

GHOSTS IN THE MACHINE: CAN ADAPTIVE MT HELP RECLAIM A PLACE FOR THE HUMAN IN THE LOOP?

HANNA MARTIKAINEN



ESIT / UNIVERSITÉ SORBONNE NOUVELLE – PARIS 3
CLESTHIA - LANGAGE, SYSTÈMES, DISCOURS - EA 7345



OUTLINE

- ✓ Introduction
- ✓ Context of the experiment
- ✓ Results



HYPE SURROUNDING NEURAL MACHINE TRANSLATION

Google's Neural Machine Translation System:
Bridging the Gap
between Human and Machine Translation

2016.

Oct 3, 2016

eBay's Machine Translation
Technology Breaks Down Borders

Achieving Human Parity on Automatic
Chinese to English News Translation

Microsoft AI & Research
2018.

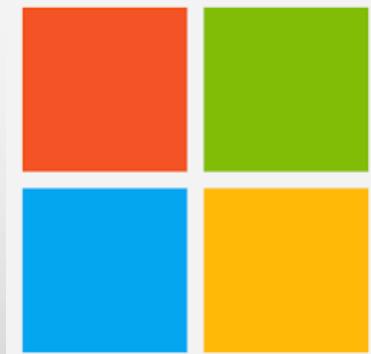
For this language direction, our translations have been declared superhuman by the WMT organizers, meaning that human evaluators preferred them over translations done by human experts.

FACEBOOK AI August 01, 2019

MT DEVELOPMENT: MILITARY AND DEFENSE PURPOSES



MT TODAY: COMMERCIAL PURPOSES



PROS AND CONS OF MT FOR PROFESSIONAL TRANSLATORS



speed = productivity

reduced typing effort

added objectivity

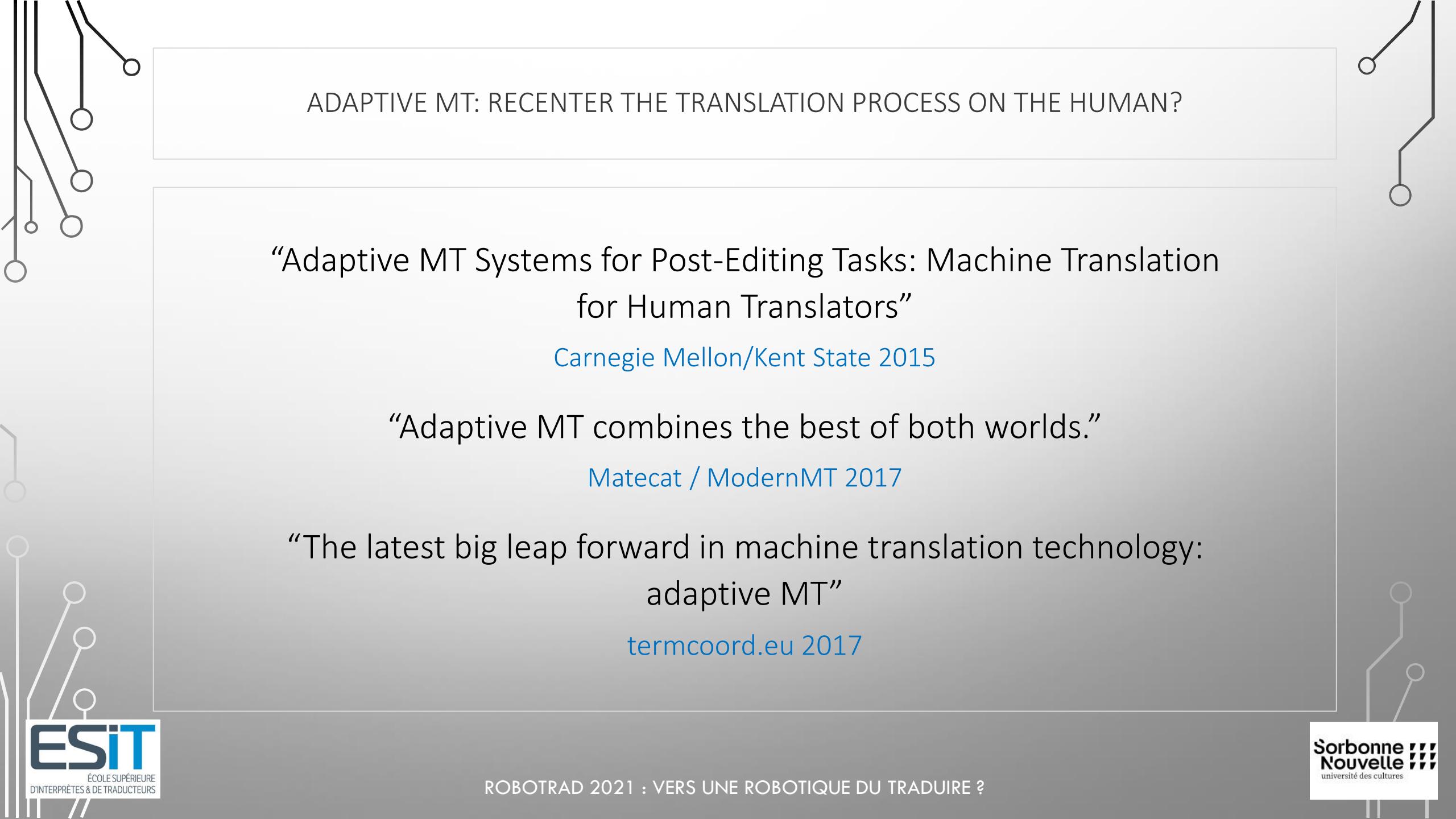
MTPE



static & tedious

lack of creativity

reduced role for human



ADAPTIVE MT: RECENTER THE TRANSLATION PROCESS ON THE HUMAN?

“Adaptive MT Systems for Post-Editing Tasks: Machine Translation for Human Translators”

Carnegie Mellon/Kent State 2015

“Adaptive MT combines the best of both worlds.”

Matecat / ModernMT 2017

“The latest big leap forward in machine translation technology:
adaptive MT”

termcoord.eu 2017

INTERACTIVE OR 'REAL-TIME ADAPTIVE' MT:



34

Syntax is normal, grammar and punctuation are correct.

La syntaxe est normale, la grammaire et la ponctuation sont correctes.

35

Aim for grammatically, syntactically and semantically correct translation.



0



Visez une traduction grammaticale, syntaxiquement et sémantiquement correcte.

36

Ensure that key terminology is correctly translated and that untranslated terms belong to the client's list of "Do Not Translate" terms.



RESEARCH QUESTIONS

Do students feel more in control when working with adaptive MT vs when post-editing static MT output?

Which parameters influence usability and enjoyability of adaptive MT?

OUTLINE

- ✓ Introduction
- ✓ Context of the experiment
- ✓ Results



human-machine interaction
translator training
ergonomics of post-editing
interactive MT
adaptive MT
adaptive MT
translator training human-machine interaction
ergonomics of post-editing adaptive MT
human-machine interaction
ergonomics of post-editing
adaptive MT interactive MT

MT & PE COURSE / S2 2020 - 2021

- Enrolled M2 students: 70

- Course program:

Module 1. Introduction to MTPE

Module 2. MT quality assessment

Module 3. Best practices in post-editing on the LSP market

→ Module 4. Group projects with adaptive MT (3 sessions)

MODULE 4 GROUP PROJECT

Fixed parameters

- MT engine: Lilt
- Min. length of text / translator
- Min. requirements for termbase
- Min. requirements for analysis

Free parameters

- Project management: groups & tasks
- Text for translation: domain, degree of specialization
- Languages: A/B/C (*Lilt limitations)

OUTLINE

- ✓ Introduction
- ✓ Context of the experiment
- ✓ Results



PROJECT LANGUAGES AND TEXTS

- Groups: 14 with 3 to 7 students / group
- Languages: 8 x EN-FR – 1 x EN - AR / IT / POR / SP / ZH – 1 x RU-EN
- Texts: general press, web pages, UN report, data sheet, EU brochure...
- Length of translated extracts: 1,000 to 5,000
- Topics: human rights, youth job program, environment and economy, meteorology, cooking, cancel culture...

MAIN OBSERVATIONS: POSITIVE ASPECTS

- ✓ Most students find interactive / adaptive MT technology promising
- ✓ Students felt less constrained by MT suggestions presented in separate space from blank target segment
- ✓ Students appreciated reduced typing effort with functional real-time adaptation:
 - MT suggestion adapts to produce the desired output with minor changes
 - Translation choices are automatically propagated
- ✓ Students appreciated Lilt interface for simplicity of project management & easy of use, e.g. click-and-add term insertion

« Les propositions de Lilt ne s'ajoutent pas directement dans la zone de traduction, mais dans une zone en dessous de celle-ci. Cela permet d'être moins influencés par le logiciel. »

« Avoir le choix entre taper toute la traduction soi-même ou se servir des suggestions (ou même travailler des deux façons à la fois) est une option intéressante et pratique »

« Lilt s'adapte très vite. Les corrections qu'il apporte à mesure de la traduction sont pertinentes, et le fait que la phrase proposée se modifie à mesure de la traduction est un vrai plus. »

MAIN OBSERVATIONS: NEGATIVE ASPECTS

- ✓ MT adaptation not always up to the task: degraded quality & introduction of errors not seen at pre-translation stage:

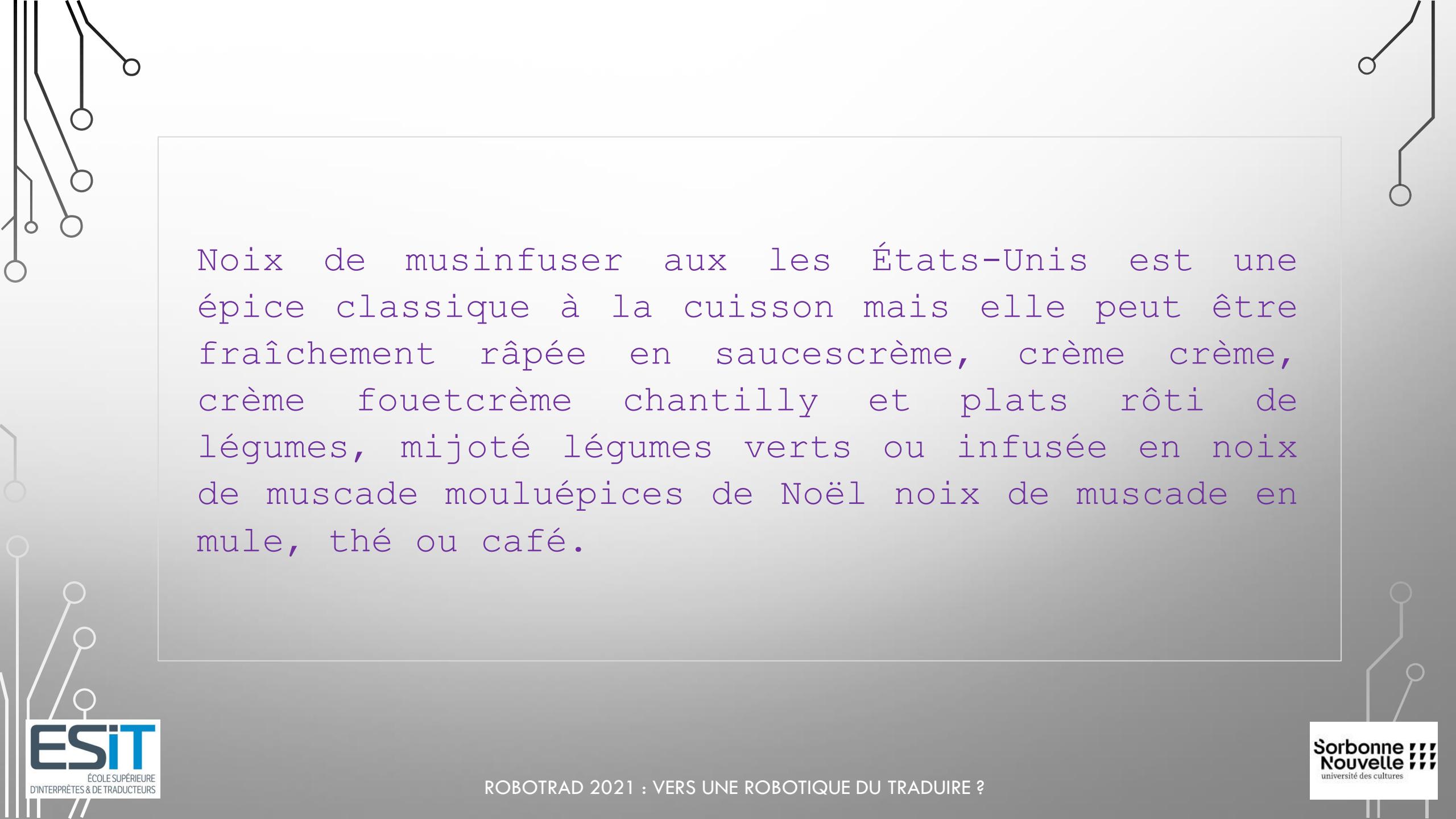
« De manière générale, chaque modification conduisait soit à une suggestion grammaticalement incorrecte, soit à une suggestion qui consistait en un déplacement de sens. »

MT QUALITY AS A KEY PARAMETER IN USER SATISFACTION

✓ Interaction with MT less important for usability than MT quality

- Quality often lacking dramatically for EN-FR
- Quality often compared negatively with competition for EN-FR (e.g. DeepL)

✓ Influence of language: higher user satisfaction for POR / AR & Lilt MT quality found superior to competition



Noix de musinfuser aux les États-Unis est une épice classique à la cuisson mais elle peut être fraîchement râpée en saucescrème, crème crème, crème fouetcrème chantilly et plats rôti de légumes, mijoté légumes verts ou infusée en noix de muscade mouluépices de Noël noix de muscade en mule, thé ou café.

« D'après l'avis de tous les membres du groupe, la traduction automatique de Lilt s'est avérée meilleure que celles de Google Translate et de DeepL, par exemple. » (EN-POR)

USUAL NMT ISSUES

- ✓ Calque / Omission / Terminological variation / Ungrammatical suggestions
- ✓ Hallucinations
 - dried rose petals: Pétathé
 - COCOVID-19
- ✓ Syntax
 - women burn survivors: les femmes brûlent les survivants
 - small women's organizations: organisations de petites femmes

MAIN MT ISSUE

- ✓ Integration of glossary terms into MT output
 - Glossary terms not always used in MT output
 - Glossary terms inserted randomly into MT output
 - Insertion of glossary terms results in degraded MT quality
- Challenge in NMT development because of lack of grammatical rules

CAT ERGONOMICS' IMPACT ON USER SATISFACTION

- ✓ CAT quality often compared negatively with competition (e.g. Trados)
- ✓ CAT ergonomics issues :
 - Unusable hyperlinks
 - Problems with shortcuts
 - Spell check not spotting all errors
 - Impossible to modify translation after revision
- ✓ High expectations for CAT: automatic declination of lemmatized glossary terms, currency conversion...

CONCLUSIONS

- ✓ Interactive / adaptive MT has the potential for better user experience
- ✓ Learning curve for interactive work but usability most with CAT & MT quality
 - Language impact on MT quality
 - Translation environment ergonomics lacking for professional use
 - Best practice workflow: integration of MT to CAT environment
- ✓ Main limitation: short experiment (MT adaptation)
- ✓ Next step: test another interactive & adaptive engine for EN-FR



MERCI POUR VOTRE ATTENTION

&

MERCI AUX ÉTUDIANTS M2 2020-2021
DE L'ESIT POUR LEUR CONTRIBUTION