universität wien

The use of speech technologies in translation, revision, and post-editing machine translation (PEMT)

Raluca Chereji, Claudia Wiesinger, Justus Brockmann, Alina Secară, Dragoș Ciobanu

universität wien

Outline

1. HAITrans research group
2. Speech recognition technology
3. Speech synthesis technology
4. HAITrans research projects

universität wien

HAITrans research group

- Human and Artificial Intelligence in Translation
  - Zentrum für Translationswissenschaft (ZTW)
  - University of Vienna

- Focus on translation technologies
  - Education
  - Industry
  - Research

https://haitrans.univie.ac.at/team/

Univ.-Prof. Dragoș Ciobanu, PhD (Lead)

Alina Secară, MA PhD (Senior Scientist) -associate member-

Miguel Rios, PhD (Post doc)

Justus Brockmann, BA MA (Prae doc)

Raluca Chereji, BA MSc (Prae doc)

Claudia Wiesinger, BA MA MA (Prae doc)

universität wien

HAITrans research group

- Education

EMT

IAMLADP International Annual Meeting on Language Arrangements, Documentation and Publications

- Industry

European Commission

GALA Globalization & Localization Association

EDUinLOC

TRANSLATION COMMONS

HAITrans research group

- Research

R1

Speech technologies in translation, revision and post-editing (PEMT)

- Speech synthesis in revision and PEMT

R2

Technology-supported translation, revision and PEMT

- NMT engine building
- Medical MT domain adaptation

R3

Translation technology for accessibility

- Bias, lexical and structural analysis of audio introductions

R4

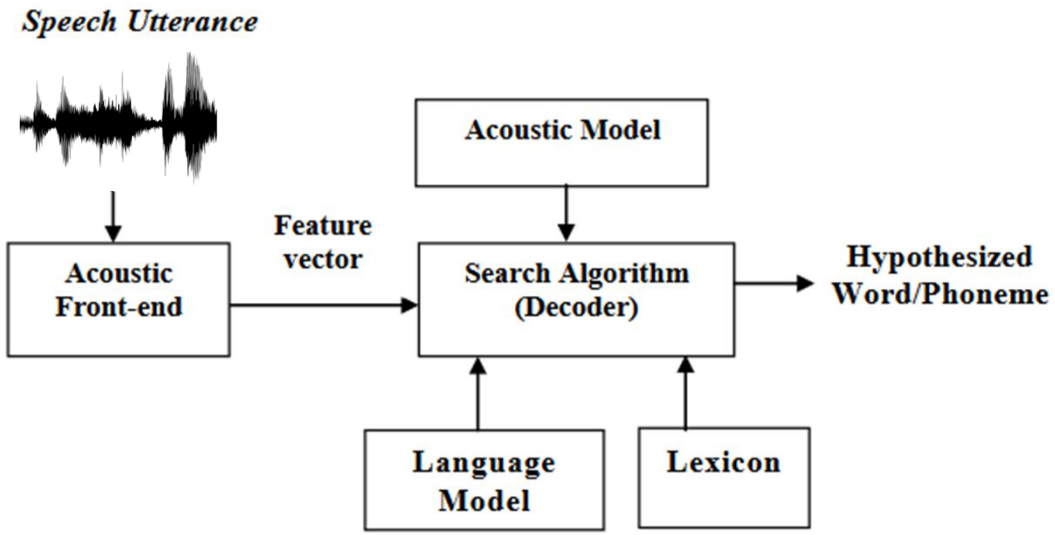
Translation technology didactics

- Improving student MT literacy levels

Automatic speech recognition (ASR)/Speech-to-text (S2T)

“allows a computer to take the audio file or direct speech from the microphone as an input and convert it into the text” (Malik *et al*., 2021: 9412)

Automatic speech recognition (ASR)/Speech-to-text (S2T)



(Karpagavalli and Chandra, 2016)

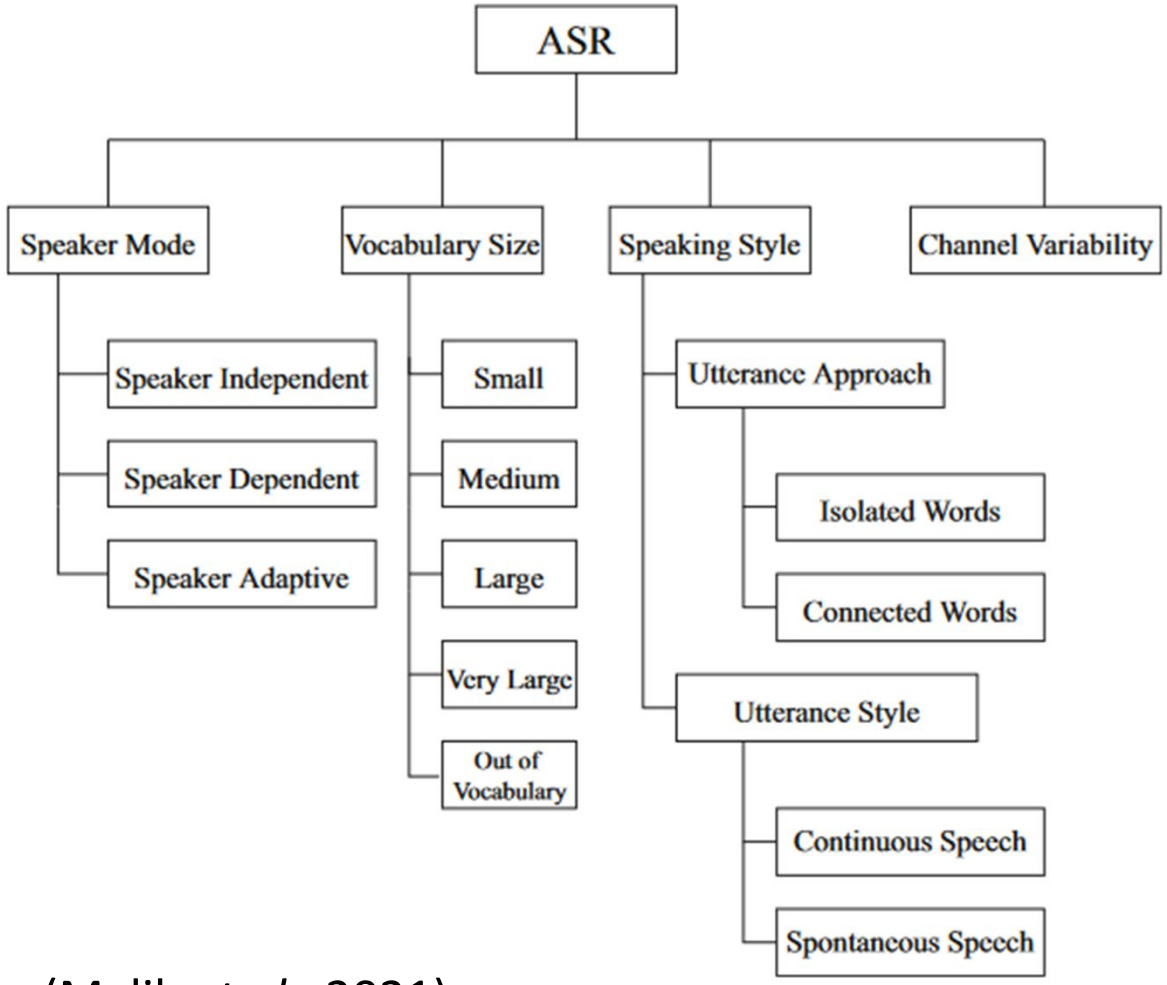
Types of ASR systems

Speaker mode

Vocabulary size

Speaking style

Channel variability



(Malik *et al.*, 2021)

ASR applications

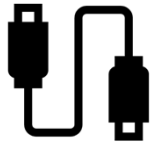
- **Wide variety of domains and industrial use cases:**

Healthcare, medical assistance and telemedicine

Telecommunications industry

Forensics and law enforcement

Defense and aviation



IT and consumer electronics

Home automation and security

Education and accessibility

(Vajpai and Bora, 2016; Kanabur, Harakannanavar and Torse, 2019; Ibrahim and Varol, 2020)

universität wien

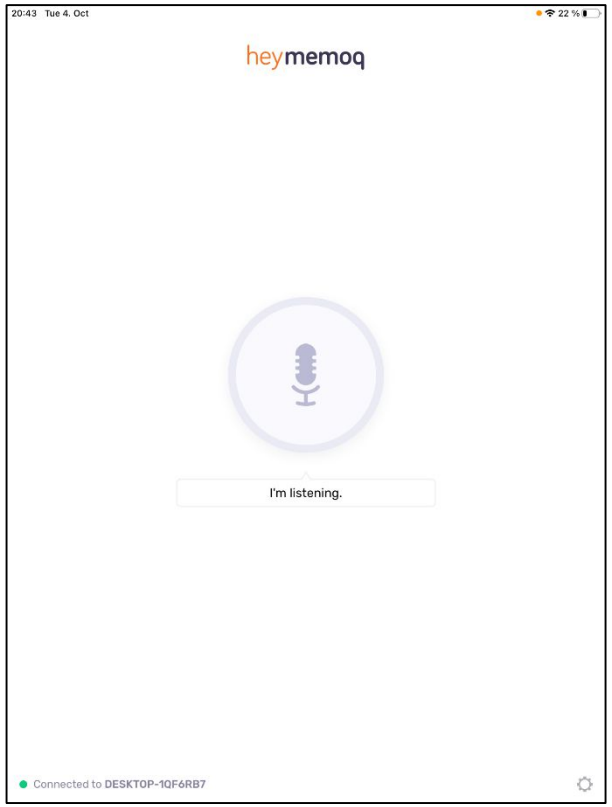
ASR applications

- **In a translation context:**
  - **Translation Process Research** on translation dictation and post-editing
    - (Ciobanu, 2014, 2016; Mesa-Lao, 2014; Zapata *et al*., 2017; Tzoukermann and Miller, 2018; Liyanapathirana, 2021)
  - ASR for **interpreting**
    - (Defrancq and Fantinuoli, 2021; Gaber and Corpas-Pastor, 2021)
  - **Respeaking** in audiovisual translation
    - (Romero-Fresco, 2020)
  - Multilingual **corpora building**
    - (Di Gangi *et al*., 2019; Iranzo-Sánchez *et al*., 2020; Salesky *et al*., 2021)

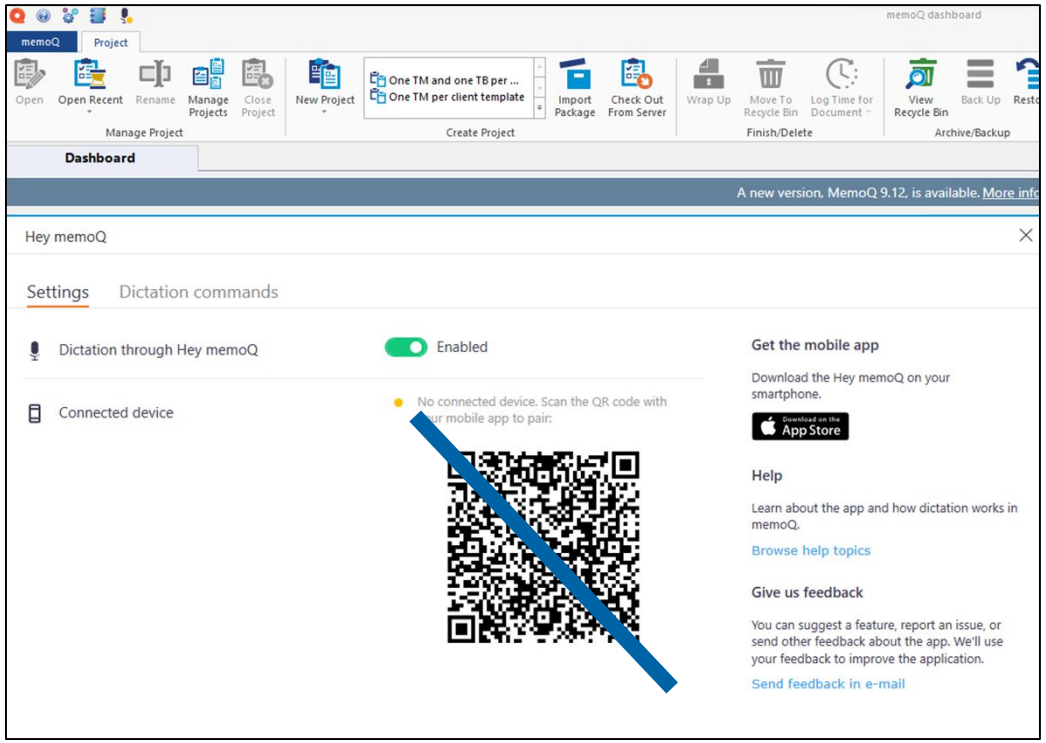
Integrating ASR into medical translation workflows – a demo

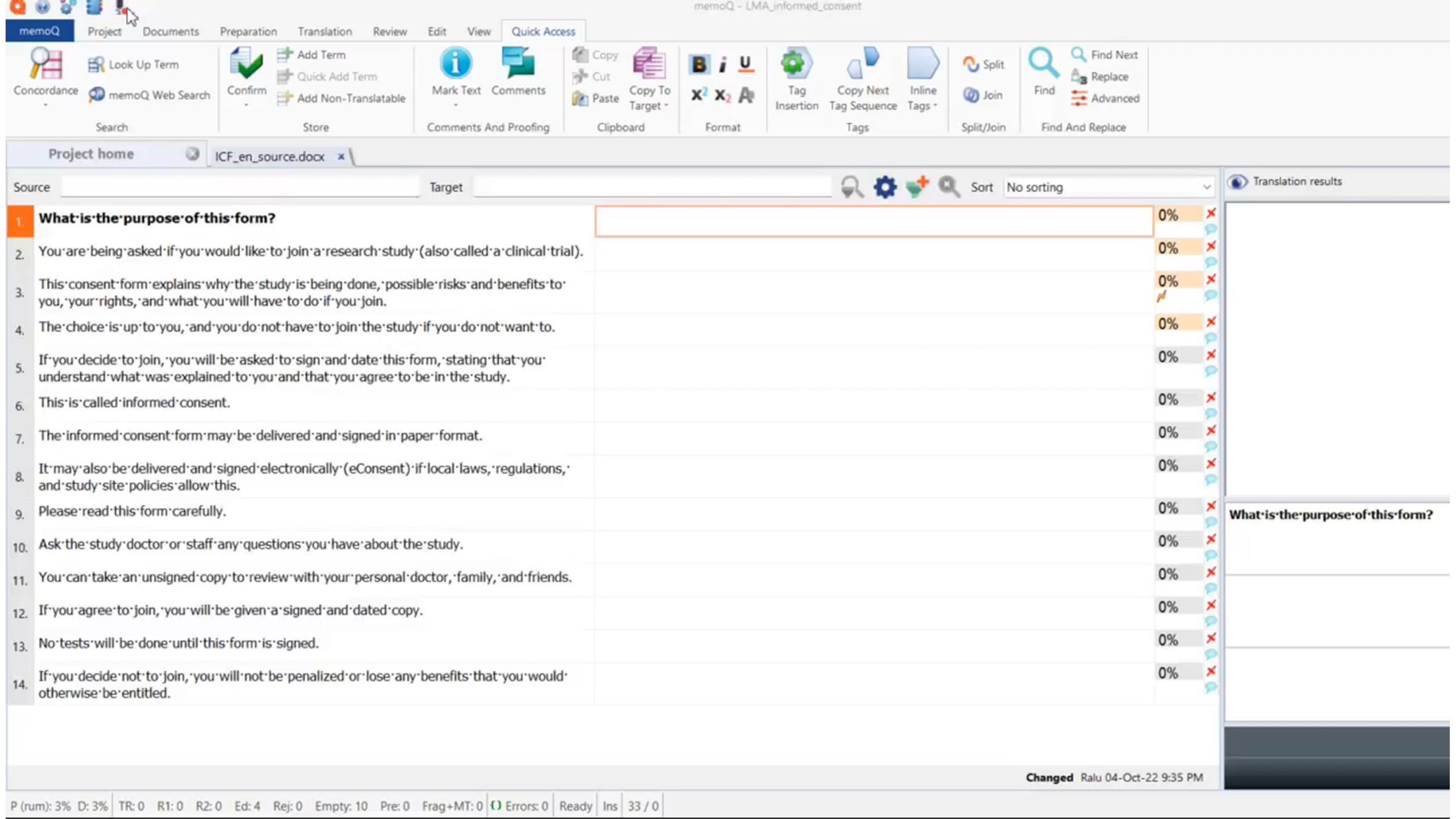
Set-up:

hey memoQ interface on iOS



hey memoQ desktop dashboard





1.	**What·is·the·purpose·of·this·form?**	0%
2.	You·are·being·asked·if·you·would·like·to·join·a·research·study·(also·called·a·clinical·trial).	0%
3.	This·consent·form·explains·why·the·study·is·being·done,·possible·risks·and·benefits·to·you,·your·rights,·and·what·you·will·have·to·do·if·you·join.	0%
4.	The·choice·is·up·to·you,·and·you·do·not·have·to·join·the·study·if·you·do·not·want·to.	0%
5.	If·you·decide·to·join,·you·will·be·asked·to·sign·and·date·this·form,·stating·that·you·understand·what·was·explained·to·you·and·that·you·agree·to·be·in·the·study.	0%
6.	This·is·called·informed·consent.	0%
7.	The·informed·consent·form·may·be·delivered·and·signed·in·paper·format.	0%
8.	It·may·also·be·delivered·and·signed·electronically·(eConsent)·if·local·laws,·regulations,·and·study·site·policies·allow·this.	0%
9.	Please·read·this·form·carefully.	0%
10.	Ask·the·study·doctor·or·staff·any·questions·you·have·about·the·study.	0%
11.	You·can·take·an·unsigned·copy·to·review·with·your·personal·doctor,·family,·and·friends.	0%
12.	If·you·agree·to·join,·you·will·be·given·a·signed·and·dated·copy.	0%
13.	No·tests·will·be·done·until·this·form·is·signed.	0%
14.	If·you·decide·not·to·join,·you·will·not·be·penalized·or·lose·any·benefits·that·you·would·otherwise·be·entitled.	0%

Automatic speech recognition (ASR)/Speech-to-text (S2T)

Advantages

- Increased **productivity**
  - (Dragsted *et al.*, 2011; García Martínez *et al*., 2014; Zapata *et al.*, 2017; Liyanapathirana and Bouillon, 2022)
- Improved **ergonomics**
  - (Ehrensberger-Dow and O'Brien 2015; Ehrensberger-Dow and Hunziker Heeb, 2016)
- Improved **accessibility**
  - (Lewis, 2015; Ciobanu and Secară, 2019; Lucía *et al*., 2020)
- Improved output **naturalness**
  - (Ciobanu, 2016)

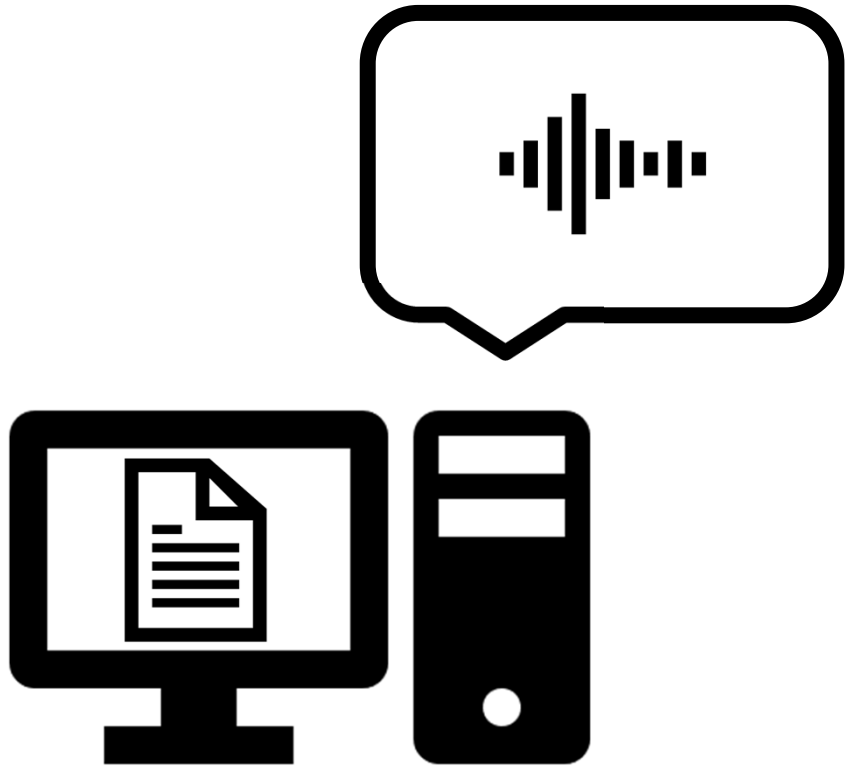
Automatic speech recognition (ASR)/Speech-to-text (S2T)

Challenges

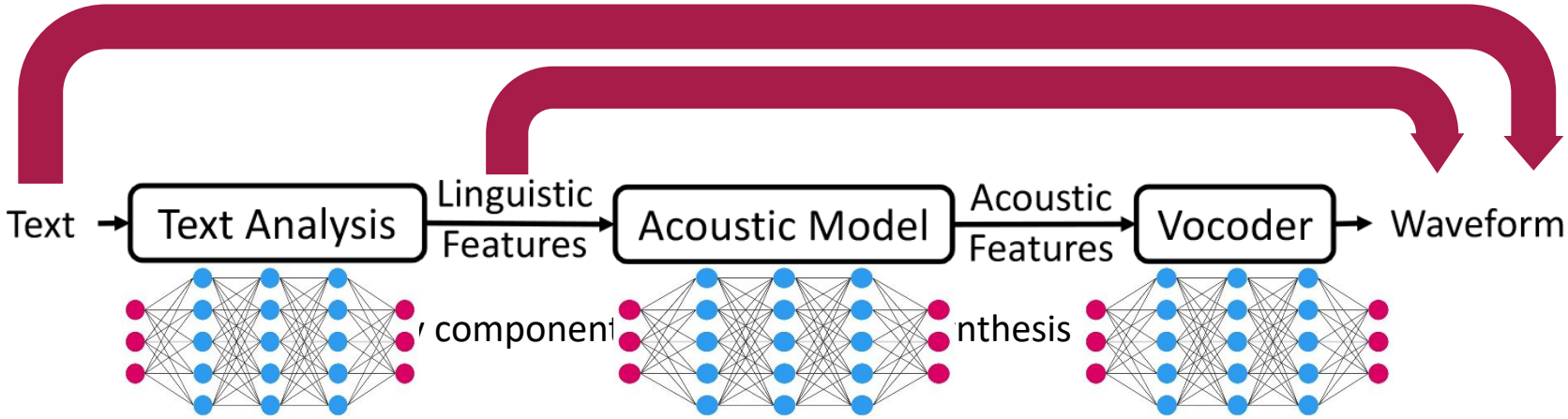
- Risk of **introducing errors**
  - ‘Speakos’ over typos (Ciobanu, 2016), e.g., homophones (Dragsted *et al*., 2011)
- **Accessibility** issues
  - (Tobin *et al*., 2022)
- Potential **increase in cognitive load**
  - (Ciobanu, 2016)
- More **colloquial and informal** translation choices
  - (Ciobanu, 2016)

Speech synthesis/Text-to-speech (T2S)

“getting computers to read out loud” (Taylor, 2009: 1)



Speech synthesis/Text-to-speech (T2S)



Advantages:

- high voice quality in terms of both intelligibility and naturalness
- less requirement on human preprocessing and feature development

(Tan et al., 2021)

universität wien

Applications

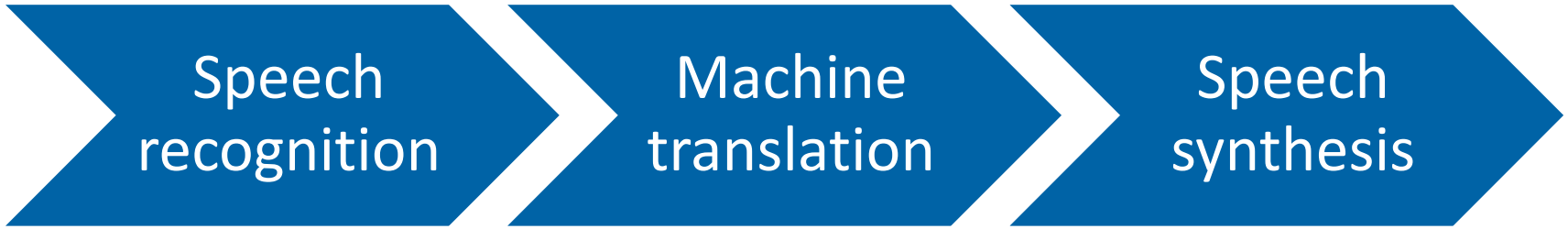
- **Variety of scenarios, including:**
  - automatic call-centre dialogue systems or announcements of travel directions (Taylor, 2009)
  - voice user interfaces in vehicles (Chen et al., 2010)
  - accessibility solutions (Freitas, 2010)
  - personal assistants (e.g. Google, n.d.)

(cf. Olohan, 2019)

universität wien

Applications

- **In a translation context:**
  - Mainly focused on automatic speech-to-speech translation (Ehsani et al., 2010; Hashimoto et al., 2011; Tan, 2014; Seligman et al., 2017)

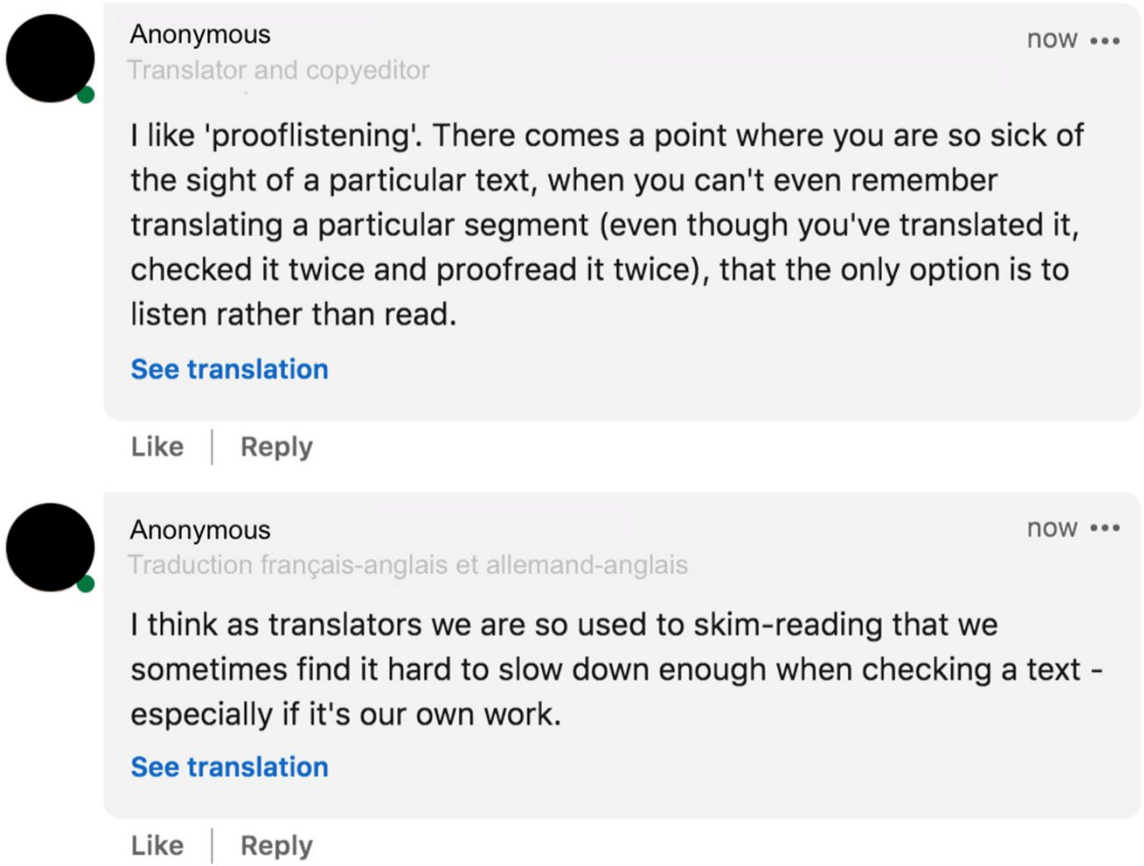


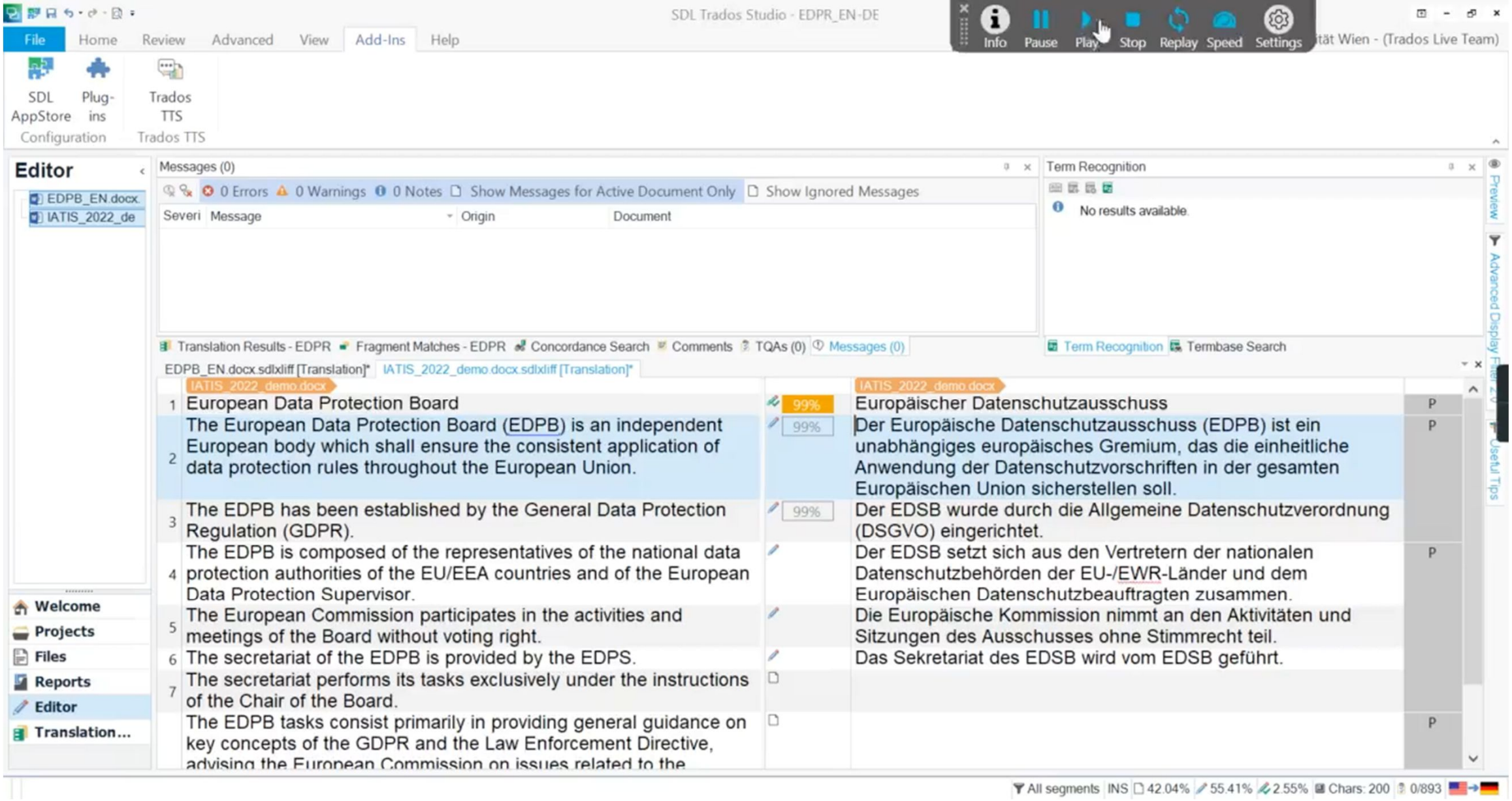
universität wien

Applications

- **In a translation context:**
  - T2S used to support revision process/quality checks of dictated text (Ciobanu, 2016)
  - Practice of reading aloud in (self-)revision (Allain, 2010; Scocchera, 2017)

- **But: No default integration with CAT tools (yet)**





Trados TTS plug-in for SDL Trados Studio

universität wien

Applications

- Initial investigations of applications in human-centred revision and post-editing workflows

1. **Speech synthesis in revision (2019)**
2. **Speech synthesis in post-editing performed by students (2021)**
3. **Speech synthesis in post-editing performed by professional translators (forthcoming)**

The experiment

5 professional translators,
6 trainees (MA level)

2 FR-EN **revision** tasks
in memoQ

in silence,
with source text sound

methods include

(1) error counts,
(2) questionnaires,
(3) eye tracking

impact of sound on

(1) quality,
(2) attitudes,
(3) viewing behaviour

Ciobanu *et al*. (2019)

The use of speech synthesis in revision – Experiment (2019)

Findings

RQ1 (quality)	• Source text (ST) sound conducive to better revision quality overall ➢ Biggest difference in relation to Accuracy errors (66% vs. 37%)
RQ2 (attitudes)	• Majority of participants (7 out of 11) preferred ST sound to silence
RQ3 (viewing behaviour)	• Attention distribution (between ST, target text (TT) and external resources) is similar in both conditions • Intensity of reading ST is decreased in ST sound condition

Ciobanu *et al*. (2019)

The experiment

17 trainees (16 BA, 1 MA)

4 EN-DE **post-editing** / **error annotation** tasks in Microsoft Word

in silence,
with source text sound,
with target text sound,
with ST and TT sound

methods include

(1) error counts,
(2) error annotation analysis,
(3) post-edited words/h,
(4) questionnaires

impact of sound on

(1) quality,
(2) error annotation,
(3) productivity,
(4) attitudes

Brockmann *et al*. (2022), Wiesinger *et al.* (forthcoming)

The use of speech synthesis in PEMT – Experiment (2021)

Findings

RQ1 (quality)	• On average, the combination of ST and TT sound appeared to be most conducive to correcting errors in line with Gold Standard • Major variation between individuals
RQ2 (error annotation)	• High disparity in annotations • More 'preferential annotations', but also fewer actual errors missed using ST and TT sound, as compared to working in silence • Highest number of Accuracy errors annotated in ST sound condition
RQ3 (productivity)	• On average, productivity increase from PEMT in silence to PEMT with ST and TT sound • Major variation between individuals
RQ4 (attitudes)	• Perceived benefits of using sound in less strict set-ups

Brockmann *et al*. (2022), Wiesinger *et al.* (forthcoming)

The use of speech synthesis in PEMT – Project (forthcoming)

- Imminent Research Grants scheme 2021
- The planned experiment:

30 professional translators

4 EN-DE **post-editing** tasks in Matecat

in silence,

with ST and TT sound

methods include

(1) eye tracking,
(2) questionnaires,
(3) editing log,
(4) error counts

impact of sound on

(1) cognitive load,
(2) productivity,
(4) quality,
(5) attitudes

Practical applications – Crisis Translation (Doctoral project Claudia Wiesinger)

"**Timely** and **accurate** communication is essential for crisis management […]" (Crisis Translation, 2020; own emphasis)

- Suspected widespread use of MT in disasters and crises (Anastasopoulos et al., 2020)
- Recommendations for MT-assisted crisis translation workflows (Parra Escartín and Moniz, 2020)
- Difficulties striking a balance between quality assurance processes and timely translations

Speech-enabled PEMT – a viable solution for crisis translation?

Practical applications – Institutional Translation (Doctoral project Justus Brockmann)

In multilingual settings, **institutions govern by translation** (Koskinen, 2014)

- High quality requirements
- Increasing productivity demands
- Use of MT by institutions increasing

Can speech synthesis support the PEMT process in a translating institution?

Practical applications – Medical Translation
(Doctoral project Raluca Chereji)

Paradigmatic shift in healthcare to **patient-centrism**

- "medical knowledge and texts must be translated intralingually […] from expert language to plain language" (Brøgger and Zethsen, 2021)
- In translation, "there is a tendency to revert to expert medical language" (Montalt, Zethsen and Karwacka, 2018)
- ASR-produced translations as less literal, more creative and natural-sounding (Ciobanu, 2016)

Q: Can ASR benefit patient-facing translations and mitigate expert-to-lay bias?

Thank you and Q&A

HAITrans - Human and Artificial Intelligence in Translation

HAITrans - Human and Artificial Intelligence in Translation - is a research group based in the University of Vienna Centre for Translation Studies. It investigates the behavioural and cognitive effects which technologies such as machine translation and automatic speech recognition and synthesis have on translators, as well as their impact on the profession, practice, training and society at large.

At present, the core research areas of the Vienna HAITrans Group are:

1. Effects of speech technologies (Speech to Text and Text to Speech) on translation, revision and post-editing machine translation (PEMT) tasks (R1)
2. Technology-supported translation, revision, and PEMT practices (R2)
3. (Translation) technology for accessibility (R3)
4. Translation technology didactics (R4)

In our qualitative and quantitative investigations we use data gathered via eye-tracking, questionnaires, focus groups, corpora, and translation environment tool metrics. We also collaborate with academic partners, international organisations, language service providers, dedicated professional associations and cultural-sector partners.

Contact

Centre for Translation Studies (ZTW)
Kolingasse 14-16
1090 Vienna
T: +43-1-4277-58080
haitrans@univie.ac.at

Raluca Chereji
raluca-maria.chereji@univie.ac.at

Justus Brockmann
justus.brockmann@univie.ac.at

Claudia Wiesinger
claudia.wiesinger@univie.ac.at

Alina Secară
alina.secara@univie.ac.at

Dragoș Ciobanu
dragos.ioan.ciobanu@univie.ac.at

HAITrans research group
https://haitrans.univie.ac.at/

References

Allain, J.-F. (2010). Repenser la révision: Défense et illustration de la relecture croisée. *Traduire*, *223*, 114–120.

Anastasopoulos, A. et al. (2020) 'TICO-19: The Translation Initiative for COvid-19', in *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. EMNLP 2020*, Online: Association for Computational Linguistics.

Brockmann J., Wiesinger C. & Ciobanu D. (2022). Error Annotation in Post-Editing Machine Translation: Investigating the Impact of Text-to-Speech Technology. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*. Ghent, Belgium: European Association for Machine Translation. pp. 249–257

Chen, F., Jonsson, I.-M., Villing, J., & Larsson, S. (2010). Application of Speech Technology in Vehicles. In F. Chen & K. Jokinen (Eds.), *Speech Technology: Theory and Applications* (pp. 195–219). Springer US.

Ciobanu, D. (2014). Of dragons and speech recognition wizards and apprentices. *Tradumàtica*, 12, pp. 524–538.

Ciobanu, D. (2016). Automatic speech recognition in the professional translation process. *Translation Spaces*, *5*(1), 124–144.

Ciobanu, D. and Secară, A. (2019). Speech recognition and synthesis technologies in the translation workflow. in *The Routledge Handbook of Translation and Technology*. Routledge.

Ciobanu, D., Ragni, V., & Secară, A. (2019). Speech Synthesis in the Translation Revision Process: Evidence from Error Analysis, Questionnaire, and Eye-Tracking. *Informatics*, *6(4)*(51), Article 4.

Crisis Translation (2020) *INTERACT: International Network on Crisis Translation*. Available at: https://sites.google.com/view/crisistranslation/ (Accessed: 24 October 2022).

Defrancq, B. and Fantinuoli, C. (2021). Automatic speech recognition in the booth : assessment of system performance, interpreters' performances and interactions in the context of numbers. *Target - International Journal of Translation Studies*, 33 (1), pp. 73–102. doi: 10.1075/target.19166.def.

Di Gangi, M. A., Cattoni, R., Bentivogli, L., Negri, M. and Turchi, M. (2019). MuST-C: a Multilingual Speech Translation Corpus. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. *NAACL-HLT 2019*, Minneapolis, Minnesota: Association for Computational Linguistics, pp. 2012–2017. doi: 10.18653/v1/N19-1202.

Dragsted, B., Mees, I. M. and Hansen, I. G. (2011). Speaking your translation: students' first encounter with speech recognition technology. *The International Journal for Translation & Interpreting Research*, 3 (1), pp. 10–43.

Ehrensberger-Dow, M. and O'Brien, S. (2015). Ergonomics of the Translation Workplace: Potential for Cognitive Friction. *Translation Spaces*, 4 (1), pp. 98–118. doi: 10.1075/ts.4.1.05ehr.

Ehrensberger-Dow, M. and Hunziker Heeb, A. (2016). 'Investigating the ergonomics of the technologized translation workplace.' in Muñoz Martín, R. (ed.) *Reembedding Translation Process Research*. Amsterdam: John Benjamins, pp. 69–88.

Ehsani, F., Frederking, R., Rayner, M., & Bouillon, P. (2010). Spoken Language Translation. In F. Chen & K. Jokinen (Eds.), *Speech Technology: Theory and Applications* (pp. 167–193). Springer US.

Freitas, D. (2010). Accessibility and Design for All Solutions Through Speech Technology. In F. Chen & K. Jokinen (Eds.), *Speech Technology: Theory and Applications* (pp. 271–299). Springer US.

Gaber, M. and Corpas-Pastor, G. (2021). 'Automatic speech recognition systems for interpreters: Spoken corpora exploitation by interpreter trainers and trainees'. Available at: https://riuma.uma.es/xmlui/handle/10630/23270 (Accessed: 1 November 2022).

García-Martínez, M., Singla, K., Tammewar, A., Mesa-Lao, B., Thakur, A., M.A., A., Bangalore, S. and Carl, M. (2014). SEECAT: ASR & Eye-tracking enabled computer-assisted translation. In *Proceedings of the 17th Annual conference of the European Association for Machine Translation*. *EAMT 2014*, Dubrovnik, Croatia: European Association for Machine Translation, pp. 81–88. Available at: https://aclanthology.org/2014.eamt-1.18 (Accessed: 2 November 2022).

Google (n.d.). *Hey Google*. Available online: https://assistant.google.com/

Hashimoto, K., Yamagishi, J., Byrne, W., King, S., & Tokuda, K. (2011). An analysis of machine translation and speech synthesis in speech-to-speech translation system. *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5108–5111.

Kanabur, V., Harakannanavar, S. S. and Torse, D. (2019). 'An Extensive Review of Feature Extraction Techniques, Challenges and Trends in Automatic Speech Recognition'. *International Journal of Image, Graphics and Signal Processing*, 11 (5), pp. 1–12. doi: 10.5815/ijigsp.2019.05.01.

Karpagavalli, S. and Chandra, E. (2016). 'A Review on Automatic Speech Recognition Architecture and Approaches'. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 9, pp. 393–404. doi: 10.14257/ijsip.2016.9.4.34.

Koskinen, K. (2014). Institutional translation: The art of government by translation. *Perspectives: Studies in Translatology*, *22*(4), 479–492.

References

Ibrahim, H. and Varol, A. (2020). A Study on Automatic Speech Recognition Systems. in *2020 8th International Symposium on Digital Forensics and Security (ISDFS)*. *2020 8th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1–5. doi: 10.1109/ISDFS49300.2020.9116286.

Iranzo-Sánchez, J., Silvestre-Cerdà, J. A., Jorge, J., Roselló, N., Giménez, A., Sanchis, A., Civera, J. and Juan, A. (2020). Europarl-ST: A Multilingual Corpus for Speech Translation of Parliamentary Debates. in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8229–8233. doi: 10.1109/ICASSP40776.2020.9054626.

Kanabur, V., Harakannanavar, S. S. and Torse, D. (2019). An Extensive Review of Feature Extraction Techniques, Challenges and Trends in Automatic Speech Recognition. *International Journal of Image, Graphics and Signal Processing*, 11 (5), pp. 1–12. doi: 10.5815/ijigsp.2019.05.01.

Lewis, W. (2015). Skype Translator: Breaking down language and hearing barriers. A behind the scenes look at near real-time speech translation. in *Proceedings of Translating and the Computer 37. TC 2015*, London, UK: AsLing. Available at: https://aclanthology.org/2015.tc-1.9 (Accessed: 3 October 2022).

Liyanapathirana, J. (2021) 'Integrating post-editing with Dragon speech recognizer: a use case at international organizations'. *Asling 43*, online, 17 November. Available at: https://www.asling.org/tc43/videos/Liyanapathirana.mp4.

Liyanapathirana, J. and Bouillon, P. (2022). *Integrating post-editing with Dragon speech recognizer: a use case in an international organization*. *Translating and the Computer 43*. Tradulex, p. 55. Available at: https://archive-ouverte.unige.ch/unige:163351 (Accessed: 3 October 2022).

Lucía, O. G., Gerlach, J., Schwab, D., Bouillon, P. and Lecouteux, B. (2020). *Building and enhancement of an ASR system for emergency medical settings: towards a better accessibility for allophone and disabled patients*. report. LIG (Laboratoire informatique de Grenoble). Available at: https://hal.univ-grenoble-alpes.fr/hal-02896653 (Accessed: 3 October 2022).

Malik, M., Malik, M. K., Mehmood, K. and Makhdoom, I. (2021). Automatic speech recognition: a survey. *Multimedia Tools and Applications*, 80 (6), pp. 9411–9457. doi: 10.1007/s11042-020-10073-7.

Mesa-Lao, B. (2014). Speech-Enabled Computer-Aided Translation: A Satisfaction Survey with Post-Editor Trainees, in *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*. Gothenburg, Sweden: Association for Computational Linguistics, pp. 99–103.

Romero-Fresco, P. (2020). *Subtitling Through Speech Recognition: Respeaking*. Routledge.

Parra Escartín, C.P. and Moniz, H. (2020) 'Ethical considerations on the use of machine translation and crowdsourcing in cascading crises', in F.M. Federici and S. O'Brien (eds) *Translation in cascading crises*. London: Routledge, pp. 132–151.

Salesky, E., Wiesner, M., Bremerman, J., Cattoni, R., Negri, M., Turchi, M., Oard, D. W. and Post, M. (2021). 'The Multilingual TEDx Corpus for Speech Recognition and Translation'. arXiv. doi: 10.48550/arXiv.2102.01757.

Scocchera, G. (2017). Translation Revision as Rereading: Different Aspects of the Translator's and Reviser's Approach to the Revision Process. *Mémoires Du Livre / Studies in Book Culture*, *9*(1).

Tan, L. (2014). Speech translation. In S.-W. Chan (Ed.), *Routledge Encyclopedia of Translation Technology* (1st ed., pp. 619–631). Routledge.

Tan, X., Qin, T., Soong, F. K., & Liu, T.-Y. (2021). *A Survey on Neural Speech Synthesis*. ArXiv, abs/2106.15561. https://arxiv.org/abs/2106.15561

Tobin, J., Li, Q., Venugopalan, S., Seaver, K., Cave, R. and Tomanek, K. (2022). 'Assessing ASR Model Quality on Disordered Speech using BERTScore'. arXiv. doi: 10.48550/arXiv.2209.10591.

Taylor, P. (2009). *Text-to-Speech Synthesis*. Cambridge University Press.

Tzoukermann, E. and Miller, C. (2018). 'Evaluating Automatic Speech Recognition in Translation'. in *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 2: User Papers)*. Boston, MA: Association for Machine Translation in the Americas, pp. 294–302. Available at: http://www.aclweb.org/anthology/W/W18/W18-1922 (Accessed: 9 August 2018).

Vajpai, J. and Bora, A. (2016). Industrial Applications of Automatic Speech Recognition Systems. *Int. Journal of Engineering Research and Applications*, 6(3), pp.88-95

Wiesinger C., Brockmann J., Secara A. & Ciobanu D. (forthcoming). Speech-enabled machine translation post-editing (PEMT) in the context of translator training. In Kornacki M. & Massey G. (eds.), *Contextuality in Translation and Interpreting: Selected Papers from the Łódź-ZHAW Duo Colloquium on Translation and Meaning 2020–2021*. Volume 70. Peter Lang. (Łódź studies in language, Volume 70).

Zapata, J., Castilho, S. and Moorkens, J. (2017) 'Translation Dictation vs. Post-editing with Cloud-based Voice Recognition: A Pilot Experiment', in *ResearchGate*. *MT Summit XVI*, Nagoya, Japan. Available at: https://www.researchgate.net/publication/319853140_Translation_Dictation_vs_Post-editing_with_Cloud-based_Voice_Recognition_A_Pilot_Experiment (Accessed: 26 February 2018).