



HAL
open science

TraduXio Project: Recent Developments and Users' Feedback

Philippe Lacour, Aurélien Béné

► **To cite this version:**

Philippe Lacour, Aurélien Béné. TraduXio Project: Recent Developments and Users' Feedback. Journal of Data Mining and Digital Humanities, In press. hal-02920044v1

HAL Id: hal-02920044

<https://hal.science/hal-02920044v1>

Submitted on 24 Aug 2020 (v1), last revised 21 Dec 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TraduXio Project: Recent Developments and Users' Feedback

Philippe Lacour¹, Aurélien Béné²

1 Universidade de Brasilia, Brazil

2 Université de Technologie de Troyes, France

*Corresponding author: traduxio@philippelacour.net

Abstract

TraduXio is a digital environment for computer assisted multilingual translation which is web-based, free to use and with an open source code. Its originality is threefold - whereas traditional technologies are limited to two languages (source/target), *TraduXio* enables the comparison of different versions of the same text in various languages; its concordancer provides relevant and multilingual suggestions through a classification of the source according to the history, genre and author; it uses collaborative devices (privilege management, forums, networks, history of modification, etc.) to promote collective (and distributed) translation.

TraduXio is designed to encourage the diversification of language learning and to promote a reappraisal of translation as a professional skill. It can be used in many different ways, by very diverse kind of people. In this presentation, I will present the recent developments of the software (its version 2.1) and illustrate how specific groups (language teaching, social sciences, literature) use it on a regular basis.

In this paper, I present the technology but concentrate more on the possible uses of *TraduXio*, thus focusing on translators' feedback about their experience when working in this digital environment in a truly collaborative way.

Keywords

translation, computer-assisted, texts base, collaborative, multilingual, users feedback, comparison

INTRODUCTION

TraduXio was designed in 2006 by the Zanchin NGO (led by Philippe Lacour) and built in close collaboration with the Tech-Cico Lab of the University of Technology of Troyes, in France (<https://recherche.utt.fr/technologies-for-cooperation-interaction-and-knowledge-in-collectives-tech-cico>). Prof. Aurélien Béné, the Head of Tech-Cico, has been developing open-source software focusing on computer assisted interpretation for years, particularly in the humanities and social sciences (<https://hypertopic.org>). *TraduXio* benefited from both

public (UNESCO, Organisation Internationale de la Francophonie, French Ministry of Culture (DGLFLF)) and private funding (IATIS NGO) for a total budget of approximately 60,000 euros (until now). This sum was mainly used to pay a computer scientist (first Diana Zambon, then Franck Eyraud) for coding. Other people participated, on a temporary basis, for web development (Assid Taounza), testing (Yuliya Goncharova) and fund raising (Any Freitas). A few graduate students from the Web Design class at the University of Technology of Troyes and from Telecom ParisTech also took part in the project.

TraduXio is a digital environment for computer assisted multilingual translation. It is a free device, with an open source code (GNU/GPL3) – technical details can be consulted online: <https://github.com/Hypertopic/TraduXio>. It is web based and not machine based which means it does not need to be downloaded. It was developed using innovative technology which is especially designed for multilingual translation and is particularly suited to tackling the challenges of cultural (non-commercial, non-repetitive) texts. It promotes the creation of common goods guided by a logic of pooling (gradual feeding of the text base).

As opposed to the vast majority of past and existing (MT, Human-assisted MT (post-edition), CAT) tools, *TraduXio* does not rely on the postulates of Natural Language Processing for which semantic units are sentences, translation consists of sending a message from a source to a target language (by pairs) and which focuses on universal rules (grammatical, statistical (including deep learning), or a combination of both). Instead it considers that texts are the real semantic units, that translation consists in transforming a text into other texts in many different languages and that singularities are important for the technology of language (allowing only for contextualized generalization). *TraduXio*'s uniqueness therefore lies in its threefold originality. First, whereas traditional technologies are limited to two languages (source/target), *TraduXio* enables the comparison of different versions of the same text in various languages. Also, its concordancer provides relevant multilingual suggestions through a classification of the source according to the history, genre, author, etc. Finally, inspired by the strong collaborative spirit of “Web 2.0”, it uses collaborative devices (privilege management, forums, networks, history of modification, etc.) to promote collective (and distributed) translation – for details, see [Lacour et al, 2010 and 2011].

TraduXio's first draft version was designed in 2007 and a version 1.0 created in 2009 using technologies such as PHP and PostgreSQL with some basic functionalities. An improved version was ready in 2011 and included more efficient code and advanced collaborative functionalities (including tags, chats and forum). Unfortunately, its performance remained somewhat questionable. In 2014, the use of CouchDB allowed for a more robust and rapid version (2.0), which immediately gave more satisfaction to the users. Collaborative dimensions were added in the version 2.1 (in 2017), like privilege management, licensing, glossaries, localization (English, French, Spanish, Portuguese), although others have yet to be implemented (forum, history of modification).

TraduXio is designed to encourage the diversification of language learning [Lacour et al, 2013] and to promote a reappraisal of translation as a professional skill. It can be used in many different ways by very diverse kinds of people. In this paper, I shall focus less on the technology and the legal aspects [Lacour and Freitas, 2011] than on the possible uses of *TraduXio*, thus focusing on translators' feedback about their experience when working in this digital environment in a truly collaborative way. I will successively study *TraduXio*'s impact in the respective fields of: language teaching, the humanities and social sciences, corpus building (and browsing), massive abstract translation, multilingual publishing management and literature.

I LANGUAGE TEACHING

First of all, language students can use the platform to propose multilingual translations of assigned texts while language teachers can easily supervise the translation, propose corrections, compare drafts and evaluate students' questions and hesitations using the online interface. I shall now discuss how a translation teacher in Korea deals with *TraduXio* both inside and outside the classroom.

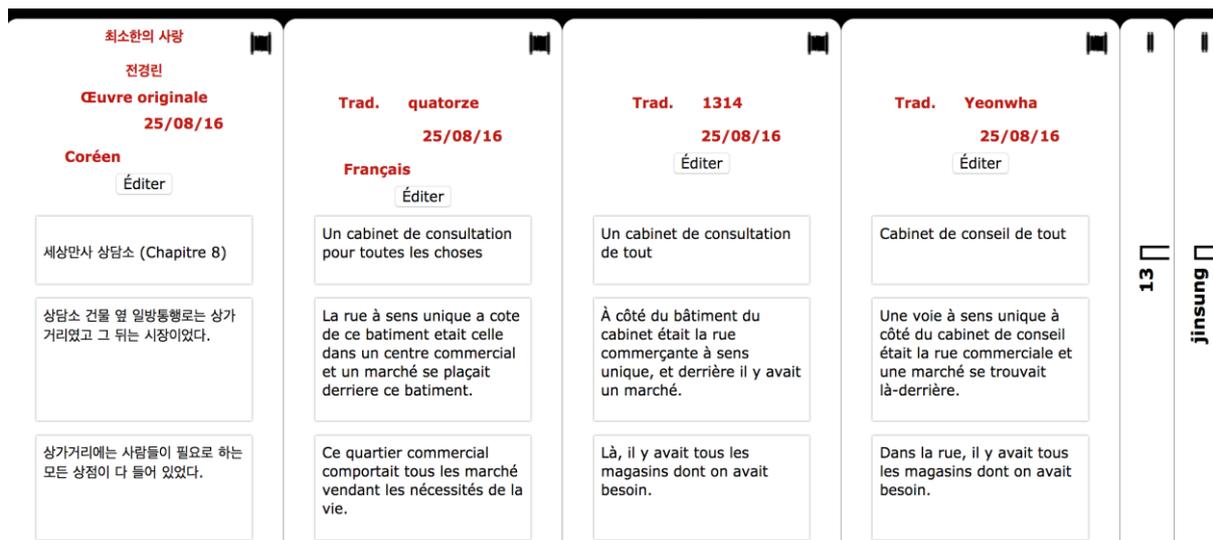
The teacher has succeeded in having students translate classical texts from French literature (Gaston Miron, Jacques Godbout, Eric-Emmanuel Schmitt, Gabrielle Roy) into Korean and English.

Figure 1: Translation of Roy's text into Korean

Original (French)	Translation (Korean)	Translation (English)
<p>Ne sachant plus que tenter pour me distraire, me consoler, m'ayant vainement offert à manger tout ce qu'elle pouvait avoir de si bon à la maison, elle finit par dire:</p>	<p>더 이상 내 주위를 둘러 울음을 멈추게 할 방법을 찾아낼 수 없었던 할머니는 별 소득 없이 그저 그녀의 집에 있을 수 있는 맛있는 것들만 모두 내어 놓다가는, 결국 이렇게 말했다.</p>	<p>이런 내 기분을 돌리고 울음을 달래기 위해 어찌해야 할지 몰라 안절부절 못하던 할머니는 집 안에 있는 맛있는 것이라든가 맛있는 것은 죄다 내왔다. 그러다 결국 할머니 입에서 이런 말이 나오곤 말았다.</p>
<p>Si tu cesse de l'irer, je vais te faire une <catin></p>	<p>"그만 훌쩍거리면 '미미'를 만들어줄게."</p>	<p>"얼른 똑 그치면 '에쁜이'를 만들어 줄게."</p>
<p>Du coup mes pleurs cessèrent.</p>	<p>갑자기 눈물이 똑 그쳤다.</p>	<p>그러자 별안간 눈물이 씻은 듯이 쏙 들어갔다.</p>
<p>Sceptique, je regardai ma grand-mère assise en sa haute chaise berceuse.</p>	<p>하지만 그 말이 의심스러웠던 나는 키가 큰 혼들의자에 앉아있는 할머니를 뻔히 바라보았다.</p>	<p>하지만 그 말이 미심쩍었던 나는 높은 혼들의자 위에 앉아 있는 할머니를 뻔히 바라보지만 할 뿐이었다.</p>

He can also translate from Korean

Figure 2: Translation from Korean into French



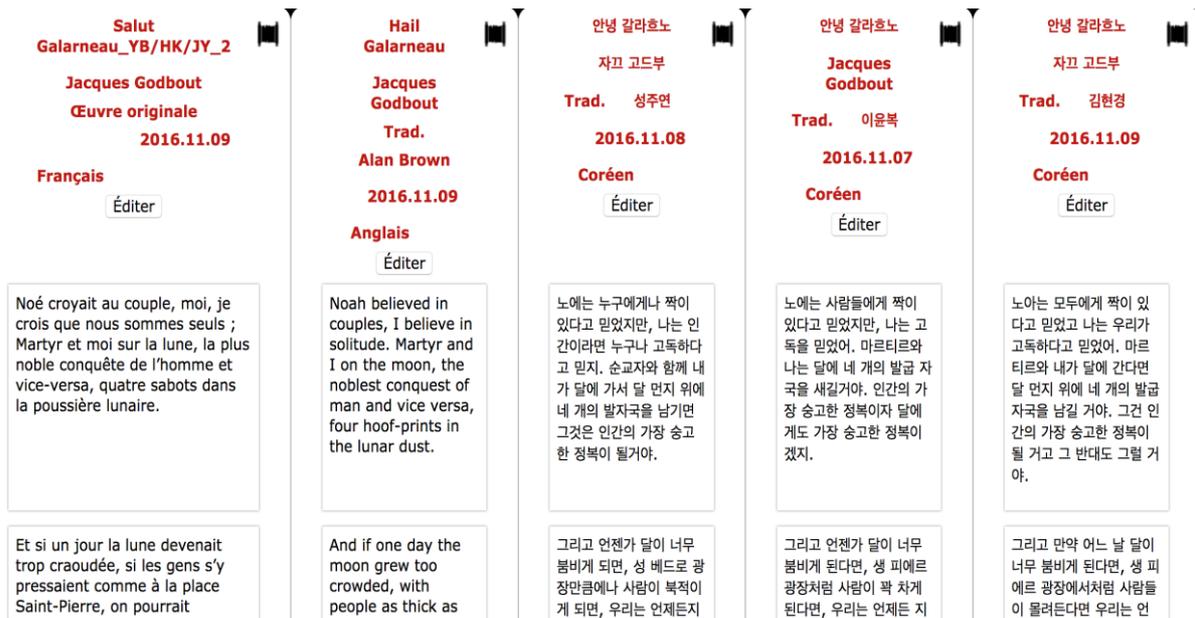
He can analyze four to six versions of the same translation on the same page, thus comparing the various efforts, focusing on the main mistakes and clever solutions. He can thus allow for hesitations and alternatives within a certain limit while at the same time ruling out mistaken interpretations.

Figure 3: Comparing, hesitating, correcting



The instructor can focus on the Korean language or draw the comparison across languages, by referring to the English version, thus evaluating the signification in a more dynamic and contrasted way.

Figure 4: Translation of Jacques Godbout's text into Korean and English



or translate in other Asian languages

Figure 5: Translating from French into various Asian languages



He also accustoms his students to comparing their own work themselves in a spirit of emulation. The language teacher therefore uses *TraduXio*'s multilingual device to its fullest potential.

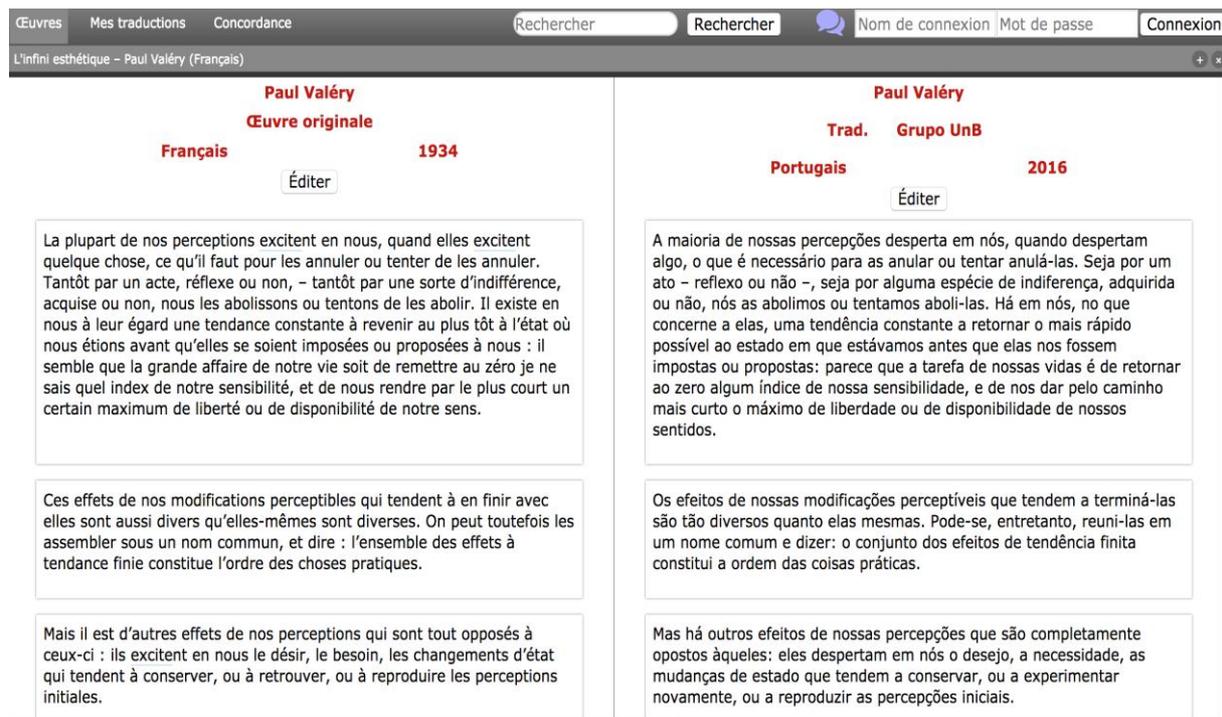
II TRANSLATING THE HUMANITIES AND SOCIAL SCIENCES

Second, *TraduXio* is also suited to use by scholars and Academic Departments particularly in the fields of Literature and the Social Sciences. Specialists can translate in a precise and critical way, respecting the highest academic standards, by creating multilingual glossaries or building a customized corpus thus creating a dedicated 'translation memory' for any topic or author. Users can manage text privileges, thus deciding which translations will be available to the public or remain private (and to what extent) and whether they can be modified or even

cancelled. Leaders can thus distribute responsibility while avoiding miscarriages of translation due to inappropriate collaboration.

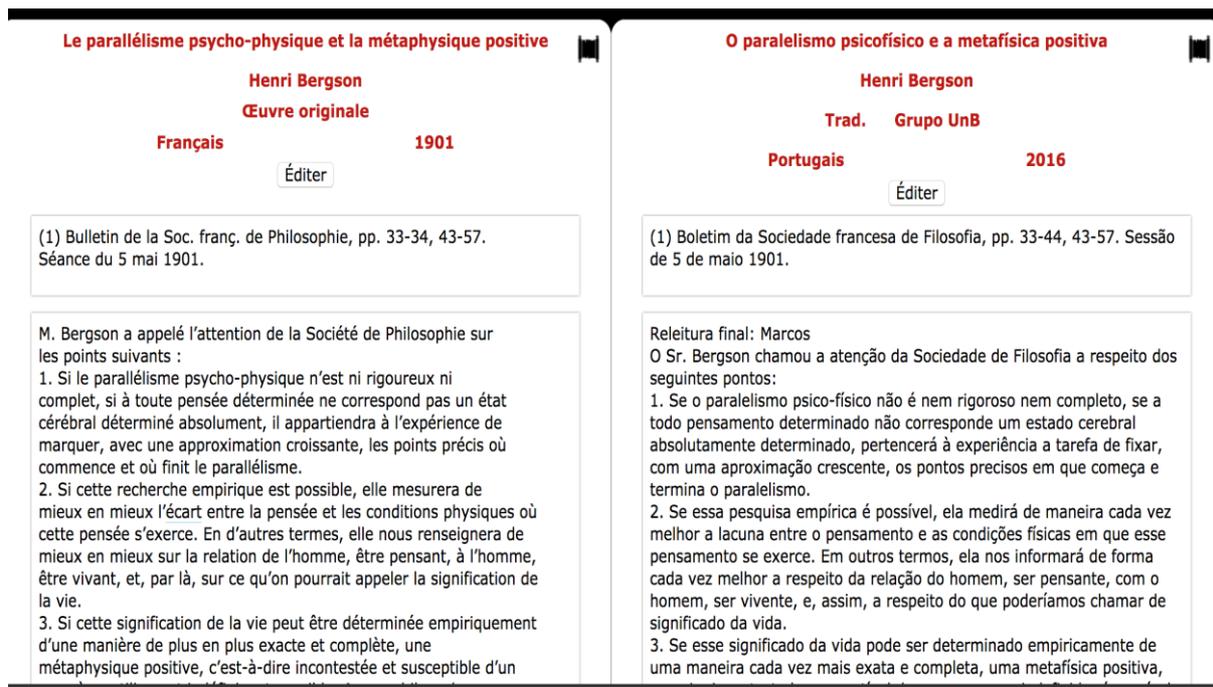
In this case, I shall refer to my own experience as a philosophy teacher at the Federal University of Brasilia, where I have been leading a workshop for the translation of French philosophical classics into Portuguese. Here is an example of an author (Paul Valéry) whose work and thoughts are often half-way between literature and philosophy.

Figure 6: Translating Valéry into Portuguese



The group is composed of different kinds of students (translation studies, philosophy, even other disciplines like International Relations) who are both Brazilian and French-speaking (from Benin). None of the participants has full mastery of the three main required skills (French language, Portuguese language, philosophy). Therefore, everyone needs to rely on each others' capacities and engage in true collaboration through constructive dialogue.

Figure 7: Translating Bergson into Portuguese



The epistemic aspect of the discussions must be must underlined, particularly how the group overcame disagreements through a process of “fine tuning” in order to come up with a satisfying solution. Also, one must stress the contrast between work inside and outside the classroom, in order to elucidate the malfunctions of collaboration and how these could be fixed from both a human and technical point of view.

III BUILDING A TRANSLATION CORPUS

Thirdly, I will also insist on the experimentation carried out with the Faculty of Literature and Translation in Porto (Portugal).

The budget is really modest and the operation is restricted to opening a special digital text base inside the main *TraduXio* website with a special access restricted to the members of the Centre for English Translation and Anglo-Portuguese Studies.

Figure 8: CETAPS’ web page

Œuvres Mes traductions Concordance



“ La traduction révèle alors le texte à lui-même : en quelque sorte, le texte semble inachevé tant qu'il n'est pas traduit. François Rastier

Traduction participative de textes culturels



ANGLAIS

- ▼ Shakespeare
 - Troilus and Cressida (pt)
 - The Tempest (fr)
 - Tempest (pt)
 - Othello
- ▼ William Shakespeare
 - As you like it (pt)

ITALIEN

- ▼ Manzoni
 - Promessi sposi

PORTUGAIS

- ▼ Shakespeare
 - Soneto 18
- ▼ William Shakespeare
 - king john
 - COMO VOS APROUVER

The project consists in constituting a text base of (half of) the Portuguese translation of Shakespeare’s plays. In so doing, the idea was to allow students to browse and look for concordances and also to foster corpus linguistics on Elizabethan theatre.

Figure 9: Portuguese translation of Shakespeare’s play “As you like it”

<p>As you like it William Shakespeare Œuvre originale</p> <p>English 1623</p> <p>Éditer</p> <p>As you Like it.</p> <p>Actus primus. Scœna Prima.</p> <p>Enter Orlando and Adam.</p> <p>Orlando.</p> <p>As I remember Adam, it was vpon this fashion bequeathed me by will, but poore a thousand</p>	<p>Como vos aproouer William Shakespeare Trad. Fátima Vieira</p> <p>Portuguese 2008</p> <p>Éditer</p> <p>Como vos aproouer</p> <p>Acto I. Cena I</p> <p>Entram Orlando e Adão.</p> <p>Orlando.</p> <p>Se bem me lembro, Adão, foi assim que em testamento me foram deixadas umas míseras mil coroas e, como tu dizes, encarregado o meu irmão, para ter a bênção de</p>
--	--

So far, the project only concerns a bilingual corpus.

Figure 10: Concordancer’s suggestions for the Portuguese translation of “entreat”

Œuvres Mes traductions Concordance entreat Search

Lang

Cha.
entri
him
pers
from

Du.
you
ther
chal
gers
will
bee
you
moo

Ori.

Ros.
ther
new

Ros.
And
Had
I sh
entri
Ere

TraduXio-IATIS

Welcome to **TI** platform

Collaborative & multilingual translation tool for
delivering an English-into-Chinese book of abstracts
for the IATIS Hong Kong Conference (July 2018)

If you are a member please **Start**

If you are not a member yet and want to learn about and/or get involved in the project, please click [here](#).




Contact

He
mu
Tr
da

languages as one desires, thus improving the multilingual dimension of the corpus. Adding another translation language is an operation that can easily be repeated over time (by inserting a supplementary column), whatever the segmentation initially chosen, be it for the source text or for any of the translations. This is a genuinely subtle characteristic that data sets cannot offer. It could also be interesting for students and teachers alike to work together in the *TraduXio* environment on the translation of the rest of Shakespeare’s plays which have yet to be retranslated. They could then constitute a reflexive and thorough glossary of the lexicon and make full use of the concordancer. This would be even more the case because they would use a multilingual corpus for inspiration.

IV MANAGING THE TRANSLATION OF ABSTRACTS FOR AN INTERNATIONAL CONFERENCE

The fourth project deals more with pragmatic or technical translation than with their “literary” counterpart although there is a link between the two aspects. The International Association for Translation and Intercultural Studies asked *TraduXio* to help them prepare their international symposium in Hong-Kong in 2018.

Figure 11: IATIS NGO customized Traduxio’s web page

This was a very challenging project requiring more funding and manpower than the others. The operation is quite complex, since it involves taking abstracts and metadata from the

EasyChair platform then translating them and sending them back in preparation for them being located on the IATIS website. The *TraduXio* environment was chosen because of its unique capacity to deal with a multilingual workflow and to distribute translation. In this project three hundred original English versions had to be translated into Chinese. This required a team of about forty properly trained translators (which involves a tutorial, the right training and the localization of the software). A supervisor was needed to act as both trainer and project and community manager with technical code skills. A set of extra-functionalities was also needed, including an automatic importing and exporting device for abstracts (from the IATIS symposium's 'comfortable' collaborative space to the customized IATIS-*TraduXio* web environment and back), a set of fixed “roles” or “tasks” (translation begun, translation completed, under revision, revision completed, finalizing (double check), finalization completed), automatic segmentation per sentence.

Figure 12: role distribution among translators working on the IATIS-*TraduXio*'s interface

#	Title	Author	Language	Type of Presentation	Thematic Panel	Keywords	Translation status	Translator	Language	Revisers
100	Incorporating Translation Reception into Translation Policy through Actor-Network Theory	Alireza Jazini	en	Panels	Panel16-i	Translation Policy, Translation Management, Translation Practice, Translation Beliefs, Translation Reception, Actor-Network Theory	1 - Translating	IATIS (iatis)	zh	
100	Incorporating Translation Reception into Translation Policy through Actor-Network Theory	Alireza Jazini	en	Panels	Panel16-i	Translation Policy, Translation Management, Translation Practice, Translation Beliefs, Translation Reception, Actor-Network Theory	3 - Revising	Julie Boéri (julie.boeri)	zh	julie.marin
101	Cognates in translated and non-translated Dutch	Lore Vandevoorde	en	Panels	Panel01	corpus-based, translated language, original language, cognates, false friends	2 - Translated	IATIS (iatis)	zh	

The fundamental asset of the operation is for IATIS to begin shaping a multilingual texts base that can be gradually fed and reused for its future symposiums.

V MANAGING MULTILINGUAL ARTICLE FLOWS FOR AN ONLINE PUBLICATION

The fifth project is editorial in nature and led by the online general-interest magazine *Sens Public* (<http://sens-public.org>).

This digital journal was created ten years ago using one of the best CMS technologies at the time which is indeed still used by the Quebec digital editing platform Erudit (<http://erudit.org>). The technology remains reliable, robust and rapid but suffers from certain limits in terms of functionalities and potential evolution. The *Sens Public* journal's plan is therefore to upgrade its CMS to the highest level using an innovative tool. While working on this, the idea arose to insert a multilingual translation device inside the information workflow to allow the editorial team to publish various different language versions of the same text and gradually build a multilingual corpus with their own articles and contributors. Once more,

TraduXio was chosen because of its singular nature, its unique capacity to manage multilingual content and propensity to harness collaborative work carried out by a group of different people worldwide who can work asynchronously.

Figure 13: Sens Public’s web page



In this case again, the idea is to carefully invent adapted functionalities to design a kind of “back office” where articles, written in different original languages, can be stored and translated into French, English or Portuguese. It also required proper training being given to the editorial and translation team in order for them to diligently manage the multilingual content directly on the CMS.

VI ORGANIZING LITERATURE TRANSLATION ONLINE

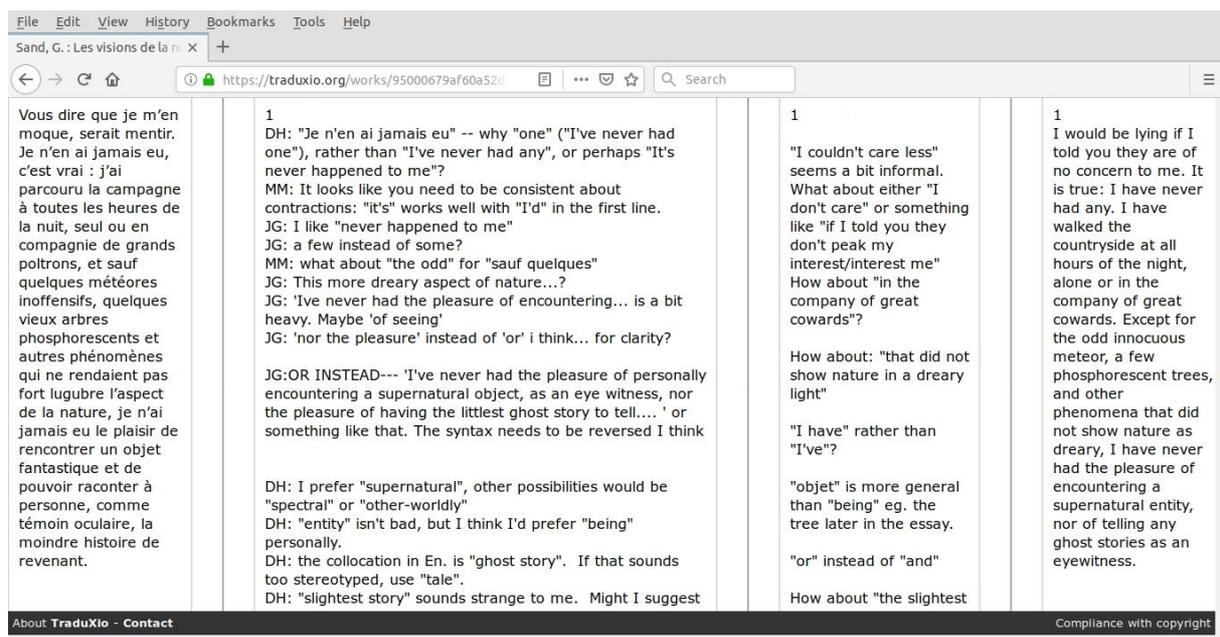
The *TraduXio* digital environment can of course finally be used to translate literature as it was specifically designed for this and particularly for semantically difficult texts (religion, philosophy, etc). Two kinds of experimentation have been led in this respect. The first project was led by a single translator, Yuliya Goncharova, interested in translating French literature into Russian [Lacour and Goncharova 2011]. It helped identify specific needs related to the length of the texts, their lexical richness and their original setting and page layout. The translator worked on Rimbaud’s famous early poem, “Le dormeur du val” using *TraduXio*’s concordancer for lexical and semantic analysis. This tool is particularly helpful when it comes to metaphors and images, when a translator needs suggestions and can use the concordancer to retrieve relevant ideas taken from other literary sources. The advantage (for truncation, homonyms and polysemic terms) and limits (lack of recognition of vowel alternation which is very common in Russian; the inability to identify suppletion) are also clearly underlined.

Figure 14: Translation of the beginning of Rimbaud’s “Le dormeur du val” into Russian

<p>C’est un trou de verdure où chante une rivière Accrochant follement aux herbes des haillons D’argent ; où le soleil, de la montagne fière, Luit : c’est un petit val qui mousse de rayons.</p>	<p>В зеленом разломе, где речка поет, Росою сплетая траву в паутину, Где с гордой вершины солнце свет льет, Щедро купая в том свете долину.</p>
--	--

The second project was led by Daniel Henkel and involved coordinating groups of translators (both graduate students and teachers) from three different academic institutions - University Paris 8 (France), University of Berkeley (USA) and University of Florence (Italy) - for six months plus one intensive week [Henkel and Lacour 2020]. This time *TraduXio* was mainly used for confronting ideas and critical discussion (also using other canals of communication) and involved many students discovering for the first time the advantages and limits of collective translation and the necessity of examining one’s own point of view. The students were organized in small groups of 3-4 people to translate English short stories into French and Italian while often comparing their ideas with other groups before finally harmonizing their work under the supervision of the faculty members.

Figure 15:

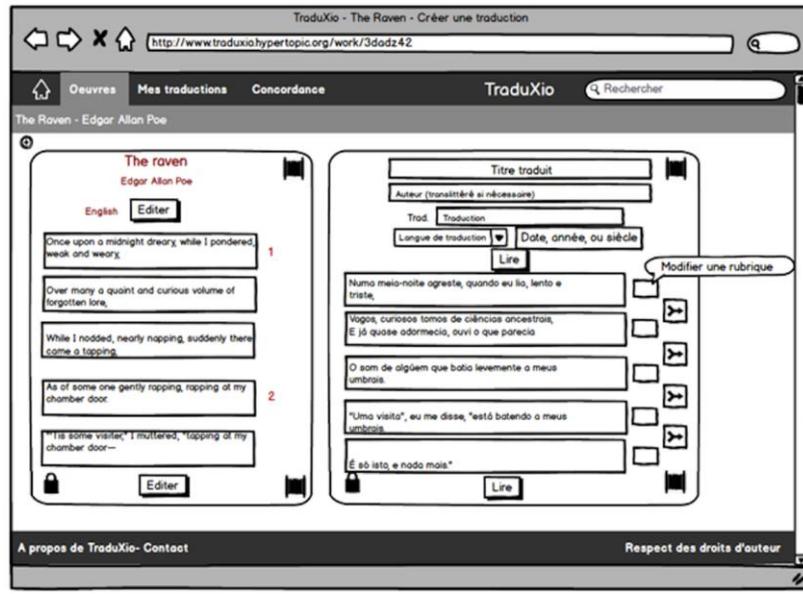


Users’ feedback was also crucial to gather information in order to imagine future developments of the TraduXio project. Students (of the Web design class) from the University of Technology of Troyes (France) also contributed to this study by participating in defining the next steps for the process, thus paving the way for future improvements (<https://github.com/Hypertopic/TraduXio/issues>).

In particular, two functionalities are of great importance and both rely on the same piece of technology. The first is more oriented towards erudition as it allows for reference to be made to the original version of the source text, for instance, Kant’s *Critique of Pure Reason: A*

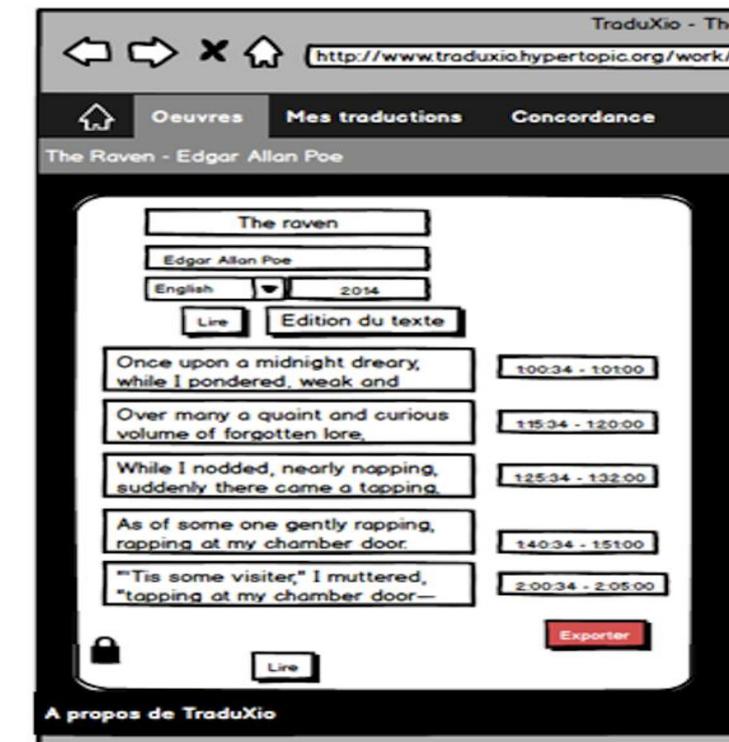
(first) or B (second edition). It could become of immense value for philological use of *TraduXio*.

Figure 16: *TraduXio*' possible reference to the original version of Poe's poem "The Raven"



Now, as often happens in computer science, one functionality can lead to very different uses and user adoption (fork). The same idea of referring to the original version of a document is relevant to mention the exact chronological location of a specific subtitle in a video. *TraduXio* could therefore also be used - this time in an industrial way (culture, entertainment) - for digital multilingual subtitling.

Figure 17: *TraduXio*'s possible use for the translation of video subtitles



References

- Lacour P., Bénel A., Eyraud F., Freitas A. and Zambon D. TIC, Collaboration et Traduction: vers de nouveaux laboratoires de translocalisation culturelle. *Meta*. 2010; 55(4):674-692. <https://www.erudit.org/fr/revues/meta/2010-v55-n4-meta4003/045685ar.pdf>
- Lacour P., Freitas A. Translation and the New Digital Commons. *Métiers et technologies de la traduction: quelles convergences pour l'avenir ?*, Proceedings of the Tralogy symposium, Paris, March 3-4, 2011. <http://odel.irevues.inist.fr/tralogy/index.php?id=150>.
- Lacour P., Goncharova Y. Traduxio: nouvelle expérience en traduction littéraire. *Traduire* (Société Française des Traducteurs). 2011; 225. <https://journals.openedition.org/traduire/94>
- Lacour P., Bénel A., Eyraud F. Towards a Collaborative Platform for Cultural Texts Translators. Pierre Maret (ed.), *Virtual Community Building and the Information Society: Current and Future Directions*. IGI Global (Hershey, Pennsylvania), 2011.
- Lacour P., Bénel A., Eyraud F., Freitas A., Zambon D. Enhancing Linguistic Diversity through Collaborative Translation. *TraduXio: an Open Source Platform for Multilingual Workflow Management in Media*. Elin Haf Gruffydd Jones and Enrique Uribe-Jongloed (eds.), *Social Media and Minority Languages. Convergence and the Creative Industries*. Multilingual Matters (Bristol, Buffalo, Toronto), 2013.
- Henkel D., Lacour P. Collaboration Strategies in Multilingual Online Literary Translation: the case of the *TraduXio* environment. Desjardin R., Larssonneur C., Lacour P. (eds.), *When Translation Goes Digital. Case Studies and Critical Reflections*. Palgrave (London/New York), 2020.