

www.jostrans.org · ISSN: 1740-367X

Perrino, S. (2009). User-generated Translation: The future of translation in a Web 2.0 environment. *The Journal of Specialised Translation, 12*, 55-78. https://doi.org/10.26034/cm.jostrans.2009.618

This article is publish under a *Creative Commons Attribution 4.0 International* (CC BY): https://creativecommons.org/licenses/by/4.0



© Silvio Perrino, 2009

User-generated Translation: The future of translation in a Web 2.0 environment Saverio Perrino, freelance translator

ABSTRACT

The purpose of this paper is to investigate how translation activities can be carried out using available online tools. In Section 1 of this paper, I intend to illustrate how mass collaboration happens in the so-called Web 2.0 – a term coined by Tim O'Reilly in 2005 to describe a number of services which enable today's Internet users to interact and share information efficiently. The innovations of Web 2.0, along with the controversial issues to which they give rise, will be presented, before being applied to translation in Section 2. Subsequently, I will introduce the concept of User-generated Translation (UGT), an umbrella term I use to define translation practices made possible by various online services. I call such services UGT tools and classify them according to the nature of the content for which they offer means of translation. In Section 3, seven tools will be introduced and reviewed according to established parameters, such as ease of use and user participation. The data thus gathered will be compiled in tables for at-a-glance review. In the final chapter, I conclude that User-generated Translation is feasible but that, at present, the available tools and their users cannot compete with the professional translation industry.

KEYWORDS

Web 2.0, User-generated Translation, Internet, Mass collaboration.

1. Introduction

While surfing the Internet, I recently stumbled upon a picture taken in China showing a gigantic restaurant sign. Near the Chinese ideograms composing the original name of the brasserie, stood the words "Translate server error" (Doctorow 2007).



It is impossible to know the reason behind such a translation disaster, but it is fair to assume it was generated by a cheap automatic tool.

A viral picture showing that somewhere in China there is a restaurant called "Translate server error" could easily be considered an excellent form of marketing for the translation industry and tends to reinforce the belief that, while good translations pass unnoticed, inaccurate ones are potentially disastrous. But it is also a misleading sight that does not do justice to the impressive achievements of the marriage between technology and translation. Indeed, the two fields are anything but incompatible, and while we translators wait for some futuristic robot to do our job, we can concentrate on how today's translators can be assisted rather than replaced by technology.

In this paper, I investigate the emerging phenomenon of User-generated Translation in a Web 2.0 environment. The purpose is to analyse current trends of online collaboration modelled on sites such as Wikipedia and assess whether they can be efficiently applied to translation practices. Indeed, a number of newly created online tools are making inroads in today's Internet, and considering the fast-developing nature of the medium, they demand immediate attention in the field of Translation Studies. In an effort to capture a glimpse of this expanding phenomenon, I shall review seven tools which allow Internet users to translate online texts. However, I do not intend to evaluate the quality of the translations obtained by means of such tools, for two reasons. First, in order to obtain exhaustive results, a considerable corpus of translations in several languages would need to be examined. Second, the concept of translation quality tends to be highly subjective. The judgement of a foreign language rendition of a text does not only take into account easily identifiable mistranslations or grammar and spelling mistakes, but also debatable linguistic nuances and subtleties. Lacking the time and expertise to carry out such an extensive qualitative study, I shall confine my reviews to more technical aspects and ultimately assess the potential of the tools.

At the end of this research, I shall summarise my findings in an attempt to answer the following questions concerning User-generated Translation (UGT):

- Is UGT feasible?
- Are UGT tools of an acceptable standard?
- Is UGT affecting the professional translation industry?

2. Mass collaboration and Web 2.0

In this chapter, I intend to provide a critical summary of current Internet trends and practices. In an attempt to offer a brief but comprehensive and diverse account of the topic, I shall introduce the main innovations of the so-called Web 2.0, i.e. the new information environment in which Internet users freely create, collaborate, and share data.

Even though the central topic of this paper is Translation, throughout the present chapter I shall carefully avoid making extensive use of the T-word, for

two reasons. First, translation activity is always "bound up with the technical environment which makes it possible" (Cronin 2005: 25) and it would therefore be preferable to examine the pros and cons of Web 2.0 before introducing the idea of translation into my argument. Second, it is essential to consider the ongoing debate among scholars and users with regard to Internet content production and distribution before addressing the same issues with regard to translation.

In this section, I shall start with a comparison which exploits the language of biologists to illustrate how mass collaboration happens in the online world. I shall then introduce the concept of Web 2.0 and use relevant examples and contributions to demonstrate how it has changed—and keeps on changing—the information environment. Subsequently, I intend to explore the debate sparked by Web 2.0 innovations, focusing on the unsolved problems of quality and linguistic accessibility of online content which are of particular relevance given the topic of this paper.

1.1 Stigmergic collaboration

spiar le file di rosse formiche ch'ora si rompono ed ora s'intrecciano a sommo di minuscole biche. (1) E. Montale

Another example of stigmergic collaboration between social insects is the building of anthills. Each termite secretes pheromones over its mud balls in order to attract other individuals. Consequently, the colony understands where the designated spot to place other "bricks" is and restlessly carries on building. This relatively simple process results in the complex structure of termite nests, a unit created by workers whose "individuality is subjugated, apparently to the welfare of the community" (*ibid*.).

It does not take an entomologist to see how the notion of stigmergy can successfully be applied to us vertebrates. A number of authors (Elliott 2007; Gregorio 2002; Heylighen 2007; Bonabeau 1999) have identified examples of

stigmergic collaboration between humans in communication networks and the Web. It is safe to say that every single time we connect to the Internet, we leave trails and contribute to the bigger picture. We either do it willingly, by editing user-generated encyclopaedias and writing blogs, or unintentionally, simply by clicking on the entries offered by Google searches. In the words of a blogger, "we all became ants building a nest we can't even see" (Gregorio 2002).

1.2 A New information environment

In 2006, *Time's* Person of the Year was "You." By honouring the (often anonymous) Internet user with this title, the magazine acknowledged the importance of online mass collaboration and officially introduced the social media revolution to mainstream culture (Grossman 2006). The unprecedented growth of communication networks and their increasing influence on our everyday lives resulted in the emergence of a new information environment, "one in which individuals are free to take a more active role than was possible in the industrial information economy of the twentieth century" (Benkler 2006, 14).

When this paradigm shift became evident and called for a new grammar, Tim O'Reilly coined an umbrella term to label the revolution before our eyes (2005). "Web 2.0" can be defined as "a set of principles and practices that tie together a veritable solar system of sites." Online encyclopaedias and dictionaries (Wikipedia, Wiktionary), social news sites and aggregators (Delicious, Digg, Current), blogging platforms (WordPress, Blogger), online marketplaces (eBay, Craigslist), social networks (Facebook, MySpace) and multimedia platforms (YouTube, Last.fm, Flickr) are just a handful of popular services allowing users to control their own data and share information in ways that were unthinkable only a few years ago. Indeed, in the so-called Web 1.0 environment, the ability to manage online content required a relatively high level of technical expertise, elevated costs, and considerable time spans. Today, publishing one's work on the Internet is made easy by intuitive platforms and high-speed connections. A blog can be opened for free in a matter of minutes, and uploading audiovisual material and pictures does not require advanced skills.

But Web 2.0 is not only about accessing and sharing information – it is also about generating ideas and putting them into practice in an environment that trusts collective intelligence.

1.3 Commons-based peer production

An important novelty made possible by Web 2.0 is the phenomenon of commons-based peer production (Benkler 2002; 2006). This is an innovative production system that allows users to collaborate in the creation of goods and services – be they wikis or web applications like the ones I shall review further on. Much like in a social insect colony, individuals contribute to the bigger

picture by adding or editing the elements that make up the "building." But roles are generally not hierarchically assigned, and users can volunteer according to their availability and their level of expertise.

Such a system functions on the basis of a non-proprietary model. Participants and developers "usually retain copyrights of their contribution, but license it to anyone" (Benkler 2006: 63). This allows users to be credited for their work and authorizes others to get involved in the different stages of production. The simplest examples of the commons-based model are, of course, Wikipedia entries and open-source software, but its possibilities are endless. Information and IT are not the sole industries being affected by the Web 2.0 revolution and it is worth mentioning a couple of successful instances to prove the point. In recent years, we have seen musicians giving away their songs for fans to create videos or remixes (Nine Inch Nails, Radiohead, Moby) (2), prestigious publishers promoting collaborative novels written by users (Penguin) (3), and even online communities set up to facilitate the development of electric cars.(4) These are just a few noteworthy examples showing that online innovations can be extended and applied to the 'offline' world.

It is easy to see how this revolution can inspire enthusiasm for a newfound democracy that supports and encourages everyone's contribution, but while its possibilities are immeasurable, so are its downsides.

1.4 Web 2.0 criticism: quality and authorship

Wikipedia is like hearing a great story in a bar: you hope it's true but never bother confirming it.

J. Feinstein

In 2007, digital media entrepreneur Andrew Keen expressed fervent criticism towards the Web 2.0 revolution in a provocative book called The Cult of the Amateur: How Today's Internet is Killing Our Culture. He argues that the democratisation of the medium is irreparably downgrading traditional cultural institutions and polluting the information environment. "Cultural gatekeepers" who achieved expertise in their fields through education and experience, are being rivalled by a horde of amateurs. Why should we choose a user-generated encyclopaedia over the prestigious Britannica? Why should we be interested in the biased opinions of an anonymous blogger when we can read articles by acclaimed professional journalists? And, as we shall see later on, why should we trust amateur translators instead of established translation agencies? But quality is but one of the issues at stake. The risk of stumbling upon subtle ideological or commercial propaganda is not one to underestimate. For instance, how can we be sure that a product review on Amazon has not been written by the very seller (Hallberg 2008) or that people with a political agenda are not hiding behind well-crafted blogs? One may argue that the same can happen in traditional media as well, but while journalists and scholars have a reputation to keep, the anonymous user has nothing to lose. According to Wired's Editor-At-Large Kevin Kelly (2006), where anonymity becomes common, the system fails. The recent taint in the honor of Wikipedia stems from the extreme ease which anonymous declarations can be put into a very visible public record. Communities infected with anonymity will either collapse, or shift the anonymous to pseudo-anonymous, as in eBay, where you have a traceable identity behind an invented nickname.

Indeed, the impossibility of verifying the effective authorship of online content is arguably one of the most criticised aspects of today's Internet. In such a chaotic environment, overcrowded by experts and amateurs alike, the information sources are often untraceable, lost in the boundaries between real and virtual. As mentioned by Kelly, Wikipedia's system is the emblem of Internet anonymity and has been raising issues of accuracy and reliability since the website's foundation. However, the popularity of the free encyclopaedia demonstrates the success of the system, which has now been adopted by countless websites and dubbed "the Wiki way" (Leuf and Cunningham 2001). Therefore, contrary to Kelly's belief, the system has not failed: in fact, it has not shown signs of weakness.

1.5 The Babel paradox

Cuando se proclamó que la Biblioteca abarcaba todos los libros, la primera impresión fue de extravagante felicidad. Todos los hombres se sintieron señores de un tesoro intacto y secreto.

- J. L. Borges

In the aforementioned *Cult of the Amateur*, Keen makes an interesting and evocative comparison between today's Internet and Jorge Luis Borges' "Biblioteca de Babel" (2007: 84). In 1939, the Argentine writer envisaged a "total library" containing the whole corpus of human knowledge. In the essay, the announcement that the library encompassed every single book ever written and ever to be written was welcome with a sentiment of "extravagant happiness." However, the initial excitement soon transformed into terror and fear for this metaphysical library was organised around the elusive logic of an infinite number of hexagonal galleries, and in truth ruled by chaos.

According to Keen, "the experience of surfing the Internet is akin to wandering around the hexagonal galleries of Borges' Library of Babel. Truth is elusive, always one click or Web site away" (2007: 84).

Interestingly, Keen failed to mention an essential feature of the Library of Babel: the availability of every single book in every single language or dialect. Indeed, Borges' visionary creation was inspired by the biblical Tower of Babel – the veritable symbol of multilingualism. Keen's omission does not come as a surprise, since the predominance of English on the Internet is a well-known fact. The Internet World Stats website estimates that the English-speaking population of 430,802,172 with access to the Internet counts for 29.4% of all global users, (6) and English entries make up more than 22% of the total of Wikipedia articles. (7)

So, where is Babel? Certainly not in the meanders of the World Wide Web, where "stigmergy" occurs mainly between English-speaking users. But it is fair to say that some action is being taken by international communities willing to make online content accessible in multiple languages. The recent development of a variety of translation services hints at a new, inspiring revolution, in which users not only create and share but also translate. In the following chapters, I shall explore some instances of translation in a Web 2.0 environment. I shall introduce the concept of User-generated Translation and provide a critical account of state-of-the-art tools aimed at breaking down linguistic barriers.

2. Translation in a Web 2.0 Environment

As I pointed out at the beginning of Section 1, the problems examined thus far are generally addressed to Web 2.0 content production and distribution rather than translation. It is indisputable that the latter is subject to the very same issues of authorship and quality but, alas, the proverbial invisibility of the translator (Venuti 1995) seems to have found a virtual equivalent. However, as I shall demonstrate, there is fervent online activity in the field of translation and a host of newly emerged practices call for popular and scholarly consideration.

In this chapter, I intend to introduce the concept of User-generated Translation (UGT), i.e. the exploitation of today's Internet tools for translation purposes. In an attempt to capture the present juncture in this fast-developing field, I shall take into consideration a selection of websites offering a variety of UGT services. Their purpose-specific application, ease of use, and quality will be reviewed and compared in Section 3. But before diving into the brave new world of translation in a Web 2.0 environment, let us recall the state of affairs in the mid-'90s, when the World Wide Web was at an early stage of development in terms of widespread public availability.

2.1 From teletranslation...

More than ten years ago, Translation scholar Minako O'Hagan envisaged the emergence of a global network where translators and clients would collaborate effectively by means of telecommunication systems (1996). In her seminal paper, O'Hagan created a new body of work and definitions by contextualising the translation industry in the paradigm shift from industrial to information society. Translation practices in the industrial society were based on print media, and communication with customers was often asynchronous due to the lack of viable and affordable technology. The advent of the Internet and more reliable networks allowed customer and provider to communicate in real time, and consequently changed the industry. O'Hagan coined the term teletranslation to describe the early marriage of "two previously unrelated branches of communication:" translation and telecommunication (1996: 13).

Teletranslation will exploit the global networking capabilities of telecommunications technology to bring language service providers (using

both human and computer resources) and their customers together. In doing so, it will help overcome many otherwise insoluable language problems.

Today, translation without electronic processing and Internet resources would be inconceivable, but in 1996 such a scenario appeared truly futuristic. Interestingly, O'Hagan's predictions already pointed at stigmergic collaboration and the author herself revisited her body of work to encompass recently developed translation practices such as Audiovisual Translation (AVT) for DVDs, Fansubbing, and Videogame and software Localisation (2005; 2006). However, I feel that in the fast-moving world of Web 2.0, a more specific definition is needed to describe the marriage of teletranslation and stigmergic collaboration between Internet users. Using O'Hagan's framework as a starting point, I intend to introduce a new umbrella term to cover the newly emerged translation practices presented in this paper.

2.2 ... to User-generated Translation

User-generated Translation (UGT) describes the harnessing of Web 2.0 services and tools to make online content – be it written, audio or video – accessible in a variety of languages. It differentiates from automatic translation for it requires human expertise and implies the collaboration between users – be they amateurs or experts. It includes:

- Professional translation networks
- Translation wikis
- User-generated dictionaries
- Online subtitling practices
- Volunteer website localisation

I chose this term with "User-generated Content" (UGC) in mind, i.e. any form of digital media created or edited by users and distributed via online platforms. At the time of writing, a Google search for "user-generated translation" produces a mere 250 entries, as opposed to the 5,880,000 entries for "usergenerated content." However, the latter is certainly not immune from linguistic controversy and it is fair to mention some critical interpretations of the term put forward by experts in the field. According to editor of *Publishing 2.0* Scott Karp, the word "user" once defined "audience," but has now become an unmerited synonym for "publisher" (2006). Karp's observations are echoed by The Guardian's journalist Jemima Kiss (2007) and CEO of Denuo Rishad Tobaccowala, who mischievously asks "User Generated Content: Since when did I become a heroin addict?" (quoted in Karp 2006). Such concerns point us to the previously examined issues of Web 2.0, which I argue that, from a purely linguistic point of view, are unconvincing. The word "user," while largely associated with the Internet, carries a neutral connotation and could possibly encompass future developments. Its marriage with "generated" is also particularly apt, as it does not exclude any form of media. Hereunder I provide a list of other viable definitions which, for various reasons, I find less appropriate and effective than UGT:

- Collaborative Translation is a widely used definition (15,000 Google entries). However, it is undoubtably vague and tautological. Indeed, collaboration should be a sine qua non to translation. Particularly in a professional environment, the final product is or, at least, should always be the result of joint forces of translator and proofreader.
- Fan-based Translation (186 entries) is mainly associated with manga and videogames. I argue that, even though User-generated Translation can be fuelled by strong interest on a given topic, the word "fan" somehow excludes the involvement of professionals.
- Amateur Translation (2,550 entries) would be belittling and points us to the quality issues of Web 2.0 content examined in the first chapter of the present paper.
- Crowd-sourced Translation (1,240 entries) is mainly used in the field of software localisation but I would argue that it hints at disorganisation and an absence of rules.
- Open Source Translation (78,100 entries) is mostly used when speaking about translation memory or software development and localisation. The meaning of "open source" became blurred with its popularisation, and a quick comparison between the English and Italian Wikipedia entries for the term (8) underlines there is still widespread confusion about its correct usage. Some simply use it as a synonym for "free" (Stallman 2007).

2.3 Introduction to UGT tools

In the following chapters, I shall examine seven UGT tools providing the means to translate digital texts: professional translation networks, translation wikis and user-generated dictionaries. Brief descriptions and examples will be provided, along with critical reviews. Several aspects will be taken into consideration:

- the type of website hosting the UGT tool;
- the type, length and licence of the material supported for translation purposes;
- the number of languages supported for the translations and the interface;
- the positive and negative aspects of the tools, with particular attention to ease of use and design, popularity, and user participation be it professional or amateur.

All the gathered data will be summarised in tables, to provide a quick at-aglance review at the end of chapter 3. It is worth pointing out that my selection of tools excludes volunteer website localisation and online subtitling practices for these topics would deserve a separate approach and more research and insight than this paper could provide.

3. UGT tools

Hereunder is a list of the seven UGT tools presented in this chapter, complete with name, URL and year of launch. In order to give an idea of the size of the websites taken into consideration, I provide Alexa (9) and Google PageRank (10) statistics – the two most widely used and trusted services to calculate Internet traffic and popularity respectively.

Website	URL	Online since	Alexa rank	Google PageRank	
ProZ	www.proz.com	1999	4,128	6	
Translators Café	www.translatorscafe.com	2002	28,880	5	
Traduwiki	traduwiki.org	2007	3,275,301	4	
Der Mundo (formerly known as Worldwide Lexicon)	www.dermundo.com	N/A	2,990,456	1	
Cucumis	www.cucumis.org	2005	44,914	1	
WikiTranslate	www.wikitranslate.com	2008	2,119,087	3	
Word Reference	www.wordreference.com	1999	374	8	

Table 1. UGT tools for digital texts.

3.1 Professional translation networks

3.1.1 ProZ and Translators Café

Websites description

In recent years, ProZ and Translators Café have become two essential tools for translators. The first one was opened in 1999 and has offices in Syracuse, NY, and La Plata, Argentina, the latter has been online since 2002 and is based in Canada. The two websites cover a wide range of services for language

professionals and their clients, and constitute a valuable ally to gain recognition in the translation industry.

The main features of the two sites are: a) job boards where linguists can apply for freelance or inhouse positions; b) multilingual moderated forums where registered users can share experiences and useful tips, and discuss topics ranging from translation rates to marketing strategies, from localisation software to literature reviews; c) advanced profiles for linguists, agencies and clients; d) polls and contests; e) user-generated dictionaries and glossaries. I shall concentrate on the last point, which is of particular interest in the context of the present paper.

UGT tool

ProZ features a section called "KudoZ term translation questions," (11) where linguists can "assist each other with translations or explanations of terms and short phrases." To describe this function, I shall use an example based on my experience as a ProZ user. (12) On October 25, 2007, an Italian member asked the community for suggestions regarding the English expression "all singing, all dancing." He provided the context by copying a passage of the source text he was translating into Italian and defined the type of client he was working for (a recording studio). His question was filed under the relevant language pair (English to Italian), broad field (Art/Literary), specific field (Music), and specific category (idiomatic expressions). In less than two days he received 5 answers from other users, including me. Everyone suggested a translation for the term and some added links to support it and prove its effective usage. The user ultimately chose my option and I was awarded 4 points ("KudoZ") out of 5 because, when answering his question, I selfassessed my level of confidence at 4 (high). With this tool, translators can give assistance to colleagues in situations where dictionaries and other Internet resources cannot help. Moreover, the asker can assess the level of trustworthiness of every contribution by checking the profile and history of the answerer - an option which comes in handy for questions requiring fieldspecific expertise. Registered users can ask up to 5 questions a day, while paying subscribers can ask up to 15. Selected contributions are awarded with "KudoZ," the website's currency. The more KudoZ translators have, the more reliable they are in the eyes of other users or possible employers. As of today, the ProZ database counts almost 2 millions questions asked and the best answers are filed to create a reliable user-generated dictionary. The same feature is available on Translators Café and applies an identical system. (13) On both websites, users can also submit glossaries or articles about languages and translation. All these data are also accessible for non-registered users through search engines.

Review

Overall, the sections of ProZ and Translators Café dealing with translation requests can be considered as very successful experiments of UGT. Fast, reliable, and user-friendly, the two websites are the perfect directory for

translators, interpreters, and language students. Then why is such a well-designed service limited to words or phrases yet does not extend to longer texts? I am convinced that ProZ and Translators Café's choice is not accidental. In fact it is a statement. The website creators, while acknowledging the importance of stigmergic collaboration between users, probably believe that translation practices should be left to professional translators. Longer translations require patience, research, time, and are expected to be performed for payment.

3.2 Translation Wikis

3.2.1 Traduwiki

Website description

Traduwiki was created by Thuan Huynh, a self-proclaimed "journalist and web designer amazed by the power of We." (14) He states that his website's idea came from the desire to share knowledge and the "frustration that many texts freely available on the Internet are accessible in only one language [and] that many of us can read a foreign language and translate." In order to solve the problem, in 2007, he launched this Wikipedia-inspired service.

UGT tool

Traduwiki applies the concept of Wikipedia to translation in a mission to make a variety of texts linguistically accessible. As stated in the Guide to Traduwiki, "a wiki translation is a collaborative translation in which anyone, members or anonymous readers, could contribute to the content." (15) Users can upload anything from scientific articles to poetry, provided that the content is under an open source licence. Therefore, users must either own the rights of the text or have the original authors' permission to freely distribute their work. When a text is online, it gets divided into small "chunks" of variable length - generally every portion constitutes a period - and becomes available for anyone to translate. Other users can then select their target language and start translating short sentences according to their expertise and availability. The result is viewable and editable by everyone, and it is also possible to add notes to explain and justify translation choices. The homepage features the articles users are currently translating and it is possible to assess how much work is needed by means of coloured bars showing the percentage of translated text. The author estimates that "a 300-word text attracting 40 users could be translated in half an hour." The FAQ section states that completed translations can be downloaded free from the website. However, I have not found any downloadable text and at the time of writing I am waiting for a reply on the issue from the staff.

Review

The site is certainly a viable alternative to sloppy automatic translation services and is very well-designed, easy to use and commercial-free. However, I have a certain reservations about the way Traduwiki works. First, the fact that a text is divided into small chunks poses some problems regarding the continuity and cohesion of the translations. The system allows users to select the sentences they wish to translate according to their self-assessed level of expertise. For instance, confident users can work on complex and convoluted fragments while the less experienced can train themselves on simpler oneliners. Of course, a final redraft can solve the continuity problem and correct possible mistranslations, but I would argue that working on a text in a nonsequential order can be a tricky process. Second, the website does not require registration or member log-in. This means that anyone can contribute and take advantage of the privilege of anonymity (see § 1.4). If users were able to register—not necessarily by providing name and surname, but at least an email address and a nickname—the site would gain in credibility and benefit from the presence of a community in which users are regarded for the quality and quantity of their contributions.

To sum up, Traduwiki is a praiseworthy experiment of User-generated Translation and a well-crafted website, which however tends to empower amateurs by placing them on equal footing with professionals. It will be interesting to see how it develops should it become increasingly popular.

3.2.2 Der Mundo (formerly known as Worldwide Lexicon)

Website description

Worldwide Lexicon (WWL) was first introduced at the O'Reilly Emerging Technologies Conference in California in 2002. Its inventor, software designer and media entrepreneur Brian McConnell, believed in the possibility of overcoming language barriers on the Internet by organising an online community of volunteer translators. Seven years later, the project has come a long way, and even though the core principles on which it was founded have not changed, its system has. In the early days of WWL, registered volunteer users were required to download a piece of software which would send instant messages asking for the translation of short blocks of texts. According to their availability, translators would work on randomly assigned pieces of the "puzzle" and contribute to the final translation of a given content. However, this system was blocked by a couple of problems. The first was of a theoretical nature, and was immediately put forward by scholars. British researchers interviewed by The New Scientist (2002) at the time of the launch, insisted on the impossibility of translating sentences taken out of context and argued that this could lead to poor quality results. The second problem of the first version of WWL was more difficult to predict, and was recently put forward by the inventor himself (2007). When creating WWL, McConnell underestimated the fact that volunteer translators would not work on any text, but only on those they were genuinely interested in. It soon became evident that, say, a rugby fan would not be willing to translate articles about fashion, and vice versa. But since matching translators with their interests was a complex task, McConnell and his team had to find alternative ways to outsource the jobs and, in 2007, they revised the WWL project and brought in a number of substantial changes. The site was relaunched to incorporate a web text editor to substitute the old software, allowing translators free choice of what to work on rather than being assigned texts they might not like. In February 2009 the site underwent its third makeover and is now called Der Mundo.

UGT tool

The website retrieves articles from selected blogs via RSS feeds. When an article is published, users are taken to the original website. The page will show a "Der Mundo" bar with a list of 59 languages to choose from. By selecting one of the languages, a pop-up window containing an automatic translation will appear. Users can then edit the automatically-generated draft and publish their own translations.

McConnell believes that this system will transform the Internet into a "truly global medium," and predicts that within three years the concept of "accidental translation" will change the way we share online information. (2008)

Accidental translation is where one person discovers an interesting article, translates it for a friend, and by doing so, creates something that the rest of the world can share too.

As a fact, the revolution predicted by the entrepreneur is already happening with the phenomenon of social bookmarking and social news. Digg, Delicious, Current, StumbleUpon, Newsvine and Netvibes are just a few sites offering users the possibility to bookmark their favourite web pages and share them with anyone or a closed community. However, at present such tools are monolingual and content can be shared only among people speaking a common language. With Der Mundo we may soon be able to discover and read content originally written in a language we do not master – thus solving the Babel paradox (see § 1.5).

Review

Judging from the achievements of Worldwide Lexicon and Der Mundo, the scenario envisaged by McConnell is still far from reality. For instance, let us consider the entrepreneur's blog entry on "accidental translation" from which I extracted the above quotation, and which was supposed to be a must-read for any user accessing the now-defunct WWL site. It was published on July 26, 2008 and three months later it had been (poorly) translated into two out of the many supported languages: French and Serbian. For another 18 languages, only an automatic translation by Google was available, and no human translator had taken the time to put those texts out of their misery.

At present, Der Mundo does not provide anything but an automatic translation service embedded into a complex and non user-friendly interface. Possibly, the website creators assumed that by letting Google do the translation draft, they could make a translator's task easier: users would simply have to correct a few mistakes, check for fluency, et voilà. On the contrary, I would argue that both translators and readers are likely to be put off by enigmatic automatic translations, which usually deserve to be deleted altogether rather than simply edited.

In conclusion, McConnell's site is an interesting project that could work effectively only if backed by a strong and active community of 'accidental translators.' With nearly a decade of history, WWL/Der Mundo has not yet shown its full potential and its inventor's vision still remains a utopia.

3.2.3 Cucumis

Website description and UGT tool

Cucumis is a community of (self-styled) translators sharing their linguistic knowledge. It was founded in Paris in 2005. Registered users can submit a text and request to have it translated by the community into one of the 75 languages supported. The users' hierarchy is based on the number and rate of completed translations. The website's administrators rate the quality of a user's work with a 10-point scale, and when one reaches an average mark of 7/10 is considered an "expert." Users who distinguished themselves with outstanding contributions can be awarded the title of "Power Cucumers." The website's currency is the points accumulated by translating material. A user who does not intend to collaborate but just wishes to have a text translated, is assigned a standard number of points to 'pay"=' the translators. Experts' translations are more 'expensive.' As of today, the website counts more than 75,000 completed translations, most of which are still to be greenlighted by administrators. The translations database features a variety of topics ranging from short biographies to national anthems, but arguably the most popular requests are love poems and fortune cookie-style aphorisms.

Review

The nature of the translated texts does not come as a surprise, considering the website's childish logo and schmaltzy design. Aesthetic issues aside, the main aspects to take account of when judging this online service must be quality and trustworthiness. Even though translations can be checked by "experts," "Power Cucumers" and administrators, most users hide behind nicknames and avatars. A quick glance at the "most active translators list" reveals profiles in which little or no information about the translator is provided. If not through charts showing stars and watermelon-shaped symbols, it is impossible to judge a user's translation activity on the website – a list of previous translations, for instance, would give readers an approximate idea of how trustworthy the

person dealing with their submitted text is. All we have access to are a user's points, and source and target languages (some members state they can translate from and into as many as six different languages – hexalingual?). As opposed to ProZ and Translators Café, where collaboration and expertise meet, Cucumis looks like the home of the empowered anonymous amateur.

3.2.4 WikiTranslate

Website description and UGT tool

WikiTranslate is a translation wiki in which, upon registration, it is possible to submit any type of text of maximum 500 characters and wait for the community to translate it. Much like in Der Mundo, the draft is provided by an automatic service. Users can submit only one request per day, but edit and translate as many texts as they wish. Those in need of a fast translation service can sign up for an annual \$25 subscription fee. The passages submitted by premium members will appear in a featured list at the top of the page, thus gaining more visibility.

Review

WikiTranslate is the youngest website examined in this paper, but seems to embody its competitors' best options and correct many of their flaws. Unlike Der Mundo, the interface is well-designed and automatic translation drafts do not represent a nuisance in the user's experience as they are very easy to manage and edit.

At the time of writing, it is impossible to judge the quality of the translations featured on WikiTranslate, as user participation is still very limited. However, I believe the website has the potential to become a popular UGT tool. The simple and intuitive interface, combined with a comprehensive selection of language pairs, make WikiTranslate a valuable service. If marketed properly, the site could become very successful.

3.3 User-generated dictionaries

3.3.1 Word Reference

• Website description

Not only is Word Reference one of the most popular instances of a UGT service, it is also one of the 500 most trafficked websites of the globe (see Table 1). According to its creator, Word Reference was launched in an effort "to provide free online bilingual dictionaries and tools to the world for free on the Internet." (16) Since 1999, the site has gradually grown in size and popularity and now includes 10 language pairs (English into French, Italian, Spanish, German, Russian, and vice versa) and a monolingual English dictionary. In 2008, Word Reference teamed up with Oxford University Press to

provide both Concise and Unabridged versions of their authoritative dictionaries for a small annual subscription fee.

UGT tool

The most interesting part of Word Reference to examine in the context of this paper is the forum section. Dictionary entries, regardless of their reliability, can sometimes raise questions about the effective usage of words and idiomatic expressions in formal or everyday situations. Word Reference adds to every dictionary entry a link to users' threads in the language forums. When searchers are not satisfied with the dictionary's results, they can read what the community has said in previous discussions about usage, spelling, grammar, translation, etc. Should the searchers not find answers to their questions, they can open a new thread and ask the community for help or request a Usergenerated Translation. Although amateurs and professionals are impossible to distinguish, readers can access each user's previous posts and assess their trustworthiness.

Review

Word Reference is more of an open debate about languages and their nuances. It is not a website for scholars and experts, but rather a meeting point – a forum – to exchange opinions. Since everyone's contribution is public and cannot be edited or voted, it is the reader's task to select the best answer. This detail makes

Unlike ProZ and Translators Café, in which for each question or translation request there is always a "winner" answer, Word Reference a very democratic UGT tool and a very interesting directory for the language-curious.

3.4 At-a-glance reviews

The following chart (Table 2) is aimed at summarising the data gathered thus far to provide an at-a-glance review of UGT tools for digital texts. Three of the table fields require clarification because of their controversial and non-scientific nature. As for the "user participation" field, the absence of precise data obliged me to make a rough estimate of the number of active users and completed translations available. With regard to the "popularity" field, I based my judgement mainly on the previously mentioned Alexa and Google PageRank statistics, and on the quantity of articles and blog entries dealing with the tools in question I found during my research. Finally, the "ease of use and design" field is intended only to represent my experience as a user.

Name	Type of website	Text supported	Long texts supported	Translations No. of languages	Interface No. of languages	User participation	Popularity	Ease of use and design	RSS support
ProZ	Network, job board	Words, idioms	×	200+	48	•	•	•	×
Translators Café	Network, job board	Words, idioms	×	200+	35	•	•	,	•
Traduwiki	Wiki	Articles (Creative Commons licence)	*	14	1 (EN)	×	×	•	×
Der Mundo (WWL)	Wiki	Blog entries	*	59	59	×	~	×	*
Cucumis	Wiki	Any (mainly words, idioms, aphorisms, poems)	•	75	34	•	•	×	*
WikiTranslate	Wiki	Any (500 characters max)	×	9 (22 language pairs)	1 (EN)	×	×	•	•
Word reference	Dictionary	Words, idioms	×	Any in the forums, 6 in the dictionaries	6	•	•	•	×

Table 2. At-a-glance reviews of UGT tools for digital texts

6. Conclusion

As anticipated in the Introduction, this research attempts to answer three questions concerning User-generated Translation.

- Is UGT feasible?
- Are UGT tools acceptable?
- Is UGT affecting the professional translation industry?

Each of the following sub-chapters focuses on one question and tries to provide a detailed answer which encompasses possible future developments of UGT activities.

6.1 Is UGT feasible?

User-generated Translation is feasible, but for lack of promotion or scholarly attention, it cannot be considered a mass phenomenon. Even though, as I pointed out in the reviews, some websites show good user participation and encouraging statistics, they are still far from the numbers of Wikipedia and other popular instances of user-generated content.

I assume that UGT tools do not tend to catch the interest and participation of the Internet masses, for translation is hardly perceived by the general public as a rewarding activity – both in terms of money and glory. While the Web is overcrowded with would-be musicians, filmmakers, photographers and journalists who pursue their dream and gladly accept no compensation to promote the results of their hobbies, the same cannot be said in the field of translation. The very nature of the activity, which is somehow performed in the shadow of the source text writer and rarely leads to fame and prestige, is not compatible with the Web 2.0 amateur, who ultimately seeks recognition (Keen 2007). However, let us recall what happened in the summer of 2007, when the final book of the *Harry Potter* saga was published in English-speaking countries. In a matter of weeks, the tome was already circulating in countless 'pirate' translations by international fans. Most impressively, the Chinese version was available after only two days, translated by 60 volunteers under the direction of a young student (Brantley 2007). This gargantuan work was organised in chatrooms and forums, thus without the use of purpose-made tools.

The UGT tools I examined have not yet been used to such an extent, but if they were to become the favourite directory of enthusiastic amateurs and fansubbers, their potential would be enormous. It is impossible to predict whether the *Harry Potter* case will remain an exception or establish a trend, but nonetheless it showed the power and feasibility of mass collaboration applied to translation.

6.2 Are UGT tools of an acceptable standard?

In the light of this paper, UGT tools are not of an acceptable standard. Indeed, none of the tools examined throughout this paper managed to meet all the criteria set. However, it must be said that the selection comprises very diverse websites and services, and the attempt to classify them according to common features proved extremely difficult.

The two professional translation networks examined (ProZ and Translators Café) present several interesting aspects: a high number of languages available for both translations and interface, simple and intuitive design and, most importantly, gooduser participation. They are very effective tools especially for 'emergency' translation needs, in situations where dictionaries or other available resources cannot help. Moreover, anonymity is not an issue on ProZ and Translators Café: translators are asked to take full responsibility for their work, and if they want to use the two sites as marketing tools, it is in their best interests to fill in all the details and use their real names instead of nicknames.

However, ProZ and Translators Café are hardly accessible tools outside the industry. Even though the dictionaries and glossaries generated by users' activity are available for everyone through search engines, their high level of specificity prevents them from being a useful resource for non-professionals. Moreover, as discussed previously, ProZ and Translators Café support exclusively short texts for translation requests.

As far as the four wikis are concerned, they are still far from being wholly satisfactory UGT tools. Either scarce user-friendliness and poor interface design (WWL/Der Mundo, Cucumis) or very limited user participation (Traduwiki, WikiTranslate) prevent the websites from being a viable alternative to professional translators.

All things considered, the only user-generated dictionary included in my selection (Word Reference) possibly represents the best instance of a UGT tool for online texts available at present. Indeed, it makes translation activities accessible to amateurs and professionals alike in the friendly environment of forums, and is a useful 'human' addition to trustworthy dictionaries.

6.3 Is UGT affecting the professional translation industry?

At present User-generated Translation is not affecting the professional translation industry because it is not a popular phenomenon. However, it would be in the best interests of professionals and academics to observe its development and to find ways to deal with it. Indeed, recent Internet history tells us how Internet revolutions have the power radically to change "offline" scenarios. When in 1998 Napster, a software for freely distributing and exchanging music, made its appearance, its potential was largely underestimated or even ignored by the music industry. But in a few years it had already become an uncontrollable phenomenon and in less than a decade has resulted in substantial financial loss, job cuts, and a plethora of lawsuits (Mann 2003).

The mistake made by the music industry reveals that online revolutions must be embraced rather than overlooked. Even though at present the UGT tools and their users cannot compete against the translation industry in terms of efficiency and adequacy, they might soon do so. I believe User-generated Translation – possibly under another name – could soon transform into a phenomenon before we even notice it, and radically change the Web. Whether it will be fuelled by the collective utopia to break down linguistic barriers or by media giants' business operations, UGT is happening, and demands large-scale quantitative and qualitative study.

^{1 &}quot;[...] to spy the red ants filing past, / breaking, then twining, massing / at the tips of the tiny sheaves." (translated by William Arrowsmith)

² Eliot Van Buskirk (2008). "Reznor vs. Radiohead: Innovation Smackdown," *Wired*, March 27. http://www.wired.com/entertainment/music/news/2008/03/reznor_radiohead (consulted 17.01. 2009).

A Million Penguins, "Welcome." http://www.amillionpenguins.com/wiki/index.php/Main_Page (consulted 17.01. 2009).

4 Sähköautot—Nyt! "Electric Cars - Now! in English." http://www.sahkoautot.fi/eng (consulted 17.01/.2009).

Sami Torma (2008). "Crowd-sourcing the e-car." *Reuters,* July 23. http://www.reuters.com/article/technologyNews/idUSL1870916120080723 (consulted 17.01.2009).

5 For a comprehensive list of popular wikis: Wikipedia contributors, "List of wikis." *Wikipedia, The Free Encyclopedia,* http://en.wikipedia.org/w/index.php?title=List_of_wikis&oldid=231780648 (consulted 19.01.2009).

6 Internet World Stats (2008). "Internet World Users by Language. June." http://www.internetworldstats.com/stats7.htm (consulted 06.01.2009).

7 Wikipedia, "Multilingual Statistics," http://en.wikipedia.org/wiki/Wikipedia:Multilingual statistics (consulted 06.01.2008).

8 Wikipedia contributors, "Open source," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Open_source&oldid=227774619 (consulted 17.01.2009).

Wikipedia contributors, "Open source," Wikipedia, The Free Encyclopedia, http://it.wikipedia.org/w/index.php?title=Open_source&oldid=17765378 (consulted 17.01.,2009).

- 9 Alexa Internet, Inc. (www.alexa.com) is a company providing information on website traffic. "The traffic rank is based on three months of aggregated historical traffic data from millions of Alexa Toolbar users and data obtained from other, diverse traffic data sources, and is a combined measure of page views and users." The closer a site gets to 1, the more popular it is. For more information, visit: http://www.alexa.com/site/help/traffic_learn_more (consulted 17.01.2009).
- Google PageRank is an algorithm which calculates the importance of websites on the basis of hyperlinks. "PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves 'important' weigh more heavily and help to make other pages 'important'." Popularity is counted on a scale of whole numbers from 1 (least popular) to 10 (most popular). For more information, visit: http://www.google.com/corporate/tech.html (consulted 17.01.2009).
- 11 ProZ, "KudoZ Term Questions Translators Helping Translators." http://www.proz.com/kudoz (consulted 17.01.2009).
- 12 ProZ, "English to Italian translations [PRO]." http://www.proz.com/kudoz/english_to_italian/music/2218414-singing_and_dancing.html (consulted 19.01.2009).
- 13 Translators Café, "TCTerms Questions." http://www.translatorscafe.com/tcterms/EN/thDefault.aspx (consulted 17.1. 2009).
- 14 Traduwiki, collaborative translation Anybody is a translator. "About." http://traduwiki.org/About/About (consulted 17.01.2009).
- 15 Traduwiki, collaborative translation Anybody is a translator. "FAQ." http://traduwiki.org/Faq (consulted 17.01.2009).

16 Word Reference, "About Word Reference." http://www.wordreference.com/english/AboutUs.htm (consulted 19.01.2009).

Bibliography

- Benkler, Yochai (2002). "Coase's Penguin, or, Linux and The Nature of the Firm." Yale Law Journal 112, n° 3 (December). http://www.yalelawjournal.org/112/3/369_yochai_benkler.html (consulted 18.01.2009).
- Benkler, Yochai (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* New Haven, Conn.: Yale University Press.
- Bonabeau, Eric (1999). "Editor's Introduction to Stigmergy." Artificial Life 5, n° 2 (Spring). http://www.mitpressjournals.org/doi/abs/10.1162/106454699568700 (consulted 18.01.2009).
- Borges, Jorge Luis (2008). Ficciones. New York: HarperCollins.
- Boutin, Paul (2006). "Web 2.0. The new Internet 'boom' doesn't live up to its name." Slate, March 29. http://www.slate.com/id/2138951/ (consulted 18.01.2009).
- Brantley, Peter (2007). "Translating Potter." *O'Reilly Radar*, July 30. http://radar.oreilly.com/2007/07/translating-potter.html (consulted 18.01.2009).
- Cronin, Michael (2003). Translation and Globalization. London-New York: Routledge.
- Dawkins, Richard (1976). The Selfish Gene. Oxford: Oxford University Press.
- Díaz-Cintas, Jorge and Pablo Muñoz Sánchez (2006). "Fansubs: Audiovisual Translation in an Amateur Environment." *JoSTrans, Journal of Specialised Translation* 6 (July). http://www.jostrans.org/issue06/art_diaz_munoz.php (consulted 18.01.2009).
- Doctorow, Cory (2008). "Chinese restaurant called TRANSLATE SERVER ERROR," Boing Boing, July 15. http://www.boingboing.net/2008/07/15/chinese-restaurant-c.html (consulted 17.01.2009).
- Elliott, Mark Alan (2007). "Stigmergic Collaboration: A Theoretical Framework for Mass Collaboration." PhD diss., Centre for Ideas, Victorian College of the Arts, University of Melbourne.
- Feinstein, Justin (2008). "Wikipedia is like..." Blank is Like Blank, January 2,. http://blankislikeblank.com/?p=14 (consulted 17.01.2009).
- Gregorio, Joe (2002). "Stigmergy and the World Wide Web." *Bitworking*. http://bitworking.org/news/Stigmergy (consulted 18.01.2009).
- Grossman, Lev. "Time's Person of the Year: You." *Time Magazine*, December 13, 2006. http://www.time.com/time/magazine/article/0,9171,1569514,00.html (consulted 18.01.2009).
- Hallberg, Garth Risk (2008). "Who Is Grady Harp? Amazon's Top Reviewers and the fate
 of the literary amateur." Slate, January 22. http://www.slate.com/id/2182002/
 (consulted 18.01.2009).
- Heylighen, Francis (2007). "Why is Open Access Development so Successful? Stigmergic organization and the economics of information." In B. Lutterbeck, M. Bärwolff & R. A.

- Gehring (eds.), *Open Source Jahrbuch 2007*, Lehmanns Media. http://arxiv.org/abs/cs/0612071 (consulted 18.01.2009).
- Karp, Scott. "Death of the User." (2006). *The Blog Herald*, December 27. http://www.blogherald.com/2006/12/27/death-of-the-user/ (consulted 18.01.2009).
- Keen, Andrew (2007). The Cult of the Amateur: How Today's Internet Is Killing Our Culture and Assaulting Our Economy. London: Brealey.
- Kelly, Kevi(2006). "On the option of being anonymous." *The Technium*, January 27. http://www.kk.org/thetechnium/archives/2006/01/on_the_option_o.php (consulted 19.01.2009).
- Kelly, Kevin (2006). "More anonymity is good." *The World Question Center*. http://www.edge.org/q2006/q06_4.html (accessed Jan 19, 2009).
- Knight, Will (2002). "Distributed program to translate many languages." *New Scientist*, April 2. http://www.newscientist.com/article.ns?id=dn2115 (consulted 19.01.2009).
- Leuf, Bo, and Cunningham, Ward (2001). *The Wiki Way. Collaboration and Sharing on the Internet.* Reading, Mass.: Addison-Wesley.
- Mann, Charles (2003). "This Year the Music Dies." Wired 11, n° 2 (February). http://www.wired.com/wired/archive/11.02/dirge.html (consulted 19.01.2009).
- McConnell, Brian (2007). "The Worldwide Lexicon Reloaded." O'Reilly, January 2. http://www.oreillynet.com/pub/a/etel/2007/02/01/the-worldwide-lexicon-reloaded.html (consulted 19.01.2008).
- McConnell, Brian. "Accidental Translation: Creating A Global Translation Community." Worldwide Lexicon Blog, July 26, 2008. http://blog.dermundo.com/original/9468.html (consulted 10.08.2008, now unavailable).
- O'Hagan, Minako (1996). *The Coming Industry of Teletranslation*. Clevedon: Multilingual Matters.
- O'Hagan, Minako. "Multidimensional Translation: A Game Plan for Aufiovisual Translation in the Age of GILT." *Proceedings of the EU High Level Scientific Conferences: Multidimensional Translation (MuTra).* 2005. http://www.euroconferences.info/proceedings/2005_Proceedings/2005_proceedings.html (consulted 19.01.2009).
- O'Hagan, Minako (2006). "Teletranslation Revisited: Futurama for Screen Translators?"
 Proceedings of the EU High Level Scientific Conferences: Multidimensional Translation
 (MuTra).
 2http://www.euroconferences.info/proceedings/2006_Proceedings/2006_O'Hagan_Minak
 o.pdf (consulted 19.01.2009).
- O'Reilly, Tim (2005). "What Is Web 2.0, Design Patterns and Business Models for the Next Generation of Software." *O'Reilly*, September 30. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1 (consulted 19.01.2009).

- Schneier, Bruce (2006). "Anonymity won't kill the Internet." *Wired*, January 12. http://www.wired.com/politics/security/commentary/securitymatters/2006/01/70000 (consulted 19.01.2009).
- Stallman, Richard. "Why 'Open Source' misses the point of Free Software." *Philosophy of the GNU Project*, 2007. http://www.gnu.org/philosophy/open-source-misses-the-point.html (consulted 19.01.2009).
- Venuti, Lawrence (1995). *The Translator's Invisibility: a History of Translation*. London-New York: Routledge.
- Wilson, Chris (2008). "The Wisdom of the Chaperones. Digg, Wikipedia, and the myth of Web 2.0 democracy." Slate, February 22. http://www.slate.com/id/2184487/ (consulted 19.01.2008).

Biography

Saverio Perrino is a freelance translator and subtitle editor. He graduated cum laude from Bologna University's Advanced School of Modern Languages for Interpreters and Translators (SSLMIT) in 2007 and holds an MA with distinction in Audiovisual Translation from Roehampton University. He has taught Italian at Exeter University, and has worked in the media industry in London and Paris. His scholarly interest focuses on the relationship between screen translation and Web 2.0.



saverio.perrino@gmail.com