scientometrics*, 108(3), 1155-1169.
- Literat, I. (2015). Implications of massive open online courses for higher education: mitigating or reifying educational inequities? *Higher Education Research & Development*, 34(6), 1164-1177.

- Macdonald, R. M., & Scott, B. (1997). A postal survey of OU students' reading skills. *Open Learning: The Journal of Open, Distance and E-Learning*, 12(2), 29-40.
- Maslin, P. (2007). Comparison of Readability and Decodability Levels across Five First Grade Basal Programs. *Reading Improvement*, 44(2), 59-75.
- McGreal, R. (2017). Special report on the role of open educational resources in supporting the sustainable development goal 4: Quality education challenges and opportunities. *The International Review of Research in Open and Distributed Learning*, 18(7).
- Oh, S. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, 35, 69–96.
- OpenLearn. (2020). Free Learning from the Open University. [OER courses website]. Retrieved March 2, 2020, from <https://www.open.edu/openlearn/>
- Sanghvi, S., Cherla, D., Shukla, P., & Eloy, J. (2012). Readability assessment of internet-based patient education materials related to facial fractures. *Laryngoscope*, 122(9), 1943-1948.
- Saylor Academy. (2020). Free online courses. [OER courses website]. Retrieved March 2, 2018, from <https://www.saylor.org/>
- Textinspector (2020). Advanced Readability Tool. Retrieved 11 March 2020, from <https://textinspector.com/>
- Uchihara, T., & Harada, T. (2018). Roles of Vocabulary Knowledge for Success in English-Medium Instruction: Self-Perceptions and Academic Outcomes of Japanese Undergraduates. *TESOL Quarterly*, 52(3), 564-587.
- UNESCO. (2012). Paris Declaration on Open Educational Resources. Retrieved 1 June 2018 from http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/Paris%20OER%20Declaration_01

- UNESCO. (2015). Guidelines for open educational resources (OER) in higher education. Commonwealth of Learning. Paris. Retrieved 8 September 2018 from <https://unesdoc.unesco.org/ark:/48223/pf0000213605>
- UNESCO. (2019). Draft Recommendation on Open Educational Resources. Retrieved 2 March 2020 from <https://unesdoc.unesco.org/ark:/48223/pf0000370936>
- Wiley, D., & Gurrell, S. (2009). A decade of development *Open Learning: The Journal of Open and Distance Learning*, 24 (1), 11–21.
- William and Flora Hewlett Foundation (2018), Open Educational Resources. Retrieved 3 April 2018 from <https://www.hewlett.org/strategy/open-educational-resources/>
- Xia, M., Kochmar, E., & Briscoe, T. (2016). Text readability assessment for second language learners. In *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 12-22).
- Xie, D. X., Wang, R. Y., & Chinnadurai, S. (2018). Readability of online patient education materials for velopharyngeal insufficiency. *International journal of pediatric otorhinolaryngology*, 104, 113-119.

Appendix 1. Excerpts of dendrograms of OpenLearn courses arising from the hierarchical clustering algorithm, with Euclidean distance that were visually inspected to decide on the number of clusters.

Figure 3. Excerpt from the dendrogram of OpenLearn courses, arising from the hierarchical clustering algorithm, with Euclidean distance



Figure 4. Dendrogram of Saylor courses, arising from the hierarchical clustering algorithm

