

Fraseologia, traduzione e *digital literacy* nel contesto universitario: riflessioni e proposte per un percorso didattico sperimentale*

Silvia Domenica Zollo, Silvia Calvi

Introduzione

La fraseologia svolge un ruolo chiave non solo nello sviluppo delle competenze traduttive, ma anche nel processo di valutazione di queste ultime (QCER 2020). Tre sono le principali sfide da affrontare durante il processo traduttivo delle unità fraseologiche (UF): la corretta identificazione e comprensione della UF nel testo di partenza, la selezione dell'equivalente adeguato e/o della strategia traduttiva da adottare e la riproduzione appropriata delle caratteristiche semantiche, pragmatiche e discorsive della UF nel testo di arrivo. Il processo traduttivo si complica ulteriormente allorché le UF sono tradotte attraverso sistemi di traduzione automatica (TA) basati sull'intelligenza artificiale (IA), in quanto capita spesso che le corrispondenze nella lingua d'arrivo non raggiungano un effetto simile a quello della lingua di partenza.

Nel presente contributo si illustrano i risultati di una sperimentazione didattica realizzata presso l'Università degli Studi di Verona, con le studentesse e gli studenti frequentanti il corso di *Langue Française 2* (LF2), finalizzata all'analisi metalinguistica e metatraduttiva dal francese all'italiano di alcune UF. La ricerca, in linea con le proposte del progetto *Artificial Intelligence for European Integration* (AI4EI)¹ all'interno del quale si inserisce, è stata svolta con l'ausilio di risorse linguistiche informatiche e di strumenti di TA. Dopo aver delineato il protocollo operativo della ricerca,

Silvia Domenica Zollo, Università di Verona, silviadomenica.zollo@univr.it

Silvia Calvi, Università di Verona, silvia.calvi@univr.it

* L'Introduzione, il paragrafo 3 e le Conclusioni sono stati redatti da Silvia Domenica Zollo. I paragrafi 1 e 2 sono stati redatti da Silvia Calvi.

¹ Per maggiori informazioni circa i macro e i microobiettivi del progetto, si rinvia all'introduzione del presente volume. Si ringrazia la prof.ssa Rachele Raus per il coinvolgimento e per il coordinamento scientifico dei lavori.

vengono analizzati il contesto di partenza, il campione di riferimento e i risultati emersi dalla somministrazione di un primo questionario conoscitivo intitolato “Intelligenza artificiale, variazione linguistica e multilinguismo”, mirato a sondare come e in che misura l’IA influisca sulle pratiche linguistiche e metalinguistiche della componente studentesca. Segue, poi, la descrizione del percorso didattico concernente l’analisi e la traduzione dal francese all’italiano di alcune UF, che le studentesse e gli studenti hanno affrontato mediante l’utilizzo di strumenti informatici. Infine, vengono presentate alcune riflessioni conclusive, con riferimento ad eventuali linee di ricerca future.

1. Metodologia e protocollo operativo della ricerca

Nell’ambito del progetto AI4EI è stata condotta, nei mesi di febbraio e maggio 2021, una prima indagine esplorativa inerente alla ricerca dal titolo “Intelligenza artificiale, variazione linguistica e multilinguismo”, che ha avuto come obiettivo principale quello di indagare le abitudini e l’utilizzo di strumenti digitali automatici nella componente studentesca universitaria, attraverso la somministrazione di un questionario *online* e la conseguente raccolta dei dati. Le domande principali che hanno mosso la ricerca sono nate dalla riflessione attorno alle seguenti questioni:

1. Quali sono le caratteristiche che connotano il rapporto tra didattica delle lingue straniere (LS) e l’IA in contesti universitari?
2. Di quale formazione hanno bisogno le discenti e i discenti per sperimentare una nuova didattica basata sull’uso delle tecnologie informatiche?

Si riportano di seguito le ipotesi di partenza:

- a) quanti più strumenti vengono messi a disposizione della comunità studentesca, tanto maggiore sarà la possibilità di garantire un uso corretto delle tecnologie;
- b) quanto più motivata sarà la componente studentesca nell’utilizzare queste risorse, tanto più si avranno risultati positivi in termini di apprendimento di una LS.

Per quanto concerne l’impianto metodologico, è stato costruito come strumento di rilevazione un questionario composto da 44 domande, sottoposto ad un *try out* al quale hanno partecipato circa 15 persone tra do-

centi, ricercatrici e ricercatori di diverse università europee. L'obiettivo del questionario è stato quello di verificare il livello di conoscenza che la componente studentesca possiede in relazione alla molteplicità degli strumenti tecnologici basati sull'IA, esistenti e impiegabili nella vita quotidiana e nella vita universitaria ed eventuali aspettative nei confronti di questi ultimi. Lo strumento di rilevazione utilizzato è stato articolato in tre sezioni:

1. i dati ascrittivi: sezione in cui sono stati raccolti dati anagrafici e una serie di informazioni riguardanti l'università di provenienza e il percorso di studi;
2. l'utilizzo e la percezione degli strumenti basati sull'IA: sezione avente lo scopo di indagare la percezione della comunità studentesca in relazione all'impiego di strumenti quali traduttori automatici, programmi per la gestione lessico-terminologica, piattaforme bilingui e/o multilingui per l'allineamento e la concordanza tra i documenti di partenza e di arrivo, scrittura intelligente e risponditori automatici;
3. rilevamento di opinioni personali sugli effetti positivi e/o negativi dell'IA nell'apprendimento delle LS.

2. Il questionario d'indagine: somministrazione e raccolta dei dati

Lungi dall'essere statisticamente rappresentativo, il quadro delineato all'interno del progetto di ricerca in oggetto intende configurarsi come un'indagine esplorativa su un piccolo campione di partecipanti e vuole essere il primo passo di uno studio pilota da trasformare, col tempo, in un osservatorio permanente. Si presentano, di seguito, il contesto di partenza, il campione di riferimento e i primi risultati rilevati.

2.1 Il contesto di partenza e il campione di riferimento

La comunità studentesca che ha preso parte all'indagine esplorativa poc'anzi menzionata², è composta da studentesse e da studenti frequentanti il corso di *Langue Française 2 (LF2)* A.A. 2020/2021, un insegnamen-

² Si è deciso di somministrare un solo questionario (denominato anche "questionario di inizio corso") in quanto la programmazione delle attività didattiche del corso di LF2 è antecedente alle riflessioni teorico-pratiche nate in occasione del progetto di ricerca in oggetto. Prima della compilazione, è stato brevemente illustrato il progetto AI4EI ed è stato altresì assicurato l'anonimato delle risposte.

to della durata di 36 ore, erogato al secondo anno del corso di laurea triennale professionalizzante in Lingue e letterature per l'editoria e i media digitali (ED-DH) presso l'Università degli Studi di Verona.

Il corso di studi in ED-DH offre un'introduzione alle discipline legate alla tradizione degli studi linguistici e umanistici già presenti presso il Dipartimento di Lingue e Letterature Straniere (i.e. linguistica, filologia, letteratura, editoria e filologia dei testi) e una formazione di base connessa al flusso di informazioni provenienti dall'universo dell'IA (i.e. informatica umanistica, teorie linguistiche, trattamento e analisi dei dati testuali), sì da formare anche future figure professionali con compiti legati alla digitalizzazione del patrimonio linguistico e culturale.

Quanto all'insegnamento di LF2, esso fornisce conoscenze introduttive sul lessico, sulla morfo-semantica della lingua francese e sugli strumenti metodologici quali-quantitativi necessari per le attività di analisi lessicale e lessicografica. Questo insegnamento non è particolarmente orientato agli studi di natura traduttologica: durante il corso vengono affrontate nozioni teoriche legate soprattutto alla linguistica generale di base, alla lessico-semantica, ai principi e alle applicazioni della lessicografia francese cartacea ed elettronica, nonché ai metodi quantitativi di base delle DH per l'introduzione all'analisi semiautomatica di fenomeni lessicali.

Il campione esaminato è costituito da 26 persone. Dal punto di vista del genere, il campione è poco bilanciato, essendo costituito dal 7,7% di uomini e dal 92,3% di donne; l'età media è di 20 anni, con un minimo di 19 e un massimo di 23. La maggior parte della componente studentesca (92,4%) ha come lingua madre l'italiano; solo una minoranza dichiara di essere madrelingua portoghese (3,8%). Le studentesse e gli studenti conoscono inoltre diverse lingue straniere, tra cui l'inglese (100%), il francese (100%), lo spagnolo (64,3%), il tedesco (53,57%), il russo (7,69%), l'arabo (3,57%) e il portoghese (3,57%).

2.2 I risultati del questionario

La prima tematica affrontata nel questionario riguarda la relazione tra la componente studentesca e l'IA: le studentesse e gli studenti mostrano di essere interessati allo sviluppo delle tecnologie e dell'IA (Fig. 1). Dato ancora più rilevante, la maggior parte ritiene che questi strumenti siano

estremamente utili per la formazione universitaria e abbastanza importanti per il proprio futuro professionale (Fig. 2, 3).

Le domande successive entrano nello specifico di diversi aspetti legati all'IA: l'utilizzo di traduttori automatici, di programmi per la gestione della terminologia, di piattaforme bilingui e multilingui, nonché di programmi per l'allineamento e la concordanza di *file source* con la rispettiva traduzione.

La maggior parte delle persone utilizza traduttori automatici gratuiti o a pagamento, quali Reverso, Google Traduttore, Pons e Papago per la traduzione da o verso diverse lingue, quali l'inglese, il francese, lo spagnolo, il tedesco, l'arabo e il greco (Fig. 4). Le studentesse e gli studenti, pur essendo all'inizio del loro percorso di studi, sono consapevoli che i risultati dei traduttori automatici, per quanto buoni, necessitano di una revisione umana: la maggior parte — il 76% — considera, infatti, solo "abbastanza affidabili" i risultati della TA (Fig. 5).

Le interessa lo sviluppo delle tecnologie e dell'intelligenza artificiale?

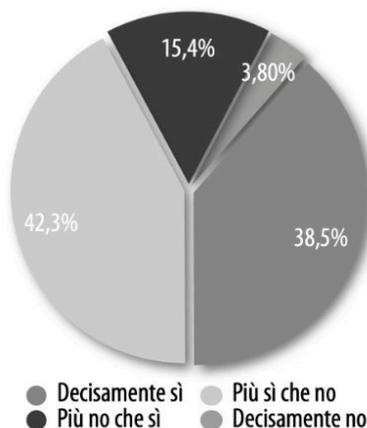
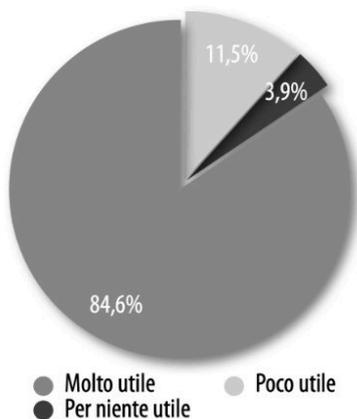


Figura 1: Interesse per lo sviluppo delle tecnologie e dell'IA

Quanto pensa che potrà essere utile nel prossimo futuro l'intelligenza artificiale nel suo campo di studio?



Pensa che l'intelligenza artificiale assumerà un ruolo importante nel suo futuro professionale?

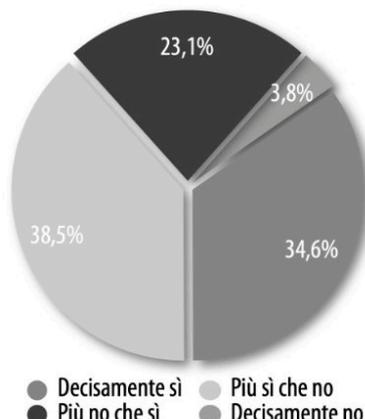


Figure 2, 3: Utilità dell'IA nello studio e nel futuro professionale

Generalmente si serve di traduttori automatici gratuiti o a pagamento (es. Google translate, Reverso, DeepL...)?

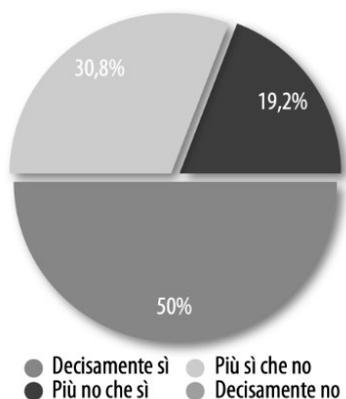


Figura 4: Utilizzo dei traduttori automatici

Pensa che i risultati forniti dai traduttori automatici siano:

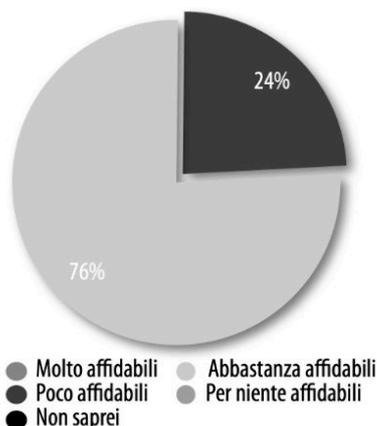


Figura 5: Affidabilità dei traduttori automatici

Quando il risultato che ottiene dal traduttore automatico o dalla piattaforma non le sembra corretto, pensa che l'errore sia dovuto:	
All'algoritmo/strumento utilizzato che è poco performante	12%
Al fatto che la locuzione ricercata sia idiomatica o troppo complessa ³	44%
Ai corpora/testi su cui si basano i traduttori automatici	24%
Non mi sono mai posto/a questa domanda	20%

Tabella 1: Percezione degli errori causati dalla TA

Per quanto riguarda le osservazioni sugli errori causati dai traduttori automatici, le risposte sono molto eterogenee, come illustrato nella tabella 1, in cui si evidenzia che solo il 20% non si è mai posto questa domanda.

Un numero inferiore di persone ha invece riflettuto sulle differenze circa il trattamento delle traduzioni, a seconda della tipologia di lingua: solo il 4% ha risposto di aver decisamente riscontrato queste differenze (Fig. 6). Tra le risposte sono tuttavia emerse delle riflessioni interes-

Ha mai riscontrato delle differenze nel trattamento di traduzione a seconda della tipologia di lingua (lingue romanze, germaniche ecc.)?

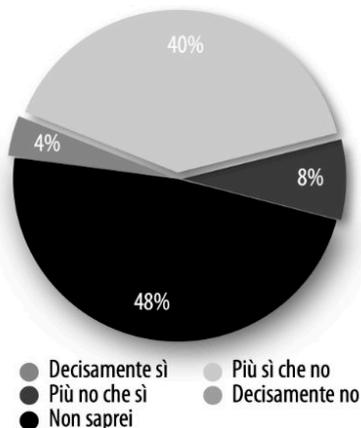


Figura 6: Differenze nel trattamento delle traduzioni a seconda della tipologia di lingua

³ Su questo preciso aspetto si è intervenuto durante la sperimentazione didattica.

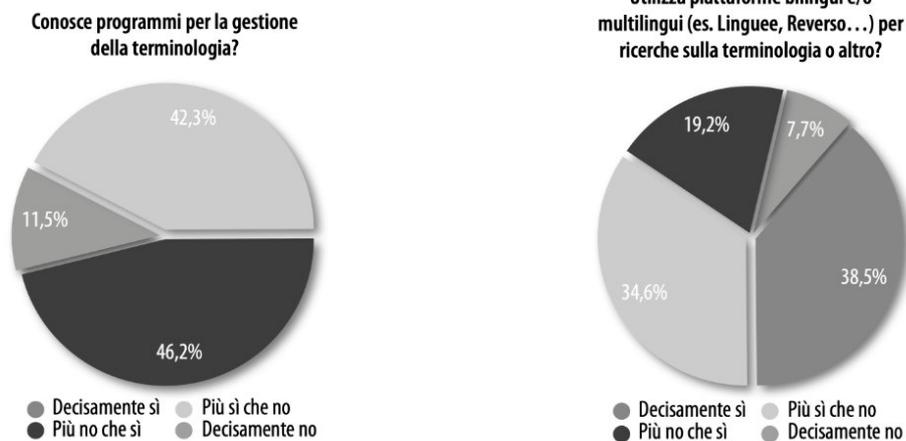


Figura 7, 8: Conoscenza dei programmi per la gestione terminologica e utilizzo di piattaforme bilingui e/o multilingui

santi: le studentesse e gli studenti hanno la percezione che le lingue romanze e l'inglese siano tradotte meglio rispetto a lingue come il tedesco e il russo che spesso risultano essere tradotte in modo troppo letterale, causando sia errori grammaticali sia errori sintattici. La componente studentesca possiede una conoscenza ben più limitata dei programmi per la gestione terminologica, rispetto ai programmi di TA. Tuttavia, le stesse persone dichiarano di utilizzare piattaforme bilingui e/o multilingui, come Linguee e Reverso, per ricerche terminologiche e lessicali (Fig. 7, 8). La sezione successiva del questionario, dedicata alla percezione della componente studen-

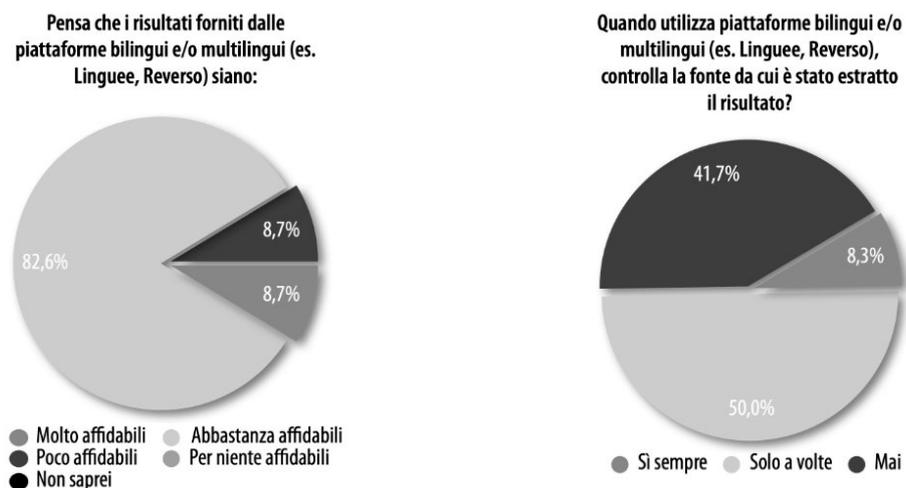


Figura 9, 10: Affidabilità e verifica delle fonti

Quando utilizza piattaforme bilingui e/o multilingui, qual è il suo criterio per scegliere il termine equivalente in lingua straniera?	
Scelgo il termine più frequente	8,3%
Scelgo il primo termine proposto dalla piattaforma	0%
Penso al contesto di utilizzo e poi scelgo il termine	91,7%
Di solito la piattaforma propone un solo termine equivalente	0%

Tabella 2: Parametri per la scelta dei termini equivalenti in LS

Conosce programmi per l'allineamento e la concordanza tra i documenti di partenza e quelli di arrivo (es. WordSmith Tools, Lingua MultiConcord, Trados WinAlign...)?

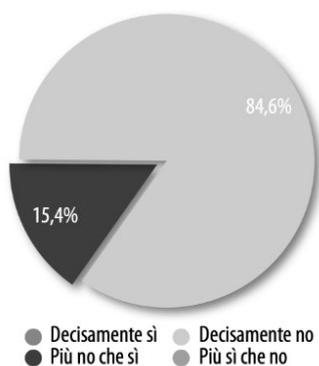


Figura 11: Conoscenza di programmi con l'allineamento e la concordanza

scelta di questi strumenti a supporto delle pratiche linguistiche e multilinguistiche, tra cui l'affidabilità e il controllo delle fonti da cui sono estratti i risultati linguistici e, infine, i parametri per la scelta dei termini equivalenti in lingua straniera. In termini generali, prevale una certa fiducia nei confronti di questi strumenti: circa l'83% ritiene che i risultati forniti da queste piattaforme digitali siano abbastanza affidabili. Per quanto concerne la verifica della fonte da cui vengono estratti i risultati, il 41,7% dichiara di non controllare mai la fonte, il 50% lo fa solo a volte e soltanto l'8,3% delle studentesse e degli studenti la controlla sempre (Fig. 9, 10).

Quanto ai criteri per la scelta dei termini equivalenti in LS, si segnala che la quasi totalità delle persone — il 92% — pone attenzione al contesto d'uso e, seppur in piccola percentuale —

tesca rispetto all'uso delle piattaforme bilingui e/o monolingui po- c'anzi menzionate, si compone di un set di domande volte a rintracciare le motivazioni e i criteri sottesi alla

l'8,3% — utilizza il criterio della frequenza (Tab. 2), un dato che, a nostro avviso, lascia ben sperare relativamente all'introduzione dei metodi statistico-quantitativi an-

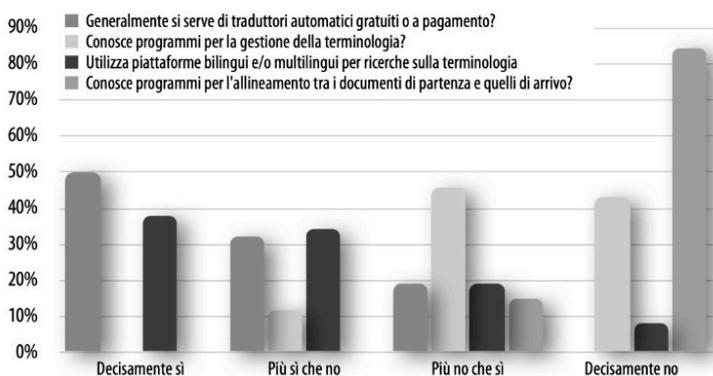


Figura 12: Confronto delle risposte circa l'utilizzo di diversi strumenti basati sull'IA

che nei corsi di linguistica d'area afferenti alle lauree triennali. Dall'indagine emerge un altro dato rilevante. La componente studentesca non conosce programmi per l'allineamento e per la concordanza tra i documenti di partenza e quelli di arrivo: in totale, solo il 15,4% dichiara di conoscerli più o meno (Fig. 11).

Il grafico (Fig. 12) riassume e confronta

le risposte della componente studentesca relativa alla conoscenza e all'utilizzo di tutti i programmi menzionati: le studentesse e gli studenti hanno una maggiore consapevolezza circa l'utilizzo dei traduttori automatici gratuiti e delle piattaforme bilingui e/o multilingui, rispetto a programmi per la gestione della terminologia e per l'allineamento e la concordanza tra i documenti di partenza e quelli di arrivo. Quanto alla percezione e all'uso dei suggerimenti di scrittura intelligente e/o automatica offerti dai *social media* o da programmi come Gmail e Google research attraverso la tecnologia del *machine learning*, si evidenziano risultati abbastanza eterogenei, come illustrato nella tabella 3.

Tale eterogeneità emerge anche nel livello di percezione dei messaggi vocali preregistrati automatizzati e dei risponditori automatici. Come si evince dalla tabella 4, il dato più interessante riguarda gli aspetti prosodici della lingua.

Quando scrive e-mail o messaggi sui social, che pensa dei suggerimenti automatici di scrittura che le vengono proposti?	
Mi sembrano corretti e li utilizzo	3,9%
Mi accerto che siano corretti tramite altra ricerca prima di utilizzarli	11,5%
Li utilizzo solo a volte, quando mi sembrano corretti	61,5%
Non ho l'abitudine di utilizzarli	23,1%
Non saprei	0%

Tabella 3: Utilizzo dei suggerimenti automatici di scrittura

Quando le capita di ascoltare risponditori automatici preregistrati automatizzati, cosa nota? (più risposte possibili)	
L'intonazione o la pronuncia non mi sembra adeguata	46,2%
La lingua non mi sembra adeguata (troppo formale, troppo inusuale)	23,1%
La presenza quasi esclusiva di voci maschili	3,8%
La presenza quasi esclusiva di voci femminili	15,4%
Non noto nulla	30,8%
Quando le capita di ascoltare risponditori automatici preregistrati automatizzati, cosa nota? (più risposte possibili)	
L'intonazione o la pronuncia non mi sembra adeguata	34,6%
La lingua non mi sembra adeguata (troppo formale, troppo inusuale)	23,1%
La presenza quasi esclusiva di voci maschili	0%
La presenza quasi esclusiva di voci femminili	19,2%
Non noto nulla	34,6%

Tabella 4: Messaggi vocali preregistrati automatizzati e i risponditori automatici

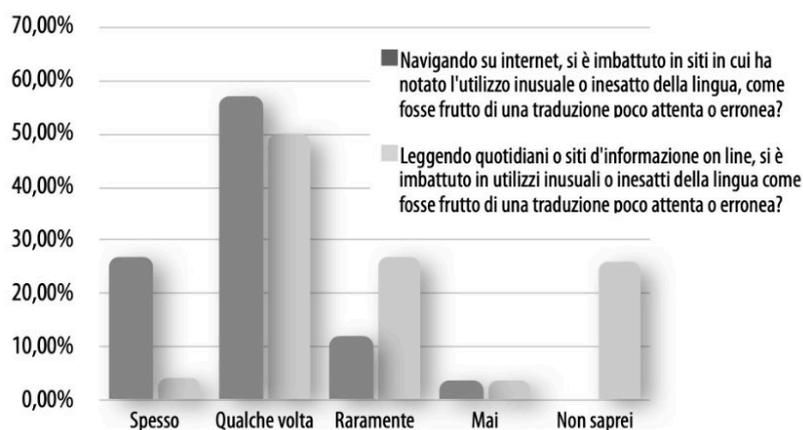


Figura 13: Rilevamento di traduzioni poco attente e/o erranee

Un altro tema, solo parzialmente analizzato dal questionario, è la capacità di rilevare eventuali inesattezze della lingua, quale frutto di traduzioni poco attente, navigando in Internet o leggendo quotidiani o siti d'informazione online (Fig. 13).

È possibile osservare che oltre la metà della componente studentesca si è imbattuta qualche volta in siti in cui ha notato inesattezze dal punto di vista linguistico, mentre quasi il 30% dichiara di aver rilevato spesso errori dovuti a traduzioni poco attente.

Gran parte delle rispondenti e dei rispondenti è, quindi, in grado di riconoscere eventuali errori dovuti a traduzioni inesatte. A tal proposito, sono interessanti le risposte aperte circa gli errori e le inesattezze linguistiche, che sono per lo più attribuiti all'assenza di *post-editing*, ai traduttori automatici e alla superficialità nel processo traduttivo; tra le motivazioni degli errori, vengono richiamate anche la poca attenzione, la scarsa revisione e la mancanza di intervento umano, da cui si deduce una consapevolezza da parte delle studentesse e degli studenti nell'uso di questi mezzi e strumenti.

Infine, nell'ultima parte del questionario, sono state raccolte le opinioni della componente studentesca circa gli aspetti positivi e negativi che potrebbero caratterizzare l'uso dell'IA relativamente allo studio delle LS. Tra gli aspetti positivi, la componente studentesca nomina la rapidità, la comodità e la possibilità di consultare banche dati lessicografiche e banche testuali elettroniche. Tra gli aspetti negativi, riscontriamo, invece, l'i-

naffidabilità, la confusione, la superficialità, le traduzioni non corrette, la delega di decisioni umane affidate agli algoritmi, la circolazione di nozioni false, nonché la perdita di spontaneità nell'espressione linguistica.

3. Il percorso didattico sperimentale

Coerentemente con i risultati del questionario e con l'obiettivo di sviluppare una maggiore consapevolezza nella componente studentesca circa l'utilizzo e il potenziale delle nuove tecnologie basate sull'IA, è stata progettata una sperimentazione didattica volta a coniugare conoscenze metalessicali e metatraduttive di base e nozioni introduttive di *digital* e di *machine translation literacy* (Bowker, Delsey 2016; Bowker 2018; Bowker, Buitrago 2019) applicate allo studio e alla traduzione delle UF.

Lo scopo del suddetto intervento didattico è stato quello di valutare se e in che misura siffatti percorsi ibridi possano rivelarsi efficaci per la formazione linguistica di studentesse e di studenti iscritti al secondo anno del corso di laurea triennale professionalizzante ED-DH. Si è partite dall'assunto che interventi interdisciplinari integranti l'uso di strumenti computazionali orientati alla TA, possano accrescere significativamente le prestazioni di apprendimento del lessico e della fraseologia di una lingua straniera, esercitando un effetto positivo anche sul processo traduttivo delle UF.

Come già anticipato nell'introduzione, le attività formative sono state condotte nel secondo semestre dell'A.A. 2020/2021, durante l'insegnamento di LF2, per un totale di otto ore non consecutive⁴. In linea con gli obiettivi dell'insegnamento in oggetto, si è ritenuto di fondamentale importanza proporre sia lezioni seminariali di carattere teorico-metodologico sia lezioni applicative, sì da far emergere le capacità analitiche e operative delle studentesse e degli studenti, facendo sperimentare loro modalità alternative per l'acquisizione di conoscenze metalessicali e metatraduttive di base.

3.1 Le nozioni metalessicali e metatraduttive di base per lo studio delle UF

La prima attività didattica è consistita in un seminario introduttivo finalizzato all'apprendimento delle nozioni metalessicali di base, quali lessia, lessema, vocabolo, unità lessicale (UL) e unità fraseologica (UF), così

⁴Le prime quattro ore di lezione hanno avuto luogo nell'ambito del corso di LF2, tenutosi a distanza su Zoom a causa dell'emergenza sanitaria Covid-19. La sperimentazione di quattro ore è stata, invece, svolta al di fuori del corso in oggetto, in quanto non pertinente con i contenuti del programma di LF2 dell'A.A. 2020/2021.

come la loro interrelazione con gli studi traduttologici. Prediligendo una prospettiva che mette l'accento principalmente sul significato (Polguère 2016), sono stati introdotti altresì i concetti di idiomacità e di non composizionalità, uniti alla fissità o stabilità sintattica e, naturalmente, al criterio della polilessicalità, con riferimento al quadro teorico della Teoria Senso-Testo (Mel'čuk *et alii* 1995). È stata messa in luce l'eterogeneità delle strutture lessicali che rientrano sotto il termine generico di unità fraseologica (UF), nonché le numerose problematiche con le quali si confronta chi intende studiare le UF; a seconda della prospettiva che si assume, è possibile imbattersi in termini che si differenziano nelle diverse lingue e tradizioni di studi scientifici. Si è, dunque, tentato di offrire una visione quanto più schematica possibile delle UF, basandosi sulla classificazione proposta da Mel'čuk (2013).

Nature des contraintes \ Compositionnalité des phrasèmes	non compositionnels	compositionnels
lexicales	LOCUTIONS	COLLOCATIONS
sémantico-lexicales	impossible	CLICHÉS

Figura 11: Classificazione delle UF (dette anche frasemi) di Mel'čuk (2013)

Ai fini della sperimentazione didattica in questione, ci si è soffermate unicamente sulle UF di tipo locuzione in lingua francese (i.e. *gagner son pain* > “guadagnarsi la pagnotta”; *huile de coude* > “olio di gomito”), sviscerandone sia le caratteristiche di natura sintagmatica e grammaticale (i.e. locuzioni verbali, nominali, aggettivali e avverbiali) sia gli aspetti più puramente semantici legati al grado di opacità e di trasparenza, prendendo in considerazione la sotto-categorizzazione mel'čukiana (i.e. locuzione forte, semi-locuzione e locuzione debole)⁵.

Per quanto concerne gli aspetti teorici di base circa la traduzione delle UF, sono state illustrate le definizioni di unità di traduzione (UT) (Vinay, Darbelnet 1958; Froeliger *et alii* 2018) e di “atomo traduzionale” (“*atome traductionnel*”, Bennett (1994)), due concetti che alimentano da sempre il dibattito scientifico tra le esperte e gli esperti in traduttologia e in fraseologia (Corpas Pastor 2003; Colson 2008; Baker, Saldanha 2011). Seppur breve, questa disamina teorica si è rivelata particolarmente utile dal punto di vista didattico: le studentesse e gli studenti hanno mostrato, da un

⁵ Per approfondimenti teorici sulla classificazione dei frasemi, vedasi Mel'čuk (2013).

lato, di essere concordi con la prospettiva descrittiva di Bennett (anziché con quella prescrittiva di Vinay, Darbelnet (1958)), secondo cui anche le più piccole unità linguistiche devono essere tradotte nell'insieme; dall'altro, gran parte del gruppo si è posta la seguente domanda di ricerca: come articolare il senso idiomatico o non compositivo delle UF di tipo locuzione in un'altra lingua (i.e. *raconter des salades* > “raccontare frottole”)? È stata, quindi, proposta un'ulteriore classificazione delle UF dal punto di vista traduttivo: le UF di tipo locuzione che hanno equivalenti totali e/o parziali e quelle che non hanno equivalenti nella lingua d'arrivo, intensificando la riflessione sull'aspetto culturale (Mejri 1997). A titolo d'esempio, per tradurre in lingua italiana la semi-locuzione *travailler pour le roi de Prusse* (“lavorare per la gloria”) bisogna tener conto non solo della dimensione linguistica, ma anche di quella culturale sia per la lingua di partenza sia per la lingua d'arrivo. Infine, sono state suggerite le seguenti fasi di lavoro per l'analisi traduttiva e metatraduttiva delle UF di tipo locuzione:

- a) identificazione della UF nel testo di partenza;
- b) analisi del significato, dell'intenzione comunicativa e classificazione della UF, secondo i criteri grammaticali e semantici prestabiliti con l'ausilio di Spiderlex (RL-fr) (Gader *et alii* 2014; Ollinger, Polguère 2020);
- c) ricerca dell'equivalente in lingua italiana con e/o senza l'ausilio della TA, tenendo conto degli aspetti lessicali, discorsivi e culturali delle due lingue di lavoro (francese e italiano);
- d) confronto e analisi delle corrispondenze tra la UF di partenza e quella di arrivo;
- e) qualora le corrispondenze non dovessero soddisfare questi criteri, ricerca, attraverso le risorse linguistiche basate sull'IA, di una UF equivalente nella lingua d'arrivo, che raggiunga un effetto perlocutorio analogo a quello della UF della lingua di partenza.

3.2 Le principali risorse informatiche per la ricerca e per la traduzione delle UF

Durante la seconda attività didattica è stata presentata una panoramica dei principali strumenti informatici accessibili gratuitamente in rete e funzionali all'attività traduttiva delle UF. Le risorse lessicografiche introdotte includono sia dizionari monolingui in lingua francese e in lingua italiana, come i dizionari *Larousse*, *Le Robert Dico en ligne*, il *Trésor de la Langue Française informatisé (TLFi)*, il *Wiktionnaire*, il *Treccani* e il *Garzanti*, sia dizio-

nari bilingui, quali il *Dictionnaire Français-Italien en ligne - Larousse*, il *WordReference Français-Italien Virtual Dictionary*, il dizionario *ReversoContext* con il *plug-in Expressio* e, infine, il *Collins Dictionary Traduttore*. Oltre allo studio dei significati e delle corrispondenze fraseologiche in altre lingue, questi strumenti permettono anche di effettuare la ricerca delle UF a partire dalla base (i.e. la base “*lapin*” > “coniglio” in UF come *poser un lapin* > “dare buca”; *coup du lapin* > “colpo di frusta”; *mariage de la carpe et du lapin* > “stare come i cavoli a merenda”).

Un'altra risorsa molto preziosa messa a disposizione delle apprendenti e degli apprendenti sono i corpora monolingui e paralleli offerti dalla piattaforma *online Sketch Engine*, più nello specifico, quelli che appartengono alla famiglia dei corpora *TenTen*, in esso inclusi: “*frTenTen17*” (5,7 miliardi di *tokens*) e “*itTenTen16*” (4,9 miliardi di *tokens*). Le studentesse e gli studenti hanno appreso alcune funzionalità di *Sketch Engine*, essenziali anche per l'attività traduttiva: trattasi della funzione “*concordance*” per l'osservazione degli esempi d'uso di una data UL o UF; del *thesaurus* per la ricerca di sinonimi o di UL appartenenti allo stesso campo semantico e, infine, della funzione “*word sketch*” per l'analisi del comportamento grammaticale e collocazionale di una data UL.

Avendo più della metà del gruppo già acquisito le conoscenze metodologiche e informatiche di base in materia di TA durante il corso di Laboratorio di *computer-assisted translation (cat tools)*, i sistemi di TA maggiormente utilizzati sul web nell'epoca attuale (i.e. *Google Traduttore*, *DeepL*, *Bing*, *Reverso* e *Systran*) sono stati rapidamente passati in rassegna.

Infine, si è ritenuto opportuno integrare la riflessione con alcuni riferimenti ai più recenti studi europei in materia di traduzione, cultura e fraseologia informatica (Corpas Pastor, Leiva Rovo 2011; Leiva Rovo 2013; Monti *et alii* 2018; Colson 2020), mostrando lo strumento *open access IdiomSearch* (Colson 2016) in fase di sperimentazione presso l'Université Catholique de Louvain. Si tratta di una piattaforma che offre la possibilità di testare algoritmi che permettono di estrarre le UF all'interno di un testo e di verificarne la frequenza e il grado di opacità e di trasparenza idiomatica.

3.3 La traduzione automatica delle UF di tipo locuzione

Affinché le discenti e i discenti familiarizzassero con i concetti teorico-metodologici finora esposti, sono stati presentati sei casi di UF di tipo lo-

cuzione in lingua francese (*avoir du pain sur la planche* > “essere indaffarato”; *moulin à paroles* > “chiacchierone”; *casser sa pipe* > “passare a miglior vita”; *à tire-larigot* > “a bizzeffe”; *écrire comme un chat* > “avere una scrittura che è una raspatura di gallina/ a zampa di gallina”; *avoir la chair de poule* > “avere la pelle d’oca”) da analizzare e tradurre in lingua italiana con l’ausilio degli strumenti poc’anzi menzionati. Nel processo di diagnosi metalinguistica e metatraduttiva delle UF di tipo locuzione, sono state richiamate le fasi di lavoro illustrate durante il primo incontro (Par. 3.1) ed è stato suggerito l’utilizzo del seguente modello di scheda per la classificazione delle caratteristiche linguistico-traduttive delle locuzioni oggetto di studio (Tab. 5).

I risultati complessivi mostrano un uso maggioritario, con cifre quasi analoghe, degli strumenti lessicografici elettronici (35%) e dei corpora *TenTen* (32%), e un uso, seppur limitato, del motore di ricerca di

UF di tipo locuzione	...
Tipo di locuzione	
Definizione e/o parafrasi	
Attestazione nei corpora	
Frequenza e fissità nel corpus web	
Traduzione letterale	
Traduzione automatica (FR>IT)	
Traduzione umana (FR>IT)	
Risorse elettroniche consultate	

Tabella 5: Modello di scheda: caratteristiche linguistico-traduttive delle UF di tipo locuzione

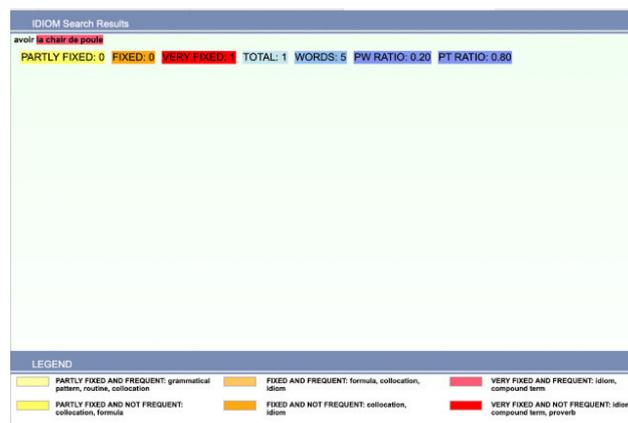


Figura 12: Riconoscimento automatico e grado di fissità della UF “avoir la chair de poule” (IdiomSearch)⁶

⁶ I colori utilizzati corrispondono al grado di fissità secondo lo *CPR-score* (Colson 2016), che va dal giallo paglierino al rosso. La legenda mostra una stima della relazione tra fissità e frequenza nel corpus di riferimento. Sotto il testo colorato appare il numero di combinazioni parzialmente fisse, fisse e molto fisse, così come il numero di *tokens*. Il rapporto *PW ratio* corrisponde al numero di frasi, diviso per il numero di *tokens*. Il rapporto *PT ratio*, invece, è un'indicazione, seppur approssimativa, della proporzione della UF nel testo.



Figura 13:
Associazione (CPR-score)
e frequenza della UF
“avoir du pain sur la planche”
(IdiomSearch)⁷

Google (17%). A seguire, emerge anche l’impiego della piattaforma Idiom-Search con il 14%, allo scopo di ottenere dati quantitativi sul grado di fissità delle UF di tipo locuzione (Fig. 12) e sulla frequenza d’uso nel corpus di riferimento (Fig. 13).

Quanto ai traduttori automatici disponibili in rete, solo il 2% della componente studentesca dichiara di averli utilizzati, in particolare Reverso e DeepL, elemento che ci si aspettava, dato l’approccio meramente lessicalista adottato durante il corso di LF2.

Alla luce di questi primi dati, è stato effettuato un secondo test al fine di promuovere un uso ragionato dei sistemi di TA presenti *online*. L’attenzione si è focalizzata su due software di TA — Reverso e Systran — ai quali è stato sottoposto un testo in lingua francese contenente trenta locuzioni di vario tipo⁸. Le studentesse e gli studenti hanno rilevato che i due sistemi di TA non riconoscono il carattere idiomatico delle locuzioni presenti nel testo e che, nel 60% dei casi, sono generate traduzioni letterali non appropriate: *être fier comme un coq* > “essere orgoglioso come un gallo”; *être chaud lapin* > “essere caldo coniglio/ coniglio caldo”; *fine mouche* > “mosca fine/ mosca sottile”; *devenir chèvre* > “diventare (una) capra”; *frais comme un gar-*

⁷ I dati sono calcolati su un corpus di 1,4 miliardi di *tokens*.

⁸ Il testo è stato rinvenuto sul forum della piattaforma ProZ, la più grande comunità mondiale per traduttori professionisti *online*. Cf. https://ara.proz.com/forum/french/15388-les_termes_emprunt%E9s_au_monde_animal_hilarious.html?print=1 (data ultima di consultazione: 28.07.2021).

don > “fresco come un cane”; *faire le pied de grue* > “fare il piede/ piedino di una gru”; *être copains comme cochons* > “essere amici di/come maiali”; *avoir du chien* > “avere un cane”; *crier comme un putois* > “urlare come una puzzola”; *ne pas casser trois pattes à un canard* > “non rompere tre zampe a un’anatra”; *plate comme une limande* > “piatta come una limanda/ lumaca”; *souffler comme une phoque* > “soffiare come una foca”; *rire comme une baleine* > “ridere come una balena”; *tirer les vers du nez* > “tirare i vermi dal naso”; *noyer le poisson* > “affogar(si) il pesce”; *avoir le bourdon* > “avere il bombo”; *filer comme un lièvre* > “filare/ scappare come una lepre”; *prendre quelqu’un pour un pigeon* > “farsi prendere per un piccione/ fare il capo espiatorio”.

Tabella 6:
Locuzioni tradotte
adeguatamente dai TA (FR>IT)

Locuzione (FR)	Traduzione automatica adeguata (IT)
Être rusé comme un renard	“Essere astuto come una volpe” (Reverso) “Essere furbo come una volpe” (Systran)
Être têtu comme une mule	“Essere testardo come un mulo” (Reverso, Systran)
Obtenir un rencard	“Dare un appuntamento” (Reverso, Systran)
Poser un lapin	“Dare buca” (Reverso, Systran)
Traiter comme un chien	“Trattare come un cane” (Reverso, Systran)
Verser des larmes de crocodile	“Versare lacrime di cocodrillo” (Reverso, Systran)
Fièvre de cheval	“Febbre da cavallo” (Reverso, Systran)

Locuzione (FR)	Traduzione automatica adeguata in modo parziale (IT)	Proposta di traduzione umana (IT)
Être myope comme une taupe	“Essere miope come una talpa” (Reverso, Systran)	“Essere cieco come una talpa”
Être fort comme un bœuf	“Essere forte come un bue” (Reverso, Systran)	“Essere forte come un toro”
Être malin comme un singe	“Essere furbo come una scimmia” (Reverso, Systran)	“Essere furbo come una volpe”
Rester muet comme une carpe	“Restare muto come una carpa” (Reverso, Systran)	“Essere muto come un pesce”
Être doux comme un agneau	“Essere dolce come un agnello” (Reverso, Systran)	“Essere docile come un agnellino”

Tabella 7: Locuzioni tradotte adeguatamente in modo parziale dai TA (FR>IT) e proposta di traduzione umana

Le apprendenti e gli apprendenti hanno altresì constatato che solo il 23% delle locuzioni presenti nel testo, è stato tradotto adeguatamente (Tab. 6), mentre il 17% è stato tradotto adeguatamente in modo parziale (Tab. 7).

Allo scopo di verificare se i sistemi di TA utilizzati non fossero realmente in grado di rilevare le locuzioni in oggetto o se la componente ludico-umoristica del testo di partenza in lingua francese avesse potuto rappre-

sentare un freno all'identificazione automatica delle UF di tipo locuzione, le studentesse e gli studenti hanno lanciato una *query*, digitando solo la stringa della locuzione, tralasciando, dunque, il contesto d'uso. I risultati, in particolare quelli di Reverso, sono stati assai diversi (Tab. 8): il 47%

Locuzione (FR)	Equivalente (IT) generato dai TA (al di fuori del contesto d'uso)
Avoir le bourdon	"Essere giù di corda" (Reverso)
Être fort comme un bœuf	"Essere forte come un toro" (Reverso, Systran)
Faire le pied de grue	"Aspettare a lungo in piedi" (Reverso)
Faire la fine mouche	"Fare la vecchia volpe" (Reverso)
Rester muet comme une carpe	"Essere muto come una tomba" (Reverso) "Essere muto come un pesce" (Reverso, Systran)
Sauter du coq à l'âne	"Cambiare argomento" (Reverso) "Saltare di palo in fresca" (Reverso)
Tirer les vers du nez	"Cavargli le parole di bocca" (Reverso)
Rire comme une baleine	"Ridere a crepapelle" (Reverso)
Noyer le poisson	"Menar il can per l'aia" (Reverso)
Frais comme un gardon	"Come nuovo" (Reverso) "Fresco e riposato come una rosa" (Reverso, Systran)
Être copains comme cochons	"Amici per la pelle" (Reverso) "Culo e camicia" (Reverso)
Ne pas casser trois pattes à un canard	"Niente di eccezionale" (Reverso)
Être myope comme une taupe	"Essere cieco come una talpa" (Reverso, Systran)
Prendre quelqu'un pour un pigeon	"Essere lo zimbello della compagnia" (Reverso)

Tabella 8: Locuzioni ed equivalenti generati dai TA (FR>IT) al di fuori del contesto d'uso

(Reverso) e il 13% (Systran) delle locuzioni è stato tradotto adeguatamente; invece, sottoponendo l'intero testo ai due software, il 60% delle locuzioni è stato tradotto letteralmente.

Seppur empirico, questo studio ha permesso alle studentesse e agli studenti di constatare che la traduzione delle UF di tipo locuzione è problematica e rappresenta un ostacolo per i sistemi di trattamento e di elaborazione naturale del linguaggio. In alcuni casi, i due software sono stati in grado di riconoscere le UF di tipo locuzione e di stabilire corrispondenze tra le due lingue di studio; in altri casi, invece, questo lavoro di automazione non è stato fatto. Tra le varie segnalazioni, le discenti e i discenti

ritengono che Systran dovrebbe implementare i dizionari della lingua di partenza, ossia della lingua francese: alcune espressioni non sono per nulla presenti e di conseguenza non sono tradotte in lingua italiana o sono tradotte male. Quanto a Reverso, bisognerebbe potenziare e perfezionare, attraverso sistemi di IA, l'integrazione delle UF di tipo locuzione nei dizionari e nelle risorse lessicografiche di base ad esso collegate (i.e. Expressio, ReversoContext, etc.), sì da sostenere la loro traduzione in modo bidirezionale.

Conclusioni

Attraverso la sperimentazione esposta, è emersa la necessità di assumere prospettive alternative per la TA delle UF di tipo locuzione. Potrebbe rivelarsi utile orientarsi verso l'elaborazione di lessici bilingui e la definizione di regole sintattiche che favoriscano il riconoscimento delle UF di tipo locuzione e, più in generale, verso la distinzione automatica tra sintagmi liberi e sintagmi legati.

La creazione di lessici bilingui permetterebbe non solo di prendere in considerazione l'aspetto traduttivo delle UF di tipo locuzione in vista di un perfezionamento attraverso gli strumenti basati sull'IA, come segnalato dalla componente studentesca, ma offrirebbe anche dati interessanti per l'insegnamento e l'apprendimento delle UF del francese come lingua straniera e aprirebbe la strada a nuove riflessioni sulle metodologie da adottare in aula. Le UF, di qualunque tipologia esse siano, traducono aspetti culturali della lingua di partenza ed è dunque fondamentale indirizzare le discenti e i discenti in questa direzione, con l'aiuto di strumenti informatici, la cui maggiore o minore utilità dipenderà solo dalla strategia adottata di fronte alla singola scelta traduttiva.

L'uso di metodologie ibride e di risorse informatiche, combinato con un approccio studente-centrico (Rivoltella 2014), consentirebbe all'apprendista traduttrice e/o traduttore di superare gli ostacoli derivanti dall'anisomorfismo tra le diverse lingue.

Appare, dunque, essenziale integrare, anche nell'ambito dei corsi di lingua straniera, conoscenze basilari di *digital literacy* orientate alla traduzione. Nel caso specifico delle UF di tipo locuzione, le risorse lessicografiche e gli strumenti come i *corpora* e i traduttori automatici potrebbero essere utilizzati prevalentemente per ricercare un equivalente fraseo-

logico *ad hoc*, la cui proposta di corrispondenza può attivare entrambe le traduzioni (letterale e figurativa), ottenendo un effetto simile a quello del testo di partenza.

Una volta acquisite queste conoscenze basilari, si potrebbero altresì introdurre alcune regole di base relative ai *parser* sintattici delle locuzioni, con la collaborazione di esperte o esperti in informatica (Clas, Gross 1998; Anscombe, Mejri 2017). Ad un livello più avanzato, un siffatto percorso consentirebbe di affrontare in aula un altro aspetto cardine per lo studio delle UF di tipo locuzione in un'ottica contrastiva: la distinzione tra l'uso puramente predicativo e l'uso idiomatico di queste espressioni, sì da disambiguare il loro significato e, di conseguenza, tradurle correttamente. A titolo d'esempio, nella locuzione forte "*tomber dans les pommes*" ("svenire"), l'attante X è un soggetto umano (*Nhum*); tuttavia, se la macchina non viene allenata con regole sintattico-semantiche precise, la frase "*Jean est tombé dans les pommes*" potrebbe essere erroneamente tradotta in modo letterale da un sistema di TA. In altri termini, se al posto di "*Jean*" (attante X soggetto umano) ci fosse stato "*La cerise*" (attante X soggetto non umano), non sarebbero sorti problemi di traduzione automatica; al contrario, con un attante X soggetto umano (*Nhum*), quale "*Jean*", si evincerebbe la difficoltà per il traduttore automatico che non è in grado di distinguere l'uso predicativo e l'uso idiomatico di questa espressione.

Infine, alla luce dei risultati della sperimentazione di cui sopra, è importante sottolineare che un insegnamento comprendente le diverse risorse informatiche applicate alla traduzione deve essere necessariamente accompagnato da una disamina delle varie strategie traduttive da adottare, in modo tale che le studentesse e gli studenti possano beneficiare al massimo di tutte le possibilità offerte da questi strumenti ed evitare di trovarsi in difficoltà di fronte alla sfida di tradurre le UF di tipo locuzione.

Bibliografia

- Anscombe Jean-Claude, Mejri Salah (eds) (2017). *Le figement linguistique: la parole entravée*. Parigi: Honoré Champion.
- Baker Mona, Saldanha Gabriela (eds) (2011). *Routledge Encyclopedia of Translation Studies*. Londra: Routledge.
- Bennett Paul (1994). "The translation unit in human and machine". *Revue Babel*, 40/1, 12-20.
- Bowker Lynne (2018). "Corpus linguistics is not just for linguists: considering the potential of computer-based corpus methods for library and information science research". *Library Hi Tech*, 36/2, 358-371.
- Bowker Lynne, Buitrago Ciro Jairo (eds) (2019). *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. UK: Emerald.
- Bowker Lynne, Delsey Tom (2016). "Translation Studies and Information Science: Adaptation, Collaboration, Integration". In: Yves Gambier, Luc Van Doorslaer (eds). *Border Crossings: Translation Studies and Other Disciplines*. Amsterdam: John Benjamins, 73-96.
- Clas André, Gross Gaston (1998). "Classes de figement des locutions verbales". In: *Le figement lexical*. Tunis: Éditions du CERES.
- Colson Jean-Pierre (2008). "Cross-linguistic phraseological studies: An overview". In: Sylviane Granger, Fanny Meunier (eds). *Phraseology: An interdisciplinary perspective*. Amsterdam: John Benjamins, 191-206.
- Colson Jean-Pierre (2016). "Set phrases around globalization: an experiment in corpus-based computational phraseology". In: Francisco Alonso Almeida, Ivalla Ortega Barrera, Elena Quintana Toledo, Margarita E. Sánchez Cuervo (eds). *Input a Word, Analyze the World. Selected Approaches to Corpus Linguistics*. Newcastle: Cambridge Scholars Publishing, 141-152.
- Colson Jean-Pierre (2020). "Automated Phraseology Extraction and Cultural Factors: An Experiment". *Phrasis, Rivista di studi fraseologici e paremiologici*, 4/1, 47-58.
- Consiglio d'Europa (eds) (2020). "Quadro comune di riferimento per le lingue: apprendimento, insegnamento, valutazione. Volume complementare". *Italiano LinguaDue*, 12/2, 8-285 (traduzione italiana a cura di Monica Barsi, Edoardo Lugarini, Anna Cardinaletti).
- Corpas Pastor Gloria (2003). "La traducción de unidades fraseológicas: técnicas y estrategias". In: Gloria Copas Pastor (ed). *Diez años de investigación en fraseología: Análisis sintáctico-semánticos, contrastivos y traductológicos*. Madrid: Iberoamericana, 213-224.

Corpas Pastor Gloria, Leiva Rojo Jorge (2011). "Placing Italian Idioms in a Foreign Milieu: a Case Study". In: Antonio Pamies, Lucia Luque Nadal, José Manuel Pazos Breña (eds). *Multi-Lingual Phraseography: Second Language Learning and Translation Applications*. Baltmannsweiler: Schneider Verlag Hohengehren, 289-298.

Froeliger Nicolas, Hewson Lance, Balliu Christian (eds) (2018). *Équivalences*, 45/1-2, numero tematico "Des unités de traduction à l'unité de la traduction". Bruxelles: Les Éditions du Ha(z/s)ard.

Gader Nabil, Ollinger Sandrine, Polguère Alain (2014). "One lexicon, two structures: so what gives?". In: Heili Orav, Christiane Fellbaum, Piek Vossen (eds). *Proceedings of the Seventh Global Wordnet Conference (GWC2014)*. Tartu: University Tartu Press, 163-171.

Leiva Rojo Jorge (2013). "Propuesta de formulario electrónico para la revisión de unidades fraseológicas traducidas". *Educatio Siglo XXI*, 31/2, 277-296.

Mejri Salah (1997). "Binarisme, dualité et séquence figée". In: Georges Kleiber, Martin Riegel (eds). *Les formes du sens*, Louvain-la-Neuve: Éditions Duculot.

Mel'čuk Igor (2013). "Tout ce que nous voulions savoir sur les phrasèmes mais...". *Cahiers de lexicologie*, 102/1, 129-149.

Mel'čuk Igor, Clas André, Polguère Alain (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Éditions Duculot.

Monti Johanna, Seretan Violeta, Corpas Pastor Gloria (eds) (2018). *Multiword Units in Machine Translation and Translation Technology*. Amsterdam: John Benjamins.

Ollinger Sandrine, Polguère Alain (2020). *Guide de navigation Spiderlex (version du 19 mai 2020)*. Nancy: ATILF-CNRS.

Polguère Alain (2016). *Lexicologie et sémantique lexicale*. Montréal: Les Presses de l'Université de Montréal.

Rivoltella Pier Cesare (ed) (2014). *Smart future: didattica, media digitali e inclusione*. Milano: Franco Angeli.

Vinay Jean-Paul, Darbelnet Jean (1958). *Stylistique compare du français et de l'anglais*. Parigi: Didier.

Variation terminologique et traduction automatique : une expérience didactique dans l'enseignement du français sur objectif spécifique (FOS)*

Jana Altmanova, Luca Bottiglieri

Introduction

La réflexion proposée dans la présente contribution se base sur l'expérience d'enseignement du français pour objectifs spécifiques (FOS) menée dans le cadre du Master 2 en Traduction spécialisée-français (désormais MTS) du *Dipartimento di studi letterari, linguistici e comparati* de l'Université de Naples « L'Orientale »¹. L'un des objectifs de ce cours, centré sur la pratique de la traduction spécialisée du français vers l'italien, est de guider les étudiant(e)s dans leur usage des outils de traduction automatique et de traduction. Une explication concernant les caractéristiques et le fonctionnement des plateformes de traduction précède l'atelier pratique durant lequel les étudiant(e)s comparent et analysent les résultats issus de la traduction humaine et de la traduction automatique (dorénavant TA) par le biais, entre autres, des instruments disponibles sur le web : Google Traduction, DeepL et Systran. Dans la première partie, nous exposons les résultats d'un questionnaire sur la perception des outils de traduction et de l'intelligence artificielle soumis aux étudiant(e)s en Traduction spécialisée² mais aussi aux étudiant(e)s du Master en Langues et Communication interculturelle de la zone euro-méditerranéenne (MLC) axé sur la communication interculturelle et moins centré sur la TA. Dans la seconde partie de cet article, nous décrivons un exemple d'activité proposée pen-

Jana Altmanova, Università L'Orientale di Napoli, jaltmanova@unior.it

* Jana Altmanova est auteure des paragraphes 1 et 2 ; le paragraphe 3 a été rédigé par Luca Bottiglieri, étudiant en MTS, sous la supervision de J. Altmanova.

¹ Le cours a été tenu entièrement en ligne sur la plateforme Microsoft Teams à cause des mesures gouvernementales liées à la pandémie de Covid-19.

² Les étudiant(e)s interviewé(e)s ne suivent pas tou(te)s le cours de deuxième année en Traduction spécialisée II (français), bien qu'il s'agisse d'étudiant(e)s se spécialisant en français du Coursus en Traduction spécialisée.

dant le cours, portant sur la traduction institutionnelle dans le domaine de l'environnement et, plus particulièrement, sur les unités terminologiques qui se sont révélées problématiques du point de vue traductologique et pour lesquelles la traduction automatique n'a pas toujours donné de sorties satisfaisantes.

1. Résultats des questionnaires

Le questionnaire sur l'intelligence artificielle, la variation linguistique et le multilinguisme a été soumis aux étudiant(e)s des formations mentionnées ci-dessus³.

Ces différences d'objectifs de formation expliquent que les étudiant(e)s de MTS (36 tests remplis) apparaissent plus conscient(e)s des instruments offerts par l'intelligence artificielle que les étudiant(e)s du Master en MLC (44 tests remplis). La première partie du questionnaire concerne les données sociologiques. Comme les études linguistiques sont poursuivies plus fréquemment par des femmes, il est assez naturel que les étudiantes soient plus nombreuses que les étudiants (11,1 % pour MTS et 13,6 % pour MLC). La langue maternelle des interviewés est l'italien, exception faite pour deux étudiantes Erasmus (une Espagnole et une Française). En ce qui concerne la connaissance d'autres langues étrangères, les étudiant(e)s déclarent maîtriser surtout l'anglais, l'espagnol, l'allemand⁴ et naturellement le français, la langue de leur spécialisation. Les autres langues étrangères connues sont : l'arabe et le russe pour les étudiant(e)s de MTS, et le hollandais, l'arabe, le russe, le turc, le berbère et le japonais pour le MCL. Dans la plupart des cas, les étudiant(e)s choisissent la combinaison entre une langue européenne et une langue extra-européenne. Ils et elles déclarent avoir un très bon niveau d'italien et estiment bien connaître le français eu égard aux quatre compétences langagières (production écrite et orale, compréhension écrite et orale). Quant à l'intérêt pour l'intelligence artificielle et le développement des technologies (Fig. 1, 2), les étudiant(e)s apparaissent très intéressé(e)s (MTS : 69,4 % ; MLC : 36,4 %), assez intéressé(e)s (MTS : 27,8 % ; MLC : 47,7 %), plus non que oui (MTS : 2,8 % ; MLC : 15,9 %) :

³ Pour les objectifs des Masters auxquels sont inscrit(e)s nos étudiant(e)s, nous renvoyons aux pages suivantes : <https://www.unior.it/didattica/17412/2/traduzione-specialistica.html> ; <https://www.unior.it/didattica/17347/2/presentazione-obiettivi-formativi-sbocchi-occupazionali.html> ; date de la dernière consultation : 12.07.2021).

⁴ L'italien dans le cas des étudiantes étrangères.

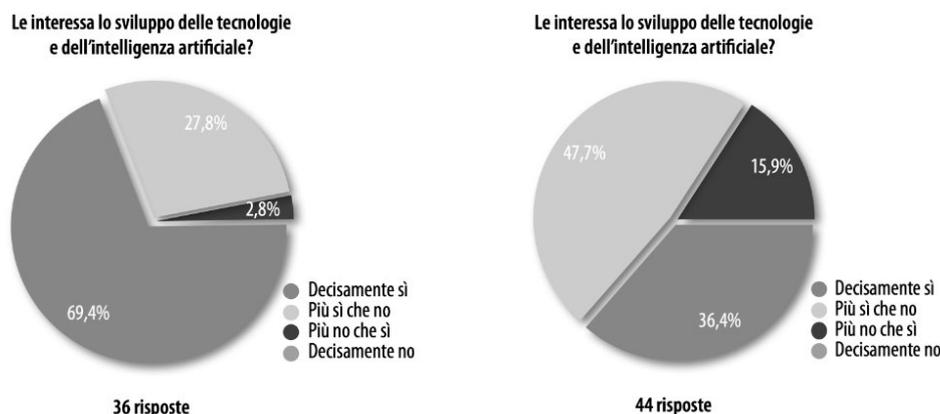


Figure 1, 2 : Intérêt des étudiant(e)s pour le développement des technologies et l'intelligence artificielle

En ce qui concerne les instruments technologiques, les étudiant(e)s déclarent utiliser au quotidien les applications et les réseaux sociaux tels que Facebook, Instagram, WhatsApp, LinkedIn, OmegaT, Amazon, Treccani, WordReference, YouTube ; les programmes d'écriture et le paquet office, tels que Microsoft Word, Office Suite, Excel et naturellement Microsoft Teams et Zoom (employés durant la période de la crise sanitaire). Les étudiant(e)s de MTS déclarent connaître aussi les plateformes de traduction automatique telles que DeepL, Reverso, Google Traducteur et WordReference, alors que les étudiant(e)s de MLC les connaissent moins, bien qu'ils/elles utilisent les sites de Reverso et WordReference.

Une dernière question porte sur l'utilité de l'intelligence artificielle dans le futur proche dans le domaine d'études des étudiant(e)s interviewé(e)s (97,2 % des étudiant(e)s de MTS les considèrent très utiles contre 90 % des étudiant(e)s en MLC). Étant donné que le cursus en MTS est davantage centré sur la traduction, y compris la traduction automatique et l'évolution technologique, cette catégorie d'étudiant(e)s trouve l'intelligence artificielle plus utile dans son domaine que les étudiant(e)s de MLC, dont 90 % estiment cependant que l'intelligence artificielle leur sera utile, même si ils/elles connaissent mal les instruments de TA.

La deuxième partie du questionnaire est consacrée aux informations plus précises concernant, entre autres, l'utilisation des traducteurs automatiques gratuits ou payants. Les étudiant(e)s de MTS utilisent ces outils de façon assez convaincue (63,9 % — oui, tout à fait + 19,4 % — plus oui que non), alors que les étudiant(e)s de MLC les utilisent de temps en temps

(47,7 % — certainement oui + 40,9 % — plus oui que non). Il est intéressant d'observer que le degré de fiabilité des traducteurs automatiques n'est pas particulièrement élevé pour les étudiant(e)s de MTS, à savoir 77,1 %, alors que 20 % les considèrent peu fiables et 2,9 % pas du tout fiables. Ces pourcentages sont presque les mêmes pour les étudiant(e)s de MLC dont 75 % estiment que les traducteurs automatiques sont assez fiables, 20,5 % les considèrent peu fiables, 2,25 % pas du tout fiables et le même pourcentage d'étudiant(e)s n'a pas d'opinion à cet égard. Néanmoins, le degré de confiance envers l'instrument varie en fonction du couple de langues utilisées, comme nous pouvons l'observer dans les figures 3 et 4 :

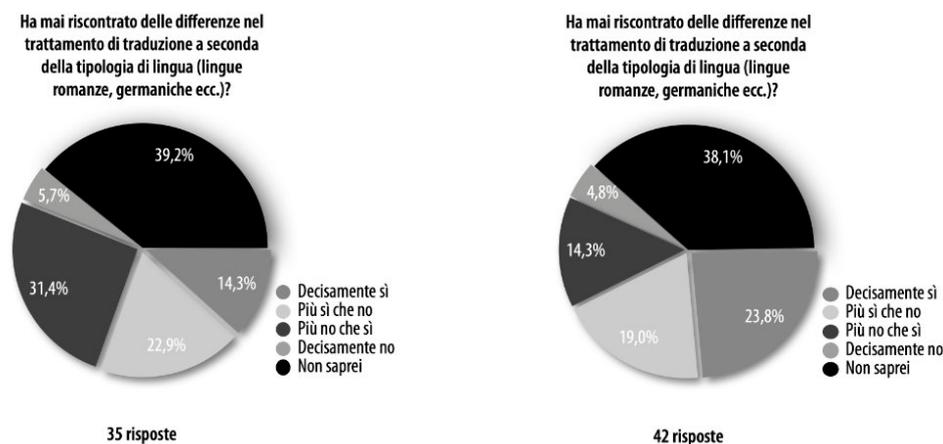


Figure 3, 4 : Différences dans la TA en fonction du couple de langues utilisées

Ainsi, les étudiant(e)s de MTS estiment que la traduction de l'anglais et des langues romanes vers l'italien apparaît plus fiable que la traduction à partir d'autres langues, telles que l'arabe, le russe et les langues slaves en général qui sont perçues comme plus problématiques. Quant aux langues romanes, les étudiant(e)s soulignent les difficultés de traduction de l'accord, du genre, du nombre et des temps verbaux. Les étudiant(e)s de MLC relèvent la difficulté de traduire les phrases figées (et certaines structures complexes) des langues germaniques et remarquent une précision moins élevée dans la traduction des temps verbaux pour les langues romanes. En outre, les étudiant(e)s estiment que certaines plateformes sont plus adéquates pour certaines langues, alors que d'autres le sont moins. Nous avons pu observer que plus de la moitié des étudiant(e)s de MTS connaissent un plus grand nombre de logiciels de gestion de la terminologie (Fig. 5, 6) (44,4 % plus oui que non et 13,9 % certainement oui) alors que les étudiant(e)s de MLC n'en connaissent presque pas du tout.

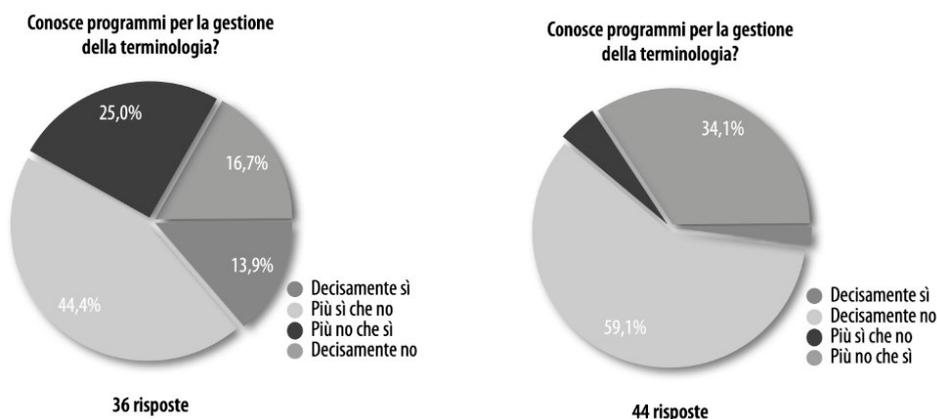


Figure 5, 6 : Connaissance de logiciels pour la gestion de la terminologie

Quant à l'utilisation des programmes de traduction, la situation entre les deux cursus est beaucoup plus équilibrée : 47 % d'étudiant(e)s de MTS l'utilisent sans aucun doute, 38,9 % plus oui que non ; juste un peu moins les étudiant(e)s de MLC.

Cependant, les étudiant(e)s de MTS se méfient davantage des plateformes que les étudiant(e)s de MLC (82,9 % d'étudiant(e)s de MTS contre 92,9 % d'étudiant(e)s de MLC) (Fig. 7, 8). Les étudiant(e)s orienté(e)s vers la traduction professionnelle ont une attitude plus critique à cet égard que les étudiant(e)s de MLC, plus intéressé(e)s à la communication et à la médiation interculturelle. En outre, la différence majeure entre les deux cursus concerne la consultation des sources, car 37,1 % des étudiant(e)s de

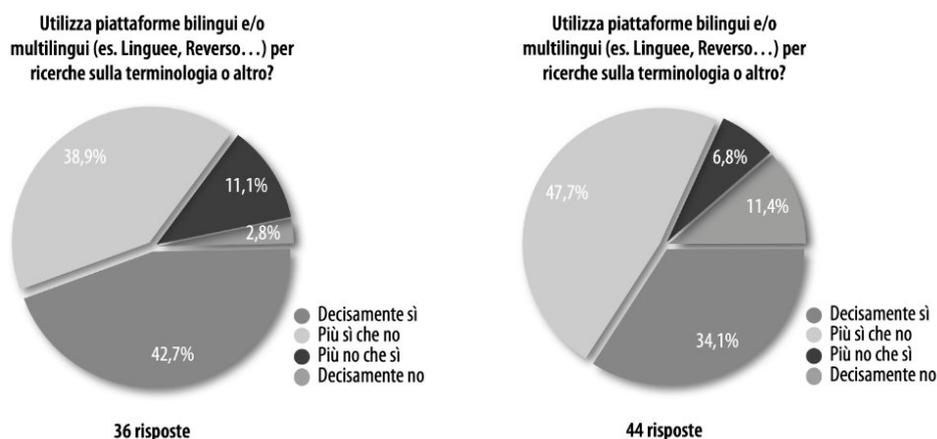
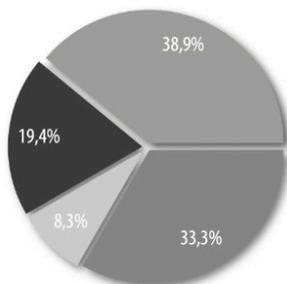


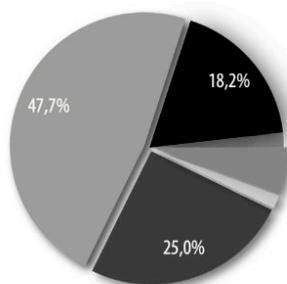
Figure 7, 8 : Utilisation des plateformes bilingues et/ou multilingues

Utilizza piattaforme bilingui e/o multilingui (es. Linguee, Reverso...) per ricerche sulla terminologia o altro?



36 risposte

Utilizza piattaforme bilingui e/o multilingui (es. Linguee, Reverso...) per ricerche sulla terminologia o altro?



36 risposte

Figure 9, 10 : Utilisation de l'écriture automatique et perception des suggestions

MTS affirment toujours contrôler la source contre 11,9 % des étudiant(e)s de MLC. En revanche, en ce qui concerne le choix du terme proposé par la plateforme, la totalité des étudiant(e)s soulignent l'importance du contexte, qui constitue le premier critère de choix.

Quant aux programmes d'alignement, moins de la moitié d'étudiant(e)s de MTS (30,5 %) affirment connaître

les logiciels d'alignement, et moins de 8 % des étudiant(e)s de MLC connaissent ces programmes. En ce qui concerne l'utilisation de l'écriture automatique (Fig. 9, 10), une perception différente apparaît. En effet, les étudiant(e)s de MTS les utilisent après vérification (38,9 % + 19,4 %), alors que 33,3 %

n'ont pas l'habitude de les utiliser. Quant aux étudiant(e)s de MLC, 47,7 % + 25 % d'entre elles et eux les utilisent après vérification et 18,2 % n'ont pas l'habitude d'y avoir recours.

Une question est consacrée à la reconnaissance vocale dans le cas des messages préenregistrés et automatisés. La plupart des étudiant(e)s remarquent que l'intonation et la prononciation ne sont pas adéquates (55,6 % pour MTS ; 47,7 % pour MLC) et que la langue utilisée par les réponders automatiques est très formelle et inusuelle (53,8 % pour MTS ; 47,7 % pour MLC). En ce qui concerne les chats et l'assistance en ligne, les étudiant(e)s (55,6 % pour MTS ; 50 % pour MLC) ne sont pas sûr(e)s qu'il s'agisse d'une voix humaine. De la même manière, elles et ils trouvent que les sites emploient une langue inusuelle et inexacte, et estiment que les raisons sont les suivantes :

Pour MTS :

- Traduction automatique sans *post-editing*, sans révision humaine, sans révision d'un traducteur natif ;
- Compétence linguistique limitée de la part du traducteur ;
- Traduction entre deux langues très distantes ;
- Instruments de traduction automatique vieilliss ;
- Superficialité de celui ou celle qui publie les contenus du site.

Pour MLC :

- Utilisation de traducteurs automatiques qui ne distinguent pas les termes polysémiques ;
- Utilisation de traducteurs automatiques sans vérification ;
- Manque de connaissance de la langue/culture (phrases idiomatiques) ;
- Traduction peu soignée des traducteurs humains peu experts (surtout dans le cas de la traduction spécialisée) ou effectuée par les traducteurs automatiques ;
- Manque de contextualisation des termes.

Plus généralement, les étudiant(e)s reconnaissent la nécessité de l'intervention d'un traducteur humain qui puisse contextualiser les expressions qui sont parfois culturellement connotées, mais aussi des néologismes et/ou des collocations. La dernière question concernant les aspects positifs et négatifs de la TA permet de relever des correspondances entre les deux cursus. Les étudiant(e)s de MTS relèvent les aspects positifs suivants :

- Augmentation de l'interaction dans l'industrie de la langue, l'enseignement des langues, etc. ;
- Rapidité d'accès à l'information, de l'exécution de la traduction ;
- Accroissement du lexique actif ;
- Capacité d'interaction entre l'IA et l'humain, qui pourrait résoudre des problèmes linguistiques tels que l'incapacité de trouver immédiatement les mots justes dans une autre langue ;
- Efficacité et multifonctionnalité dans des délais plus brefs.

Parmi les aspects positifs, les étudiant(e)s de MLC classent les éléments ci-dessous :

- Rapidité d'exécution de la traduction ;

- Traducteurs automatiques plus efficaces, sans être parfaits ;
- Corpus textuels plus riches ;
- Amélioration des résultats dans l'apprentissage, notamment en ce qui concerne les personnes ayant des difficultés d'apprentissage ;
- Exploitabilité 24h sur 24 ;
- Exploitation des programmes de transcription (but didactique, ludique, professionnel, communicatif) ;
- Repérage plus rapide et plus facile des équivalents traductifs pour un emploi standard de la langue.

Parmi les aspects négatifs les étudiant(e)s de MTS relèvent :

- Erreurs de traduction et malentendus possibles ;
- Inexactitude, imprécision d'expression due à une faible connaissance de la langue ;
- Manque de contextualisation ;
- Réduction de la motivation de la part du traducteur ;
- Réduction de l'apprentissage des langues (compétence lexi-culturelle limitée) ;
- Dévalorisation du travail du traducteur humain ;
- Dépendance de la technologie et réduction des opportunités de travail ;
- Manque de créativité, erreurs d'évaluation à cause de l'utilisation des algorithmes.

Les aspects négatifs relevés par les étudiant(e)s de MLC :

- Tendance à ne pas vérifier les informations fournies par la traduction automatique ;
- Ne plus mettre en œuvre ses propres connaissances ;
- Erreurs de traduction parfois dues à une révision incomplète de la part d'un traducteur humain ;
- Dévalorisation du travail du traducteur humain ;
- Cristallisation de l'évolution de la langue qui devient plus contrôlée et moins naturelle ;
- Travail humain aliénant ;
- Réduction de la créativité humaine ;
- Utilisation et apprentissage erronés de la langue ;
- Automatisation du travail humain ;
- Réduction de la capacité d'améliorer nos compétences linguistiques.

2. Remarques pertinentes concernant le questionnaire

Pour conclure, les résultats du questionnaire montrent que les étudiant(e)s sont conscient(e)s de l'utilité de l'IA, bien qu'elles et qu'ils en reconnaissent certaines limites. Elles et ils estiment par exemple que l'IA, en particulier les plateformes de TA, ne contextualisent pas toujours de façon satisfaisante les termes à traduire. En général, l'ensemble des étudiant(e)s partage ces considérations, bien que les étudiant(e)s en Traduction spécialisée II (français) semblent plus averti(e)s non seulement quant au fonctionnement des moyens de TA mais aussi quant aux autres ressources terminologiques et dictionnairiques utiles à la traduction. En revanche, les étudiant(e)s en MLC, très averti(e)s en ce qui concerne l'analyse linguistique, possèdent moins d'informations quant au fonctionnement des plateformes de traduction, car cette compétence ne rentre pas *a priori* dans les objectifs du cours. Pourtant les compétences en TA, tout comme la compréhension des mécanismes qui régissent la traduction automatique et les autres ressources informatiques pertinentes pour l'industrie de la langue, peuvent se révéler utiles pour toutes celles et tous ceux qui travaillent dans le domaine de la communication interlinguistique.

L'exemple qui suit illustre le travail pratique de traduction mené à bien par le biais des plateformes de TA fait par les étudiant(e)s de MTS.

3. Expérience en traduction automatique : le point de vue des étudiant(e)s

3.1 Présentation du corpus et méthodologie

L'étude présentée ici est axée sur la traduction du français vers l'italien et vice versa, à partir d'un corpus de 88 476 occurrences⁵, composé de textes officiels d'ordre juridico-normatif portant sur la question environnementale, sélectionnés sur différents sites Internet, sur ceux de l'Union européenne et de l'Unesco entre autres. Nous y avons identifié les termes et les expressions susceptibles de poser problème pour les traducteurs automatiques.

⁵ À l'aide de TermoStat (<http://termostat.ling.umontreal.ca/interfaceTermostat.php>), développé par Patrick Drouin, professeur agrégé à l'Université de Montréal, nous avons établi notre corpus en travaillant sur un échantillon de termes simples et de termes complexes nominaux qui nous paraissaient d'intérêt aux fins de notre analyse traductionnelle : pour le premier groupe de termes, nous en avons sélectionné cinq ; pour le deuxième, trois ; pour le troisième, quatre ; enfin, pour le quatrième, onze (au total, vingt-trois), raison pour laquelle cette étude ne vise pas à être exhaustive mais seulement à offrir des pistes de réflexion sur les points forts et les points faibles de la TA de nos jours.

Unité terminologique (fr.)	Phrase française	Équivalents possibles	Google Traducteur	DeepL	Systran
Durabilité de l'océan	Rio + 20 : <u>durabilité de l'océan</u> et des zones côtières.	Sviluppo sostenibile degli oceani.	Rio + 20 : <u>sostenibilità dell'oceano</u> e delle zone costiere	Rio + 20 : <u>sostenibilità dell'oceano</u> e delle zone costiere	Rio + 20 : <u>sostenibilità degli oceani</u> e delle zone costiere
En état de Surpêche	La moitié des réserves de poisson situées au large des côtes de l'Afrique de l'Ouest sont considérées comme en <u>état de surpêche</u> [...].	In stato di sovrasfruttamento delle risorse ittiche/ Interessate da eccessivo sfruttamento delle risorse ittiche.	[...] considerata <u>sovrasfruttata</u> , in parte a causa della pesca illegale.	[...] considerati <u>sovrasfruttati</u> , in parte a causa della pesca illegale.	[...] considerata <u>in stato di sovrasfruttamento</u> , in parte a causa della pesca illegale.
Herbiers marins	[...] en y ajoutant certains des plus grands estrans du monde, y compris des <u>herbiers marins</u> et des marais salés.	Letti di alghe marine/ Praterie di posidonia.	[...] tra cui <u>alghe</u> e paludi salmastre.	[...] tra cui <u>letti di fanerogame</u> e paludi salate.	[...] tra cui <u>praterie marine</u> e paludi salmastre.

Tableau 1 : Unités terminologiques et leurs équivalents proposés par la machine

Notre argumentation s'articule autour de deux points fondamentaux : l'un portant sur les difficultés de traduction dues soit à une instabilité au niveau de la signification soit à une complexité conceptuelle, l'autre tenant aux caractéristiques morphologiques des termes en question. Nous avons ainsi considéré les problèmes liés aux limites de l'outil de traduction en tant que tel, par exemple la méconnaissance des variantes socioculturelles des termes ou bien l'incapacité à considérer l'unité terminologique dans son ensemble comme une unité traductionnelle.

Pour une plus grande clarté de présentation, les unités terminologiques retenues ont été divisées en quatre catégories en fonction des sorties et des noyaux traductionnels problématiques pour un traducteur automatique, à savoir :

- traductions erronées ;
- synonymies terminologiques ;
- difficultés de traduction des mots contenant des préfixes d'origine gréco-latine ;
- difficultés de traduction des dénominations dérivant de l'anglais.

3.1.1 Erreurs de traduction des termes spécialisés

Le tableau 1 propose quatre exemples d'unités terminologiques relevant du lexique spécialisé de l'environnement, accompagnées de leurs

équivalents respectifs en contexte issus des traducteurs en ligne (Google Traducteur, DeepL et Systran) ainsi que des équivalents possibles en italien. Les deux premières catégories concernant les erreurs de traduction et la synonymie terminologique, que nous analyserons ensemble, présentent un intérêt pour ce qui concerne l'aspect conceptuel des termes traduits automatiquement, car elles relèvent de la sélection de l'unité terminologique par les outils de TA pendant la phase de transfert interlinguistique (Lavault-Olléon 2007).

3.1.2 Cas de synonymie terminologique

Nous illustrerons les cas de synonymie terminologique par trois exemples d'unités terminologiques traduites automatiquement en contexte au moyen des trois outils susmentionnés ; des choix alternatifs d'équivalents additionnels en italien pour ces mêmes termes sont proposés. Cette catégorie se caractérise notamment par une variabilité des dénominations dans le système de la langue d'arrivée (Humbley 2018).

Unité terminologique (fr.)	Phrase française	Équivalents possibles	Google Traducteur	DeepL	Systran
Durabilité de l'océan	Rio + 20 : <u>durabilité de l'océan</u> et des zones côtières.	Sviluppo sostenibile degli oceani.	Rio + 20 : <u>sostenibilità dell'oceano</u> e delle zone costiere	Rio + 20 : <u>sostenibilità dell'oceano</u> e delle zone costiere	Rio + 20 : <u>sostenibilità degli oceani</u> e delle zone costiere
En état de Surpêche	La moitié des réserves de poisson situées au large des côtes de l'Afrique de l'Ouest sont considérées comme <u>en état de surpêche</u> [...].	In stato di sovrasfruttamento delle risorse ittiche/ Interessate da eccessivo sfruttamento delle risorse ittiche.	[...] considerata <u>sovrasfruttata</u> , in parte a causa della pesca illegale.	[...] considerati <u>sovrasfruttati</u> , in parte a causa della pesca illegale.	[...] considerata <u>in stato di sovrasfruttamento</u> , in parte a causa della pesca illegale.
Herbiers marins	[...] en y ajoutant certains des plus grands estrans du monde, y compris des <u>herbiers marins</u> et des marais salés.	Letti di alghe marine/ Praterie di posidonia.	[...] tra cui <u>alghe</u> e paludi salmastre.	[...] tra cui <u>letti di fanerogame</u> e paludi salate.	[...] tra cui <u>praterie marine</u> e paludi salmastre.

Tableau 2 : La TA et les synonymes terminologiques

Il faut souligner que parmi les unités contenues dans les tableaux 1 et 2, certaines n'ont pas d'équivalent direct (c'est-à-dire littéral) en italien : c'est le cas de « cindyniques » (Tab. 1) et du syntagme prépositionnel « en état de surpêche » (Tab. 2). Le premier terme se traduit généralement en italien par l'emprunt anglais « *risk analysis* » (3 150 occurrences sur

Google), utilisé notamment dans les contextes de communication spécialisée, ou par le calque correspondant « *analisi del rischio* », certainement plus générique (1 610 000 résultats), ou bien par une autre formulation tout aussi valable, « *scienza della sicurezza* » (80 600 occurrences sur Google). Ici, les traducteurs automatiques n'ont pas su s'adapter aux normes terminologiques de l'italien, préférant opter pour des traductions strictement calquées sur le terme français, tantôt le traitant en tant qu'adjectif (Google Traducteur), tantôt le considérant comme un substantif (DeepL et Systran). Parallèlement, « en état de surpêche » (Tab. 2) a posé beaucoup de difficultés aux outils de TA car les sorties de ceux-ci sont incomplètes, notamment à cause du terme « surpêche », qui n'a pas d'équivalent littéral en italien. Le traducteur automatique a donc proposé des mots simples, c'est-à-dire des adjectifs (Google Traducteur : « *sovrasfruttata* » ; DeepL : « *sovrasfruttati* »), ou une paraphrase qui ne recouvre pas le concept auquel la dénomination française renvoie (Systran : « *in stato di sfruttamento* »). Ce qui manque ici est la référence à la pêche et au domaine halieutique ; pour cela, le traducteur humain pourrait opter pour des unités certainement plus longues mais plus précises : « *interessate da eccessivo sfruttamento delle risorse ittiche* » (2 080 occurrences sur Google) ou bien « *in stato di sovrasfruttamento delle risorse ittiche* » (3 610 occurrences sur Google). En outre, il est important de faire mention de tous ces concepts complexes, difficiles à délimiter et à circonscrire dans les bornes d'une définition exhaustive, d'où les pièges que leurs dénominations représentent pour le traducteur automatique : il s'agit des syntagmes nominaux tels qu'« aménagement touristique » et « aménagements fluviaux ». La première unité terminologique est plutôt obscure du point de vue conceptuel, d'où la difficulté des traducteurs automatiques à la gérer de façon ponctuelle. En effet, chaque traducteur propose une sortie qui ne recouvre pas l'extension conceptuelle du terme : ici, les résultats sont plutôt décevants (Google Traducteur : « *sviluppo turistico* » ; DeepL : « *sviluppo del turismo* » ; Systran : « *struttura turistica* »). En revanche, le corpus de Google invite à choisir des unités polylexicales telles que « *pianificazione turistica* » (12 100 occurrences sur Google) avec sa variante « *pianificazione turistica territoriale* » (1 810 occurrences sur Google). De même, « aménagements fluviaux » pose un problème aux traducteurs automatiques dans la mesure où il n'existe pas d'équivalent en italien qui soit parfaitement correspondant à la structure notionnelle du terme français, raison pour laquelle les sorties

des outils de TA ne sont pas fiables (Google Traducteur : « *strutture fluviali* » ; DeepL : « *impianti fluviali* » ; Systran : « *impianti fluviali* ») ; à cet égard, il faudrait proposer une correspondance plus complexe mais quelque peu récurrente dans les textes techniques italiens (appartenant à cette sphère d'activité), à savoir « *opere di sistemazione fluviale* » (2 350 occurrences sur Google) ou bien la forme abrégée « *sistemazioni fluviali* » (6 440 occurrences sur Google). Enfin, nous signalons les cas où un terme de la langue de départ peut donner lieu à une multitude d'équivalents dans la langue d'arrivée : nous nous référons aux syntagmes nominaux « crédits carbone » et « herbiers marins ». La première unité terminologique est traduite de façon pertinente par tous les traducteurs en ligne que nous avons consultés (« *crediti di carbonio* », 48 300 occurrences sur Google) ; cependant, aucun d'entre eux n'a opté pour la variante, tout à fait acceptable, de « *crediti verdi* » (6 280 occurrences sur Google). Précisons par ailleurs que les deux unités, française et italienne, sont calquées sur la forme anglaise « *carbon credits* ». Pareillement, « herbiers marins » renvoie à une notion floue, d'où les différentes variantes acceptables : les opérations des traducteurs automatiques ont abouti à des équivalents simples (« *alghe* » pour Google Traducteur) ou à des formulations complexes (« *letti di fanerogame* » pour DeepL ou « *praterie marine* » pour Systran). Ces outils auraient pu également fournir d'autres variantes, tout aussi acceptables en italien, telles que « *letti di alghe marine* » (672 occurrences sur Google) ou bien « *praterie di posidonia* » (42 000 occurrences sur Google).

3.1.3. Difficultés de traduction des mots contenant des préfixes d'origine gréco-latine

Pour cette troisième catégorie de termes, la sortie de TA de quatre mots dérivés à partir du préfixe savant « éco- » fait l'objet de l'analyse ; à cet égard, sont proposées, le cas échéant, des traductions équivalentes de ces termes (avant une évaluation des traits terminologiques les plus saillants). Ces formations lexicales en *éco-* peuvent être définies comme autochtones puisqu'elles sont typiques de la langue française en dépit de leur dérivation savante (issue du grec).

Une fois entrés dans les différents outils de TA, les termes français ont entraîné des inexactitudes en raison de leur idiomatie (au sein du système de la langue d'arrivée, nécessitant ainsi une adaptation terminologique au contexte cible). Il s'agit par exemple des termes « éco-réfugiés » et « éco-participation » : le premier est traduit de façon littérale par les traducteurs

Variation terminologique et traduction automatique : une expérience didactique dans l'enseignement du français sur objectif spécifique (FOS)

Unité terminologique (fr.)	Équivalents possibles	Phrase française	Google Traducteur	DeepL	Systran
Éco-réfugiés	Profughi ambientali/ Rifugiati ambientali	[...] amplifient ou au contraire limitent le nombre des <u>éco-réfugiés</u> climatiques.	[...] il numero degli <u>eco-rifugiati</u> climatici.	[...] Il numero di <u>eco-rifugiati</u> climatici.	[...] Il numero di <u>eco-rifugiati</u> climatici.
Écoparticipation		Ils sont donc à ce titre soumis à <u>écoparticipation</u> [...]	In quanto tali, sono soggetti a <u>eco-partecipazione</u> [...]	Come tali, sono soggetti a un' <u>ecotassa</u> [...].	In quanto tali, essi sono quindi soggetti ad <u>ecopartecipazione</u> [...].
Écogestes citoyens	Azioni ecologiche dei cittadini/ Azioni ecologiche cittadine	Les experts n'en doutent plus : les <u>écogestes</u> citoyens et les modèles de recyclage [...]	Gli esperti non ne dubitano più: le <u>azioni eco-friendly</u> e i modelli di riciclo più innovativi [...]	Gli esperti non dubitano più che le <u>eco-azioni</u> dei cittadini e i modelli di riciclaggio più innovativi [...]	Gli esperti non ne dubitano più: i <u>cittadini ecologici</u> [sic] e i modelli di riciclaggio più innovativi [...].
Éco-organisme		Le décret oblige les récupérateurs à contractualiser avec un <u>éco-organisme</u> .	Il decreto obbliga i raccoglitori di rifiuti a contrattare con un' <u>organizzazione ecologica</u> .	Il decreto obbliga le aziende di riciclaggio a firmare un contratto con un' <u>organizzazione ecologica</u> .	Il decreto obbliga i recuperatori a contrattualizzare con un <u>ecoorganismo</u> [sic].

Tableau 3 : La TA des mots contenant des préfixes d'origine gréco-latine

en ligne (« *eco-rifugiati* »). Néanmoins, ceux-ci ont ignoré l'existence de désignations italiennes plus appropriées au domaine des migrations liées au dérèglement climatique : « *profughi ambientali* » (29 400 occurrences sur Google) ou bien « *rifugiati ambientali* » (8 480 occurrences sur Google). Dans le même ordre d'idées, « écoparticipation » est un terme culturellement connoté lié au contexte juridique français, c'est pourquoi un traducteur humain opérerait sans aucun doute pour une adaptation culturelle (Gaudin 1993). Toutefois, deux traducteurs automatiques sur trois ont donné en sortie des calques du français qui ne sont pas du tout idiomatiques en italien (Google Traducteur et Systran : « *eco-partecipazione* »), tandis que l'autre (DeepL) s'est avéré extrêmement efficace en ce qui concerne l'adaptation du terme (« *eco-tassa* », 347 000 occurrences sur Google)⁶.

⁶ Citons aussi d'autres exemples de problèmes d'aisance expressive : le terme « écoconception », calque de l'anglais « *ecodesign* » dont l'équivalent italien est l'unité polylexicale « *progettazione ecocompatibile* », que Google Traducteur et Systran ont reconnu lors de la phase de transfert contrairement à DeepL, qui, lui, a offert en sortie le terme anglais « *ecodesign* », très peu commun en italien ; le terme « écoquartier », calqué sur l'anglais « *eco-district* » et sur sa variante « *eco-neighbourhood* », pour la traduction duquel l'italien choisit d'habitude la forme « *ecoquartiere* », presque identique à la dénomination française : Google Traducteur et Systran traduisent ce mot correctement, tandis que DeepL opte pour le calque de l'anglais « *eco-district* », à savoir « *eco-distretto* », dont l'usage est assez rare en italien.

Deuxièmement, quant au syntagme « écogestes citoyens », nous constatons une situation particulière : ici, la problématique majeure tient au fait qu'il n'y a pas de véritable équivalent en italien. Pour résoudre ce problème traductionnel, dans la traduction humaine on pourrait avoir recours à des paraphrases explicatives du type « *azioni ecologiche compiute dai cittadini* » mais les trois traducteurs automatiques optent pour des désignations synthétiques telles que « *azioni eco-friendly* », cette fois-ci en mélangeant italien et anglais (Google Traducteur), « *eco-azioni* » (DeepL), gommant la référence aux citoyens, et « *cittadini ecologici* » (Systran), désignation erronée faisant abstraction du sens de « geste », « action ». N'ayant pas de solution toute faite et établie du point de vue institutionnel (Raus 2013), chaque traducteur a donc mobilisé les ressources qui l'alimentent pour essayer de fournir une traduction dont le sens est finalement approximatif.

Troisièmement, nous rappelons la présence d'un noyau traductionnel problématique pour tout traducteur automatique : la polysémie (Poibeu 2019). En effet, le terme « éco-organisme » pose un problème à ces outils à cause du caractère polysémique du terme « organisme » en français tout comme en italien. Toutefois, dans ce cas, l'usage effectif de l'italien renvoie à une différence linguistique où la forme italienne s'éloigne du signifiant français, raison pour laquelle les sorties de Google Traducteur et de DeepL sont plus attentives à la réalité sociolinguistique de l'italien (« *organizzazione ecologica* ») que celle de Systran (« *ecoorganismo* » [sic]). Il est également à signaler que les deux premiers outils recourent à une unité bilinguistique obtenue par ajout de l'adjectif, alors que le français a recours à une unité monolingue obtenue par dérivation par le biais du préfixe italien « *eco-* ».

3.1.4 Dénominations dérivées de l'anglais

La dernière catégorie présentée ici concerne les unités terminologiques issues de l'anglais, et donc créées dans un contexte de communication multiculturelle (Guidère 2008), mais qui ont pris une forme différente en fonction des pays d'accueil de ces termes. La quasi-totalité des termes repérés sont des formations lexicales en « éco- » ou « bio- » qui ont comme point de départ la langue anglaise. En d'autres termes, le préfixe en question a été d'abord employé en anglais et il a rejoint par la suite le système linguistique des langues telles que le français et l'italien sous la forme d'un calque.

Variation terminologique et traduction automatique : une expérience didactique dans l'enseignement du français sur objectif spécifique (FOS)

Unité terminologique (fr.)	Expression anglaise de départ	Phrase française	Google Traducteur	DeepL	Systran
Biodéchets	Biowaste	[...] se distingue des déchets biodégradables, dans la mesure où les <u>biodéchets</u> n'incluent pas le papier.	[...] i <u>rifiuti organici</u> non includono la carta.	[...] i <u>rifiuti biologici</u> non includono la carta.	[...] i <u>rifiuti organici</u> non comprendono la carta.
Biobarbotage	Biosparging	Le <u>biobarbotage</u> est une technologie de restauration in situ [...]	Il <u>bio-sparging</u> è una tecnologia di ripristino in situ [...]	Il <u>biobarking</u> [sic] è una tecnologia di bonifica in situ [...]	Il <u>biobarbotaggio</u> [sic] è una tecnologia di restauro in situ [...]
Écocivisme	Environmental citizenship	[...] les mesures visant à encourager la prise de conscience des problèmes environnementaux, à favoriser l' <u>écocivisme</u> [...]	[...] favorire la sensibilizzazione problematiche ambientali, favorendo l' <u>eco-cittadinanza</u> e promuovendo la gestione ambientale.	[...] le misure di promozione della consapevolezza ambientale, della <u>cittadinanza ambientale</u> e della gestione ambientale.	[...] le misure volte a promuovere la sensibilizzazione ai problemi ambientali, a favorire l' <u>ecocivismo</u> [sic] e a promuovere la gestione ambientale.
Écoservices	Ecosystem services	[...] l'importance de nombre de ces « <u>écoservices</u> » demeure souvent invisible [...]	[...] l'importanza di molti di questi " <u>eco-servizi</u> " [sic] spesso rimane invisibile, provocando la scomparsa delle specie e il degrado degli ecosistemi.	[...] l'importanza di molti di questi " <u>servizi ecosistemici</u> " rimane spesso invisibile, causando la perdita di specie e il degrado dell'ecosistema.	[...] l'importanza di molti di questi " <u>ecosistemi</u> " [sic] resta spesso invisibile, provocando la scomparsa delle specie e il degrado degli ecosistemi.
Écosensible	Ecosensitive/ Ecologically sensitive/ Environment ally sensitive	[...] analyses environnementales dans les secteurs <u>écosensibles</u> [...]	[...] nelle aree <u>sensibili dal punto di vista ambientale</u> e rafforzare le azioni penali contro i crimini ambientali.	[...] nelle aree <u>sensibili dal punto di vista ambientale</u> e di rafforzare il perseguimento dei crimini ambientali.	[...] nei settori <u>ecosociali</u> [sic] e rafforzare le azioni penali contro i crimini ambientali.
Écocité	Ecocity/ Sustainable city	En 2009, se met en place l' <u>écocité</u> Alzette-Belval [...]	Nel 2009, l' <u>ecocity</u> Alzette-Belval è stata costituita [...]	Nel 2009, l' <u>ecocity</u> Alzette-Belval è stata creata [...]	Nel 2009, si è insediato l' <u>Ecocità</u> [sic] Alzette-Belval quale ente pubblico [...]

Tableau 4 : La traduction des unités terminologiques issues de l'anglais

En ce qui concerne les deux unités monolexicales en « bio- », c'est-à-dire « biodéchets » et « biobarbotage », deux termes techniques posant plusieurs défis lors de la traduction humaine et surtout dans le cas de la TA, une remarque traductionnelle s'impose. En effet, en prenant le cas de *biodéchets* — terme calqué sur l'anglais « *biowaste* » — les traducteurs automatiques n'en ont délivré que des traductions partielles et parfois inexactes (Google Traducteur : « *rifiuti organici* » ; DeepL : « *rifiuti biologici* » ; Systran : « *rifiuti*

organici »), alors que, comme le rappelle un document de la Commission européenne de 2009⁷, l'équivalent officiel de cette unité en italien est « *ri-fiuti organici biodegradabili* », une unité polylexicale composée d'un nom et de deux adjectifs, tandis que la forme française est dérivée par le biais du préfixe « bio- ». L'autre terme hautement spécialisé dont nous nous sommes occupés est « biobarbotage », dont l'origine attestée est anglaise (*biosparging*). Là aussi, les sorties des traducteurs automatiques ne sont pas du tout satisfaisantes (Google Traducteur : « *bio-sparging* » ; DeepL : « *biobar-king* » [sic] ; Systran : « *biobarbottaggio* » [sic]), le premier du point de vue de l'orthographe en raison d'un trait d'union non nécessaire et les deux derniers parce qu'il s'agit de dénominations erronées, voire inexistantes dans le contexte sociolinguistique de l'italien. Dès lors, les systèmes de TA n'ont pas été à même de récupérer la désignation ponctuelle dont il est fait mention dans les textes techniques de la langue d'arrivée (« *biospargimento* », 91 occurrences sur Google) ni de proposer la forme anglaise correcte, souvent utilisée en tant qu'emprunt en italien (« *biosparging* », 1 560 occurrences sur Google), peut-être à cause du concept extrêmement sectoriel auquel renvoie la dénomination prise en compte. De plus, des résultats similaires ont été constatés pour certains mots en « éco- » : *ecoservices*, *écosensible* et *écocité*. Ainsi, le terme « *ecoservices* », qui provient de l'anglais « *ecosystem services* », s'est avéré difficile à gérer pour Google Traducteur et Systran, lesquels offrent des traductions erronées (« *eco-servizi* » et « *ecosistemi* », respectivement) alors que DeepL reconnaît la forme exacte du terme en italien, « *servizi ecosistemici* », peut-être en raison de la présence de cette information dans le corpus qui alimente le système de ce traducteur automatique ; encore, lors du traitement du terme « *écosensible* », issu de l'anglais « *ecosensitive* » avec ses variantes « *ecologically sensitive* » et « *environmentally sensitive* », Google Traducteur et DeepL ont offert des traductions valables (« *sensibili dal punto di vista ambientale* »), respectueuses de l'usage effectif de cette expression dans le contexte socioculturel italien, tandis que Systran a proposé une traduction totalement hors sujet (« *ecosociali* ») ; en outre, la traduction du terme récent « *écocité* », calqué sur l'anglais « *ecocity* » avec sa variante « *sustainable city* », a été un véritable échec pour tous les traducteurs automatiques car Google Traducteur et DeepL ont opté pour l'emprunt de l'anglais (*ecocity*), peu courant en italien, tandis que Systran a délivré un mot agrammatical (*Ecococità* [sic]), là où

⁷<https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=LEGISSUM:ev0009&from=SL>

l'usage habituel de l'italien aurait orienté la traduction vers le syntagme « *città ecologica* », comme bon nombre de documents de l'Union européenne l'attestent⁸. Ceci vaut également pour le terme « écocivisme », calqué sur l'expression anglaise « *environmental citizenship* », traduit par les outils de TA en tenant compte des usages réels de l'italien pour ce qui concerne Google Traducteur (« *eco-cittadinanza* ») et DeepL (« *cittadinanza ambientale* »), tandis qu'il est mal traduit par Systran, qui a opté pour une dénomination qui reprend à la lettre la structure morphologique de la dénomination française.

Conclusion

Comme les étudiant(e)s l'ont bien relevé durant l'atelier de traduction, la TA peut se révéler très utile dans la traduction des termes ou de portions de textes, notamment des expressions bien établies dans la langue, mais pose parfois des problèmes de traduction dans des contextes spécialisés. Grâce à une analyse interprétative des données métalinguistiques et traductionnelles relatives à certaines unités terminologiques jugées complexes, les étudiant(e)s ont mis en lumière, d'une part, la capacité des outils de TA à gérer au mieux les unités terminologiques qui constituent des dénominations attestées et validées par des institutions et des organisations telles que la Commission européenne ou l'Unesco, et d'autre part, les limites de ces instruments dans le cas des néologismes ou des termes culturellement connotés, qui n'ont pas la même saillance selon les différentes langues. Bien que les résultats des traducteurs automatiques puissent varier en fonction des termes sélectionnés, il est évident que des noyaux durs, comme ceux que nous avons décelés, apparaîtront. Ce sont ces traits communs des sorties problématiques qui peuvent aider les développeurs des outils de TA à améliorer leur fonctionnement, en intégrant des méthodes d'analyse plus proches de la réalité socioculturelle et discursive dans laquelle les termes sont employés.

⁸ https://europa.eu/learning-corner/what-are-the-eus-energy-sources_it

Bibliographie

Drouin Patrick, *Termostat. Logiciel d'extraction terminologique*, <http://termostat.ling.umontreal.ca/interfaceTermostat.php>

Gaudin François (1993), *Pour une socioterminologie : des problèmes sémantiques aux pratiques institutionnelles*. Mont-Saint-Aignan : Publications de l'université de Rouen.

Guidère Mathieu (2008). *La Communication multilingue : traduction commerciale et institutionnelle*. Bruxelles : De Boeck Université.

Humbley John (2018). *La néologie terminologique*, préface de Rita Temmerman. Limoges : Lambert-Lucas.

Lavault-Olléon Elisabeth (éd.) (2007). *Traduction spécialisée : pratiques, théories, formations*. Bern : Peter Lang.

Patrick Drouin, *Termostat. Logiciel d'extraction terminologique*, <http://termostat.ling.umontreal.ca/interfaceTermostat.php>

Poibeau Thierry (2019). *Babel 2.0 : Où va la traduction automatique ?* Paris : Odile Jacob.

Raus Rachele (2013). *La terminologie multilingue. La traduction des termes de l'égalité H/F dans le discours international*. Bruxelles : De Boeck Université.

Sources pour les travaux en classe

La date de dernière consultation de toutes les références sitographiques est le 12.07.2021.

AA.VV. (2008). « Gestione dei rifiuti organici biodegradabili nell'Unione europea ». Dans : *EUR-Lex Access to European Union Law*, <https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=LEGISSUM:ev0009&from=SL>

AA.VV. (2009). « Gestion des biodéchets dans l'Union Européenne ». Dans : *EUR-Lex Access to European Union Law*, <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=LEGISSUM%3Aev0009>

AA.VV. (2009). « Renseignements communiqués par les États membres sur les aides d'État accordées conformément au règlement (CE) no 1857/2006 de la Commission concernant l'application des articles 87 et 88 du traité aux aides d'État accordées aux petites et moyennes entreprises actives dans la production de produits agricoles et modifiant le règlement (CE) no 70/2001 ». Dans : *Journal officiel de l'Union européenne*, <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:075:0028:0030:FR:PDF>

AA.VV. (2010). « Étude sur les écosystèmes et la biodiversité : rendre les ressources naturelles visibles à l'économie mondiale ». Dans : *CORDIS Commission européenne*, <https://cordis.europa.eu/article/id/32681-teeb-study-make-natural-resources-visible-to-global-economy/fr>

- AA.VV. (2021). « Océan : changeons de cap ! ». Dans : *Le Courrier de l'UNESCO*. janvier-mars 2021, <https://www.ecoledecarlepont.fr/attachment/2093148/>
- Cossardeaux Joël (2014). « Déchets électriques : les volumes collectés doivent doubler d'ici à 2019 », *Les Échos*. URL : <https://www.lesechos.fr/2014/08/dechets-electriques-les-volumes-collectes-doivent-doubler-dici-a-2019-308463>
- Duarte Carlos, Atwood Trisha *et alii* (2021). « Le patrimoine mondial marin de l'UNESCO : gardien des réserves mondiales de carbon bleu ». Dans : *UNESDOC Digital Library*, https://unesdoc.unesco.org/ark:/48223/pf0000375565_fre
- Lassailly-Jacob Véronique, Desse Michel (2009). *Migrations et vulnérabilités. Éleveurs sahéliens confrontés aux sécheresses et populations des littoraux antillais face aux cyclones*, https://www.lmd.jussieu.fr/~lshan/petites_antilles/FR/Changeement%20Climatique%20Lassailly-Jacob%20Desse.pdf
- Molga Paul (2014). « Le 'low-tech', nouvelle source d'innovation », *Les Échos*. URL : <https://www.lesechos.fr/2014/10/le-low-tech-nouvelle-source-dinnovation-1104246>
- Samson Réjean, Bage Gontran (2004). « Développement d'un outil de gestion intégrant les risques économiques pour la réhabilitation des sites contaminés ». Dans : *Rapport de mai 2004*, https://inis.iaea.org/collection/NCLCollectionStore/_Public/47/106/47106358.pdf
- Santos Isabel (2020). « PROJET DE RAPPORT sur les rapports 2019-2020 de la Commission concernant l'Albanie (2019/2170(INI)) ». Dans : *Rapports de la Commission Européenne*, https://www.europarl.europa.eu/doceo/document/AFET-PR-647074_FR.pdf
- Tu Ngo Anh (2014). *Évaluation environnementale du risque d'inondation dans le delta du fleuve Ha Thanh (centre Viêt-Nam)*. Thèse de doctorat en géographie : Université d'Orléans, https://tel.archives-ouvertes.fr/file/index/docid/1004411/filename/anhtu.ngo_3538.pdf
- Unesco - Organisation des Nations Unies pour l'éducation, la science et la culture, <http://www.unesco.org>
- Wassenberg Birte, Reitel Bernard (2015). *La coopération territoriale en Europe, une perspective historique*. Bruxelles : Office des publications de l'Union Européenne, https://ec.europa.eu/regional_policy/sources/information/pdf/brochures/interreg_25years_fr.pdf

Les genres textuels

La traduzione automatica neurale: uno strumento di sensibilizzazione per la formazione universitaria in lingua e traduzione francese

Maria Margherita Mattioda, Ilaria Cennamo

Introduzione

L'esperienza didattica oggetto del presente articolo ha coinvolto gli studenti e le studentesse iscritti e iscritte al corso di Lingua francese della terza annualità del corso di studio triennale in Scienze della mediazione linguistica dell'Università degli Studi di Torino.

I dati raccolti durante le lezioni sperimentali e attraverso la somministrazione di due questionari (a inizio e a fine corso) hanno evidenziato come il ricorso alla traduzione automatica (TA) contribuisca a sensibilizzare l'apprendente sia rispetto alle difficoltà traduttive poste dai testi oggetto della traduzione sia rispetto alle criticità legate alle prestazioni traduttive automatiche. In particolare, i risultati finora ottenuti permettono di gettare le basi per una riflessione didattica più ampia, volta all'elaborazione di una programmazione incentrata su una progressione di obiettivi formativi specifici per ogni annualità, nell'ottica di una continuità, dal corso di studio triennale al corso di studio magistrale.

La sperimentazione didattica è stata concepita alla luce dei contributi traduttologici incentrati sullo studio dell'interazione tra l'apprendente traduttore e le tecnologie per la traduzione (O'Brien 2020; Schumacher 2019; Ehrensberger-Dow, Massey 2019; Martikainen 2019; Lavault-Oléon, Zimina 2019; Massey, Kiraly 2019; Kenny 2020; Looock 2019) che sottolineano l'interesse legato all'adozione di un approccio riflessivo in cui l'apprendente assume un ruolo attivo e consapevole nel processo di apprendimento:

Maria Margherita Mattioda, Università di Torino, marita.mattioda@unito.it

Ilaria Cennamo, Università di Torino, ilaria.cennamo@unito.it

Au-delà de la connaissance, c'est une approche raisonnée prenant en compte les enjeux qui nous semble particulièrement importante¹ : que sait faire la machine? Que ne sait-elle pas faire? Quels sont les risques? [...] Pour cela, [...] il est crucial d'avoir conscience que la performance des outils de TA est en lien direct avec la qualité des données rassemblées en corpus qu'ils exploitent. Les étudiants pourront ainsi déterminer si leur utilisation est pertinente². (Loock 2019: 56)

La definizione di buone pratiche didattiche incentrate sull'adozione della traduzione automatica come ausilio all'apprendimento della traduzione, e non necessariamente ai fini di una formazione professionalizzante, resta la principale sfida posta dalle applicazioni nate nell'era dell'intelligenza artificiale (IA) all'attenzione dei formatori di "lingua e traduzione" dell'Università italiana.

Il dibattito sulle modalità e sulle tempistiche legate all'integrazione delle tecnologie per la traduzione resta aperto e, a questo proposito, si considera interessante concepire questa integrazione secondo un *continuum*, fin dal primo anno del corso di studio, in linea con quanto esposto da Loock (2019).

Questa riflessione di natura didattica necessita di una revisione approfondita e condivisa dei programmi di insegnamento, così come emerso da altre indagini recenti (Rossi 2019: 101) che si riferiscono a questa evoluzione didattica come a una sinergia (Hernandez Morin 2019: 250-252) e a un dialogo (Massey, Kiraly 2019).

Il presente contributo comprende tre sezioni principali (e relativi sottoparagrafi): nel paragrafo 1, viene descritta la sperimentazione didattica condotta; nel successivo, vengono presentati i risultati ottenuti nel corso della prima fase dell'indagine, l'analisi dei dati preliminari (Questionario 1); e, infine, il paragrafo 3 è dedicato all'analisi dei dati di fine corso (Questionario 2).

1. La sperimentazione didattica

La sperimentazione didattica descritta è stata prevista come parte dell'insegnamento di Lingua francese che rientra nell'ambito della formazione linguistica del corso di studio triennale in Scienze della mediazione linguistica (L-12). Il corso di studio è finalizzato all'acquisizione di competenze linguistiche di livello avanzato (B2-C1, secondo i descrittori del QCER³)

¹ Nostra sottolineatura.

² Nostra sottolineatura.

³ Quadro comune europeo di riferimento per le lingue: <https://rm.coe.int/168045bc72> (consultato il 23 luglio 2021).

che sono applicate alla pratica della traduzione e della mediazione linguistica e culturale. Come si evince dal Regolamento del corso di studio L-12, tra gli obiettivi formativi qualificanti si segnala in particolare l'acquisizione di:

competenze linguistico-tecniche orali e scritte sorrette da adeguato inquadramento metalinguistico; [...] specifiche conoscenze relative alla struttura delle lingue naturali e una adeguata formazione di base nei metodi di analisi linguistica; essere in grado di utilizzare gli strumenti per la comunicazione e la gestione dell'informazione⁴.

I testi selezionati nell'ambito della sperimentazione sono rappresentativi dei contesti economico-istituzionali, conformemente a quanto previsto dal Regolamento del corso di studio L-12. Oltre al consolidamento di conoscenze e competenze propriamente linguistiche, l'insegnamento considerato mira perciò all'acquisizione di competenze traduttive applicate ai diversi contesti della mediazione (Leoncini Bartoli 2019; Cordonnier 1995; Lederer 2006). Nello specifico, la traduzione viene intesa come processo strategico e decisionale (Amparo Hurtado *et alii* 2020) che permette lo sviluppo di una riflessione critica fondata sull'attivazione di macro e micro strategie messe in atto sulla base del progetto traduttivo (Gambier 2008). In quest'ottica, l'integrazione di risorse linguistiche e tecnologiche (dizionari cartacei e digitali, banche dati terminologiche, *corpora*, traduttori automatici) viene concepita in modo funzionale alla realizzazione dei singoli progetti di traduzione assegnati. Per il raggiungimento di tali obiettivi il corso di Lingua francese unisce alla parte teorica, traduttologica, esercizi di traduzione finalizzati al consolidamento e all'acquisizione di competenze linguistiche, pragmatiche e testuali relativamente alla mediazione scritta. In tal modo si intende favorire negli e nelle apprendenti lo sviluppo di un'autonomia di giudizio fondata su capacità analitiche e decisionali volte ad operare scelte comunicative, traduttive e redazionali consapevoli.

La sperimentazione didattica condotta si pone come obiettivo specifico quello di integrare nel processo di apprendimento della traduzione (previsto dal corso di Lingua francese), l'uso consapevole delle risorse di traduzione automatica allo scopo di sensibilizzare l'apprendente rispetto alle criticità poste da questi strumenti. La durata della sperimentazione è stata di 54

⁴https://www.lingue.unito.it/do/documenti.pl/ShowFile?_id=8506;field=file;key=jtJNb9OxLJ60FYA1drj8sPciDwWNUBa6SUqRDUhKtaQcEhdITUYFCUUEM50AeQLIR5Bl4B1L36kJaRWUQIcAAB2Lk6QgMBk1ztzbA1nuIA;t=2390 (consultato il 23 luglio 2021).

ore totali, suddivise in incontri da 3 ore che si sono svolti nel secondo semestre (A.A. 2020-2021), da inizio febbraio a fine aprile 2021. Sono stati presi in esame due generi discorsivi (Maingueneau 2004) differenti: il discorso giornalistico (Krieg-Planque 2000) e il discorso promozionale (Charaudeau 2009). Le attività didattiche sono state svolte secondo questo ordine: in una prima fase, le e i partecipanti hanno tradotto il testo all'interno di un documento Word, avvalendosi di dizionari cartacei o digitali e consultando Internet. Successivamente, hanno tradotto il testo automaticamente ricorrendo a Google Traduttore⁵. La terza ed ultima fase è stata incentrata su un'attività di categorizzazione degli errori di traduzione individuati all'interno della traduzione automatica di Google, con l'ausilio di una griglia apposita (Par. 1.3). La categorizzazione degli errori ha previsto, infine, l'elaborazione di una proposta di correzione finalizzata a una post-edizione "ragionata" del testo tradotto, in linea con altre recenti riflessioni d'interesse didattico, incentrate sulla pratica traduttiva (Bowker, Jairo Buitrago 2019; Cennamo 2018; Peraldi 2014; Hernandez Morin 2019; Rossi 2019).

1.1 TA e discorso giornalistico

Nell'ambito del corso di Lingua francese è stata proposta al gruppo di apprendenti una riflessione sulla natura del testo economico di carattere informativo-divulgativo secondo una prospettiva traduttologica. Le lezioni frontali sono state affiancate da esercitazioni pratiche incentrate sullo sviluppo delle competenze linguistiche e traduttive dal francese all'italiano a partire da testi giornalistici selezionati nei siti web dei principali organi di stampa francesi. L'analisi del discorso dell'informazione mediatica (Charaudeau 1997) ha riguardato in particolare l'attualità economica diffusa sia attraverso il canale della stampa scritta, sia attraverso il canale audiovisivo. In particolare, si è scelto di non affrontare la pluralità delle tematiche socio-economiche presenti nel dispositivo mediatico e di creare un corpus di lavoro incentrato sull'economia circolare⁶ per meglio delimi-

⁵<https://translate.google.com/?hl=fr&sl=fr&tl=it&op=translate> (consultato il 24 luglio 2021).

⁶ "L'economia circolare è un modello di produzione e consumo che implica condivisione, prestito, riutilizzo, riparazione, ricondizionamento e riciclo dei materiali e prodotti esistenti il più a lungo possibile. In questo modo si estende il ciclo di vita dei prodotti, contribuendo a ridurre i rifiuti al minimo". <https://www.europarl.europa.eu/news/it/headlines/economy/20151201STO05603/economia-circolare-definizione-importanza-e-vantaggi> (consultato il 14 settembre 2021).

tare l'ambito di documentazione utile al processo traduttivo, per una maggiore coerenza dei contenuti e per una osservazione più puntuale dei fenomeni linguistici e testuali caratterizzanti. Ai fini della sperimentazione didattica, sono stati forniti in un primo tempo alcuni articoli di stampa di carattere introduttivo (unità testuali specifiche dell'articolo giornalistico); in un secondo momento, sono stati introdotti documenti audio e audiovisivi presenti sui siti web di alcune reti di informazione francesi (*France culture, France Inter, France 24*), appartenenti a vari generi giornalistici (intervista, reportage) e rispondenti prevalentemente ad una funzione informativo-esplicativa (Adam 2005; Charaudeau 1997, 2006; Krieg-Planque 2000). La scelta si è orientata su testi audio e video relativi ad un ambito specialistico di recente diffusione caratterizzato da una terminologia specifica per designare con precisione concetti tecnici in cui si intersecano economia, ecologia e sostenibilità (Mattioda 2019). Il passaggio dal discorso economico ad alto grado di specializzazione al discorso di divulgazione giornalistica ha permesso di osservare fenomeni di banalizzazione (Moirand 2007) particolarmente interessanti (riformulazioni, definizioni, esplicitazioni, ...) e di definire strategie di traduzione atte a rendere conto di un discorso ibrido in cui sono imbricate tracce dell'enunciazione formale legata all'*expertise* in un'enunciazione divulgativa, a tratti didattica, propria dei media il cui pubblico risulta ampio e generalista (Charaudeau 1997: 78-80). L'approccio a questo tipo di testi non è risultato sempre immediato da parte di studenti e studentesse in formazione poiché una limitata conoscenza del settore e del suo lessico specialistico ha spesso richiesto approfondimenti e ricerche puntuali. Tra gli aspetti caratterizzanti possono essere segnalati:

- a livello lessicale: la terminologia economica di base e il lessico della "circolarità", neologismi, prestiti e calchi, sigle e acronimi, fraseologismi;
- a livello semantico: anisomorfismi⁷, slittamenti semantici, linguaggio metaforico;
- a livello pragmatico: funzione del testo di partenza, aspetti socioculturali del contesto di arrivo.

⁷ Ossia il *décalage* semantico e concettuale tra lessemi di diverse lingue. In traduzione implica un processo di ricostruzione dell'equivalenza interlinguistica e interculturale che definisce il perimetro mutevole della traducibilità.

Per gli e le apprendenti, che hanno dovuto affrontare la doppia sfida della comprensione del testo scritto e orale (col supporto di una trascrizione parziale) e del transfert dall'orale (pseudo-oralità) allo scritto, si è dunque trattato di produrre un testo di arrivo terminologicamente corretto e pragmaticamente funzionale allo scopo comunicativo prefissato (Reiss, Vermeer 2013) anche in base alla commutazione del codice in uso nel testo di partenza, per poi confrontarlo con la traduzione automatica.

1.2. TA e discorso promozionale

Il corso ha previsto quattro incontri dedicati all'analisi, alla traduzione e alla post-edizione "ragionata"⁸ di testi pragmatici (Delisle 1980: 22) caratterizzati da una finalità promozionale dai tratti divergenti (Charau-deau 2009). L'interesse legato a questi testi riguarda, *in primis*, i tratti linguistici comuni.

Sul piano lessicale, sono caratterizzati dall'adozione di un lessico comune che comprende alcuni vocaboli di specialità inerenti alla tematica trattata: il contesto museale storico, per quanto riguarda il primo testo, e l'ambito dei blog di moda, per quanto riguarda il secondo.

Il registro standard è tendenzialmente informale in entrambi i testi e contribuisce alla costruzione di una postura comunicativa diretta (Rabatel 2002; Vion 1992) nei confronti del lettore (utente del web).

Sul piano sintattico, inoltre, la paratassi accomuna i due testi, costituiti da brevi frasi, separate da una punteggiatura essenziale che, oltre ad assicurare la linearità della progressione dell'informazione, contribuisce alla produzione di un discorso la cui lettura rievoca l'oralità (Stark 2015; Marcocchia, Gauducheau 2007).

Al contempo, i tratti linguistici che contraddistinguono ogni testo inducono l'apprendente a orientare l'approccio traduttivo in modo funzionale (Nord 2018) rispetto a tali specificità.

Più precisamente, i tratti distintivi riguardano l'enunciazione della finalità promozionale. Il primo testo⁹ è rappresentativo dell'azione, istituzionale e promozionale, attribuita alle notizie di attualità diffuse attraverso

⁸ Attuata mediante il ricorso alla griglia per la classificazione degli errori rilevati nella traduzione automatica (Par. 1.3).

⁹ Estratto dall'articolo pubblicato sul sito web istituzionale della *Ville de Paris*: <http://www.paris.fr/actualites/le-musee-carnavalet-se-refait-une-beaute-4113> (consultato il 2 luglio 2021).

so il web da parte dell'istituzione pubblica interessata, la *Ville de Paris*. Il comune di Parigi informa la cittadinanza in merito alle iniziative e ai progetti avviati dall'amministrazione e, allo stesso tempo, intende valorizzare l'attrattività dell'offerta culturale parigina agli occhi dei suoi cittadini e dei potenziali turisti (utenti del web), nell'intento prettamente politico di mantenere o alimentare il consenso. In questo senso, la finalità promozionale del testo si colloca nella "*scène générique*" (Maingueneau 2016: 84) del discorso istituzionale (Krieg-Planque 2012) diffuso attraverso il web, concepito nell'interesse collettivo della cittadinanza.

Il secondo testo¹⁰ è stato estratto da un articolo pubblicato dalla rivista *online Journal des femmes* allo scopo di presentare alle lettrici (amanti della moda) le blogger del momento: in questo contesto, la finalità promozionale si realizza in modo più esplicito rispetto al testo 1, in ragione dell'identità commerciale della rivista.

Quest'ultima, infatti, si rivolge a un profilo di destinatari mirato, a fini persuasivi tali da attribuire al testo considerato i tratti identitari di un discorso pubblicitario (Charaudeau 2009: 4). Il ricorso a vocaboli di specialità (appartenenti all'universo della moda) contribuisce, infine, alla realizzazione della strategia discorsiva volta al coinvolgimento del consumatore, *target* dell'operazione pubblicitaria. Date le caratteristiche testuali descritte, la sfida traduttiva posta sia alla traduzione automatica sia all'attenzione delle e degli apprendenti consiste nella selezione di traduttori corretti e funzionali allo *skopos* (Reiss, Vermeer 2013) del singolo testo.

1.3. Categorizzazione degli errori di traduzione

Le attività didattiche incentrate sulla classificazione degli errori individuati all'interno delle traduzioni automatiche hanno previsto l'adozione di una griglia, concepita al fine di favorire lo sviluppo di un approccio riflessivo da parte dell'apprendente. Tale strumento permette, inoltre, di accedere all'analisi metalinguistica e metatraduttiva che sottende le scelte traduttive dell'apprendente in fase di post-edizione.

Nello specifico, si tratta di un adattamento della "*grille multicritères pour l'évaluation de la qualité en traduction spécialisée*" elaborata da Daniel

¹⁰ Estratto dall'articolo pubblicato dalla rivista online *Journal des femmes*, all'interno della rubrica di moda: <http://www.journaldesfemmes.com/mode/magazine/les-blogueuses-mode-a-suivre-pour-rester-a-la-page/les-blogueuses-mode-a-suivre-et-pourquoi-pas-coline.shtml> (consultato il 2 luglio 2021).

Toudic *et alii* (2014: 12). L'adattamento proposto ha lo scopo di rendere la griglia accessibile al profilo dei e delle partecipanti.

In effetti, l'obiettivo della nostra esperienza didattica non è stato quello di valutare la qualità delle prestazioni traduttive automatiche, bensì quello di studiare le criticità e il potenziale legati all'integrazione della traduzione automatica in questo preciso contesto di formazione. La griglia è stata proposta, dunque, come ausilio a una post-edizione ragionata¹¹.

Testo di partenza	Traduzione automatica	Tipo di errore	Correzione	Commento

Tabella 1:
La griglia: ausilio alla post-edizione

L'esercizio richiesto consiste nell'inserire all'interno della griglia, in corrispondenza del segmento di partenza, le traduzioni automatiche errate (o intese come tali da parte dell'apprendente). Occorre associare l'errore a una categoria linguistica oppure a una combinazione di categorie linguistiche, tra le sette indicate: ortografia; omissione/ aggiunta non pertinente; registro; senso; lessico; terminologia; grammatica o sintassi. Tale attività di classificazione induce l'apprendente a interrogarsi non solo sulla natura linguistica dell'errore, ma anche sul problema traduttivo conseguente (nel testo di arrivo). Un calco sintattico prodotto all'interno di un testo promozionale, ad esempio, può avere un impatto sulla fluidità discorsiva del testo di arrivo oppure sulla comprensibilità dei contenuti, o ancora sulla mancata equivalenza semantica rispetto al testo di partenza. Al fine di favorire questo tipo di riflessione metatraduttiva, viene richiesto allo studente o alla studentessa di correggere l'errore individuato e di inserire un "commento" per spiegare la natura dell'intervento di post-edizione (cf. Tab. 1).

1.4. I riscontri dei e delle partecipanti

Le studentesse e gli studenti che hanno partecipato alla sperimentazione didattica hanno preparato una sintesi degli esercizi svolti durante il corso che sono stati in parte presentati durante un workshop organizzato il 24 aprile 2021.

¹¹ In conformità con la concezione della celebre *Traduction raisonnée* di Jean Delisle (2013).

In occasione del seminario, la prima parte della comunicazione è stata incentrata sull'analisi e sulla categorizzazione degli errori di traduzione automatica osservati a lezione. La seconda parte è stata finalizzata, invece, alla presentazione delle percezioni sull'efficacia della traduzione automatica applicata ai diversi generi discorsivi oggetto degli esercizi di traduzione previsti durante il corso¹².

La sintesi presentata evidenzia alcune tipologie di errore considerate imputabili al funzionamento del traduttore automatico utilizzato: la traduzione non sempre pertinente degli anglicismi determinata dall'interferenza dell'inglese come lingua ponte in coppie di lingue che non comprendono l'inglese, l'approccio traduttivo letterale e la scarsa qualità redazionale del testo di arrivo risultante dalla traduzione automatica. Inoltre, alcuni esempi di categorizzazione degli errori derivanti dalla traduzione automatica mostrano la necessità di un intervento di post-edizione su più livelli (lessicale, terminologico, sintattico e pragmatico) volto sia all'effettiva correzione degli errori sia al miglioramento della qualità redazionale in lingua d'arrivo. In ultimo, si può rilevare una considerazione particolarmente significativa: gli errori vengono compresi e dunque corretti dai e dalle partecipanti solo dopo una fase di lettura attenta (e guidata) in quanto, spesso, a una prima lettura il testo di arrivo appare coeso sul piano testuale¹³. Un dato che conferma l'interesse pedagogico legato a un'azione di sensibilizzazione del pubblico di apprendenti rispetto all'utilizzo della traduzione automatica.

2. Analisi dei dati preliminari (Questionario 1)

In questa sezione verranno presentati i dati raccolti a seguito della somministrazione del primo questionario, avvenuta all'inizio del corso. Di seguito, sono riportati i dati ritenuti significativi rispetto all'obiettivo specifico dell'indagine condotta nel corso della sperimentazione didattica descritta.

¹² Nella sezione "Allegati" è consultabile un riassunto delle due parti costitutive della comunicazione presentata.

¹³ In linea con quanto emerso relativamente all'adozione di Google Traduttore, Eisele (2018: 7) ha dimostrato come la traduzione automatica neurale, nel contesto dell'Unione europea, fornisca *output* più scorrevoli a una prima lettura, ma non necessariamente più corretti sul piano interlinguistico rispetto alle prestazioni traduttive statistiche della generazione precedente.

2.1 La profilatura delle e dei partecipanti

Il questionario 1 è stato somministrato, all'inizio del corso, a 43 studenti e studentesse del corso di Lingua francese indicato in precedenza (Par. 1), di cui 42 apprendenti di genere femminile e 7 di madrelingua diversa dall'italiano oppure bilingui¹⁴.

In particolare, tenuto conto delle quattro principali abilità linguistiche (di comprensione e produzione scritta, e di comprensione e produzione orale), la maggioranza dei e delle rispondenti dichiara di possedere un ottimo livello di competenze in lingua italiana (Fig. 1)¹⁵. Si registra un unico riscontro negativo ("livello basso di produzione orale") che è stato fornito dall'unica partecipante di madrelingua francese.

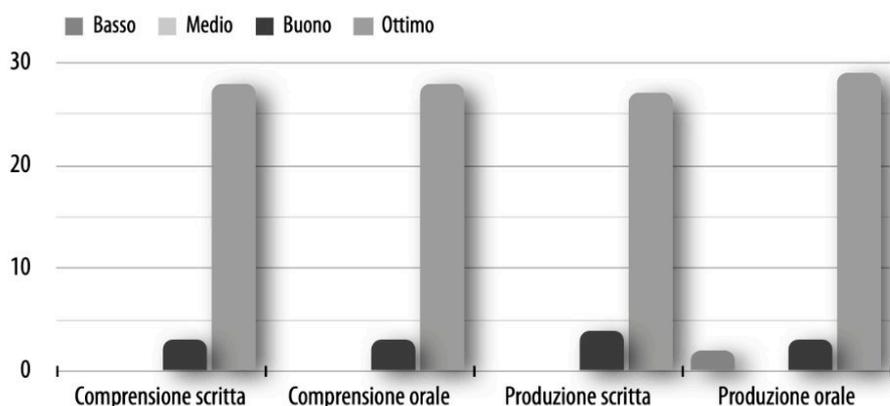


Figura 1: Competenze linguistiche in lingua italiana

In riferimento alle competenze in lingua francese (Fig. 2), dal primo questionario emerge che la maggioranza dei e delle rispondenti ritiene di possedere "buone" competenze in lingua francese¹⁶.

¹⁴ Riportiamo di seguito le 7 dichiarazioni menzionate: bilingue italiano/ arabo; di madrelingua cinese; di madrelingua polacca; di madrelingua rumena; bilingue rumeno/ italiano; di madrelingua francese; bilingue italiano/ inglese.

¹⁵ Riportiamo di seguito le percentuali corrispondenti. Per la comprensione scritta e orale, il 9,37% dichiara di possedere un livello buono; il 90,62% dichiara di possedere un ottimo livello. Per la produzione scritta, il 12,5% dichiara di possedere un livello buono; l'87,5% dichiara di possedere un ottimo livello. Per la produzione orale, il 3,12% dichiara di possedere un livello basso; il 6,25% dichiara di possedere un livello buono; il 90,62% dichiara di possedere un ottimo livello.

¹⁶ Come si evince dal grafico in figura 2, per quanto concerne la comprensione scritta si rilevano le percentuali seguenti: il 4,65% dichiara di possedere un livello medio; il 60,46% dichiara di possedere un livello buono; il 34,88% dichiara di possedere un ottimo

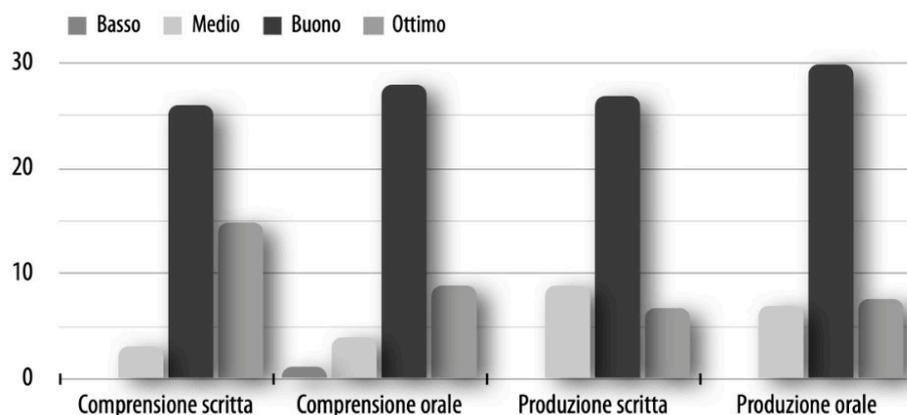


Figura 2: Competenze linguistiche in lingua francese

Sulla base dei dati raccolti in questa prima fase, le competenze linguistiche dichiarate in riferimento alle due lingue coinvolte nell'esercizio traduttivo (oggetto della sperimentazione didattica descritta dal paragrafo 1 al paragrafo 1.4) risultano essere omogenee e coerenti rispetto all'annualità considerata.

2.2 Le conoscenze e competenze tecnologiche progresse

Il primo questionario ha permesso di raccogliere informazioni inerenti all'adozione delle tecnologie. In generale, la maggioranza si dichiara interessata allo sviluppo dell'IA¹⁷ e ritiene che essa presenti un'utilità rispetto all'evolversi della formazione in lingue¹⁸ nonché un'importanza significativa per il futuro dell'industria delle lingue¹⁹. In particolare, si rileva che,

livello. Per quanto riguarda la comprensione orale: il 2,32% dichiara di possedere un livello basso; il 9,30% dichiara di possedere un livello medio; il 67,44% dichiara di possedere un livello buono; il 20,93% dichiara di possedere un ottimo livello. In riferimento alla produzione scritta: il 20,93% dichiara di possedere un livello medio; il 65,11% dichiara di possedere un livello buono; il 13,95% dichiara di possedere un ottimo livello. Infine, i riscontri sulla produzione orale evidenziano che: il 13,95% dichiara di possedere un livello medio; il 69,76% dichiara di possedere un livello buono; il 16,27% dichiara di possedere un ottimo livello.

¹⁷ Al quesito "Le interessa lo sviluppo delle tecnologie e dell'intelligenza artificiale?", il 58,1% ha risposto "più sì che no", il 30,2% "decisamente sì" e l'11,6% "più no che sì".

¹⁸ Al quesito "Quanto pensa che potrà essere utile nel prossimo futuro l'intelligenza artificiale nel suo campo di studio?", il 90,7% ha risposto "molto utile" mentre solo il 9,3% ha risposto "poco utile".

¹⁹ Al quesito "Pensa che l'intelligenza artificiale assumerà un ruolo importante nel suo futuro professionale?", il 34,9% ha risposto "decisamente sì", il 55,8% "più sì che no" e solo il 9,3% "più no che sì".

secondo le e i partecipanti, l'IA applicata alle lingue permette di velocizzare le operazioni di ricerca linguistica e interlinguistica²⁰ e rappresenta un ausilio all'apprendimento delle lingue: al quesito a risposta aperta "Scriva un aspetto positivo che potrebbe caratterizzare l'uso dell'intelligenza artificiale relativamente alle lingue (materna, straniera)", il 27,90% sottolinea questa specifica utilità per l'apprendimento di una lingua.

È interessante notare che, sulla base delle risposte raccolte, sebbene tra le risorse linguistiche disponibili in rete si parli anzitutto di Reverso²¹, la maggioranza dichiara di non conoscere programmi per la gestione della terminologia²² e che chi dichiara invece di conoscerli, cita come esempio dizionari *online* tra cui WordReference, il dizionario monolingue Larousse e il dizionario monolingue Cambridge, nonché Reverso. Inoltre, la totalità dei e delle rispondenti afferma di non conoscere programmi per l'allineamento tra testi di partenza e testi di arrivo o programmi per l'analisi delle concordanze²³.

Si segnala, infine, che il controllo delle fonti di provenienza dei dati non risulta essere una pratica di ricerca linguistica o interlinguistica consolidata²⁴. Al contempo, emerge un dato significativo: il 100% delle e dei rispondenti attribuisce alla contestualizzazione del dato linguistico un ruolo determinante per la selezione dell'equivalente in lingua straniera²⁵.

Si osserva dunque che il gruppo di rispondenti al questionario di inizio corso è interessato alla problematica delle nuove tecnologie nel campo delle lingue e della traduzione, ma che al contempo tale interesse non risulta essere accompagnato da una specifica formazione all'uso di questi strumenti

²⁰ Al quesito a risposta aperta "Scriva un aspetto positivo che potrebbe caratterizzare l'uso dell'intelligenza artificiale relativamente alle lingue (materna, straniera)", il 53,48% dei riscontri mette in luce questo aspetto positivo.

²¹ Menzionata nel 94,87% delle risposte al quesito: "Generalmente si serve di traduttori automatici gratuiti o a pagamento (es. Google Translate, Reverso, DeepL...)? Se sì, specificare quali strumenti?".

²² Al quesito "Conosce programmi per la gestione della terminologia?", il 46,5% risponde "decisamente no" e il 41,9% "più no che sì".

²³ Al quesito "Conosce programmi per l'allineamento e la concordanza tra i documenti partenza e quelli di arrivo (es. WordSmith Tools, Lingua MultiConcord, Trados WinAlign...)?", l'83,7% risponde "decisamente no" e il 16,3% "più no che sì".

²⁴ Al quesito "Quando utilizza piattaforme bilingui e/ o multilingui (es. Linguee, Reverso), controlla la fonte da cui è stato estratto il risultato?", il 26,2% risponde "mai", il 59,5% "solo a volte" mentre solo il 14,3% "sì, sempre".

²⁵ Al quesito "Quando utilizza piattaforme bilingui e/ o multilingui, qual è il suo criterio per scegliere il termine equivalente in lingua straniera?", il 100% risponde "Penso al contesto di utilizzo e poi scelgo il termine".

(traduzione assistita e automatica) come ausilio all'apprendimento delle lingue e della traduzione.

2.3 Le percezioni iniziali rispetto all'intelligenza artificiale e alle tecnologie per la traduzione

Il primo questionario ha incluso alcuni quesiti riguardanti le percezioni delle e dei partecipanti rispetto alle prestazioni traduttive automatiche. Ai fini dell'analisi dei risultati ottenuti nel corso dell'esperienza didattica, vengono qui presentati i dati che emergono dalle risposte a tre quesiti mirati (Fig. 3, 4, 5):

“Pensa che i risultati forniti dai traduttori automatici siano:”

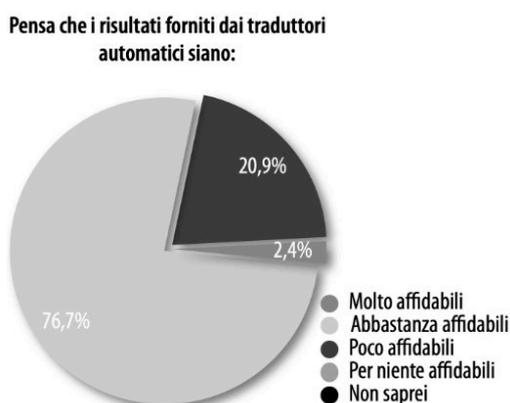


Figura 3: Affidabilità dei traduttori automatici

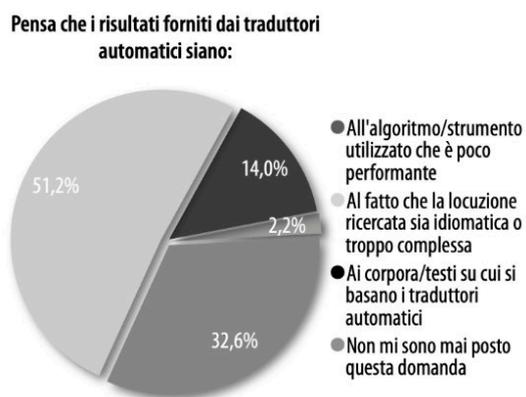


Figura 4: Le cause di errore in TA

Il grafico in figura 3 mostra che se è vero che soltanto una minoranza pari al 2,4% dei e delle partecipanti considera che le prestazioni traduttive automatiche siano “molto affidabili”, l'affidabilità dei traduttori automatici viene valutata negativamente (“poco affidabili”) solo dal 20,9% dei e delle rispondenti. Infatti, la maggioranza (del 76,7%) ritiene siano “abbastanza affidabili” (un giudizio che viene inteso come positivo nella legenda di valori prevista).

“Quando il risultato che ottiene dal traduttore automatico o dalla piattaforma non le sembra corretto, pensa che l'errore sia dovuto:”

Nel grafico in figura 4 si osserva, inoltre, che gli errori di traduzione vengono attribuiti da una maggioranza (del 51,2%) alla complessità del dato linguistico oggetto della traduzione automatica. Una percentuale significativa (del

32,6%) invece associa l'errore al livello di accuratezza dell'algoritmo di traduzione, mentre una minoranza del 14% individua nei corpora interni al sistema, la principale causa di errore.

Ha mai riscontrato delle differenze nel trattamento di traduzione a seconda della tipologia di lingua (lingue romanze, germaniche ecc.)?

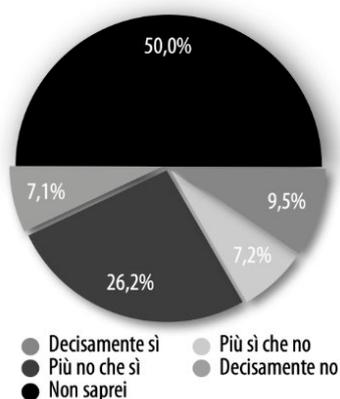


Figura 5: La correlazione tra lingue e prestazione TA

“Ha mai riscontrato delle differenze nel trattamento di traduzione a seconda della tipologia di lingua (lingue romanze, germaniche ecc.)? Se sì, quali?”

Quanto alla correlazione tra la coppia di lingue in traduzione e la prestazione del traduttore automatico neurale, l'83,3%²⁶ converge nella mancata rilevazione di dati in tal senso. Infine, si osserva che alla seconda parte del quesito (“se sì, quali?”²⁷) viene fornita soltanto una risposta esaustiva²⁸ che riguarda le lingue slave:

Lingue slave: difficoltà nel tradurre correttamente i casi di sostantivi e aggettivi (in particolare mi riferisco ai casi russi nominativo, genitivo, accusativo, ecc. che solitamente non hanno la giusta traduzione in italiano a livello di funzione grammaticale che ricoprono).

Alla luce dei dati raccolti in questa fase dell'indagine, è possibile concludere che le percezioni iniziali dei e delle partecipanti rispetto alle prestazioni traduttive automatiche risultano apparentemente positive. Al contempo tali percezioni non sembrano essere sostenute da un bagaglio esperienziale significativo, ossia da conoscenze e/ o da esperienze d'uso (o di semplice consultazione) che abbiano permesso agli apprendenti di valutare l'effettivo ausilio fornito dal sistema di TA.

²⁶ Questa percentuale è data dalla somma tra le percentuali corrispondenti ai riscontri “non saprei”, “decisamente no”, “più no che sì”.

²⁷ A questa parte del quesito avrebbe dovuto rispondere il 16,7% delle e dei rispondenti, ossia coloro che hanno dichiarato “decisamente sì” e “più sì che no”.

²⁸ Tra le quattro risposte formulate in totale.

3. Analisi dei dati di fine corso (Questionario 2)

L'obiettivo della nostra analisi è di mettere a confronto i riscontri forniti all'inizio dell'esperienza didattica (Par. 2.1; 2.2; 2.3) rispetto a quanto emerso al termine dell'esperienza (Par. 3.1; 3.2; 3.3) al fine di studiare l'evoluzione delle percezioni di studenti e studentesse e delle loro modalità di interazione con i traduttori automatici.

Sulla base dei riscontri ottenuti alla fine del corso mediante la somministrazione del questionario 2, abbiamo potuto rilevare la partecipazione effettiva di 31 studenti e studentesse. Si sottolinea, tuttavia, che i dati raccolti dal secondo questionario risultano essere rappresentativi di un gruppo di apprendenti relativamente omogeneo, quanto alle conoscenze e alle competenze linguistiche e tecnologiche dichiarate.

3.1 La traduzione automatica: una revisione delle percezioni

Un primo dato di interesse che emerge dal confronto puntuale²⁹ tra i riscontri al primo e al secondo questionario riguarda la modalità d'interazione tra l'apprendente e il traduttore automatico neurale.

Si osserva, innanzitutto, una revisione generale delle percezioni circa l'affidabilità della traduzione automatica: se il 76,7% dei e delle rispondenti al questionario 1 aveva considerato tale strumento come "abbastanza affidabile" e solo il 20,9% lo aveva definito "poco affidabile" (Fig. 3), a seguito dell'esperienza didattica si assiste a un'inversione di questi valori percentuali: il questionario 2 evidenzia in effetti un 29% di "abbastanza affidabile" e un 64,5% di "poco affidabile", e compare un 6,5% di riscontri negativi ("per niente affidabile" e "non saprei") che non era stato registrato in riferimento allo stesso quesito posto nel questionario 1.

Le attività didattiche basate sull'integrazione della TA potrebbero aver favorito l'instaurarsi di una modalità di interazione utente-macchina più consapevole, in cui l'apprendente è incoraggiato a verificare con maggiore attenzione le prestazioni traduttive automatiche. La traduzione automatica sarebbe dunque sottoposta, in questo caso, a un processo di valutazione in cui l'apprendente non si limiterebbe a una passiva (o fiduciosa) consultazione.

²⁹ Con "riscontro puntuale" si intende la comparazione dei riscontri formulati in risposta allo stesso quesito proposto in entrambi i questionari, prima e dopo il corso sperimentale.

Inoltre, si può notare una ripartizione più equa dei riscontri in risposta al quesito sulle cause di errore in traduzione automatica. Rispetto a quanto evidenziato dal quesito analogo posto all'inizio del corso (Fig. 4), la principale causa di errore non viene più associata alla complessità del linguaggio umano. Dopo aver seguito il corso, infatti, gli errori sono imputati in modo comparabile ai tre fattori proposti all'interno del quesito³⁰: l'algoritmo, i *corpora* e il carattere idiomatico delle locuzioni.

L'esperienza didattica potrebbe dunque aver contribuito a una prima presa di coscienza relativamente all'interdipendenza che lega le diverse variabili interessate dal processo di traduzione automatica.

Se si analizzano i riscontri in merito al controllo della fonte da cui viene estratto il risultato proposto dalla piattaforma di traduzione automatica, si può rilevare un altro confronto significativo, illustrato dall'evoluzione dei valori percentuali all'interno della tabella 2.

Questionario 1	Questionario 2
26,2% Mai	6,4% Mai
59,5% Solo a volte	45,2% Solo a volte
14,3% Sì, sempre	48,4% Sì, sempre

Tabella 2: Il controllo della fonte

Il controllo della fonte, e quindi l'operazione di valutazione dell'*output* della macchina, sembrerebbe essere stato favorito dalle attività didattico-sperimentali proposte. Questo dato risulta dunque convergere verso l'osservazione di un riposizionamento, più attivo, da parte dell'apprendente umano nei confronti della macchina.

Tuttavia, si rilevano alcune considerazioni in risposta (aperta) ai quesiti relativi all'utilità dell'intelligenza artificiale applicata alle lingue³¹ che meriterebbero un ulteriore approfondimento in quanto mettono in luce diversi elementi di potenziale utilità e/ o di criticità, ma necessiterebbero

³⁰ “Dopo aver seguito il corso, quando il risultato che ottiene dal traduttore automatico o dalla piattaforma non le sembra corretto, pensa che l'errore sia dovuto: “all'algoritmo/strumento utilizzato che è poco performante” (32,3%); “ai corpora/testi su cui si basano i traduttori automatici” (35,5%) e “al fatto che la locuzione ricercata sia idiomatica o troppo complessa” (32,3%).

³¹ “Dopo aver seguito il corso, quale pensa possa essere un aspetto positivo che potrebbe caratterizzare l'uso dell'intelligenza artificiale relativamente alle lingue (materna, straniera)?”.

“Dopo aver seguito il corso, quale pensa possa essere un aspetto negativo che potrebbe caratterizzare l'uso dell'intelligenza artificiale relativamente alle lingue (materna, straniera)?”.

di un maggior grado di precisione nella loro formulazione (un dato da tenere presente per una futura fase sperimentale).

Analogamente a quanto emerso dal primo questionario (Par. 2.2), il vantaggio offerto dall'IA che è stato registrato nella maggioranza dei riscontri (pari al 48,38%) è la velocizzazione del processo interlinguistico che sottende l'esercizio traduttivo e, nello specifico, la ricerca degli equivalenti in lingua d'arrivo.

Dall'analisi dei riscontri al primo questionario, l'IA applicata alle lingue era stata considerata un utile ausilio all'apprendimento delle lingue. A seguito dell'esperienza didattica, questa considerazione permane, ma l'utilità viene segnalata più dettagliatamente in quanto si traduce nell'accessibilità ai contenuti testuali in più lingue a vantaggio dell'utenza globale del web³², e nella possibilità di prendere nota di sinonimi e di varianti lessicali e terminologiche³³. Si registra, inoltre, un unico riscontro in cui si associa all'adozione delle applicazioni linguistiche dell'IA un'utilità in termini di stimolo alla riflessione metalinguistica e metatraduttiva³⁴.

È interessante analizzare, in particolare, una serie di riscontri (pari al 16,1%) secondo i quali l'IA potrebbe fornire un ausilio alla comprensione dei contenuti testuali. All'interno di questi riscontri, però, non figura chiaramente la distinzione tra l'utente profano e l'utente competente in ambito linguistico, da intendersi come l'utente in possesso di conoscenze e competenze linguistiche atte a un uso consapevole di tali tecnologie. Di conseguenza, resterebbe da esplicitare chi sia considerato come il principale beneficiario di questo specifico ausilio alla comprensione del testo.

Ciononostante, all'interno di alcune risposte al quesito sulle criticità legate all'applicazione dell'intelligenza artificiale nel settore delle lingue³⁵, viene menzionata la necessità di una valorizzazione delle competenze linguistiche e traduttive umane. Grazie a questi riscontri (pari al 38,70%), vengono distinti tre diversi ruoli: il traduttore umano (professionista), l'utente/lettore/fruttore di servizi web, tra cui le applicazioni nate dall'IA, e

³² Dato che emerge dal 9,67% dei riscontri.

³³ Dato che emerge dal 25,80% dei riscontri.

³⁴ "paradossalmente il fatto che, per controllare la correttezza della traduzione proposta dal traduttore automatico, il traduttore umano abbia uno stimolo di riflessione in più".

³⁵ "Dopo aver seguito il corso, quale pensa possa essere un aspetto negativo che potrebbe caratterizzare l'uso dell'intelligenza artificiale relativamente alle lingue (materna, straniera)?"

il traduttore automatico accessibile in rete. Si riportano di seguito alcune risposte rappresentative in tal senso:

Fare affidamento in modo eccessivo sulle risposte prodotte dal traduttore automatico può trascinarci in errori più o meno gravi. La traduzione come mestiere richiede delle competenze cognitive e di tipo enciclopedico, una capacità di riflessione e di ragionamento sul tipo di testo che si sta traducendo, sul tipo di lettore al quale ci stiamo rivolgendo e sul contesto culturale in questione, tutte capacità che per il momento può sviluppare solo il traduttore umano.

Un aspetto negativo è pertanto la quantità di errori che i traduttori automatici possono commettere durante l'operazione e che l'utente può non riconoscere subito, utilizzando una traduzione errata.

Chiunque può accedere alla traduzione automatica di una lingua straniera, ma solo chi possiede determinate competenze riesce a captarne gli errori meno evidenti.

Infine, un altro gruppo di riscontri significativi al questionario 2 (pari al 45,16%) insiste sulla presenza di errori di senso e sulla rilevazione di traduzioni non pertinenti rispetto al contesto³⁶, e quindi, sull'inattendibilità delle traduzioni automatiche proposte.

In conclusione, da questo primo confronto tra i risultati ottenuti a seguito della somministrazione dei due questionari, è possibile individuare un primo insieme di dati che parrebbe convergere verso la rilevazione di un approccio più attento e consapevole da parte dell'apprendente nei confronti delle prestazioni di TA, a seguito dell'esperienza formativa.

3.2 L'efficacia della TA: percezioni e contraddizioni

Il questionario 2 ha previsto 11 quesiti sull'efficacia della traduzione automatica, i cui riscontri vengono presentati sinteticamente all'interno della tabella 3.

L'analisi delle percentuali riportate in tabella 3, permette di osservare come soltanto in riferimento al "lessico generale" siano stati registrati riscontri positivi rilevanti (61,29% "abbastanza efficaci" e 6,45% "molto efficaci"). Nessun'altra categoria di elementi linguistici è stata infatti considerata come traducibile dalla macchina in modo "molto efficace".

³⁶ Un concetto, quello di "contesto", su cui torneremo nel paragrafo 3.2.

Dopo aver seguito il corso, quanto pensa che siano efficaci le piattaforme di traduzione automatica (ad es. DeepL, Google Traduttore) nel tradurre i seguenti fenomeni linguistici?					
	Per nulla efficaci	Poco efficaci	Parzialmente efficaci	Abbastanza efficaci	Molto efficaci
Lessico generale	-	3,22%	29,03%	61,29%	6,45%
Lessico specialistico	9,67	35,48%	51,61%	3,22%	-
Neologismi	19,35%	51,61%	25,80%	3,22%	-
Collocazioni	6,45%	41,93%	51,61%	-	-
Parole polisemiche	9,67%	61,29%	25,80%	3,22%	-
Giochi di parole	48,38%	38,70%	12,90%	-	-
Fenomeni di non equivalenza lessicale	38,70%	48,38%	12,90%	-	-
Parole connotate culturalmente	32,25%	51,61%	16,12%	-	-
Sintagmi nominali complessi	32,25%	51,61%	16,12%	-	-

Tabella 3: L'efficacia della TA

Inoltre, sono state individuate quattro categorie che vengono classificate principalmente come traducibili dalla macchina in modo “poco efficace” e “parzialmente efficace”: lessico specialistico³⁷, neologismi, collocazioni e parole polisemiche.

Infine, si può constatare una percentuale significativa di riscontri decisamente negativi (“per nulla efficaci”) e negativi (“poco efficaci”) in corrispondenza delle seguenti categorie: giochi di parole, fenomeni di non equivalenza lessicale, parole connotate culturalmente e sintagmi nominali complessi.

All'analisi delle percentuali indicate in tabella 3, è interessante unire l'analisi delle risposte aperte al quesito seguente:

“A seguito del corso, ritiene che i sistemi di traduzione automatica forniscano risultati più affidabili con parole/stringhe isolate o intere sequenze testuali? Motivi la sua risposta”.

È possibile constatare come sia i riscontri sia le motivazioni esposte divergano in modo significativo: se il 54,83% dei e delle rispondenti afferma

³⁷ Per quanto riguarda il lessico specialistico, anche la percentuale di riscontri decisamente negativi risulta abbastanza significativa.

che i sistemi di TA forniscono risultati più affidabili a partire da parole isolate o brevi stringhe di testo perché in questi casi si riducono sia la complessità del dato sintattico sia la necessità di “contestualizzare” la parola o la stringa breve in un testo più ampio, al contempo il 45,16% associa la maggiore affidabilità della TA a intere sequenze testuali sostenendo che il traduttore automatico possa trattare una maggiore porzione di testo e quindi selezionare con maggiore esattezza gli equivalenti più pertinenti rispetto al “contesto”.

Il dato significativo che risulta da questa divergenza di concezione della prestazione traduttiva automatica riguarda il riferimento improprio all’operazione di contestualizzazione in traduzione. Più precisamente, tale operazione, oggetto di ricerche traduttologiche teoriche e applicate (Gambier 2008: 181; Durieux 2009) che ne hanno dimostrato la complessità cognitiva, linguistica e circostanziale, viene menzionata erroneamente come un processo attribuibile al sistema di traduzione automatica neurale.

Infatti, se è vero che quest’ultima integra il *Deep Learning* e le relative applicazioni nate dall’IA (Forcada 2017; Sun 2017: 1; Le Cun 2019) alle operazioni di calcolo estremamente sofisticate che sottendono il processo traduttivo, è pur vero che la prestazione del sistema resta tuttora dipendente dalla fase di addestramento (*training*) che, ad oggi, consiste nel fornire grandi quantità di dati di natura testuale (e cotestuale), non necessariamente contestualizzata all’interno di una situazione di comunicazione reale.

Spetta al traduttore umano, infatti, valutare la pertinenza della traduzione automatica proposta rispetto al contesto nel corso dell’umana operazione di *transfer* (Pym 1992: 172-173; Le Disez 2013; Gütt 2000), specialmente se, come nel caso della sperimentazione didattica condotta, si trova ad integrare al proprio progetto di traduzione ordinario un sistema di TA fruibile (e modificabile) dal grande pubblico del web, senza beneficiare di un applicativo personalizzato³⁸ come invece accade nel mondo della professione.

La ricerca della pertinenza interlinguistica e interculturale è un aspetto che emerge dai riscontri al quesito sulla post-edizione (Par. 3.3).

³⁸ Systran offre ai professionisti del settore della traduzione questo servizio, basato appunto sulla distinzione tra “*generic NMT*” e “*specialized NMT*”. Cfr. Jean Senellart, CEO of SYSTRAN: <https://www.systransoft.com/systran/translation-technology/neural-machine-translation-nmt> (consultato il 23/7/2021).

3.3 La post-edizione

Il questionario 2 ha incluso un solo quesito inerente alla post-edizione, che si aggiunge agli 11 quesiti sulle percezioni dell'efficacia delle traduzioni automatiche: "Quali degli aspetti linguistici menzionati nella domanda precedente³⁹ (o altri di sua conoscenza) ritiene che richiedano maggior lavoro di *post-editing* e perché?".

Le risposte si collocano in modo coerente⁴⁰ rispetto ai riscontri analizzati nel precedente paragrafo, ma trattandosi di risposte aperte, è interessante osservare i motivi che stanno alla base di una necessità di post-edizione secondo le persone che hanno risposto.

I riscontri, nel loro insieme, segnalano che il traduttore umano interviene perché:

- 1) è necessario contestualizzare questi elementi linguistici;
- 2) occorre riformulare le relative parti di testo per garantire una restituzione corretta e funzionale in lingua d'arrivo.

Secondo i e le partecipanti, contestualizzare significa selezionare l'equivalente più pertinente rispetto al testo considerato, ma anche rispetto al contesto di arrivo, non solo sulla base del fattore "frequenza d'uso" del lessema o dell'espressione nella lingua di arrivo⁴¹. L'operazione di conte-

³⁹ Si tratta degli elementi linguistici indicati in tabella 3.

⁴⁰ Precisando che si tratta di risposte aperte che possono comprendere più elementi linguistici, di seguito si riportano i riscontri registrati sul piano quantitativo: le parole connotate culturalmente sono gli elementi più ricorrenti (menzionate nel 41,93% dei riscontri); seguono i giochi di parole (32,25%); i neologismi (22,58%); i sintagmi nominali complessi e i fenomeni di non equivalenza lessicale (19,35%); il lessico specialistico e le espressioni idiomatiche figurano solo in un 6,45% e, in linea con i dati riportati in tabella 3, il lessico generale viene segnalato da una minoranza del 3,22% come elemento di difficoltà traduttiva per la macchina.

⁴¹ Si riportano alcuni riscontri rappresentativi: "perché devono essere tradotti in maniera inerente al testo e non a seconda di quale termine è più utilizzato o meno"; "perché i termini di un linguaggio settoriale devono avere un significato esclusivamente connotativo ed è bene che quello proposto sia corretto e adatto al contesto"; "[i neologismi] spesso il traduttore automatico non li conosce o non li riconosce e quindi deve intervenire il traduttore per contestualizzarli"; "il traduttore [automatico] non riconosce il contesto di utilizzo e dunque sceglie un significato della parola, tra quelli disponibili, non pertinente"; "i traduttori automatici non sempre riconoscono [le parole connotate culturalmente] e di conseguenza traducono letteralmente, dando un risultato non corretto nella lingua di arrivo"; "[le parole connotate culturalmente] perché non richiedono una traduzione letterale, ma un lavoro più complesso".

stualizzazione viene descritta come un adattamento alla lingua e alla cultura di arrivo. Per quanto riguarda la riformulazione e dunque l'adozione di una strategia di traduzione — un'espressione che non figura mai in questi riscontri —, le e i rispondenti descrivono l'intervento umano come un'elaborazione di dati, una riflessione di natura interlinguistica e interculturale che appartiene solo al processo decisionale del tradurre umano. Si riportano alcuni riscontri rappresentativi: “perché [i neologismi e il lessico generale] sono i più utilizzati e richiedono una maggiore attenzione”; “i neologismi, i giochi di parole e i fenomeni di non equivalenza lessicale perché richiedono una riflessione aggiuntiva che una macchina non può effettuare”; “le parole connotate culturalmente perché se non si conosce la cultura è difficile trovare un'equivalenza”; “le parole connotate culturalmente perché la cultura del testo di arrivo potrebbe distaccarsi ampiamente da quella del testo di partenza”; “le parole connotate culturalmente o i fenomeni di non equivalenza lessicale, che i traduttori automatici non sempre riconoscono e di conseguenza traducono letteralmente, dando un risultato non corretto nella lingua di arrivo”; “le parole connotate culturalmente perché la *lexiculture* della lingua di partenza e quella della lingua di arrivo non corrispondono sempre”.

Conclusioni e prospettive di ricerca

Questa prima esperienza didattica è stata caratterizzata dall'integrazione della traduzione automatica neurale nel percorso di formazione in Lingua e traduzione francese del Corso di Laurea triennale in Scienze della mediazione linguistica. Le attività didattiche proposte sono state definite come esercizi di “post-edizione ragionata” che sono stati svolti con l'ausilio di una griglia per la classificazione e per la correzione degli errori risultanti dalla traduzione automatica. L'obiettivo principale di questo corso sperimentale è stato quello di sensibilizzare l'apprendente ad un uso consapevole della traduzione automatica, e non ad un uso propriamente professionale di questo strumento. Si ritiene infatti di fondamentale importanza, nell'era dell'intelligenza artificiale, programmare un'integrazione mirata e progressiva delle tecnologie di ultima generazione nei curricula universitari per la formazione in lingue e in traduzione. Alla luce dei dati raccolti nel corso dell'indagine qui esposta, la sensibilizzazione ad un uso consapevole delle risorse linguistiche e traduttive esistenti nell'industria

delle lingue si configura come un obiettivo formativo centrale, da perseguire in modo coerente rispetto all'orientamento specifico del corso di studio considerato. In particolare, i dati relativi all'evoluzione delle percezioni dei e delle partecipanti circa l'efficacia e l'affidabilità della traduzione automatica (Par. 3.1 e 3.2) e la conseguente ridefinizione della modalità di interazione utente-macchina (Par. 2 e 3.3) potrebbero confermare l'interesse pedagogico legato a un approccio riflessivo e graduale. I risultati finora ottenuti aprono la strada a ulteriori quesiti più mirati che meriterebbero di essere approfonditi nell'ambito di indagini future. Si può citare, a titolo esemplificativo, la correlazione tra prestazione traduttiva automatica e caratterizzazione linguistico-testuale dei testi sottoposti alla traduzione automatica: la tipologia testuale considerata può determinare una variazione della percezione dell'apprendente rispetto all'affidabilità della TA? Quali configurazioni può assumere l'interazione utente-macchina nei vari contesti di formazione in lingua e traduzione?

Restano dunque da esplorare le strade per la definizione di una formazione in lingue e in traduzione basata su obiettivi formativi specifici, ai fini di un uso corretto e consapevole delle innumerevoli risorse che sono state e che continueranno ad essere sviluppate nel mondo dell'IA applicata alle lingue.

Bibliografia

- Adam Jean-Michel (2005). *Les textes : types et prototypes. Récit, description, argumentation, explication et dialogue*. Parigi: Armand Colin.
- Amparo Hurtado Albir, Galán-Mañas Anabel, Kuznik Anna, Olalla-Soler Christian, Rodríguez-Inés Patricia, Romero Lupe (2020). "Translation competence acquisition. Design and results of the PACTE group's experimental research". *The Interpreter and translator trainer*, 14/2, 95-233.
- Bowker Lynne, Jairo Buitrago Ciro (2019). *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley: Emerald Publishing.
- Cennamo Ilaria (2018). *Enseigner la traduction humaine en s'inspirant de la traduction automatique*. Roma: Aracne.
- Charaudeau Patrick (2009). "Le discours de manipulation entre persuasion et influence sociale". *Actes du colloque de Lyon*, 1-14. URL: http://www.patrick-charaudeau.com/IMG/pdf/Discours_Manipulation_-_Texte_Lyon_.pdf (Consultato il 6 settembre 2021).
- Charaudeau Patrick (2006). "Discours journalistique et positionnements énonciatifs. Frontières et dérives". *Semen* [En ligne], 22. URL: <http://journals.openedition.org/semen/2793> (Consultato il 6 settembre 2021).
- Charaudeau Patrick (1997). *Le discours d'information médiatique. La construction du miroir social*. Parigi: Nathan.
- Cordonnier Jean-Louis (1995). *Traduction et culture*. Paris: Les Éditions Didier.
- Delisle Jean (1980). *L'analyse du discours comme méthode de traduction. Initiation à la traduction française de textes pragmatiques anglais : théorie et pratique*. Ottawa: Éditions de l'Université d'Ottawa.
- Delisle Jean (2013). *La traduction raisonnée*. Ottawa: Les Presses de l'Université d'Ottawa.
- Durieux Christine (2009). "Vers une théorie décisionnelle de la traduction". *Revue LISA/LISA e-journal*, 7/3, 349-367.
- Ehrensberger-Dow Maureen, Massey Gary (2019). "Le traducteur et la machine. Mieux travailler ensemble". In: Élisabeth Lavault-Olléon, Maria Zimina (eds). *Des mots aux actes, n° 8 – Traduction et technologie : regards croisés sur des nouvelles pratiques*. Paris: Éditions Classiques Garnier, 47-62.
- Eisele Andreas (2018). *Data collection for neural MT within the Connecting Europe Facility e-Translation*, URL: <https://blogs.helsinki.fi/language-technology/files/2018/11/SMART-Select-4-Eisele.pdf> (Consultato il 7 settembre 2021).

- Forcada Mikel (2017). "Making sense of neural machine translation". *Translation Spaces*, 6/2, 291-309.
- Gambier Yves (2008), "Stratégies et tactiques en traduction et interprétation". In: Gyde Hansen, Andrew Chesterman, Heidrun Gerzymisch-Arbogast (eds). *Efforts and Models 32 in Interpreting and Translation Research*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 63-82.
- Gütt Ernst-August (2000). *Translation and relevance: Cognition and Context*. Londra: Routledge.
- Hernandez Morin Katell (2019). "Évolution des technologies et des usages en traduction". In: Élisabeth Lavault-Olléon, Maria Zimina (eds), *Des mots aux actes, n° 8 – Traduction et technologie : regards croisés sur des nouvelles pratiques*. Parigi: Éditions Classiques Garnier, 239-255.
- Kenny Dorothy (2020). "Technology in Translator Training". In: Minako O'Hagan (eds) *The Routledge Handbook of Translation and Technology*. Londra/New York: Routledge, 498-515.
- Krieg-Planque Alice (2012). *Analyser les discours institutionnels*. Parigi: Armand Colin.
- Krieg-Planque Alice (2000). "Analyser le discours de presse. Mises au point sur le « discours de presse » comme objet de recherche". *Communication*, 20/1, 75-97. URL: <http://journals.openedition.org/communication/6432> (Consultato il 7 settembre 2021).
- Lavault-Olléon Élisabeth, Zimina Maria (eds). *Des mots aux actes. 2019, n° 8 - Traduction et technologie, regards croisés sur de nouvelles pratiques*. Parigi: Éditions Classiques Garnier.
- Le Cun Yann (2019). *Quand la machine apprend : La révolution des neurones artificiels et de l'apprentissage profond*. Parigi: Odile Jacob.
- Le Dizez Jean-Yves (2013). *Le bon sens en traduction*. Rennes: Presses universitaires de Rennes.
- Lederer Marianne (2006). *La traduction aujourd'hui. Le modèle interprétatif*. Parigi: Lettres Modernes Minard.
- Leoncini Bartoli Antonella (2019). "La médiation à la lumière de la traduction ou 'dénouer des sacs de nœuds'". In: Michele De Gioia, Mario Marcon (eds) *L'essentiel de la médiation. Le regard des sciences humaines et sociales*. Bruxelles: Peter Lang, 209-229.
- Loock Rudy (2019). "La plus-value de la biotraduction face à la machine: Le nouveau défi des formations aux métiers de la traduction". *Traduire*, 241, 54-65. URL: <https://journals.openedition.org/traduire/1848> (Consultato il 6 settembre 2021).

- Maingueneau Dominique (2016). *Analyser les textes de communication*, Parigi: Armand Colin.
- Maingueneau Dominique (2004). "La situation d'énonciation entre langue et discours". *Dix ans de S.D.U*, Editura Universitaria Craiova, 197-210. URL: <http://dominique.maingueneau.pagesperso-orange.fr/pdf/Scene-d-enonciation.pdf> (Consultato il 5 settembre 2021).
- Marcoccia Michel, Gauducheau Nadia (2007). "L'Analyse du rôle des smileys en production et en réception : un retour sur la question de l'oralité des écrits numériques". *Glottopol. Revue de sociolinguistique en ligne*, 10, 39-55.
- Martikainen Hanna (2019). "Qualité fonctionnelle et distorsion d'interprétation". In: Élisabeth Lavault-Olléon, Maria Zimina (eds), *Des mots aux actes*, n° 8 – *Traduction et technologie : regards croisés sur des nouvelles pratiques*. Parigi: Éditions Classiques Garnier, 256-270.
- Massey Gary, Kiraly Don (2019). "The future of translator education: A dialogue". *Cultus: The Journal of Intercultural Mediation and Communication*, 12, 15-34. URL: https://www.cultusjournal.com/files/Archives/Cultus_2019_12_002_Kiraly_Massey.pdf (Consultato il 7 settembre 2021).
- Mattioda Maria Margherita (2019). "Economie circulaire: problèmes en boucle". *RiCOGNIZIONI. Rivista Di Lingue E Letterature Straniere E Culture Moderne*, 6/12, 239-255.
- Moirand Sophie (2007). *Les discours de la presse quotidienne. Observer, analyser, comprendre*. Parigi: Presses Universitaires de France.
- Nord Christiane (2018). *Translating as a purposeful activity: functionalist approaches explained*. Londra/New York: Routledge.
- O'Brien Sharon (2020). "Translation, human-computer interaction and cognition 1". In: Fabio Alves, Arnt Lykke Jakobsen (eds), *The Routledge Handbook of Translation and Cognition*, Londra: Routledge, 376-388.
- Peraldi Sandrine (2014). "La Post-édition à la portée du traducteur", *Tralogy [En ligne]*, *Tralogy I, Session 2 - Translation as a profession / Le métier du traducteur, Lodel I-Revues*. URL: <http://lodel.irevues.inist.fr/tralogy/index.php?id=107> (Consultato il 7 settembre 2021).
- Pym Anthony (1992). "The relations between translation and material text transfer". *Target*, 4, 171-189.
- Rabatel, Alain, "Sur-énonciateurs et construction dissensuelle des savoirs". In: *Actes de la DESCO*, IUFM de Montpellier, CRPE, 2002, 89-100.
- Reiss Katharina, Vermeer Hans Josef (2013). *Towards a General Theory of Translational Action: Skopos Theory Explained*. Londra/ New York: Routledge.

Rossi Caroline (2019). “L'apprenti traducteur et la machine. Des connaissances aux perceptions de la traduction automatique”. In: Élisabeth Lavault-Olléon, Maria Zimina (eds), *Des mots aux actes, n° 8 – Traduction et technologie : regards croisés sur des nouvelles pratiques*. Parigi: Éditions Classiques Garnier, 9-105.

Schumacher Perrine (2019). “Avantages et limites de la post-édition”. *Traduire*, 241, 108-123. URL: <http://journals.openedition.org/traduire/1887> (Consultato il 7 settembre 2021).

Stark Elisabeth (2015). “De l'oral dans l'écrit? – Le statut variationnel des SMS (textos) et leur valeur pour la recherche linguistique”. In: Kirsten Jeppesen Kragh, Jan Lindschouw (eds), *Les variations diasystématiques et leurs interdépendances dans les langues romanes*. Strasburgo: Éditions de linguistique et de philologie, 395-408.

Sun Rémy (2017). “Apprentissage profond et acquisition de représentations latentes de séquences peptidiques”, tesi di laurea magistrale (*mémoire de master*), Rennes: INRIA, 1-21. URL: <https://hal.inria.fr/hal-01406368> (Consultato il 7 settembre 2021).

Toudic Daniel, Hernandez Morin Katell, Moreau Fabienne, Barbin Franck, Phuez Gaëlle (2014). “Du contexte didactique aux pratiques professionnelles : proposition d'une grille multicritères pour l'évaluation de la qualité en traduction spécialisée”. *ILCEA*, 19, 1-19. URL: <http://journals.openedition.org/ilcea/2517> (Consultato il 7 settembre 2021).

Vion Robert (1992). *La communication verbale : analyse des interactions*, Parigi: Hachette.

Sitografia

Google Traduttore, <https://translate.google.com/?hl=fr&sl=fr&tl=it&op=translate> (Consultato il 23 luglio 2021).

Quadro comune europeo di riferimento per le lingue, <https://rm.coe.int/168045bc72> (Consultato il 23 luglio 2021).

Regolamento didattico, https://www.lingue.unito.it/do/documenti.pl/ShowFile?_id=8506;field=file;key=jtJNb9OxLJ60FYA1drj8sPciDwWNUBa6SUqRDU-hKtaQcEhdITUYFCUUEM50AeQLIR5B14B1L36kJaRWUQIcAAB2Lk6QgM-Bk1ztzbA1nuIA;t=2390

Systran Soft, <https://www.systransoft.com/systran/translation-technology/neural-machine-translation-nmt> (Consultato il 23 luglio 2021).

Ville de Paris, <http://www.paris.fr/actualites/le-musee-carnavalet-se-refait-une-beaute-4113> (Consultato il 2 luglio 2021).

Allegati

Sintesi delle diapositive presentate in occasione del workshop del 24 aprile 2021 dalle studentesse Veronica Audisio e Chiara Brizio, portavoce del gruppo di partecipanti all'esperienza didattica

Prima parte

Di seguito, la presentazione sintetica di alcuni esempi di categorizzazione e degli errori risultanti da Google Traduttore.

Testo di partenza	TA	Tipi di errore	Considerazioni
De lundi à vendredi, ils se sont réunis au bord du lac Léman pour le premier meeting international de <u>Fridays for future</u> [...]	Dal lunedì al venerdì, si sono riuniti sulle rive del Lago di Ginevra per il primo incontro internazionale <u>dei venerdì per il futuro</u> [...]	-Lessicale e di conseguenza -Semantico (errore di senso)	Mancato riconoscimento del nome proprio del movimento.
Et si aujourd'hui <u>vous preniez</u> une décision importante: devenir <u>écolo</u> .	E se oggi <u>prendessi</u> una decisione importante: diventare <u>verde</u> .	-Registro -Terminologia	Traduzioni non pertinenti rispetto al discorso in cui si collocano.
Pour partager sa joie de vivre et sa vision de la mode, elle n'hésite pas à poster ses derniers <u>coups de cœur photo, mode, déco ou beauté</u> . <u>Une grande</u> , on vous dit.	Per condividere la sua gioia di vivere e la sua visione della moda, non esita a pubblicare <u>le sue ultime foto, moda, decorazione o bellezza preferiti</u> . <u>Fantastico</u> , te lo diciamo.	-Omissione e di conseguenza -Sintassi in lingua d'arrivo -Lessico: Variazione di genere.	La TA procede calcando la sintassi di partenza. Il genere femminile non viene applicato nel testo d'arrivo.

Tabella 1: Esempi di categorizzazione e degli errori risultanti da Google Traduttore

Seconda parte

Di seguito, un elenco delle principali considerazioni formulate dalle portavoce dei e delle partecipanti al corso sperimentale.

- Le traduzioni di Google Traduttore sono rapide e più corrette rispetto al passato.
- Gli errori e le imprecisioni restano numerosi.
- L'approccio traduttivo è letterale.
- La variazione linguistica non è trattata correttamente.

- Realia, neologismi e vocaboli di specialità vengono restituiti dal sistema in modo errato o non pertinente in molti casi.
- La lingua inglese, lingua ponte per Google Traduttore, è all'origine di tradimenti errati o non pertinenti in lingua d'arrivo (diversa dall'inglese).
- La qualità redazionale del testo di arrivo è scarsa quando si tratta di tradurre automaticamente testi giornalistici e promozionali (adottati durante il corso).
- Gli errori commessi non sono di facile comprensione a una prima lettura del testo tradotto automaticamente, occorre un'analisi linguistica più attenta.
- La creatività del processo traduttivo umano resta un elemento necessario ai fini di una traduzione di qualità.
- L'esperienza didattica è stata formativa e professionalizzante.

Artificial Intelligence and Machine Translation: perceptions, opinions and experiences of Italian Graduate Students of English as a Foreign Language

Alessandra Molino

Introduction

This article presents the results of two questionnaires administered to Italian graduate students of English at the beginning (Q1) and at the end (Q2) of a teaching module on Machine Translation (MT) in academic year 2020-2021 at the University of Torino (Italy). The MT module (18 hours) was offered as part of a course on corpora and computer tools for language analysis. Due to the COVID-19 pandemic, the whole course was held *via* video conference software and using Moodle. All activities were synchronous and required oral interaction or over the chat function. The objectives of the MT module were the following:

- 1) To familiarize the students with the underlying mechanisms of MT;
- 2) To raise their awareness of the potential risks of MT, such as inadequate output, language flattening and the so-called “algorithmic bias” (see Baeza-Yates 2018; Vanmassenhove *et alii* 2021);
- 3) To assess the effectiveness of MT applied to texts belonging to different genres;
- 4) To critically compare the output of different MT systems;
- 5) To perform adequate post-editing of the target text (TT).

In order to achieve these goals, theoretical classes were complemented with translation assignments (individual or in group) and subsequent discussions. At the end of the module, the students undertook an assessed activity consisting in the analysis and post-editing of MT output accompanied by a commentary on their revision choices.

In the following sections the participants' profile is described based on the answers obtained in both questionnaires (Section 1); section 2 deals with Artificial Intelligence (AI) in general, and has the purpose of understanding the students' knowledge and use of digital tools in their daily and academic life; in section 3, the answers obtained in Q1 and Q2 in relation to MT are compared, focusing in particular on the challenges that MT still needs to face and the benefits and drawbacks of AI for natural language processing.

1. The students' profile

The two questionnaires obtained a rather consistent number of responses, namely 125 for the preliminary questionnaire and 128 for the end of course one¹ which approximately correspond to the number of students regularly attending the module. The sample is mainly composed of female students (86%) enrolled in the second year of the MA degree programme in either International Communication or International Communication for Tourism². They are mostly aged between 23 and 25 years old and their mother tongue is almost invariably Italian (98%). The respondents declared to be able to speak a range of foreign languages (Fig. 1), with English being mentioned by nearly 100% of the participants, followed by French and Spanish with similar percentages (59% and 57%), and German (30%). Other languages mentioned in the open answers are, in order of frequency, Russian, Portuguese, Japanese, Arabic, Czech, and Italian Sign Language.

Since the module focused on MT applied to the English-Italian language pair, only the students' perceived competence in English is reported here. Overall, the participants are self-confident about their proficiency, predominantly reporting excellent knowledge at the level of receptive abilities and good competence at the level of productive skills (Fig. 2).

¹ The figures reported in this paragraph are the average of the data obtained in the two questionnaires. The slightly higher number of responses in Q2 is because some students answered the end-of-module questionnaire despite the request not to participate in the survey if they had not completed the preliminary questionnaire.

² Both programmes are offered by the Department of Foreign Languages, Literature and Modern Cultures at the University of Torino.

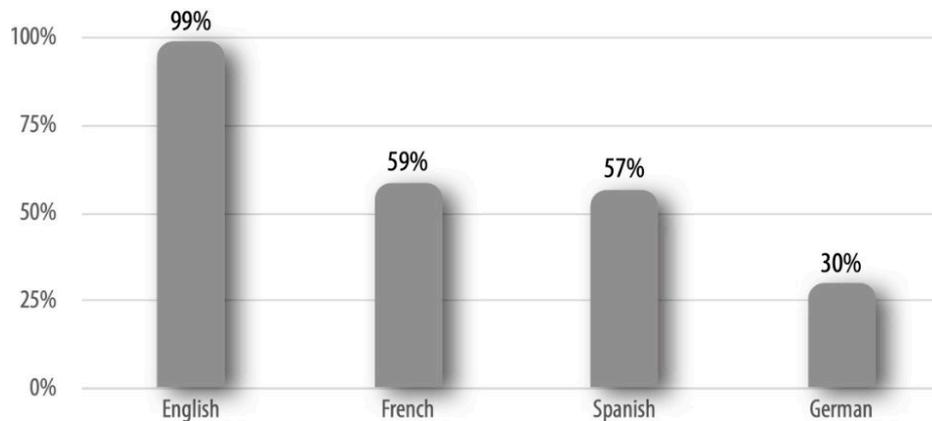


Figure 1: Main foreign languages spoken

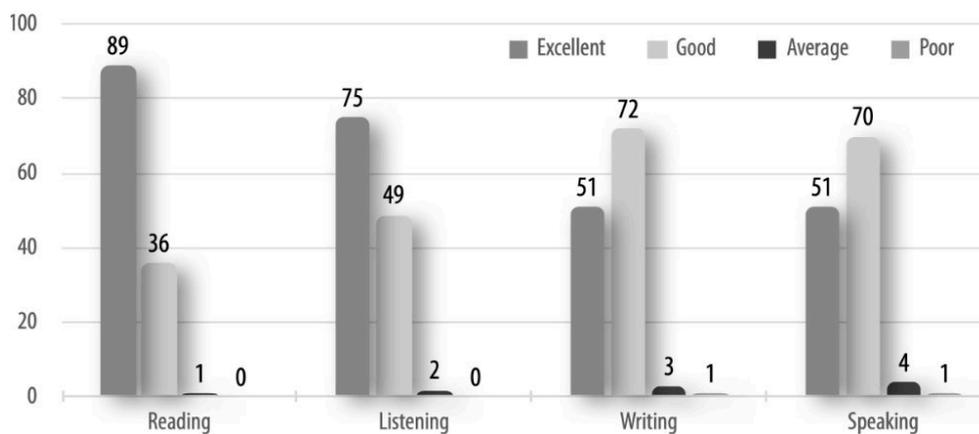


Figure 2: Self-reported ability in the four language skills

2. Knowledge and use of AI

Q1 included four questions on AI and digital tools. The first and second questions aimed at gathering information about the interest in AI and the range of instruments commonly used by the students. The two subsequent questions were about the role of AI in the respondents' field of study and future profession. These last two questions were maintained in Q2 to assess whether the students' perception changed after the activities in class.

According to the data obtained in Q1, the sample under scrutiny is interested in the development of AI, with 52% of the participants paying attention to this phenomenon and 42% expressing a very strong interest. Figure 3 provides a picture of the target population through their use of digital tools in their daily and academic life. First, the young age of the respondents is reflected in the references to social networks, cited 116 times. Their position as university students is also clear, as evidenced by the frequent mention of word processing and spreadsheet programmes (37). But it is the respondent's identity as foreign language learners that emerges strikingly compared to other student populations (see other papers in Part II of this book): the second most often cited tools are, indeed, MT systems and bilingual platforms (43). Online dictionaries appear quite frequently, too (30), and even corpora have been cited a few times (7). Finally, references to online conference platforms are present, clearly related to the pandemic and the experience of distance learning (29).

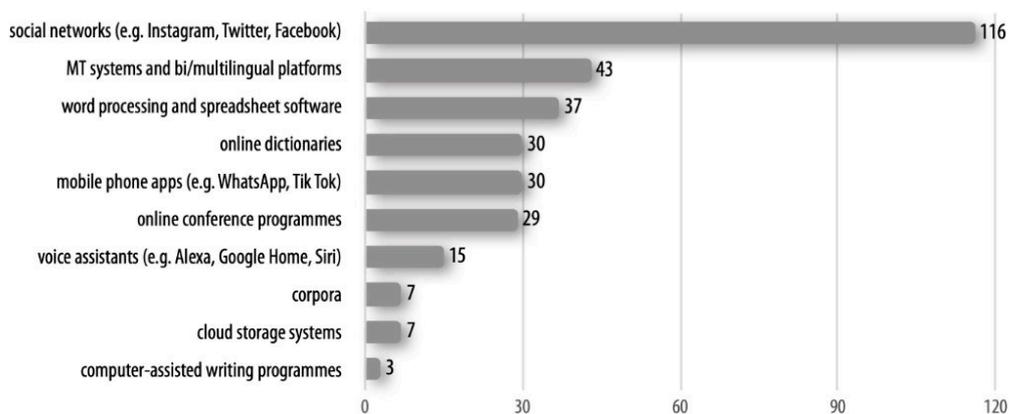


Figure 3: Digital tools most often used by students in their daily and academic life

In both questionnaires (Fig. 4 and 5), the students declared that AI will greatly contribute to the development of their field of studies. Overall, they also believe that AI will play an important role in their future profession. However, a small decrease in their degree of conviction is recorded in Q2 for both answers. Such more cautious attitude regards just a handful of respondents and is difficult to explain. One hypothesis is that the focus on MT during classroom activities led some Q2 participants to interpret these questions as referring to translation, a profession which may not be the aspiration of all the students in the sample.

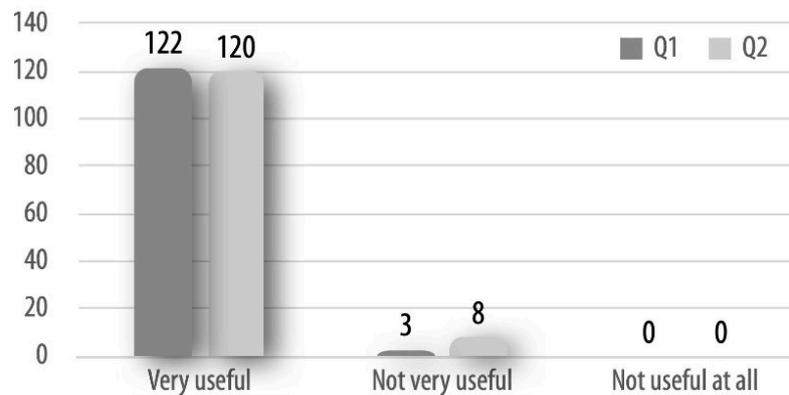


Figure 4: How useful do you think AI will be for your field of study?

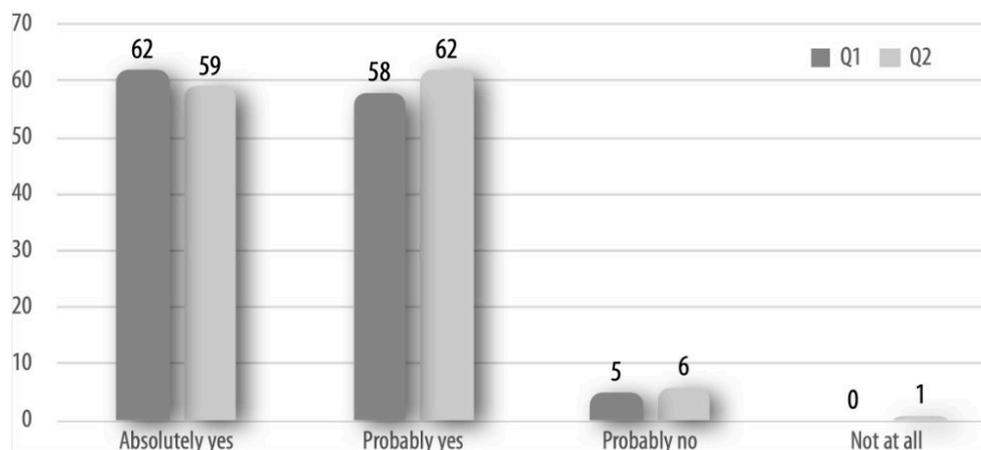


Figure 5: Do you think AI will play a major role in your future profession?

3. Machine translation: Opinions and experiences

Classroom discussions and activities aimed to help students recognize shortcomings in machine translated texts, learn how to amend inadequacies, and better understand the underlying mechanisms of MT. Therefore, in this section, only the questions related to MT will be analysed because the impact of the course can only be assessed in relation to this topic. Figure 6, from Q1, shows that the participants regularly resort to MT systems (54.4% “often”; 36.8% “sometimes”), with only 8% using them rarely and less than 1% never using them. Before the module, the most often cited system was Reverso (114 mentions), followed by Google Translate (41) and DeepL (15).

Figure 7 compares the students' degree of confidence in MT before and after the classroom activities. At the end of the module, the majority is more confident about the output produced by MT systems, which is considered "quite reliable" by 91% of the participants (116 students). During the course, the students could compare different MT programmes. For the texts analysed and the specific language pair and direction, i.e., English-Italian, there was general consensus on the greater overall effectiveness of DeepL and Google Translate than Reverso. The possibility of using and comparing new tools, some of which being relatively unfamiliar to many, could have contributed to this shift in perception. In addition, before the module, as some students declared in the open questions, MT systems were

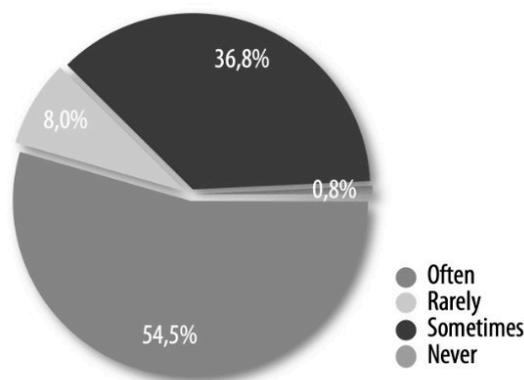


Figure 6: How often do you use free of charge or paid for MT systems?

used to check the meaning of words, as an alternative to a dictionary, or to understand the gist of a short string or passage. The increase in confidence may therefore be connected with first-hand observations of the overall efficacy of automated systems with full texts, reaching "a quality that is (arguably) close to that of human translations" (Vanmassenhove *et alii* 2020: 2203), and to having experimented the usefulness of MT in terms of productivity (see Toral 2019).

The subsequent question was about the potential source of mistakes in MT output (Fig. 8). Before and after the course, most respondents identify the main cause for errors and inadequacies in the idiomaticity or complexity of source language expressions. It is interesting to observe, however, that in Q2 a higher number of students mention the role played by the algorithm used. These data are in line with expectations, as classroom discussions aimed at raising awareness of the challenges that MT still has to face in terms of correctness and fluency, and the implications of algorithmic bias (although it was a less central topic).

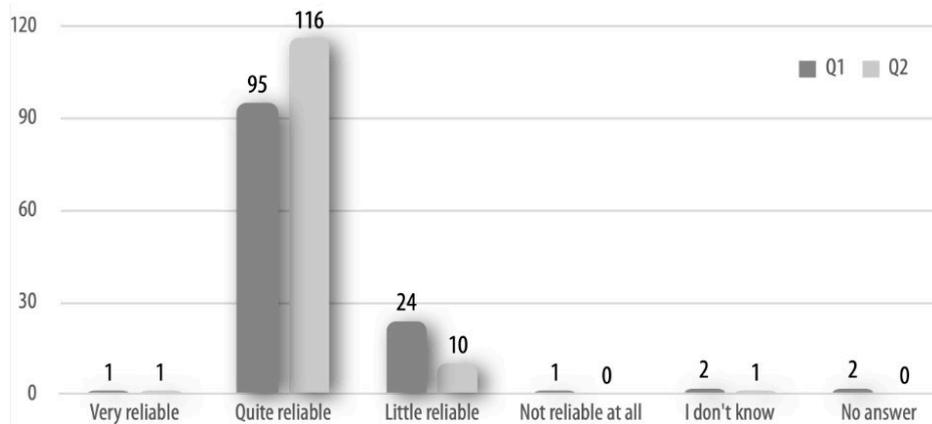


Figure 7: How reliable is the output produced by MT systems in your opinion?

3.1 Challenges for MT

Given the module's focus on the challenges for MT in terms of correctness, accuracy and fluency, a specific question on this topic was included in the end-of-course questionnaire. The students were asked which language features are more likely to be translated effectively based on the evidence gathered during activities and assignments. The selected phenomena were general lexis, specialized terminology, neologisms, polysemic words, lexical non-equivalence, collocations, long noun phrases, wordplays and culture-bound words or expressions. The results, reported in Figure 9, indicate that the students expect MT systems to be quite

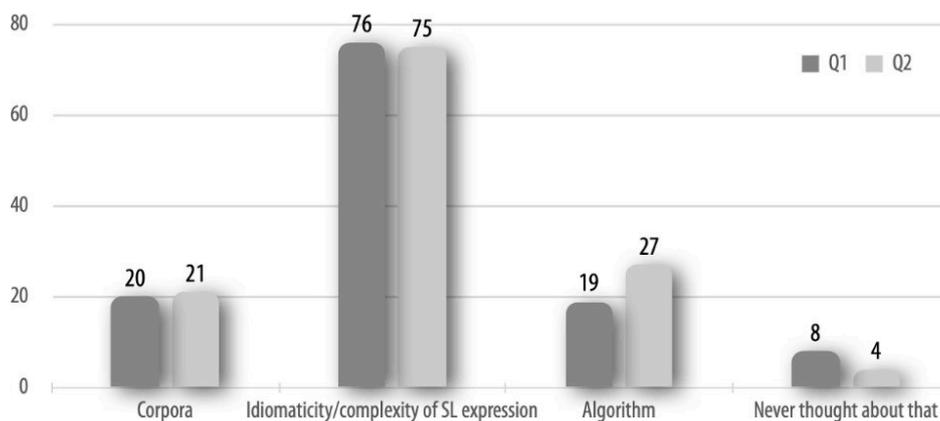


Figure 8: When MT output does not seem correct, what might be the reason for the mistake?

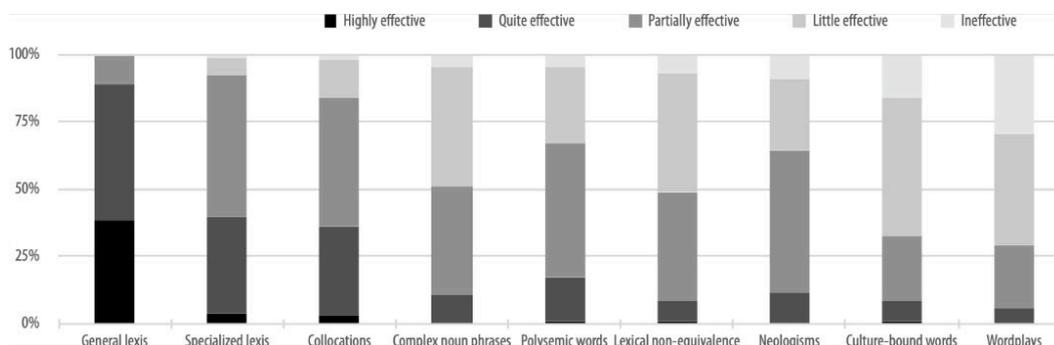


Figure 9: After attending the course, how effective do you think MT is in translating these phenomena?

effective, if not highly effective, with general-purpose lexis. They believe that MT is moderately effective with specialized terminology and collocations, while when it comes to complex noun phrases, polysemic words, lexical non-equivalence and neologisms, most respondents expect only partially or little reliable outcomes that will almost certainly require post-editing. Finally, regarding culture-bound expressions and wordplays, the students tend to believe that these phenomena do not lend themselves to be translated through current MT systems, which produce only minimally effective or totally ineffective output.

The students were asked to elaborate on this question through an open comment on what language aspects are likely to require more post-editing and why. Often cited reasons are the tendency of MT systems to translate expressions literally or the inability to correctly interpret the co-text. Inadequate sources of textual data to train MT systems and lack of creativity are also mentioned. Here are some extracts from students' answers:

- 1) *Perché i sistemi di traduzione automatica non riescono a riconoscere quando una frase o una coppia di parole è idiomatica ma traducono letteralmente.* [Because machine translation systems are not capable of understanding when a sentence or a pair of words is idiomatic, and they just translate it literally.]
- 2) *Neologismi, perché probabilmente non sono presenti nei database delle AI, quindi necessitano di una revisione.* [Neologisms, because they are probably not included in the databases of AI systems; hence they need post-editing.]

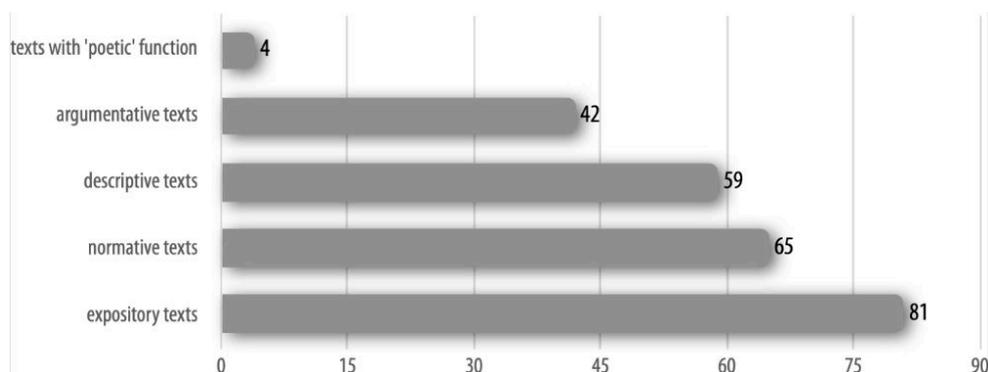


Figure 10: After taking the course, which text typologies in your opinion are effectively translated by MT systems?

- 3) *Sicuramente i giochi di parole e le parole connotate culturalmente [...] per cui serve una creatività che l'IA non possiede.* [Definitely wordplays and culturally connoted words... for which creativity is needed, which is not possessed by AI.]

Another question introduced *ad hoc* in Q2 intended to assess whether the students could form reasonable expectations about the fluency and accuracy of MT output across different text typologies. Figure 10 shows that the students believe that argumentative texts and those with a “poetic” function are more likely to require post-editing; by contrast, more codified and less “creative” typologies, like the normative and the expository ones, tend to yield better results. Although sweeping generalizations should be avoided, the students’ perception seems realistic and finds support in previous studies:

MT performs better for standard, repetitive texts featuring characteristic terms and phrases (which is typical for legislative texts), while it does not give equally good results for texts containing new terminology or rare words, including idioms, metaphors or proper names (which occur more often in non-legislative texts). (Stefaniak 2020: 266).

The following question, only present in Q2, is about the degree of reliability of MT systems when translating isolated words *versus* full textual sequences. This was an open question, and the students were asked to motivate their answer. Of the responses obtained, 73% state that more reliable results are obtained when translating full textual sequences. Almost invariably,

the students refer to the importance of the context for MT to better disambiguate meanings and offer more appropriate results within a given string. However, students are also aware that the length of textual sequences may affect the output, with long stretches not always being translated consistently and cohesively. On the other hand, among the students who answered that MT gives better results with isolated words some argued that this way of using MT is more effective because online platforms offer the user a range of possible options to choose from, which promotes greater autonomy for the human translator. Following are examples of student responses:

- 4) *Ritengo funzionino meglio con sequenze testuali, non necessariamente troppo lunghe, in modo da poter fornire all'AI maggiore contesto [...].* [I believe they work better with full text sequences, albeit not necessarily too long, because they provide AI with more context [...]]
- 5) *Se si cerca una singola parola, di questa sono date più versioni, traduzioni (soprattutto in Google translate), che possono essere utili a una persona con un livello linguistico avanzato.* [If you are looking for a single word, you are given more versions, more translations (especially in Google Translate), which may be useful for someone with an advanced language level.]

3.2 Advantages and disadvantages of AI and MT

In both Q1 and Q2, participants were asked to list the positive and negative implications of the use of AI for natural language processing. The most frequently cited advantage is undoubtedly the speed with which computers can perform operations (Fig. 11). In Q1 the respondents tended to mention speed without referring to any particular task (e.g., “*velocità di risposta*”); by contrast, in Q2 the general understanding — albeit not encouraged by the lecturer — was that the advantages had to be interpreted in relation to MT, following the activities carried out during the module. Hence, most answers make explicit reference to translation and are more elaborate than those in Q1, showing a more nuanced and critical perspective:

- 6) *L'uso dell'intelligenza artificiale può essere utile per avere una prima traduzione veloce del testo di partenza e per potersi concentrare maggiormente sui passaggi che si ritengono più complicati.* [The use of artificial intelligence can be useful to have a quick first translation of the source text and to be able to concentrate more on complex passages.]

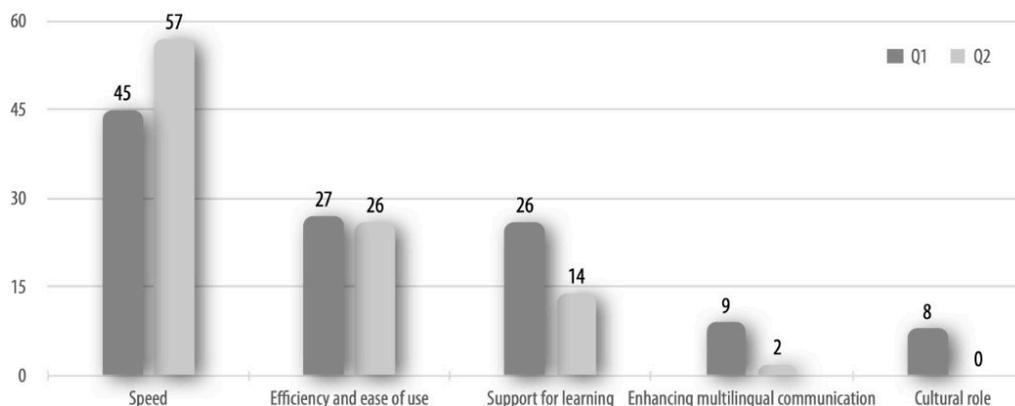


Figure 11: Indicate a positive aspect of AI in relation to language

- 7) *Sicuramente la velocità di esecuzione e di emissione di output sufficientemente corretti. A parte questo, se i database delle lingue target non sono aggiornati, non vedo molte note positive. Sono strumenti utili, sì, ma comunque ancora parzialmente fallaci.* [Definitely processing speed and sufficiently correct outputs. Apart from that, if the databases of the target languages are not up to date, I do not see many positive aspects. Of course, they are useful tools, but they remain partially flawed.]

In Q1 and Q2, a similar number of responses mention the efficiency and ease of use of AI tools as other advantages. In Q1, efficiency is related to the ability of machines to process large amounts of data, to make it easier for people to perform tasks and to do so through user-friendly applications. The fact that machines do not suffer drops in productivity is also mentioned. In Q2, efficiency is predominantly understood as increased productivity for translators, with MT systems allowing them to simplify their job and better concentrate on details:

- 8) *Un aspetto positivo è il fatto che siano un ottimo punto di partenza per traduzioni. La mano dell'uomo è fondamentale, ma comunque i traduttori riescono quasi sempre a fornire l'idea generale.* [A positive aspect is that they are a very good starting point for translation. The role of the human translator remains crucial, but MT systems almost always manage to provide the general idea.]

Much like the answers about speed, those referring to efficiency, too, appear more elaborate in Q2, underscoring the fundamental part of human post-editing and the importance of language competence to identify inaccuracies and mistakes:

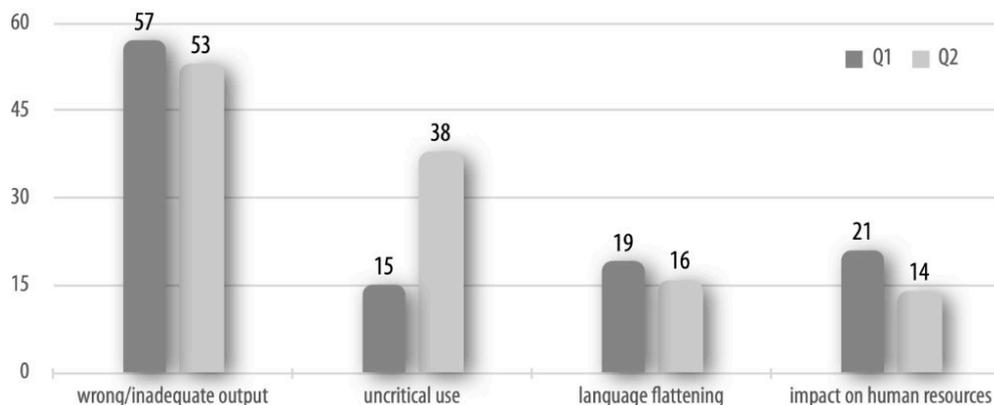


Figure 12: Indicate a negative aspect of AI in relation to language

- 9) *Penso possa essere sempre utile, soprattutto se si conosce la lingua straniera in cui si va a tradurre, altrimenti si rischia di copiare e incollare testi dove le frasi sono collegate senza una logica. Quando si ha una certa padronanza della lingua invece possono essere utili perché si individuano gli errori.* [I think it can always be useful, especially if you know the foreign language into which you are translating, otherwise you risk copying and pasting texts in which sentences are linked without any logic. When you have a certain command of the language, on the other hand, they can be useful because you can spot mistakes.]

The third advantage of AI, according to the students, is that it can support foreign language learning. This theme recurs more frequently in Q1 due to the broader understanding of AI in the preliminary questionnaire (examples 10-12).

- 10) *Potrebbe essere un valido supporto all'apprendimento di una nuova lingua.* [It could be a valuable support in learning a new language.]
11) *Correzione della pronuncia, miglioramento del lessico.* [Correcting pronunciation, improving vocabulary.]
12) *Imparare in maniera più efficace una lingua straniera e poterla confrontare con la lingua madre.* [Learning a foreign language more effectively and being able to compare it with your mother tongue.]

Finally, especially in Q1, some respondents suggested that AI can enhance communication across languages and perform what may be considered a “cultural” function, by granting access to information and promoting education:

- 13) *Ampliare la comunicazione tra parlanti di lingue diverse.* [Extending communication between speakers of different languages.]
- 14) *Può facilitare l'accesso all'informazione e all'educazione.* [It can facilitate access to information and education.]

Turning to the negative implications of AI (Fig. 12), the respondents identified four macro-issues: the production of incorrect or inadequate output; the uncritical use of AI systems; the risk of language flattening; and the negative impact that AI may have on certain professions. Output shortcomings are the most frequently cited problem in both questionnaires although the focus on this aspect slightly diminishes in Q2. In general, erroneous or inadequate translations are considered a drawback because they may disseminate incorrect information or because they still require considerable human effort.

In Q2 greater attention is paid to “how” AI is used. Responses underline that AI may cause mental laziness and affect the development of analytical skills if not supported by adequate training or knowledge (examples 15 and 16); it may encourage a shift of responsibility to the machine and a dependence on its services (example 17); it may also lead to superficial or erroneous uses, especially if accompanied by lack of language competences (example 18).

- 15) *Un maggior impigritimento di chi studia lingue, soprattutto nei novizi. Una tale velocità di esecuzione non credo sia utile per chi deve ancora sviluppare le proprie competenze linguistiche.* [Greater laziness among language learners, especially novices. I do not think such processing speed is useful for those who have yet to develop their language skills.]
- 16) *Diminuzione nella capacità di ragionamento ed analisi autonome.* [Diminished ability to think independently and carry out autonomous analyses.]
- 17) *Deresponsabilizzazione umana, totale affidamento all'intelligenza artificiale.* [Shift of responsibility away from human beings; complete reliance on artificial intelligence.]
- 18) *Mancanza di competenze linguistiche appropriate da parte della persona che traduce, nell'utilizzo di questi strumenti, poiché è necessario un livello*

alto di conoscenza della lingua. [Lack of adequate language skills on the part of the person who is translating through these tools, as a high level of language proficiency is required.]

AI and MT are believed to have a negative impact on languages themselves. In both questionnaires, the respondents expressed concerns for diminished lexical variety, excessive standardization, and the spread of forms of “machine translationese” (Vanmassenhove *et alii* 2021) or “post-edited” (Toral 2019).

- 19) *L'AI normalizza e rende standard molte frasi, e anche molte parole che potrebbero avere sinonimi utilizzabili, in contesti leggermente diversi potrebbero essere tradotte allo stesso modo [...].* [AI normalizes and standardizes many sentences; and many words that could have usable synonyms might be translated always in the same way even in slightly different contexts [...]]
- 20) *[...] il rischio potrebbe essere perdere le caratteristiche peculiari del singolo parlante tendendo a una lingua "globalizzata".* [[...] the risk might be that of losing the individual characteristics of each single speaker producing an increasingly “globalized” language.]
- 21) *Il traduttore potrebbe in qualche modo essere influenzato dalla traduzione proposta dalla MT e dunque l'esito potrebbe risolversi in un appiattimento della ricchezza lessicale.* [The translator might somehow be influenced by the version offered by the MT system, so the output may be characterized by poorer lexical variety.]

Another negative consequence of AI is that it may render human work unnecessary, thus affecting professions such as that of translator. Attention to this risk is more marked in Q1 than in Q2. This result may be partly due to the greater focus on uncritical uses of AI in Q2, and perhaps also to a heightened awareness of how much human post-editing in MT is still required, particularly in the case of professional translation.

4. Discussion

The students’ profile emerging from both questionnaires is that of respondents for whom AI and MT are an integral part of their daily and study activities. The sample comprises students who have a good or advanced proficiency level of English (i.e., the source language) at least

based on their self-assessed abilities. Being foreign language specialists, they possess adequate metalinguistic knowledge and feel quite equipped with the skills needed to avoid an uncritical usage of AI technology and MT. Already in Q1, students showed to be conscious of potential risks of AI and MT for both users and languages themselves. Faced with this population, the teaching activities did not have an “eye-opening” impact but helped the students fine-tune their initial perceptions, dispel some misconceptions, and better articulate their views in the open answers.

The learning goals that seem most fully achieved are the recognition of the challenges that MT has yet to meet in relation to output correctness, appropriateness and fluency, as well as the awareness that different genres may be translated with varying levels of accuracy. The students also became more familiar with the mechanisms underlying MT, which made them attentive to the corpora selected to train programmes and to the algorithm used. One result worth mentioning is the students’ change of perspective in Q2 regarding the negative sides of AI and MT, which are seen as more strongly connected to inexperienced or superficial usage of these tools and less dependent on intrinsic factors.

Post-editing was practiced during classroom activities although, admittedly, the limited number of hours allocated to the module (i.e., 18) only allowed initial familiarization with this task. Hence, in the future, more time should be devoted to the training of post-editing skills and specific objectives should be set also considering the students’ future professional profile. Similarly, it was only possible to scratch the surface of sensitive issues such as the implications of MT for socio-linguistic variation and gender representation, two topics that deserve greater attention in next editions of the module.

Conclusion

As the questionnaires showed, present-day foreign language university students exploit AI technology as part and parcel of their academic activities. This finding is confirmed by Jiménez-Crespo (2017: 190), who observes that software tools such as Google Translate, DeepL or Reverso “are here to stay; current language learners who are digital natives [...] already use them”. However, they may not be fully aware of the risks and limitations of such digital resources; they may employ MT programmes in

a way that negatively affects the development of their language skills, textual analytical abilities, or translation competence. Therefore, it becomes essential to face the challenges posed by AI and MT to foreign language learning and translation training turning these resources into useful pedagogical allies. For instance, Lee (2020) reports that the integration of MT in the foreign language classroom may have positive effects on the development of writing skills, helping students think of writing as a process. Similarly, Kenny and Doherty (2014) and Mellinger (2017) argue for the inclusion of MT in the translation curriculum in a way that empowers rather than marginalises translators. For these reasons, ignoring today's role of AI and MT in foreign language teaching and translation training means losing the opportunity of teaching students how to best exploit this technology or not fully preparing them for an evolving language market which increasingly incorporates AI and MT in its workflows (Gaspari *et alii* 2015).

References

All the online resources in the References were last accessed on 29 July 2021.

Baeza-Yates Ricardo (2018). "Bias on the web". *Communications of the ACM*, 61/6, 54-61, DOI:10.1145/3209581.

Gaspari Federico, Almaghout Hala, Doherty Stephen (2015). "A survey of machine translation competences: Insights for translation technology educators and practitioners". *Perspectives, Studies in Translation Theory and Practice*, 23/3, 333-358, DOI: 10.1080/0907676X.2014.979842.

Jiménez-Crespo Miguel (2017). "The role of translation technologies in Spanish language learning". *Journal of Spanish Language Teaching*, 4, 181-193.

Kenny Dorothy, Doherty Stephen (2014). "Statistical machine translation in the translation curriculum: Overcoming obstacles and empowering translators". *The Interpreter and Translator Trainer*, 8/2, 276-294, DOI: 10.1080/1750399X.2014.936112.

Lee Sangmin-Michelle (2020). "The impact of using machine translation on EFL students' writing". *Computer Assisted Language Learning*, 33/3, 157-175, DOI: 10.1080/09588221.2018.1553186.

Mellinger Christopher D. (2017). "Translators and machine translation: Knowledge and skills gaps in translator pedagogy". *The Interpreter and Translator Trainer*, 11/4, 280-293, DOI: 10.1080/1750399X.2017.1359760.

Stefaniak Karolina (2020). "Evaluating the usefulness of neural machine translation for the Polish translators in the European Commission". In: André Martins, Helena Moniz, Sara Fumega, Bruno Martins, Fernando Batista, Luisa Coheur, Carla Parra, Isabel Trancoso, Marco Turchi, Arianna Bisazza, Joss Moorkens, Ana Guerberof, Mary Nurminen, Lena Marg, Mikel L. Forcada (eds). *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, Lisboa, Portugal: European Association for Machine Translation, 263-269, <https://eamt2020.inesc-id.pt/proceedings-eamt2020.pdf>

Toral Antonio (2019). "Post-editeese: An Exacerbated Translationese". Eprint: arXiv:1907.00900, <https://arxiv.org/abs/1907.00900v2>

Vanmassenhove Eva, Shterionov Dimitar and Gwilliam Matthew (2021). "Machine translationese: Effects of algorithmic bias on linguistic complexity in machine translation". In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, April 19-23, 2021, 2203-2213, <https://aclanthology.org/2021.eacl-main.188.pdf>

