



Fattori umani nella traduzione automatica: prospettive professionali e didattiche

Randy Scansani, *Dipartimento di Interpretazione e Traduzione, Università di Bologna (Campus di Forlì)*



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

Wu et al., 2018

Achieving Human Parity on Automatic Chinese to English News Translation

Hassan et al., 2018

La *rivoluzione digitale* potrebbe infatti determinare, nell'arco di pochi anni, la diffusione di processi di doppiaggio *automatizzati*, grazie all'evoluzione dell'*intelligenza artificiale* (basti pensare al salto di qualità, negli ultimi anni, di software come *Google Translator*), che determineranno espulsioni progressive dei professionisti umani dal mercato del lavoro.

Angelo Z. Teodosi, Presidente Istituto italiano per l'Industria Culturale – IsICult (<https://bit.ly/30FHxxq>)

Potenzialità della NMT

Rispetto all'architettura phrase-based:

- Meno errori morfologici
- Meno errori sintattici
- Output più scorrevole
- Risultati migliori in lingue storicamente ostiche per la MT (ad es. tedesco)

*Castilho et al. (2018), Toral & Sánchez-Cartagena (2017),
Bentivogli et al. (2016)*

Errori comuni della NMT

Scelte lessicali sbagliate	Before welding on the unit, ...	Vor dem Einfüllen der Einheit ...
Negazione assente	Verwenden Sie dieses Gerät nicht.	Use this device.
Numeri sbagliati	0-2000	0-9
	10-1 mS/cm	10-ENBT/cm
	150	1500
Parole mancanti	Reason for conveyor congestion?	Motivo del trasportatore?
Terminologia errata o incoerente	Security	Sûreté du système de surveillance.
		Sécurité des informations
Ripetizioni	Der Transport von Milch und Molke	The transport of milk and milk
Nomi propri o elementi do- not-translate tradotti	Deutsche Bank	German Bank
Maiuscole/minuscole errate	Follow the Safety Precautions.	Respecter les Mesures de sécurité.

Errori comuni della NMT

**Fonte: Post-editing
Guidelines TAUS**

Scelte lessicali sbagliate	Before welding on the unit, ...	Vor dem Einfüllen der Einheit ...
Negazione assente	Verwenden Sie dieses Gerät nicht.	Use this device.
Numeri sbagliati	0 1 1	
Parole mancanti	R congestion?	ore?
Terminologia errata o incoerente	Security	Sûreté du système de surveillance. Sécurité des informations
Ripetizioni	Der Transport von Milch und Molke	The transport of milk and milk
Nomi propri o elementi do- not-translate tradotti	De	
Maiuscole/minuscole errate	Follow the Safety Precautions.	Respecter les Mesures de sécurité.

**Light + Full post-
editing**

Full post-editing

Scrambled in translation? Norway Olympics team orders 15,000 eggs by mistake

Chefs with the Norwegian Winter Olympic team found themselves with 13,500 unwanted eggs after what they describe as an error while ordering for their hungry athletes from their South Korean base.

The chefs, who are catering for their country's 109 competitors at the Pyeongchang games, ordered 15,000 eggs to be delivered to their kitchen instead of 1,500 needed.



<https://www.theguardian.com/world/2018/feb/08/norway-olympics-team-orders-15000-eggs-by-mistake-south-korea>

Source (English): *You recently notified us of the possibility that copyrighted material was being made **available** through our website.*

Machine translation (German): *Sie haben uns vor Kurzem von der Überzeugung in Kenntnis gesetzt, dass urheberrechtlich geschütztes Material auf unserer Website **kostenlos verfügbar** ist. [You recently notified us of a belief that copyrighted material was being made **available at no cost** through our website.]*

Source (Russian): *Наш хостел расположен в деревне Туришкино, которая находится в 60 км от Санкт-Петербурга.*

*[Our hostel is located in the village of **Turishkino**, which is 60 km away from St.Peterersburg]*

Machine translation (English): *Our hostel is located in village **Tururushkaino**, which is 60 km away from St.Peterersburg.*



[...] a French economics business magazine such as Capital (August 2013) was able to baldly state that two out of three translating jobs will disappear by the year 2025. The reason given was as follows: 'De plus en plus fiables, les logiciels de traduction automatique écrits et oraux rendront leurs services inutiles'.

Katan, D. 2016. "Translation at the cross-roads: Time for the transcreational turn?". *Perspectives*, 24:3, 365-381

[...] a French economics business magazine such as Capital (August 2013) was able to baldly state that two out of three translating jobs will disappear by the year 2025. The reason given was as follows: 'De plus en plus fiables, les logiciels de traduction automatique écrits et oraux rendront leurs services inutiles'. If we actually use a 'traduction automatique', such as Google Translate, to translate what is written, we have (as of January 2014) the following ungrammatical translation: 'More reliable the automatic translation software written and oral make their unnecessary services'.

Katan, D. 2016. "Translation at the cross-roads: Time for the transcreational turn?". *Perspectives*, 24:3, 365-381 (my emphasis)



Increasingly reliable, written and oral machine translation software will make their services useless (GT, May 2019)

Increasingly reliable, written and oral machine translation software will make their services unnecessary (DeepL, May 2019)

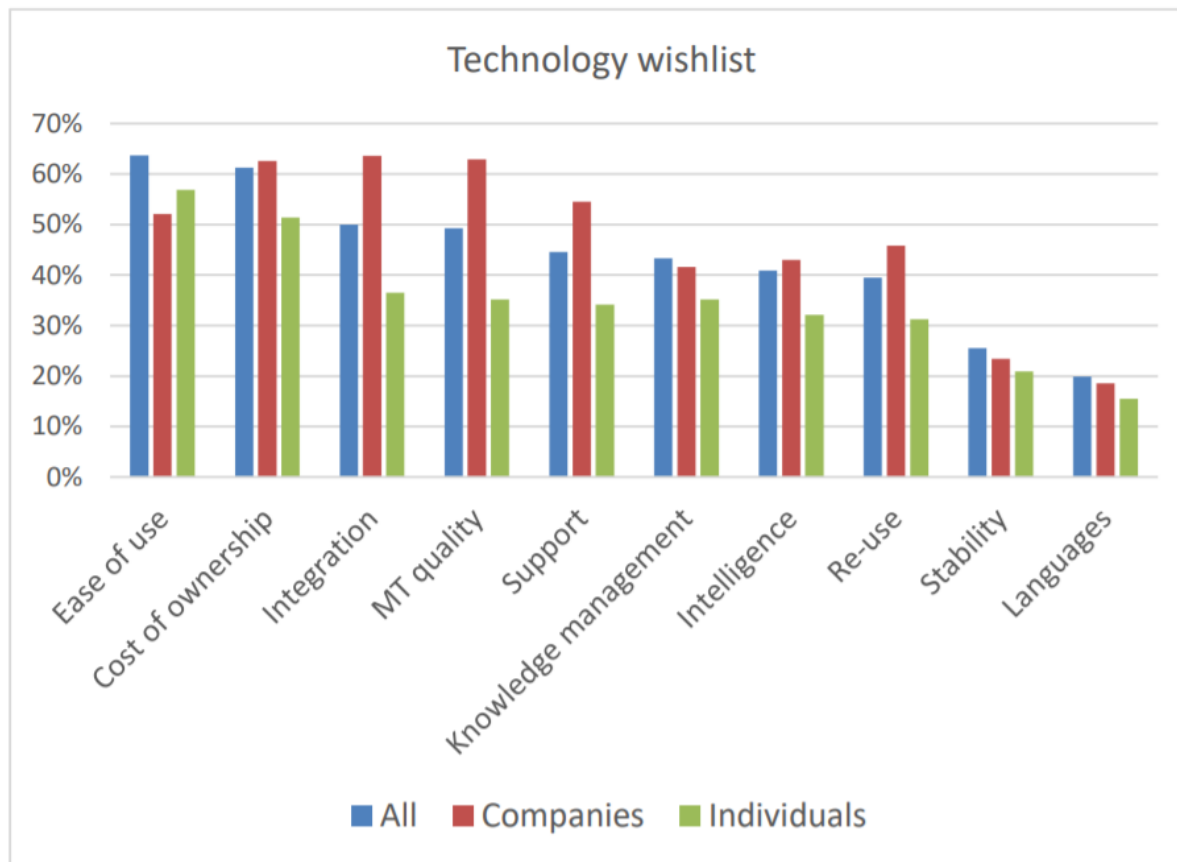


Software di traduzione automatica sempre più affidabile, scritto e orale renderanno i loro servizi inutili (GT, May 2019)

Sempre più affidabile, il traduzione automatica, scritta e orale, farà sì che la loro servizi non necessari (DeepL, May 2019)

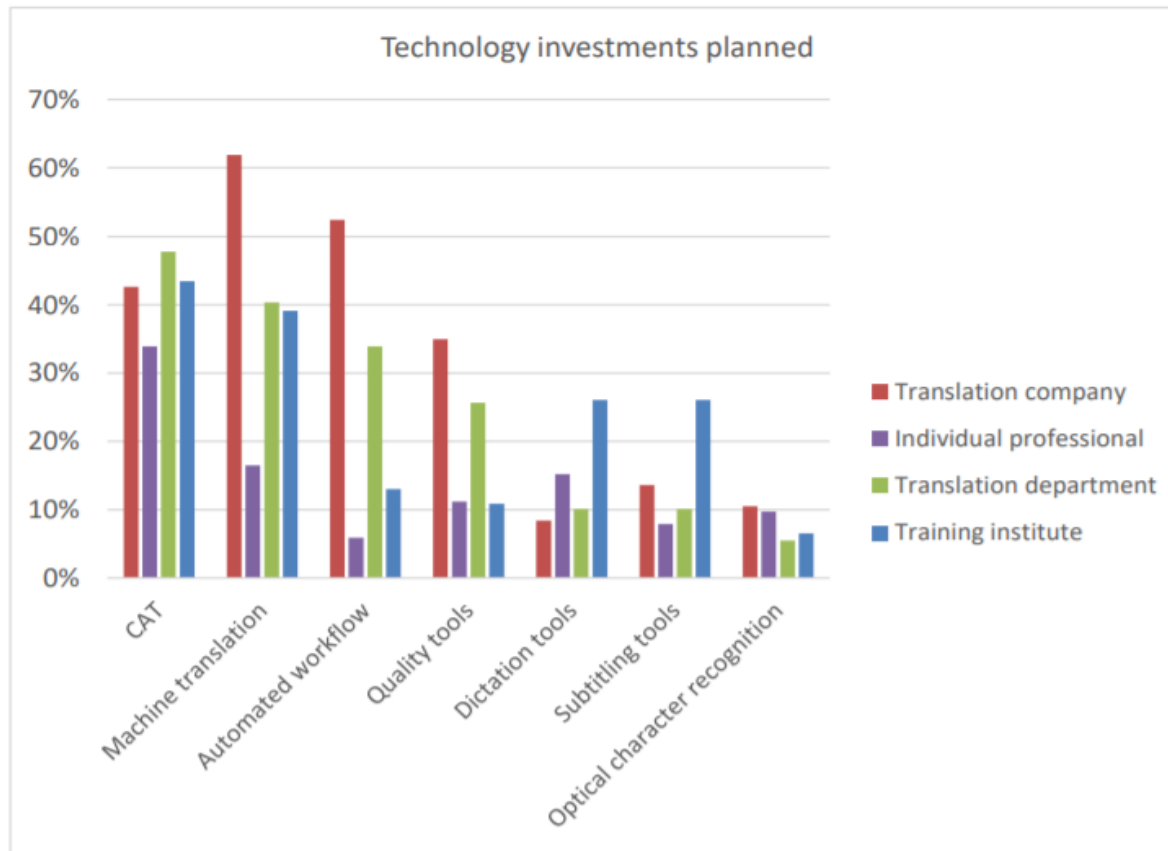
- A che punto siamo?
- Cosa serve nel mondo della traduzione?

MT: oltre i sensazionalismi



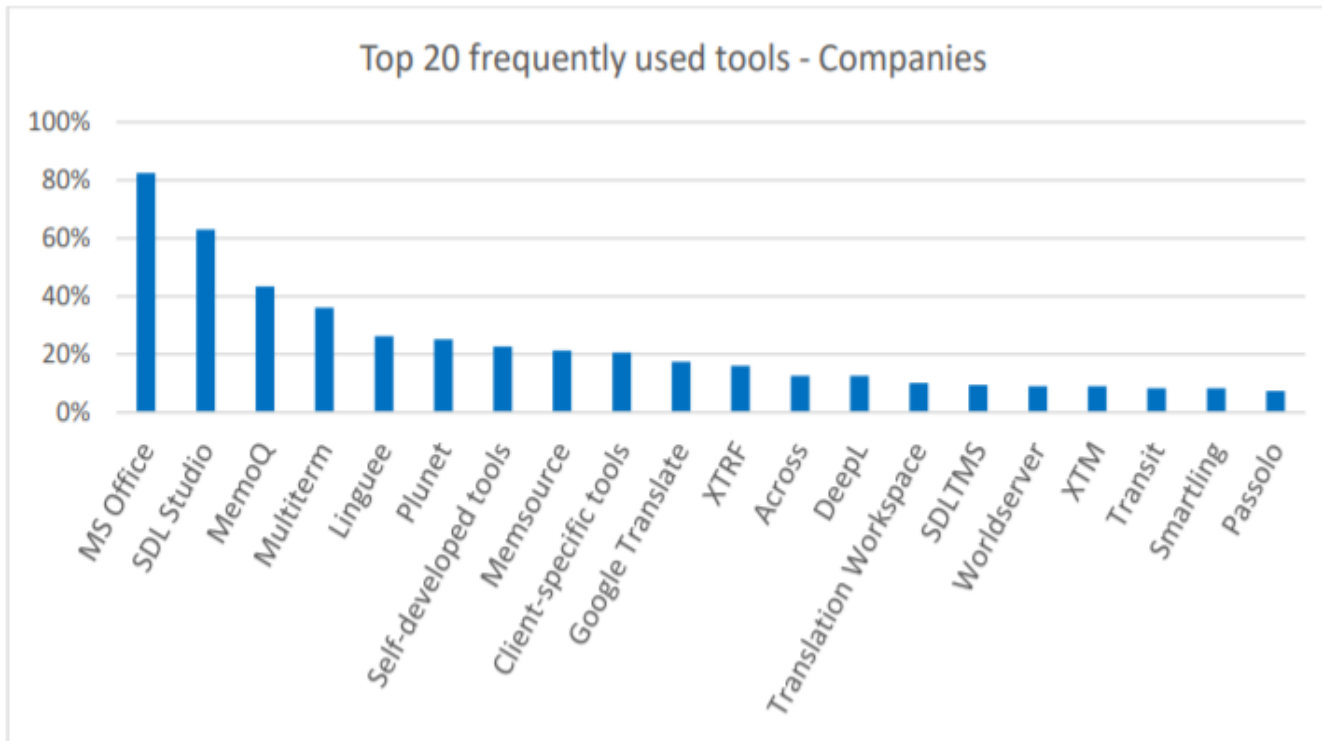
Aziende
richiedono
migliore
qualità
della MT

MT: oltre i sensazionalismi



Aziende
disposte ad
investire
nella MT

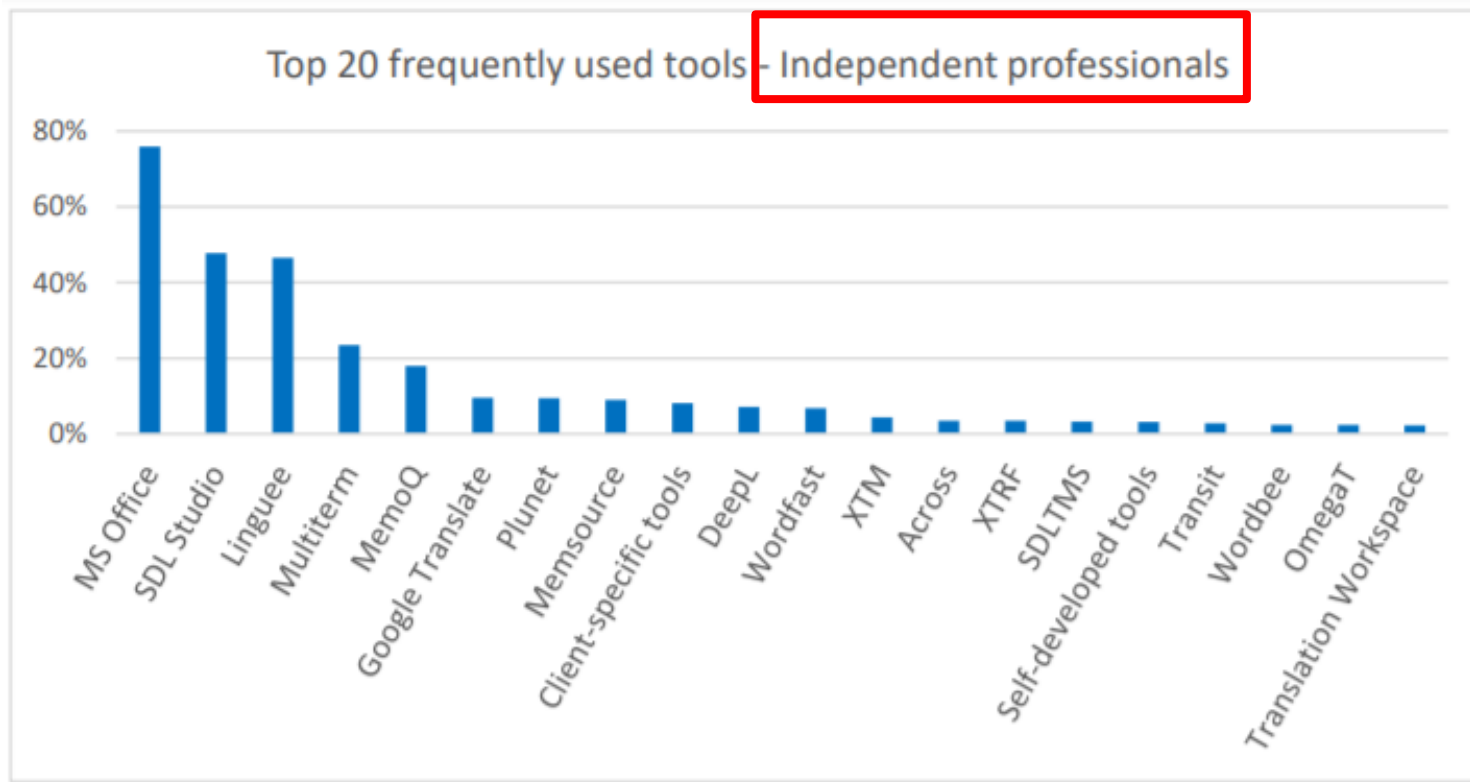
MT: oltre i sensazionalismi



Ma nella realtà:

- MT poco utilizzata
- Uso limitato ai motori generici (DeepL e GT)

MT: oltre i sensazionalismi



Basta migliorare la qualità?

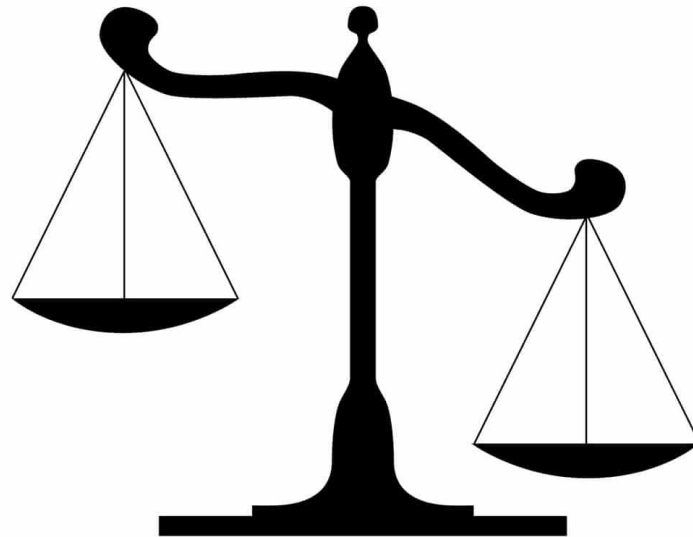
Qualità bassa forse dovuta all'uso di sistemi generici:

- Supporto alle aziende (soprattutto PMI) che vogliono iniziare ad usare la MT
 - **Progetto ANTE**
- Facilità d'uso e trasparenza (uso dei dati)
- Supporto ai professionisti
 - **Life-long learning**
 - **Offerte didattiche**

L'umano serve ancora(?)



L'umano serve ancora(=?)

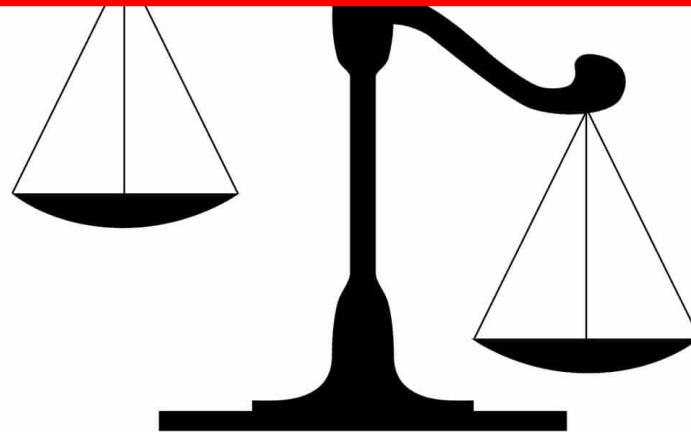


L'umano serve ancora(=?)

Traduzione con
MT

**Serve (anche) una
nuova capacità di
interagire con la
macchina**

Traduzione
umana



1. Il traduttore

Interazione traduttore-macchina: scenari futuri

Competenze del traduttore:

- Architetture MT (fondamenti)
- Errori tipici dei sistemi MT
- Post-editing
- Negoziare equilibrio produttività/qualità
- Traduzione di contenuti non adatti alla MT
- Scrittura in lingua A
- Competenze culturali

Interazione traduttore-macchina: scenari futuri

Competenze del traduttore:

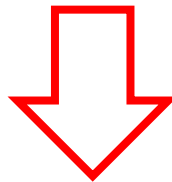
- **Architetture MT (fondamenti)**
- **Errori tipici dei sistemi MT**
- **Post-editing**
- Negoziazione equilibrio produttività/qualità
- Traduzione di contenuti non adatti alla MT
- Scrittura in lingua A

**Superare
preconcetti
vs MT**

Implicazioni professionali

Traduttore:

- Traduzione come auditing
- Traduzione come transcreation



- Dispendio di tempo \neq numero di parole nel testo:
 - Come quantificare il lavoro? Servono modelli per tenere conto della qualità dell'output (QE?)



1. L'esperto di tecnologie per il linguaggio



Interazione traduttore-macchina: scenari futuri

L'esperto di tecnologie per il linguaggio. Competenze:

- Informatica e programmazione
- Tecnologie per la traduzione (CAT, MT, gestione database)
- Problematiche relative alla protezione dei dati
- Gestione dei motori di MT
 - **Open-source vs soluzioni off-the-shelf**
 - **Quali/quante risorse per quali esigenze?**
 - **Interazione con vendor e project management**

Interazione traduttore-macchina: scenari futuri

L'esperto di tecnologie per il linguaggio. Competenze:

- Informatica e programmazione
- Tecnologie per la traduzione (CAT, MT, gestione database)
- Problematiche relative alla protezione dei dati
- Gestione dei motori di MT
- Capacità di valutare l'output
 - **Quali metriche/parametri?**
 - **Quale soglia di qualità per usare MT in produzione?**

Implicazioni professionali

Esperto di tecnologie:

- Dove reperirli?
- Dove vengono formati?
- Equilibrio tra conoscenze informatiche/linguistiche?
 - Dipende, ma conoscenze linguistiche e del workflow imprescindibili



Implicazioni didattiche

Implicazioni didattiche

- Nuovi percorsi formativi, come:
 - Máster Oficial - Tradumática: Tecnologías de la Traducción (UAB – Universitat Autònoma de Barcelona)
 - Translation Science and Technology – Mater (Universität des Saarlandes, Saarbrücken)
 - Italia? Tecnologie in ambito umanistico, ma non per la traduzione (Informatica Umanistica, Università di Pisa)

Implicazioni didattiche

Conoscenze informatiche, buone capacità di scrittura, abilità di post-editing, competenze linguistiche e culturali:

- Possono coesistere in un corso solo?
- Servono davvero?
- Esiste uno studente per questo tipo di corso?

Implicazioni didattiche

Conoscenze informatiche, buone capacità di scrittura, abilità di post-editing, competenze linguistiche e culturali:

- Possono coesistere in un corso solo?
- Servono davvero?
- Esiste uno studente per questo tipo di corso?

Scissione delle due figure anche in ambito didattico



La situazione attuale

La situazione attuale

Master's in Specialised Translation (Università di Bologna, Campus di Forlì):

- *Traduzione e revisione per l'editoria (verso lingua A)*
- *Traduzione tecnica verso lingua A e verso 2 lingue B*
- *Corpus linguistics, MT e post-editing, terminologia, CAT, localizzazione giochi e software, sottotitolaggio*

A= Italiano

B & C= **Inglese**, francese, cinese, tedesco, russo, spagnolo

La situazione attuale

Master's in Specialised Translation (University of Bologna, Campus di Forlì)

- *Traduzione e revisione per la lingua A)*
- *Traduzione tecnica verso lingue B*
- *Corpus linguistics, MT e post-editing, terminologia, CAT, localizzazione giochi e software, sottotitolaggio*

Combinare nozioni di traduzione e di tecnologie rende l'interazione traduttore-macchina efficace?

A= Italiano

B & C= **Inglese**, francese, cinese, tedesco, russo, spagnolo

La situazione attuale

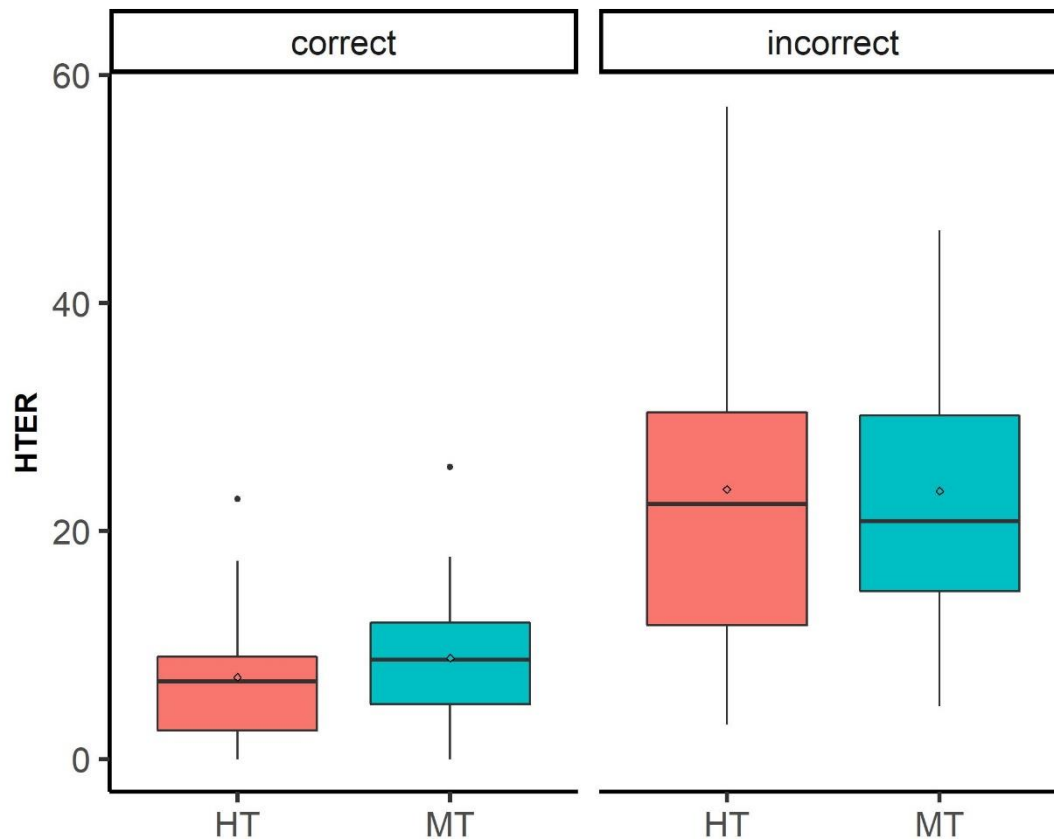
*Quali professionisti crea il modello attuale?
Esperimento di post-editing con gli studenti del
Master. Obiettivo:*

- Capire se il lavoro dei partecipanti è influenzato dal tipo di testi su cui pensano di lavorare.

Esperimento - metodo

- *Stesso testo a tutti i partecipanti (47) IT>EN:*
 - 50% pensa di fare revisione di una traduzione umana
 - 50% pensa di fare post-editing di un output
- Divisione dei segmenti in:
 - **Correct:** Segmento target non ha bisogno di modifiche
 - **Incorrect:** Segmento target ha bisogno di modifiche, ma il messaggio del testo source è rispettato nel testo target
- Misura **HTER** (num modifiche / num di parole) a livello di segmento
- Misura **WPS** (words per second, num di parole / secondi trascorsi) a livello di segmento

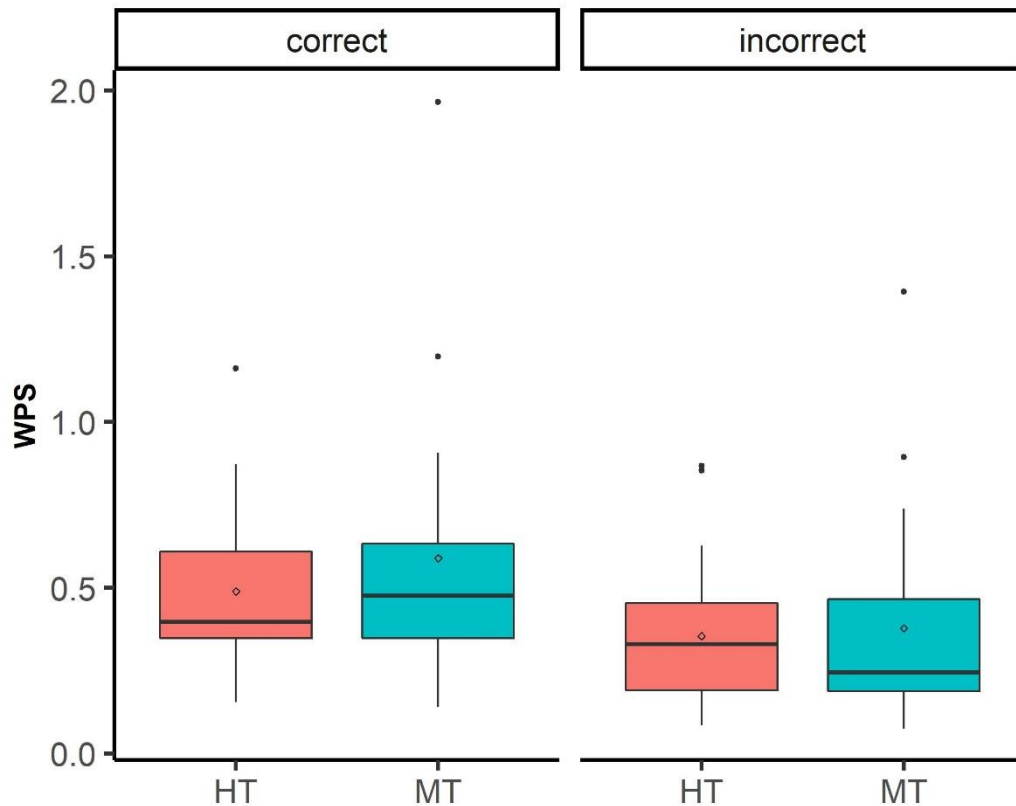
Risultati HTER



Stesso numero di modifiche fatte dai partecipanti quando pensavano di fare post-editing e quando pensavano di fare una revisione

HT = Human Translation
MT = Machine Translation

Risultati Parole/Secondo



HT = Human Translation
MT = Machine Translation

Stessa produttività dei partecipanti quando pensavano di fare post-editing e quando pensavano di fare una revisione.

Risultati - Considerazioni

- I partecipanti non hanno pregiudizi verso la MT
- Possedere conoscenze di MT e post-editing aumenta efficienza dell'interazione uomo-macchina

Riassumendo

- Professionista ancora imprescindibile, ma con nuove abilità
- Dal punto di vista professionale:
 - **Nuove figure professionali da ricercare**
 - **Nuovo modo di lavorare = nuovo modo di quantificare il lavoro**
 - **Più supporto ad aziende e professionisti**
- Dal punto di vista didattico:
 - **Ampliare la formazione sul post-editing**
 - **Potenziare la parte di scrittura creativa**
 - **Formare figure ibride**

Conclusioni

- Riflessioni sulla qualità della MT ad oggi
- Non solo qualità, ma anche fruibilità/accessibilità
- Più cooperazione tra le parti coinvolte (sviluppatori, aziende, professionisti, università)
- Fattore umano ancora fondamentale, ma solo se si cambia



Grazie per l'attenzione!

Randy Scansani – *Dipartimento di Interpretazione e Traduzione, Università di Bologna (Campus di Forlì)*

randy.scansani@unibo.it