

LES USAGES ACTUELS DE LA TRADUCTION AUTOMATIQUE (TA)

Caroline Rossi
Univ. Grenoble Alpes
ILCEA₄ (GREMUTS)

Atelier Digit_Hum 2019

Les humanités numériques en langues

JEUDI 17 OCTOBRE 2019



Dessins de Saint-Oma, reproduits
avec la permission de l'artiste
© Saint-Oma | Digit_Hum | 2017-
2019. Licence CC BY-NC-ND 4.0.

POURQUOI LES USAGES ?

« *MT quality is now good enough that millions of people are using it every day* »

A. Way (Traditional and Emerging Use-Cases for Machine Translation, 2013)

~~UX (expérience utilisateur)~~

✓ Usages et perceptions comme révélateurs d'une valeur d'usage



- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



LE MIRACLE DE LA TA NEURONALE

Miracle

Fait extraordinaire qui porte à l'étonnement et à l'admiration, réussite exceptionnelle (Etym.) Empr. au lat. *miraculum* «prodige, merveille, chose extraordinaire»

Source : <https://www.cnrtl.fr/definition/miracle>

- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Newest GNMT: time to witness the miracle of Google Translate

Published on October 4, 2016

Google's AI translation system is approaching human-level accuracy

But there's still significant work to be done

by Nick Statt | @nickstatt | Sep 27, 2016, 2:07pm EDT

SHARE TWEET LINKEDIN

DeepL Translator – The New Wondrous Translating Machine

Jan 28, 2019

What Exactly is DeepL?

The DeepL Translator has taken off like a rocket and is going strong. It is the new miracle in AI (artificial intelligence) translation because it translates as fast as other translation machines like Google Translate and Microsoft Bing Translator. However, its translations sound more natural. So, what is DeepL?

theguardian

home > UK > education media society law scotland wales northern ireland > all

Languages The case for language learning

Tech is removing language barriers - but will jobs be lost in translation?

Could Microsoft's Star Trek-inspired translation service ever replace professional human translation?

Google Translate update makes it pretty much as good as a human translator

By Emma Boyle September 28, 2016 Computing

Getting chatty around the globe just got easier

f t p e

TAUS Blog

Jaap van der Meer

Wednesday, 24 February 2016

The Future Does Not Need Translators*

Electric Translator Is Latest 'Miracle'

The age of miracles is only now dawning, it appeared to self, then typed out:

day, with the introduction of "Vyelyichvina uzla oprye."

Oklahoman, 25 Jan 1954: 'Electrical brain can now translate foreign languages', by Alice Hughes [syndicated]

Owosso Argus-Press (Michigan), 14 Jan 1954: 'Modern machines do wonders, but there's possible errors' Pittsburgh Press, 8 Jan 1954: 'Translators, beware! You can be replaced'

Presque Isle Star Herald (Maine), 28 Jan 1954: 'New IBM electronic machine changes Russian into English'

Scholastic, 20 Jan 1954: 'Translating machine'

Science, 122, 21 October 1955, pp.745-748: 'Mechanical translation: new challenge to communication', by Jacob Ornstein [*]

Science News Letter, 23 Jan 1954: 'Language translation by electronic "brain" [*]

Scranton Times (Pennsylvania), 8 Jan 1954: 'Device hurdles word barriers'

fed into the machine and their magnetic drum.

The "brain" also was "taught" rules of grammar so it could change the order of words when necessary to make a smooth intelligent English sentence.

Prof. Leon Dostert and Dr. Paul Garvin, who directed the 18-month job of creating the translator, said it has a vocabulary of only 250 words but it potentially could store up hundreds of thousands.

"Mi pyeryedayem mislyi posryedstvom ryechiyi," she punched.

"We transmit thoughts by means of speech," the brain translated on a high-speed printer.

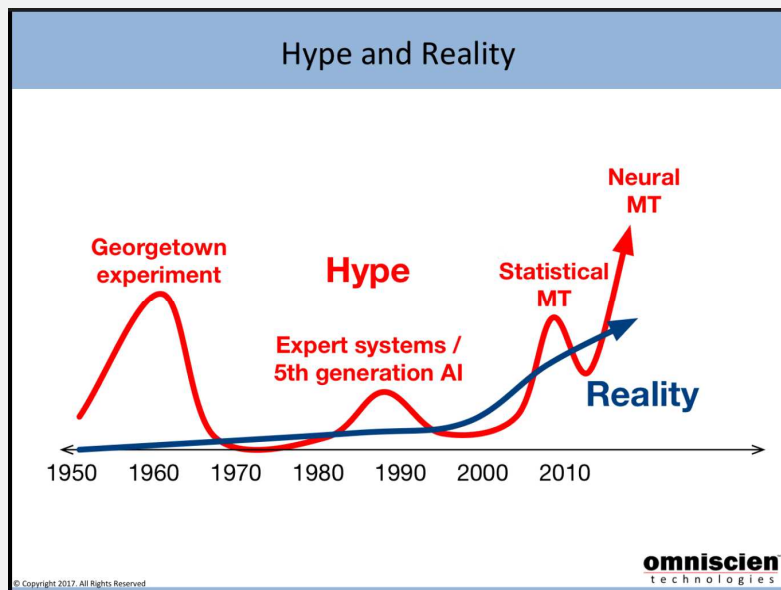
Mrs. Polle, who does not

Caroline Rossi
Univ. Grenoble Alpes

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



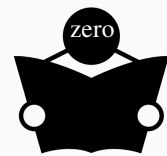
UNE CONSTRUCTION DISCURSIVE



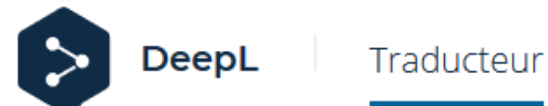
Koehn (2017)



Les données sont une ressource naturelle (Kenny 2011 : 9)



La TA sans apprentissage est une machine intuitive : *zero shot* (Rossi 2017, 2019)



La machine est un traducteur

- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



UN DISCOURS DÉTERMINISTE

« Une conception déterministe du développement technologique repose sur deux présupposés, le premier veut que les technologies se développent de façon autonome, indépendamment du social, du politique, et même de l'économie, et le second veut que le changement social découle du progrès technologique ». Olohan (Technology, translation and society, 2017)



Source : <https://medium.com/@vorapilailuckthi/technological-determinism-social-media-9232614920ef>

- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



LES USAGES ACTUELS



- ✓ Lecture et recherche d'informations : « gisting » (domaines et pratiques très diverses)
- ✓ Utilisation des sorties de TA dans l'interaction orale (Google Translate World Cup)
- ✓ Aide à la communication écrite (y compris scientifique) dans des langues inconnues
- ✓ Pratique (y compris universitaire) des langues étrangères

MT Stories (blog d'une doctorante finlandaise, Mary Nurminen) : <https://mt-stories.com/>

Machine Translation Stories

Real people finding real uses for Google Translate and other MT tools

- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Caroline Rossi
Univ. Grenoble Alpes

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



IMPLICATIONS DE CES USAGES



- ✓ Lecture et recherche d'informations :
« gisting » (domaines et pratiques très
diverses)

**Sans disparition des documentalistes,
journalistes etc.**

- ✓ Utilisation des sorties de TA dans
l'interaction orale (Google Fifa World
Cup)

**Sans disparition des interprètes (même
naturels)**

- ✓ Aide à la communication écrite (y compris
scientifique) dans des langues inconnues

Sans disparition des traducteurs spécialisés

- ✓ Pratique (y compris universitaire) des
langues étrangères

**Sans disparition des enseignants en langue
et des vocations d'étudiants**

I. Le miracle
« neuronal »

II. Les usages
actuels

III. Risques et
besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



LES USAGES ACTUELS

1. Etudiants de différentes filières (306)

Dirand et Rossi (Biotraducteur et traducteur automatique : l'homme outillé et la machine-homme, *Des mots aux actes* 2020)

Question	Réponse des étudiants en langues (219)	Réponse des autres étudiants (87)
Quand vous entendez le mot « traducteur », à quoi pensez-vous en premier ?	Une personne dont le métier consiste à traduire (76%)	Un logiciel de traduction (56%)
Pensez-vous qu'un outil technologique peut totalement remplacer un humain pour traduire un document diffusable au grand public de nos jours ?	Non, jamais (56%) Oui mais dans 10 ans (41%) Oui (3%)	Oui, mais dans 10 ans (60%) Non, jamais (31%) Oui (9%)
Avez-vous déjà utilisé un logiciel de traduction automatique comme Google Traduction ou une application mobile ?	Oui (98%) Pour comprendre une langue étrangère (82%) Pour écrire dans une langue étrangère (66%) Dans le cadre d'un travail universitaire (60%)	Oui (99%) Pour comprendre une langue étrangère (90%) Pour écrire dans une langue étrangère (29%) Dans le cadre d'un travail universitaire (19%)

- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Caroline Rossi
Univ. Grenoble Alpes

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



LES USAGES ACTUELS

1. Etudiants de différentes filières (306)

Dirand et Rossi (Biotraducteur et traducteur automatique : l'homme outillé et la machine-homme, 2020)

2. Universitaires

Bowker et Ciro *Machine Translation and Global Research* 2019

2.1 Etat des lieux

Choix majoritaire pour publier : l'anglais

Communication multilingue : aide possible de la TA

Difficultés des locuteurs d'autres langues pour publier en anglais : la TA n'est pas une solution

« If you've ever used Google Translate, you'll know how easy it is to structure sentences badly in another language. Direct translation can result in non-native sounding sentences; while this is forgivable when spoken, it's a no-no in an academic article. At best, it could make your work look sloppy, but at worst, it could change the meaning of what you're trying to say. This could mean you run the risk of your submission being rejected because it's unclear. » (Blog Elsevier http://webshop324.rssing.com/chan-52469744/all_p1.html)

2.2 Vers de nouveaux usages de la TA ?

Ecrire dans une langue simplifiée, pour faciliter la traduction (Bowker et Ciro 2019 : 72)

Réparer les sorties de traduction automatique : la post-édition (par les auteurs ou par d'autres scientifiques du domaine : experts linguistiques).

I. Le miracle
« neuronal »

II. Les usages
actuels

III. Risques et
besoins

Caroline Rossi
Univ. Grenoble Alpes

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



RISQUES ET BESOINS

The risks of Google in Polish

A Polish worker was seriously injured in a dumper truck accident after following safety instructions that had been translated into "gibberish" using Google.

Grzegorz Krzyzak, 32, who speaks pidgin English, was using a one-ton articulated dumper truck to remove soil at a nursery and garden centre near Thorpe-le-Soken, Essex. As he was tipping a load, the truck overturned and crushed his right leg.

Parker's Nurseries was convicted of breaching health and safety laws and fined GBP5,000 after bosses at the firm converted instructions and health and safety manuals using Google Translate, the search engine's free translation service. The Polish that resulted was gibberish, Colchester magistrates' court heard.

The Pole sustained multiple fractures to his shinbone and foot and required four operations. Doctors inserted pins to help repair the bones and gave him a skin graft but it is not known whether he will regain full use of his leg.

Geoff Parker, director of Parker's Nurseries, said the **firm now had an external translator to translate health and safety documents**. He added: "We are obviously very sorry that an employee was injured in our workplace"

Daily Telegraph (fev. 2013) cite par A.Way '2013)

Quelles vérifications possibles ?

La traduction inverse (back translation) est devenue plus fiable, cf. <https://www.translationparty.com/>

L'usage de l'anglais comme langue pivot

I. Le miracle
« neuronal »

II. Les usages
actuels

III. Risques et
besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



RISQUES ET BESOINS (2)

Test de Turing



Craintes (ne plus distinguer la machine de l'homme) > Vigilance accrue nécessaire

Effet d'évidence et de vérité (fiabilité) lié à :

Immédiateté (Vitali-Rosati, *On Editorialization: Structuring Space and Authority in the Digital Age* 2018)

Gratuité, et facilité d'utilisation

Invisibilité des processus (« boîte noire »)

Risque : contresens et propagation de ces erreurs dans les textes

- I. Le miracle « neuronal »
- II. Les usages actuels
- III. Risques et besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



VERS UNE NOUVELLE HABILITÉ NUMÉRIQUE

“Machine Translation Literacy”

Enseignement à développer

Pour et par les acteurs dans la communauté universitaire : enseignants chercheurs et documentalistes

Ethique de l’IA et de la TA

PHILOS. Science qui traite des principes régulateurs de l'action et de la conduite morale.(TLFi)

Tenir compte de tous les acteurs, à toutes les étapes (Kenny 2011)

Bon usage des données

Science ouverte

Respect du droit d’auteur

Pas de traduction automatique de documents contenant des informations confidentielles ou protégées

I. Le miracle
« neuronal »

II. Les usages
actuels

III. Risques et
besoins

Caroline Rossi
Univ. Grenoble Alpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019



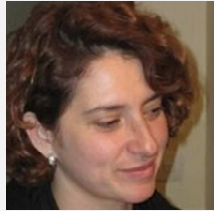
MERCI !



Dorothy Kenny
(DCU)



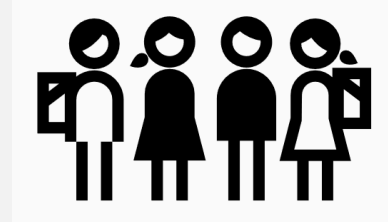
Joss Moorkens
(DCU)



Olga Torres
Hostench
(UAB)

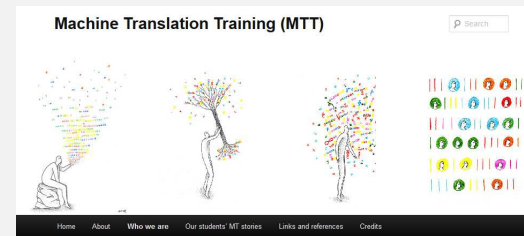


Elisabeth
Lavault-Olléon
(UGA)



Et tous les étudiants du
Master TSM de l'UGA

<https://mtt.hypotheses.org/>



Caroline Rossi
Univ. GrenobleAlpes)

Atelier Digit_Hum 2019
Les humanités numériques en langues
JEUDI 17 OCTOBRE 2019

