

The Impact of Machine Translation on a Masters Course in Web Translation: From Disrupted Practice to a Qualitative Translation/Revision Workflow



Translating and the Computer 41

London, 21-22 November 2019

Christopher GLEDHILL

Maria ZIMINA-POIROT

CLILLAC-ARP EA3967

75013, Paris, France



université
PARIS
PARIS 7
DIDEROT



Outline

- Website translation project TSA
- Challenges of MT-enhanced translation:
‘disrupted practice’ → ‘qualitative revision workflow’
- Project impact:
The role of metalanguage in students’ reporting forms
- Conclusion & References

Master 2 Pro Industrie de la langue et traduction spécialisée (M2 ILTS), Paris Diderot

- Master's degree in **Specialized Translation & Language Industry**
- **Work-study scheme**
- **Teaching: professionals** who are familiar with the actual conditions of work
- Translation courses in **English-French, French-English, French-Spanish and German-French (700 h)**
- Coord. Geneviève BORDET & Nicolas FROELIGER)



The EMT is a network of MA study programmes in translation

- **Translation** (technical, scientific, specialised...)
- **Intercultural communication**
- **Project management**
- **Terminology**
- **Revision**
- **Localization**
- **Professional integration**

The Website Translation Project

Traduction de site vers l'anglais (TSA)

- Translation from **French into English** of **webpages** taken from a French university or other institution (1-2 pages per student, 30-35 participants)
- Example: Institut de Physique du Globe de Paris (IPGP)
<http://www.ipgp.fr>
- **Genres** involved **range** from scientific reports to administrative and procedural announcement

TSA: a simulated environment

Roles

- **Project Manager:** communication between the teachers, translation team and website owners
- **Translator:** translation/post-editing
- **Reviser:** post-editing/revision *pair*
- **Auditor:** an audit of the website, manageable tasks
- **Contract writer:** an agreement with the website owners
- **Terminologist:** glossary of terms, menu items, etc.
- **Task manager:** managing tasks on the translation platform

Project deliverable

Address Book Access Contacts Intranet Webmail En | Fr Search... 

 **INSTITUT DE PHYSIQUE DU GLOBE DE PARIS**
Earth - Planets - Environment - Natural hazards

INSTITUTE | RESEARCH | OBSERVATION | EDUCATION | OUR PLANET

OBSERVATION

Volcanological and seismological observatories

Magnetic Observatory National in Chambon-la-Forêt

GEOSCOPE Observatory

Observatory of water and erosion in the Caribbean

IPGP data center

Observation

The IPGP equips and maintains two worldwide geophysics networks to survey magnetic field variations (INTERMAGNET network) and the seismic activity of the planet (GEOSCOPE network).

It observes and studies active volcanoes in France and its territories as well as the tectonic context in Martinique, Guadeloupe and La Reunion Island in order to provide, thanks to its permanent observatories, improved monitoring of the natural hazards due to the internal activity of the globe.

The Observatories carry out their own research activities in collaboration with scientific teams from countries all over the world.



Translation / revision carried out by the Masters 2 ILTS team

TSA course challenges

- **Transversality:** a wide variety of training outcomes as defined by EMT
- **Web-tectonics:** a translation training model within a permanently shifting multimodal context
- **Non-linear Task Management:** a complex pedagogical workflow using a professional translation platform
- **Process-oriented Assessment:** a qualitative self-assessment of the translation/revision workflow

Competence Framework 2017



[Competency No.]

“4 [...] mastering [...] **presentation standards, terminology and phraseology**, specialised sources etc.)”

“10 [...] **using the appropriate metalanguage** and applying appropriate theoretical approaches”

“11 [...] **review and/or revise their own work and that of others**”

“18 Master the **basics of MT** and its impact on the translation process”

“19 Assess the relevance of **MT systems in a translation workflow**”

“23 **Work in a team**, [...] using current communication technologies”

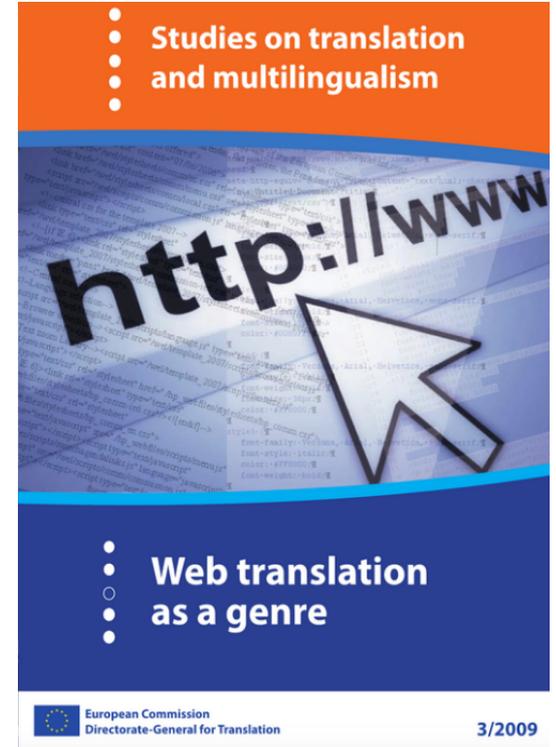
“29 Clarify **the requirements, objectives and purposes of the client**”

TSA course challenges

- **Transversality:** a wide variety of training outcomes as defined by EMT
- **Web-tectonics:** a translation training model within a permanently shifting multimodal context
- **Non-linear Task Management:** a complex pedagogical workflow using a professional translation platform
- **Process-oriented Assessment:** a qualitative self-assessment of the translation/revision workflow

The translation of web content requires flexibility and sensitivity to many different genres

- High visibility
- Short text paragraphs
- Standalone (webpage)
- Keywords for search engines, hyperlinks
- Labels and navigation menus (self-service)
- Always updated (never-ending process)
- Concise, clear wording



TSA course challenges

- **Transversality:** a wide variety of training outcomes as defined by EMT
- **Web-tectonics:** a translation training model within a permanently shifting multimodal context
- **Non-linear Task Management:** a complex pedagogical workflow using a professional translation platform
- **Process-oriented Assessment:** a qualitative self-assessment of the translation/revision workflow

SYSTRANLinks interconnected with **SYSTRAN Pure Neural Server**

- Manages collaborative translation workflow
- Identifies website updates
- Maintains consistency
- Instantly shows how translations will look online
- Enables Neural MT specialisation with custom resources
- Uses Translation Memory & User Dictionary
- Outputs translation in HTML format
- Helps with localization

MT Engine(s)

SYSTRANLinks

Use this box to translate text or to search the Dictionary and the Translation Memory...

Maria Zimina (zimina) ▾

http://www.lpgp.fr (20.↵)

Source Language:
French

Target Language:
English ▾

Visitor Preview

Browse & Edit

Hide Toolkit

Dashboard

▾ Settings

Configuration

Translation Settings

Display Mode

▾ Content

Pages

Sentences

Images & Links

▾ Resources

User Dictionary

Translation Memory

Tasks

Contributors

Review and manage translations

Validate Sentences

Publish Sentences

Select Tag ▾

Create Task

Export

View in context

<input type="checkbox"/>	▾	Context ▾	Source Sentence ▾	Usage ▾	Translation	Status ▾	Visi... ▾	▾
<input type="checkbox"/>	2017-0...	segment	Séisme des îles du Dodeca...	1	Earthquake of the Dodecan...			0
<input type="checkbox"/>	2017-0...	segment	Le séisme au large de Bodr...	1	The earthquake off Bodrum...			0
<input type="checkbox"/>	2017-0...	segment	Etude des volcans géants ...	1	Study of giant Kamchatka v...			0
<input type="checkbox"/>	2017-0...	segment	Il observe et étudie les volc...	1	It observes and studies the ...			0
<input checked="" type="checkbox"/>	2017-0...	segment	Une partie des équipes de ...	2	Some of IPGP's research t...			0

25018 sentences - 177827 words

Previous

Next



Source Sentence:

Une partie des équipes de recherche de l'IPGP seront également localisées à Paris Rive Gauche, proche de l'UFR STEP.

Terminology:

Translated Sentence: Machine Translation (SES8) (history)

Some of IPGP's research teams will also be located in Paris Rive Gauche, near UFR STEP.



Alternatives:

Diff.

TM

MT 1

Suggestions

Other languages

Machine Translation (SYSTRAN)

Part of the research teams of the IPGP will be also localised in Paris Rive Gauche, near to UFR STEP.

13

Custom Neural MT: User Dictionary (UD)

SYSTRAN Pure neural SERVER Outils de traduction

Traduction de texte Traduction de fichiers Traduction de la parole Traduction de pages Web

Français → Anglais

Sélection du profil : IPGP Propriétaire Generic M

Le Piton de la Fournaise

Le Piton de la Fournaise s'est édifié sur le flanc sud-est du Piton des Neiges. Durant au moins 500 000 ans, les deux volcans ont fonctionné simultanément. La Fournaise forme un cône de 30 km de diamètre, couvrant 26 % de la superficie de l'île.

Cet édifice se distingue du premier par l'uniformité de ses produits : il s'agit de laves fluides à l'origine d'éruptions principalement effusives. Il est formé par l'empilement sur des milliers de mètres d'épaisseur de coulées de laves, aa et pahoehoe, alternant avec des couches de scories, le tout recoupé par un réseau de dykes plus ou moins interconnectés.

Piton de la Fournaise

Piton de la Fournaise was built on the southeast **flank** of the **Piton des Neiges**. For at least 500,000 years, the two volcanoes operated simultaneously. La Fournaise forms a cone 30 km in diameter, covering 26% of the island's surface area.

This building stands out from the former because of the uniformity of its products: these are fluid lava that produces mainly effusive **eruptions**. It is formed by stacking thousands of meters thick lava flows, aa and **pahoehoe**, alternating with layers of **scoria**, all intersected by a network of **dikes** more or less interconnected.

Different phases of activities are distinguished:

Afficher les résultats pour : Le Piton de la Fournaise Le Pi...
Français -> Anglais

- flanc *nom*
- flank *👤*
- side *📏*

(↔) côté, partie, face, camp, aspect, bord, face avant

flanc de la montagne	side of the mountain
flanc très escarpé	very steep side
flanc gauche	left side
flanc d'une montagne	side of a mountain

flank *📏*

flanc sud	southern flank
flanc de colline	flanks of a hill
flanc sud-est	south-eastern flank
flanc ouest	western flank

SYSTRAN Pure Neural® Server 9.2 | SYSTRAN © 2018 tous droits réservés | Accord de licence

OBSERVATION

Volcanological and seismological observatories

- The Paris team
- Volcanological and Seismological Observatory of Martinique
- Volcanological and Seismological Observatory of Guadeloupe (OVSG)
- Observatory of the Piton de la Fournaise Volcano
- Crisis management
- Observatories instrumentation
- Magnetic observatories
- GEOSCOPE Observatory
- Observatory of Water and Erosion in the Antilles
- IPGP data center



Volcanological and seismological observatories

The volcanological and seismological observatories of the Antilles and Piton de la Fournaise at Réunion in association with local universities and regional and departmental authorities, constantly observe the humors of Soufrière, Mountain Pelee and the Piton of the Fournaise with numerous instruments and techniques.

This has enabled us to warn the authorities and the general population in advance before each of the eruptions of Piton de la Fournaise, which have been more frequent this last decade. The volcanos of the Antilles

Maria Zimina

Original Text:
 La prévention et l'information qui sont de la responsabilité des autorités nationales et territoriales restent les outils les plus importants,
 Translation: *TM Exact Match* Not published
 It remains essential for national and territorial authorities to provide preventative measures and information. The IPGP endeavours to contribute to this work. within the means placed at its disposal.

It remains essential for national and territorial authorities to provide preventative measures and information. The IPGP endeavours to contribute to this work, within the means placed at its disposal. The service of volcanological and seismological observatories collaborates in surveillance activities in foreign countries, such as Montserrat, the Republic of Djibouti or the Comoros.



Installation of a GPS station on Piton de la Fournaise (UnderVolc - IPGP)

ena are ongoing, but current

Concordancing, Terminology, Tasks...

Concordance

☑ Magnitude 9, profondeur 30km, zone de **subduction**

Magnitude 9, depth 30 km, subduction zone

☑ Deformation continentale, **subduction** aismique et glissement transitoire dans les Andes du Nord

Continental deformation, aseismic subduction and transitory landslip in the North Andes

☑ Sismicite et geodynamique de la zone de **subduction** du Vanuatu

Seismicity and geodynamics of the Vanuatu subduction zone

☑ Sismicite et geodynamique de la zone de **subduction** du...

Seismicity and geodynamics of the zone of subduction of...

Source Language: French | Target Language: English | Visitor Preview | Browse & Edit | Hide Toolkit

Review and manage translations

Validate Sentences | Publish Sentences | Select Tag | Create Task | Export | View in context

<input type="checkbox"/>	Context	Source Sentence	Usage	Translation	Status	Visibility	
<input type="checkbox"/>	2015-11-03 ... list/opti...	suprasubduction zone	1	☉ suprasubduction zones	⚙	➡	0
<input type="checkbox"/>	2015-11-03 ... list/opti...	Subduction-zone;	1	Sort: ↑ Ascending ↓ Descending	⚙	➡	0
<input type="checkbox"/>	2015-11-03 ... list/opti...	Subduction	1	Filter by URL: /en/publications/processes	⚙	➡	0
<input type="checkbox"/>	2015-11-03 ... list/opti...	Subduction zone processes;	1	Vanuatu subduction zone	⚙	➡	0
<input type="checkbox"/>	2015-11-03 ... list/opti...	Vanuatu subduction zone ;	1	Clear Filter	⚙	➡	0

Manage translation tasks

Edit | Information | Cancel

<input type="checkbox"/>	Name	Assigned to	Content	Type	Due date	Source Lang	Target Lang	Description	Progress	Status	Cre...
<input type="checkbox"/>	A10	Ines Couly ...	Sentence Set	Full Review (Edit, ...	2020-06-30	French	English		<div style="width: 20%;"></div>	In Pr...	coul...
<input type="checkbox"/>	B 49	Maissa Ha...	Sentence Set	Full Review (Edit, ...	2020-08-29	French	English		<div style="width: 40%;"></div>	In Pr...	Mai...
<input type="checkbox"/>	A31	Sarah Reh...	Sentence Set	Post-Editing only	2019-11-28	French	English		<div style="width: 60%;"></div>	In Pr...	Reh...
<input type="checkbox"/>	B11	Ines Couly ...	Sentence Set	Full Review (Edit, ...	2020-06-30	French	English		<div style="width: 20%;"></div>	In Pr...	coul...
<input type="checkbox"/>	A 20	Maissa Ha...	Sentence Set	Full Review (Edit, ...	2020-08-15	French	English		<div style="width: 40%;"></div>	In Pr...	Mai...
<input type="checkbox"/>	D55	Sarah Reh...	Sentence Set	Post-Editing only	2019-11-29	French	English		<div style="width: 60%;"></div>	In Pr...	Reh...

Revision history (tracing) gives an idea of how many revisions have been made

Entry history



User	Date	Change
Chris Gledhill (ChrisGledhill)	13 days ago	Status changed from 'validated' to 'rejected' Visibility changed to Not Published
Yuxuan Guo (yuxuan.guo)	a month ago	Status changed from 'edited' to 'validated' Visibility changed to Published Sentence changed from 'Commercialization of the IPGP' to 'Research valuation'
Sylvie Kepez (skepez43)	a month ago	Status changed from 'mt' to 'edited' Sentence changed from 'Knowledge Exchange' to 'Commercialization of the IPGP'
unknown	a month ago	Status changed from 'edited' to 'mt' Sentence changed from 'Commercialization of the IPGP' to 'Knowledge Exchange'
Sylvie Kepez (skepez43)	a month ago	Status changed from 'mt' to 'edited' Sentence changed from 'Knowledge Exchange' to 'Commercialization of the IPGP'
unknown	a month ago	Status changed from 'edited' to 'mt'
Karolin Voigt (KarolinVoigt)	a month ago	Status changed from 'mt' to 'edited'
unknown	a month ago	Status changed from 'validated' to 'mt'

TSA course challenges

- **Transversality:** a wide variety of training outcomes as defined by EMT
- **Web-tectonics:** a translation training model within a permanently shifting multimodal context
- **Non-linear Task Management:** a complex pedagogical workflow using a professional translation platform
- **Process-oriented Assessment:** a qualitative self-assessment of the translation/revision workflow

Project Reporting Form

Report on the task

Demonstrate knowledge of certain features

Comment on the project

M2 ILTS

Teaching team: Chris Gledhill, Maria Zimina

Fiche d'évaluation

Traduction de site vers l'anglais (TSA)

PROJECT REPORTING FORM

Fill in all the cells in the right-hand column of this form in English

NAME, Given names	
Speciality (IL, TS)	
Task (Genre / Register)	
Title	
Source	
Style Guide	
Volume	
Revision cycles	
Revision edits	
Key examples of translation issues 1. Context of situation 2. Discourse-semantics 3. Lexico-grammatical structure 4. Phono-graphic substance	
Remarks on bibliography, sources, reference material	
Project workflow issues 1. Editing workflow & workflow engine 2. Interacting with Revisers and Website Editor	
SystranLinks issues 1. Translation Memory (TM) 2. User Dictionary 3. Sentence List Resources 4. Rules (if applicable)	
Corpus creation & analysis tools Briefly explain the role of tools in website translation/revision. Give examples of specific problems that can be solved using CAT tools and corpora.	

Language issues

1. Context of situation
2. Discourse-semantics
3. Lexico-grammatical structure
4. Phonic-graphic substance



Context of situation: This is the “Latest news” section of the website. So the key info needs to be given at the beginning of the article, rather than making reader search for it. Have to wait until 3rd paragraph of French for the info. I began with “A new study has been published”, info that we don’t get in the French version until mid-way through the text. This is more appropriate to the **context**.

Discourse-semantics: The first sentence of the web page is: **Malgré la magnitude du séisme meurtrier du 12 janvier 2010, le contexte sismotectonique et les failles actives dans la zone épiscopentrale étaient encore mal connus.**

It is then immediately followed by a long paragraph of scientific detail. This leads to a problem of meaning and begs the question ‘encore mal connus’ until what? What has happened, what is the new development? In the French we don’t find out until later on, which confuses the reader. **Malgré la magnitude du séisme meurtrier du 12 janvier 2010, le contexte sismotectonique et les failles actives dans la zone épiscopentrale étaient encore mal connus.**

Changed the English to be clearer and avoid misunderstandings:

A new study has been published that helps to explain the deadly earthquake of 12 January 2010. Until recently the seismotectonics and active faults in the epicentral zone of this upheaval were poorly understood.

Contexte sismotectonique du sud d'Haïti : un nouveau modèle pour le séisme de magnitude Mw7.0 du 12 janvier 2010.

Malgré la magnitude du séisme meurtrier du 12 janvier 2010, le contexte sismotectonique et les failles actives dans la zone épiscopentrale étaient encore mal connus.

Le sud de l'île d'Haïti est traversé d'est en ouest par une faille décrochante majeure : la faille d'Enriquillo-Plantain-Garden (EPGF). Les modèles proposés pour le séisme de 2010 excluaient jusqu'alors une rupture le long de cette faille. Cette difficulté à comprendre l'origine et le mécanisme du séisme est principalement due au fait que ce séisme, malgré sa magnitude, n'a pas été associé à des ruptures en surface.



Seismotectonic context in southern Haiti: a new model for the Mw7.0 magnitude earthquake of 12 January 2010.

Key information first

A new study has been published that helps to explain the deadly earthquake of 12 January 2010. Until recently the seismotectonic context and active faults in the epicentral zone were poorly understood.

The Enriquillo-Plantain-Garden fault (EPGF), a major lateral strike-slip fault, cuts across southern Haiti from east to west. The models put forward for the 2010 earthquake previously excluded the possibility of the occurrence of a rupture along this fault. The difficulty in understanding the origin and physical mechanism of this earthquake was largely due to the fact that, in spite of the earthquake's magnitude, it was not associated with surface ruptures.



Reporting Forms as a CORPUS: statistically salient examples of students' metalanguage: 'main', 'tool', 'search', 'information', 'meaning', 'collocations'

Regular *lexico-grammatical patterns* revealed by **textometric analysis** :
Ex. : <using (a tool) to **search** (examples of language)>

N° ▲	Contexte
3	is to use the quote marks functionality in the google search engine, which allows users to search verbatim and to
4	in the google search engine, which allows users to search verbatim and to thus find other pages or articles where
5	different texts). I also used to corpus to search collocations. For instance, I typed "fault" to see
7	." 3. The system list resources were used to search for occurrences of terms in other documents/places on the
8	uses a lot of specialized terminology. The ability to search the entire website through SystranLinks can be particularly helpful for
12) I used the translation memory (in the concordance search) to translate but had to double check some names
13	played a crucial role. It was very useful to search the texts already translated on the website to check the
15	translate "Obsera". 3. I use the search bar in the Sentence window to look for the translations

Correspondence Analysis, Characteristic Elements: Distinctive Respondents' Profiles

vagueness or uncertainty:

'I' + adverb of uncertainty

'I mostly agreed with the changes made by...'

'I' + cognitive process

'I realized I could get a more accurate translation.'

'...once I had the topic in mind, I could focus more on style during the...'

positive learning outcome

'This' as a full pronoun

This helped with my translation choice.

'This' as a determiner relating to a tool (or a process)

'This tool gives us a broad view of...'

'This corpus allowed me to prove that the...'

Conclusion: a more equipped translator's workstation results in assisted but also disrupted translation practice

- Changing web environment
- More or less ergonomic interfaces
- Multiple windows, over-segmentation
- Blurring of communicative situation, comprehension processes, contextual interpretation of meaning
- Structural rigidity of TM-based text segmentation
- Lack of support for dynamic XML/HTML
- TBD: annotation schemes, messaging workflow
- Student expectations, teaching time/learning curve

Towards ‘editing’ and ‘revision’ within a ‘qualitative translation/revision workflow’: *editorial history...*

MT EN1: ‘Despite the magnitude of the deadly earthquake of 12 January 2010, the seismotectonic context and active faults in the epicentral zone were still poorly understood.’

Post-edited EN2: ‘Despite the magnitude of the deadly earthquake of 12 January 2010, its seismotectonic environment and the active faults in the epicentral zone were poorly understood until now.’ – *better but key info still not to the fore.*

Revision EN2: *Better to turn sentence round and start with ‘Until recently...’ so we insist on the fact there’s a new development.*

Revision EN3: ‘**A new study has been published that helps to explain the deadly earthquake of 12 January 2010.** Until recently the seismotectonics and active faults in the epicentral zone of this upheaval were poorly understood.’

References

Directorate-General for Translation (European Commission). 2013-12-16. **Web translation as a genre**. 2013-12-16.

European Master's in Translation. 2017. **Référentiel de compétences**.

Gledhill, Christopher. 2011. **A lexicogrammar approach to checking quality: Looking at one or two cases of comparative translation**. Depraetere, Ilse (éd.), *Perspectives on Translation Quality (Text, Translation, Computational Processing 9)*. Berlin : Mouton de Gruyter, 71-98.

Gledhill, Christopher & Natalie Kübler. 2015. **How Trainee Translators Analyse Lexico-Grammatical Patterns**. *Journal of Social Sciences – Special issue on Phraseology, Phraseodidactics and Construction Grammar(s)* (ed. Maria Isabel González-Rey.), 11/3: 162-178.

Gledhill, Christopher & Maria Zimina. 2018. **Teaching Project “Website Translation into English**. *CLARIN Workshop "Translation memories, corpora, termbases: Bridges between translation studies and research infrastructures"*, 8-9 February 2018, Vienna, Austria.

Lavault-Olléon, Elisabeth & Maria Zimina (forthcoming). **Traduction et technologie: regards croisés sur de nouvelles pratiques**. *Des mots aux actes*, n°8, Société Française de Traductologie (SOFT) et de la Société d'Études des Pratiques et Théories en Traduction (SEPTET).

Zimina, Maria & Serge Fleury. 2018. **Mémoire de traduction [MT] : approche paradigmatique**. *Equivalences* 45/1-2, “Des unités de traduction à l’unité de traduction”. Sous la direction de C. Balliu, N. Froeliger, L. Hewson. 26