

6 Building bridges not walls – Wikipedia in Translation Studies

Ewan McAndrew¹ and Lorna Campbell²

Abstract

Translation Studies MSc students at the University of Edinburgh take part in a Wikipedia translation assignment as part of their independent study component. The students make use of the free and open encyclopaedia's Content Translation tool which enables them to create translations side-by-side to the original article and automates the process of formatting the page. By providing a more user-friendly experience, translators can focus on creating high-quality content that reads naturally and fluently. Course leaders were keen that the students undertake much-needed published translation practice each semester to bridge the gap between academic study and the world of work ahead of their dissertations. This chapter explains how the project was structured and delivered.

Keywords: Wikipedia, translation, University of Edinburgh, content translation.

1. Context of the project

“Wikipedia is about building bridges, not walls” (Wales, 2016).

Translation Studies is a one-year, full-time taught Masters programme at the University of Edinburgh which aims to enhance “practical skills in, and

1. University of Edinburgh, Edinburgh, Scotland; ewan.mcandrew@ed.ac.uk

2. University of Edinburgh, Edinburgh, Scotland; lorna.m.campbell@ed.ac.uk; <https://orcid.org/0000-0001-6767-856X>

How to cite this chapter: McAndrew, E., & Campbell, L. (2019). Building bridges not walls – Wikipedia in Translation Studies. In A. Comas-Quinn, A. Beaven & B. Sawhill (Eds), *New case studies of openness in and beyond the language classroom* (pp. 85-100). Research-publishing.net. <https://doi.org/10.14705/rpnet.2019.37.968>

theoretical understanding of, translation as an activity” (The University of Edinburgh, 2019a, n.p.). Between 2016/2017 and 2018/2019, 20 to 30 students registered annually in the Translation Studies MSc, which supports a wide variety of languages (Arabic, Chinese, Danish, French, German, Japanese, Norwegian, Spanish, Swedish, and Turkish in 2018/2019).

The Wikipedia assignment is an elective component of the programme’s independent study course, which runs as a semester-long assignment in the first and second semesters. Students are free to choose which language pair they work in but are encouraged to change language direction in Semester 2. The students are supported by the University of Edinburgh’s ‘Wikimedian in residence’; a digital skills trainer employed by the university’s information services division in partnership with, and supported by, Wikimedia UK as the national chapter of the non-profit Wikimedia Foundation. The residency is a free service to support staff and students to learn about, contribute to, and benefit from the Wikimedia Foundation’s family of Open Knowledge projects, of which Wikipedia, the free and open encyclopaedia, is by far the best known. The assignment came about as a result of a practical translation workshop at the Wikimedia Foundation’s annual conference, Wikimania, held in Esino Lario in Summer 2016. During the workshop, the University of Edinburgh’s Wikimedian in residence, Ewan McAndrew, was introduced to Wikipedia’s Content Translation tool, which is designed to better enable knowledge exchange by taking the headache out of formatting pages when translating between different language Wikipedias. The Wikimedian in residence shared this knowledge with course leaders on the Translation Studies MSc, who agreed to trial a Wikipedia translation assignment on the course programme in the first semester of the 2016/2017 academic year. This assignment has continued for the last six semesters.

2. Intended outcomes

The formatively assessed independent study course is intended to give students much-needed authentic translation experience of 2,000-4,000 words each semester before they enter the world of work. Course leaders were keen to

motivate students to complete this translation practice as it was a core objective of the Masters programme. The prospect of the students being able to actually *publish* their work online as a clearly demonstrable and lasting output of their studies, as well as one that could be added to and improved over time as a community project, was thought to be a great opportunity in keeping with the [University of Edinburgh's \(2019b\)](#) vision “to make a significant, sustainable and socially responsible contribution to the world” (n.p.).

A study by [Selwyn and Gorard \(2016\)](#) found that 87.5% of students were finding Wikipedia ‘academically useful’ in an introductory or clarificatory role. Course leaders were mindful that students were *already* using Wikipedia and therefore should be supported in developing good habits in terms of the necessary digital research skills ahead of undertaking their dissertation.

Wikipedia is the largest reference work on the internet, with 49 million articles in 302 different languages. Google’s algorithm ranks Wikipedia articles so they routinely appear in the first page of search results. There is real agency to Wikipedia editing in terms of being able to surface knowledge between different languages and cultures. Addressing areas of underrepresentation and building understanding between different languages is particularly important when one considers how unevenly knowledge is spread between the approximately 302 different language Wikipedias (see [Figure 1](#) below).

Following recent Wikipedia translation projects elsewhere ([Al-Shehari, 2017](#); [Martínez Carrasco, 2018](#)), this Wikipedia translation assignment was introduced to help motivate a diverse group of Masters students, working on a wide range of language combinations, to translate 2,000 words individually in each semester. The objective was to let them see the value of sharing their scholarship in a published context as a significant and demonstrable output of their studies that would last beyond the life of the assignment. This chapter further evidences how moving to a student-centred learning environment can improve motivation and further explores how “introducing collaborative projects with genuine outcomes, we can allow students to coherently develop the competences required for professional translators” ([Al-Shehari, 2017](#), p. 371).

Figure 1. Screenshot excerpt from the List of Wikipedias in March 2019³

Language	Language (local)	Wiki	Articles	Total	Edits	Admins	Users	Active users	Images	Depth
English	English	en	5,817,966	47,283,382	881,932,743	1,181	35,837,885	141,047	881,615	947.45
Cebuano	Cebuano	ceb	5,379,102	9,087,479	25,633,974	6	56,028	181	0	1.34
Swedish	svenska	sv	3,751,809	7,701,516	45,135,306	66	660,377	2,813	0	6.5
German	Deutsch	de	2,279,339	6,392,173	185,121,871	189	3,138,185	20,448	129,451	94.29
French	français	fr	2,086,956	10,048,896	156,864,382	157	3,390,333	19,413	57,305	227.2
Dutch	Nederlands	nl	1,959,911	4,059,675	53,183,907	44	980,934	4,323	22	15.04
Russian	русский	ru	1,533,261	5,831,384	98,213,350	84	2,483,999	11,977	216,029	132.35
Italian	italiano	it	1,511,695	6,121,983	102,869,792	111	1,792,104	8,799	140,079	156.29
Spanish	español	es	1,509,273	6,599,133	113,972,121	73	5,345,110	16,998	0	196.42
Polish	polski	pl	1,323,172	2,984,382	55,886,431	103	944,189	4,588	265	29.52
Waray	Winaray	war	1,263,531	2,877,124	6,197,874	3	39,479	86	42	3.51
Vietnamese	Tiếng Việt	vi	1,203,627	14,481,130	50,198,752	21	653,533	1,741	30,179	421.83
Japanese	日本語	ja	1,142,242	3,383,476	71,704,601	42	1,451,631	13,808	85,376	81.59
Chinese	中文	zh	1,048,137	5,702,234	53,181,011	80	2,691,937	8,061	51,087	183.89
Portuguese	português	pt	1,018,719	4,796,116	54,286,590	81	2,219,321	5,631	51,622	155.57
Ukrainian	українська	uk	891,048	2,689,912	24,549,960	45	444,514	3,303	98,954	37.2
Arabic	العربية	ar	704,453	4,566,995	33,444,179	29	1,632,204	4,392	32,422	220.16
Persian	فارسی	fa	668,474	4,294,017	25,595,448	33	807,083	5,089	55,055	175.34

3. Nuts and bolts

3.1. Setting the context – empowering language activism

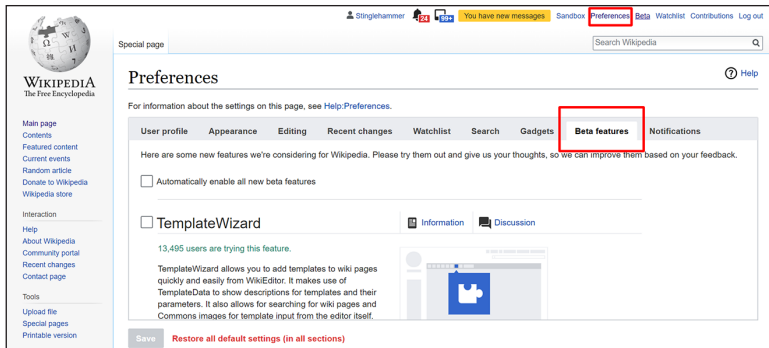
Two workshops at the beginning of the semester were facilitated by the Wikimedian in residence. The first workshop session introduces the students to the work of the Wikimedia Foundation, Wikipedia’s main policies and guidelines, and how to create a user page using Wikipedia’s new easy-to-use What-You-See-Is-What-You-Get (WYSIWYG) visual editor interface. Next, the Content Translation tool must be enabled using the beta menu at the top right of any Wikipedia page (see [Figure 2](#) below).

The first workshop closes with best practice on how to select a good article to work on so students can select one ahead of the next workshop. The students must *each* produce a translation of 2,000 words but can decide whether they will work individually, in pairs, or in groups to achieve this. They choose the

3. https://en.wikipedia.org/wiki/List_of_Wikipedias

language pair and subject matter, which should be one that interests them and will benefit readers when the translation is published.

Figure 2. Enabling content translation in the beta menu in preferences



3.2. Article selection

There are a number of criteria to be met to ensure the completion of the assignment goes smoothly. The chosen article needs to be:

- one that exists in the source Wikipedia and not in the target language Wikipedia (due diligence is required to check that the article does not already exist under another title);
- 2,000 words or more in length OR multiple articles totalling 2,000 words;
- well-referenced, with inline citations throughout (poorly referenced articles are likely to be deleted if they are published in a new Wikipedia);
- of suitable content and subject matter; and
- reasonably well structured and using language that represents a worthy level of linguistic challenge.

Chapter 6

To this end, students are asked to look first at the featured articles on Wikipedia (the highest quality standard on Wikipedia) and the good articles (the second highest level of quality) to see if any are suitable for translation into another language. These featured and good articles are community reviewed to ensure they are of sufficient quality to appear on Wikipedia’s front page (see [Figure 3](#) below).

Figure 3. Featured articles are accessible to view from Wikipedia’s front page⁴



The students are also shown how to use other open tools to find potential articles, such as:

- Wikipedia categories, such as ‘articles needing translation from foreign-language Wikipedias’⁵;
- Wikipedia portals;
- ‘Gapfinder’ tool⁶;

4. https://en.wikipedia.org/wiki/Main_Page and https://en.wikipedia.org/wiki/Wikipedia:Featured_articles

5. https://en.wikipedia.org/wiki/Category:Articles_needing_translation_from_foreign-language_Wikipedias

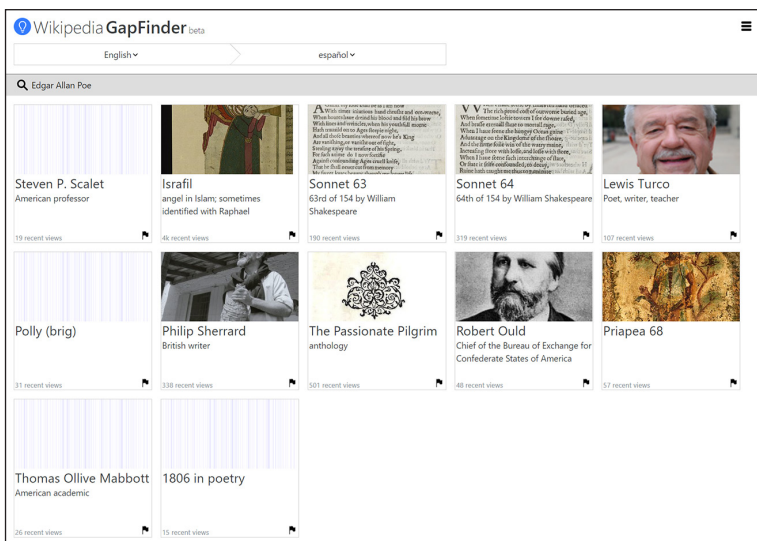
6. <http://recommend.wmflabs.org/>

- ‘Not in the other language’ tool⁷.

The goal of the Gapfinder tool’s developers was to encourage Wikipedia editors to create important articles that are missing in their field of expertise/interest and in the languages they speak. The Gapfinder system (see [Figure 4](#) below) consists of three components:

- finding which articles are missing in the first place by comparing content for every language pair in Wikipedia;
- ranking them by their importance, by predicting pageviews that the article would have, had it existed in the destination language; and
- recommending the most important articles to the best-suited editors.

Figure 4. The Gapfinder tool recommending missing articles from English to Spanish Wikipedia for the search term ‘Edgar Allan Poe’

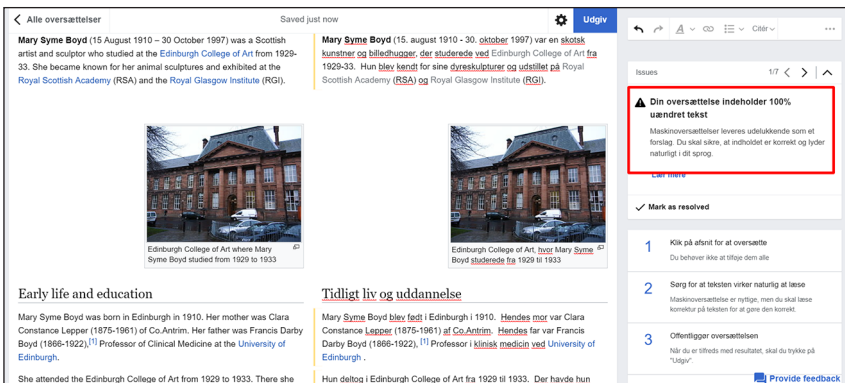


7. <https://tools.wmflabs.org/not-in-the-other-language/>

So that no student makes a poor choice of article to translate, they are given advice about how to assess article quality through paying attention to the use of inline citations throughout and the article’s rating on Wikipedia’s quality scale (stipulated on every article’s Talk page). Potential articles are forwarded to the Wikimedian in residence and the teacher who approve them in collaboration, with the teacher assessing the suitability and linguistic complexity of each on a case-by-case basis. Once approval is received from both, then the article link and language pairs are added to the assignment page so the resident, teachers, and other students know which article each student is working on so any confusion (or duplication) is avoided.

The second two-hour workshop – normally held seven to ten days after the first workshop – is when the students get introduced to the Content Translation tool (see Figure 5 below) and begin the process of translating their article. This tool allows editors to create translations right next to the original article and automates the boring steps: copying text across, looking for corresponding links and categories, etc.

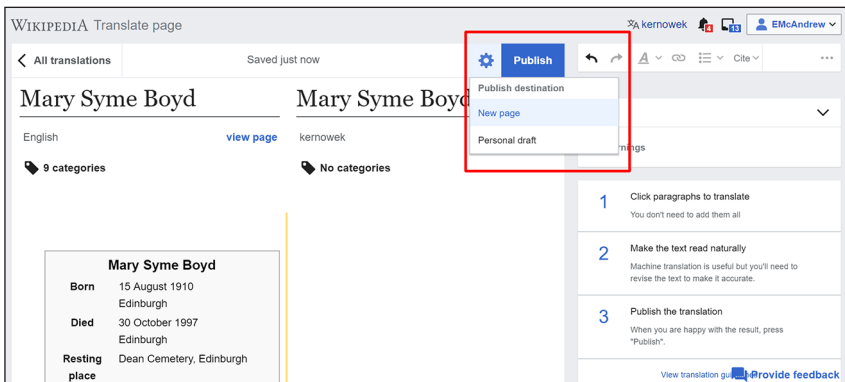
Figure 5. The Content Translation tool



The tool does a lot of the formatting work, pulling in text from the source text and automatically providing a machine translation version for most, though currently not all, language pairs, thus allowing students to focus on editing

the text to create a translation that will read naturally in the target language. It also warns when the text contains too much machine translation (as shown at 100% in Figure 5 above) as it is Wikipedia consensus that an unedited machine translation is worse than having no article at all. Unedited machine translations are likely to be deleted on publishing. The tool does not have access to the visual editor dropdown menus at present, so students are encouraged to complete all the automated formatting migration, paragraph-by-paragraph, and then publish the resulting article to their personal draft space in the target Wikipedia so that more editing checks can be done ahead of publishing to the live space (see Figure 6 below). The cog wheel icon allows users to publish directly to the live article space in the target language Wikipedia or a draft space so that more editing work can be done there before publishing.

Figure 6. Publishing from the Content Translation tool



The second workshop provides students with an opportunity to see how to begin translating, check whether machine translation is available for their language pairs, and set up a draft space on the target language Wikipedia for working on the article outside of the Content Translation tool.

Drop-in clinics are offered every two weeks in a computing lab for students to ask the Wikimedian in residence formatting questions, otherwise students are left to work independently on the assignment.

The students are required to publish their translations by the end of the penultimate week of the semester, and to notify the Wikimedian when they are ready to do this, in case there are any issues at the point of publication. The students also upload Microsoft Word documents of their source article and translated article, complete with word count, to the institution's virtual learning environment.

3.3. Learning points – developing a successful methodology

To ensure a successful assignment, it is important to have someone who is 'Wikipedia-literate' to support the students and ensure they focus their efforts on *translating* rather than formatting articles. The Wikimedia Foundation has national chapters around the world who can be contacted as a first port-of-call for discussing the Wikimedian in residence model or in order to identify local Wikimedia volunteers who could support such an assignment.

Selecting an appropriate article at the beginning of the assignment is important to ensuring the translation does not run into issues on publishing, e.g. students publishing articles that already exist in the target Wikipedia, that are very similar to other articles, that have been worked on by their classmates, or have unreferenced sections. Following the criteria outlined above and making use of featured articles, good articles, and the Gapfinder tool are recommended to avoid such issues.

Checking the word count of the source article is useful for students to ascertain whether the translation will be an appropriate length for the assignment. While the Search tool can be used to look up articles, it includes references in its word count so it is not accurate enough for the purposes of the assignment. Students are advised instead to copy the main body of the article's text, excluding notes, references, bibliographies etc., into a Word document to determine a more accurate word count. Alternatively, the Word Count tool (see [McAndrew, 2017](#)), created by Dr Alex Chow of the University of Edinburgh, can be used to count the words of the main text of any article on English Wikipedia.

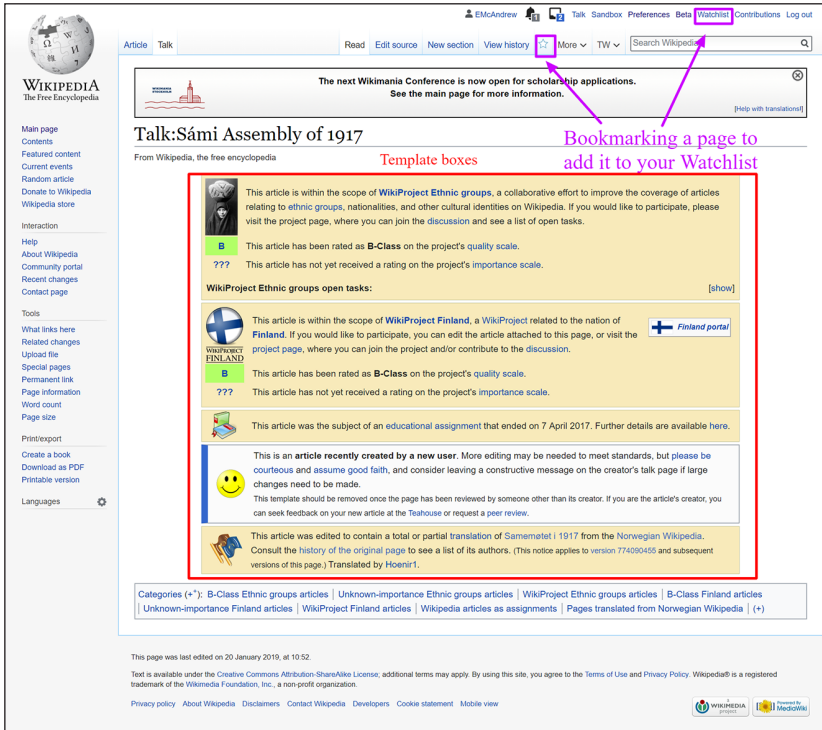
Although preparing the assignment in the right way is important, it is also helpful to ensure the article is published in accordance with Wikipedia's norms and conventions, so an 'aftercare' methodology has also been developed. Publishing newly translated articles is an important juncture for anyone supporting translation workshops or assignments. Although the Content Translation tool helps with removing a lot of the headaches associated with formatting new Wikipedia pages, a number of elements may need to be checked over, in particular:

- Does the article contain too much machine translation?
- Does it read naturally?
- Is it well structured and coherent?
- Is it well-referenced? Are there sufficient inline citations throughout the article?
- Are any paragraphs misaligned?
- Has the tool struggled with copying across citations, tables, infoboxes, etc.?
- Have images been copied across correctly? (This includes checking licenses).
- Are there any typos, grammatical errors, or empty headings?

If any of these issues have occurred the article should not be published onto the target Wikipedia until they are corrected.

Once students are logged in to Wikipedia, they can tick the blue star icon on a page to add it to their Watchlist (see [Figure 7](#) below). This means they will be notified of any changes to the articles and can monitor them during the important first few days of their infancy on the target Wikipedia.

Figure 7. The template boxes added to the Talk page newly created for the new Wikipedia article for the Sami Assembly of 1917⁸



Template boxes can be added to the article to flag up that the student editor is a new user, the article is still being worked on, is a translated page, or is part of an education assignment, etc.

3.4. Feedback from students and staff

An online survey was emailed to students at the end of the assignment to complete anonymously. The seven questions consisted of a mix of Likert scale responses and free text answers. This had 19 student responses in 2016/2017:

8. https://en.wikipedia.org/wiki/Talk:Sami_assembly_of_1917

- ten out of 19 respondents selected *Very Much So* when asked if their topic was interesting and relevant to their course;
- seven out of 19 respondents selected *Somewhat* when asked if their topic was interesting and relevant to their course; and
- only two respondents of the 19 selected *Not very much* and *Not at all* when asked if their topic was interesting and relevant to their course.

There were ten responses in 2017/2018. The 2017/2018 results included an extra Likert scale question on digital literacy which revealed that:

- seven out of ten agreed with the statement: *I better understand how to evaluate online information*;
- seven out of ten agreed or strongly agreed with the statement: *I better understand how to use and share online information*; and
- six out of ten agreed with the statement: *I have become better at creating and contributing to online information*.

Upon completion of the second semester's Wikipedia translation assignment, video interviews were also conducted with the course leaders and students to elicit feedback on their experience of the assignment (see [supplementary materials](#) for more detailed data). Feedback from students indicated that:

- they were engaged and enthusiastic about Wikimedia's mission to share knowledge globally;
- they were selecting texts they were interested in;
- they were getting much-needed published translation practice which they could use when getting a job;

- they were learning new skills and developing information and digital literacy; and
- they were enjoying the assignment.

Participating teachers were pleased that the students were:

- getting the necessary practical experience they needed;
- engaging in problem solving and critical thinking;
- engaging with how knowledge is shared around the world;
- writing neutrally for a Wikipedia audience;
- considering the verifiability of the information they were presented with;
- evaluating to what extent the translator should ever intervene; and
- learning academic research and writing skills which should stand them in good stead for their dissertation.

4. Conclusion

Far from being anathema in academia, this case study has demonstrated that Wikipedia has a lot to offer to teaching and learning, particularly in the context of teaching languages and translation. Translating between different language Wikipedias is a really impactful way to help build understanding between language communities and helps students to: (1) understand how knowledge is created, curated, and contested online; (2) create a new open educational resource that lasts beyond the lifetime of their assignment and can be added to and improved as a community project over time; and (3) achieve much-needed and *meaningful* published translation practice ahead of entering the world of work.

Acknowledgements

We would like to thank the Translation Studies MSc programme directors: Dr Charlotte Bosseaux, Dr Hephzibah Israel, and Dr Şebnem Susam-Saraeva. We would also like to thank the course teachers, our colleagues at information services and the Wikimedians who helped support the project including Melissa Highton, Anne-Marie Scott, Dr Richard Nevell, Leila Zia, and Amir Aharoni among many others. Finally, we would like to record our special thanks to the Translation Studies MSc students without whose industry and enthusiasm for the project none of this would have been possible.

Supplementary materials

- More detailed evaluation data: <https://research-publishing.box.com/s/eid6c0b9ijm4oi3k7ls4f59dzgoqr5lf>
- Lesson plan for how to conduct Wikipedia editing training: [https://commons.wikimedia.org/wiki/File:How_to_conduct_Wikipedia_Editing_Training_\(lesson_plan\).pdf](https://commons.wikimedia.org/wiki/File:How_to_conduct_Wikipedia_Editing_Training_(lesson_plan).pdf)
- Sample assignment page for the Wikipedia translation project: https://en.wikipedia.org/wiki/Wikipedia:University_of_Edinburgh/Events_and_Workshops/Translation_Studies_-_2018/2019_Semester_One
- Screencast demonstrating the Content Translation tool: https://upload.wikimedia.org/wikipedia/commons/e/ee/Content_Translation_Screencast_%28English%29.webm
- SPLOT Wikipedia Translation workshop resource: <https://thinking.is.ed.ac.uk/wikitranslation/>
- Wikimedia UK website and contact details: https://wikimedia.org.uk/wiki/About_us

References

- Al-Shehari, K. (2017). Collaborative learning: trainee translators tasked to translate Wikipedia entries from English into Arabic. *The Interpreter and Translator Trainer*, 11(4), 357-372. <https://doi.org/10.1080/1750399x.2017.1359755>
- Martínez Carrasco, R. (2018). Using Wikipedia as a classroom tool — a translation experience. *Proceedings of the 4th International Conference on Higher Education Advances (HEAD'18)*. <https://doi.org/10.4995/head18.2018.8112>
- McAndrew, E. (2017, May 10). *Word Count tool - counting the prose text in a Wikipedia article*. [Online Video]. 10 May 2017. https://media.ed.ac.uk/media/1_kfx9b4q5.
- Selwyn, N., & Gorard, S. (2016). Students' use of Wikipedia as an academic resource — patterns of use and perceptions of usefulness. *The Internet and Higher Education*, 28, 8-34. <https://doi.org/10.1016/j.iheduc.2015.08.004>
- The University of Edinburgh. (2019a). *Degree finder* [online]. <https://www.ed.ac.uk/studying/postgraduate/degrees/index.php?r=site/view&edition=2019&id=251>
- The University of Edinburgh. (2019b). *Vision and mission*. [online]. <https://www.ed.ac.uk/governance-strategic-planning/content-to-be-reused/vision-and-mission>
- Wales, J. (2016). *Wikimania conference, Esino Lario, Italy*. https://wikimania2016.wikimedia.org/wiki/Main_Page



Published by Research-publishing.net, a not-for-profit association
Voillans, France, info@research-publishing.net

© 2019 by Editors (collective work)
© 2019 by Authors (individual work)

New case studies of openness in and beyond the language classroom
Edited by Anna Comas-Quinn, Ana Beaven, and Barbara Sawhill

Publication date: 2019/07/26

Rights: the whole volume is published under the Attribution-NonCommercial-NoDerivatives International (CC BY-NC-ND) licence; **individual articles may have a different licence.** Under the CC BY-NC-ND licence, the volume is freely available online (<https://doi.org/10.14705/rpnet.2019.37.9782490057511>) for anybody to read, download, copy, and redistribute provided that the author(s), editorial team, and publisher are properly cited. Commercial use and derivative works are, however, not permitted.

Disclaimer: Research-publishing.net does not take any responsibility for the content of the pages written by the authors of this book. The authors have recognised that the work described was not published before, or that it was not under consideration for publication elsewhere. While the information in this book is believed to be true and accurate on the date of its going to press, neither the editorial team nor the publisher can accept any legal responsibility for any errors or omissions. The publisher makes no warranty, expressed or implied, with respect to the material contained herein. While Research-publishing.net is committed to publishing works of integrity, the words are the authors' alone.

Trademark notice: product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Copyrighted material: every effort has been made by the editorial team to trace copyright holders and to obtain their permission for the use of copyrighted material in this book. In the event of errors or omissions, please notify the publisher of any corrections that will need to be incorporated in future editions of this book.

Typeset by Research-publishing.net
Cover design by © Raphaël Savina (raphael@savina.net)

ISBN13: 978-2-490057-51-1 (Ebook, PDF, colour)

ISBN13: 978-2-490057-52-8 (Ebook, EPUB, colour)

ISBN13: 978-2-490057-50-4 (Paperback - Print on demand, black and white)

Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

British Library Cataloguing-in-Publication Data.

A cataloguing record for this book is available from the British Library.

Legal deposit, UK: British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: juillet 2019.
