

Alma Mater Studiorum Università di Bologna

SCUOLA DI LINGUE E LETTERATURE, TRADUZIONE E
INTERPRETAZIONE

Sede di Forlì

Corso di Laurea magistrale in Interpretazione (classe LM - 94)

Tesi di Laurea in

Traduzione tra l'italiano e l'inglese I (seconda lingua)

Google Translate e Microsoft Translator – Valutazione di due applicazioni per la traduzione automatica del parlato e analisi di una tecnologia in evoluzione.

CANDIDATA:

Federica Russo

RELATRICE:

Prof.ssa Silvia Bernardini

CORRELATRICE:

Prof.ssa Claudia Lecci

Anno Accademico 2015/2016

Primo Appello

Indice

Introduzione	1
---------------------	---

Capitolo I

La traduzione automatica o Machine Translation	5
1.1 Introduzione.....	5
1.2 Cenni di storia della traduzione automatica	7
1.3 Architetture dei sistemi di traduzione automatica	12
1.4 Utilità e impieghi della traduzione automatica.....	16
1.5 I principali utenti dei servizi di traduzione automatica	20
1.5.1 Introduzione	20
1.5.2 Istituzioni, organizzazioni internazionali e agenzie governative	20
1.5.3 Grandi aziende, multinazionali e agenzie di traduzione	24
1.5.4 Traduttori professionisti	27
1.5.5 Traduttori occasionali e utenti del pubblico generale.....	29
1.6 I limiti della traduzione automatica	30
1.6.1 Premessa.....	30
1.6.2 Come affrontare i limiti della traduzione automatica: pre-editing, post-editing e linguaggio controllato.....	33
1.7 La qualità nella traduzione automatica.....	35
1.7.1 Premessa.....	35
1.7.2 Metodi per la valutazione della qualità di una traduzione automatica ..	36
1.7.3 Valutare la qualità di un sistema di traduzione automatica	39
1.8 Dalla traduzione automatica alla traduzione automatica del parlato.....	40

Capitolo II

La traduzione automatica del parlato o Speech-to-Speech Translation	41
2.1 Introduzione.....	41
2.2 La traduzione automatica del parlato o Speech-to-Speech Translation	42
2.3 Le principali tappe storiche della traduzione automatica del parlato	44
2.4 Funzionamento dei sistemi di traduzione automatica del parlato	49

2.5	Problematiche e sfide della traduzione automatica del parlato.....	53
2.5.1	Analisi delle problematiche e delle sfide della traduzione automatica del parlato	53
2.5.2	La cooperazione dell'utente come strumento per sopperire ai limiti dei sistemi di traduzione automatica del parlato	56
2.6	Applicazioni pratiche della tecnologia SST.....	59
2.7	Prospettive e sfide future per la Speech-to-Speech Translation	62
2.8	La valutazione della qualità nella traduzione automatica del parlato	63

Capitolo III

<i>Google Translate e Microsoft Translator: Valutazione di due applicazioni per la traduzione automatica del parlato</i>		67
3.1	Introduzione	67
3.2	Obiettivi del test.....	68
3.3	Applicazioni oggetto di valutazione	70
3.3.1	Applicazioni scelte.....	70
3.3.2	Google Translate	70
3.3.3	Microsoft Translator	74
3.4	Materiale utilizzato per il test	78
3.5	Partecipanti	80
3.6	Svolgimento del test.....	81
3.7	Valutazione dei risultati	83
3.7.1	Premessa	83
3.7.2	Valutazione dell'output del modulo ASR	84
3.7.3	Valutazione dell'output del modulo MT.....	85
3.7.3.1	Scelta dei valutatori	87
3.7.3.2	Scelta dei parametri di valutazione e preparazione dei questionari di valutazione.....	88
3.8	Presentazione dei risultati	93

Capitolo IV

Risultati del test di valutazione		95
4.1	Introduzione	95

4.2	Compilazione dei risultati relativi al riconoscimento vocale in GT e MsT	.95
4.3	Risultati del riconoscimento vocale in GT e MsT	99
4.4	Valutazione dell'output tradotto da GT e MsT	104
4.5	Accordo fra i valutatori	105
4.5.1	Analisi dell'accordo fra i valutatori	105
4.5.2	Discussione delle potenziali cause di disaccordo fra i valutatori e trattamento delle risposte indicanti disaccordo	107
4.6	Risultati della valutazione degli output tradotti	109

Capitolo V

	Discussione dei risultati	111
5.1	Introduzione	111
5.2	Discussione dei risultati del riconoscimento vocale	111
5.2.1	Introduzione	111
5.2.2	Confronto fra le performance dei moduli ASR di GT e MsT	112
5.2.3	La punteggiatura nell'output ASR prodotto da GT e MsT	114
5.2.4	Il riconoscimento vocale dei toponimi	116
5.2.5	Il riconoscimento vocale per il partecipante 1 e il partecipante 2	118
5.3	Discussione dei risultati relativi alla valutazione degli output tradotti	121
5.3.1	Introduzione	121
5.3.2	<i>Fluency</i> ed equivalenza semantica negli output tradotti da GT e MsT	122
5.3.3	Incidenza del riconoscimento vocale sulla qualità dell'output tradotto	124
5.3.4	<i>Fluency</i> ed equivalenza semantica nell'output inglese e nell'output tedesco	126
5.3.5	Analisi delle principali problematiche riscontrate nelle traduzioni	129

Capitolo VI

	Conclusioni	133
6.1	Contenuti principali dell'elaborato	133
6.2	Conclusioni raggiunte	135
6.3	Prospettive future	138

Bibliografia	141
Abstract EN	145
Abstract DE	146
Appendice 1	149
Appendice 2	157
Appendice 3	181
Appendice 4	207
Ringraziamenti	217

Introduzione

Certamente la traduzione ha sempre rivestito un ruolo importante nella storia dell'umanità, in virtù della sua capacità di soddisfare il fondamentale bisogno di comunicazione dell'uomo. Oggi, in un mondo sempre più globalizzato e interconnesso, dove secondo le stime dell'UNESCO sono circa 7000 le lingue parlate, essa sembra essere più fondamentale che mai.

Pur mantenendo invariata la sua importanza nel corso degli anni, parallelamente alla rapida evoluzione tecnologica iniziata con la nascita e la diffusione dei primi sistemi informatici, anche la traduzione in quanto disciplina ha subito un profondo cambiamento. Se un tempo erano carta, penna e dizionario ad accompagnare il traduttore nel suo lavoro, oggi sono il pc e la *translator's workbench* ad assistere chi si appresta a tradurre un qualsiasi testo. Ma non solo: a fianco della traduzione assistita, che trova la sua espressione tipica nell'utilizzo di software per la traduzione e di memorie di traduzione, anche la traduzione automatica si è fatta strada nel panorama del mondo traduttivo.

Nata dall'iniziativa di alcuni gruppi di ricerca intorno alla metà del ventesimo secolo, la traduzione automatica o Machine Translation ha attraversato fasi di crescita, dettate dal grande entusiasmo dei ricercatori attivi in questo settore, e periodi di arresto, per affermarsi oggi come una tecnologia dalle prestazioni decisamente superiori rispetto al passato e che, seppur lontana dalla perfezione, trova applicazione in diversi scenari ed è sempre più diffusa.

Dall'iniziale diffidenza diffusa fra traduttori e addetti ai lavori nei confronti dei primi software di traduzione automatica, si è passati oggi ad un crescente interesse, dettato da un lato dall'evidente miglioramento in termini di qualità delle traduzioni prodotte, e dall'altro dalla consapevolezza del fatto che nessuna macchina potrà mai sostituire completamente l'uomo nel processo di traduzione. Al contrario, la traduzione automatica, portando con sé nuovi bisogni, sembra aprire nuove possibilità di guadagno per il traduttore professionista, certamente la figura più qualificata per occuparsi di post-editing. Ma ancor più dei traduttori, sono organizzazioni internazionali e

aziende multinazionali a nutrire un interesse spiccato nei confronti della traduzione automatica, scorgendo in questa tecnologia un'enorme possibilità di risparmio in termini di tempistiche e costi. Infine, sono sempre di più i traduttori occasionali e gli utenti del pubblico generale che fanno ricorso alla Machine Translation per avere accesso a contenuti in lingua straniera o per poter comunicare con interlocutori stranieri. Per questa categoria di utenti sono i servizi gratuiti di traduzione automatica online quali Google Translate o Bing Translator, per citarne alcuni, o le applicazioni per la traduzione automatica, a rappresentare la principale fonte di interesse.

Il presente elaborato è dedicato proprio alla Machine Translation, una tecnologia in evoluzione che si declina ormai in diverse forme. In particolare, la prima parte dell'elaborato sarà incentrata sulla traduzione automatica di testi scritti, ovvero la Machine Translation nella sua forma originale. Di questa tecnologia verranno descritte l'evoluzione in termini storici, il funzionamento, i possibili impieghi, nonché le principali categorie di utenti. Verranno analizzati i vantaggi offerti dalla possibilità di tradurre automaticamente testi scritti, ma anche i limiti e le problematiche da essa derivanti. Infine ci si concentrerà sulla qualità, un aspetto di fondamentale importanza per qualsiasi traduzione, sia essa "manuale" o automatica. A questo proposito verranno esposte le diverse modalità di valutazione della qualità nel campo della traduzione automatica, una pratica fondamentale per garantire il progresso nella ricerca sulla Machine Translation.

Dalla traduzione automatica di testi scritti si passerà alla traduzione automatica di testi orali, seguendo la naturale evoluzione della Machine Translation e approdando così alla Speech-to-Speech Translation, ovvero la traduzione automatica del parlato, una tecnologia che consente di tradurre automaticamente input vocali, che vengono poi a loro volta sintetizzati, ovvero riprodotti vocalmente.

Questa tecnologia altamente innovativa rappresenta un ulteriore sviluppo della traduzione automatica, nonché il vero e proprio fulcro del presente studio. Il secondo capitolo di questo elaborato fornirà una definizione di Speech-to-Speech Translation, ne tratterà le principali tappe storiche e ne descriverà il funzionamento. Verranno poi analizzate le principali problematiche inerenti alla traduzione automatica della lingua parlata, in

contrapposizione a quelle tipiche della traduzione automatica “tradizionale”, incentrata su testi scritti, e i potenziali scenari di utilizzo, presenti e futuri, dei software di Speech-to-Speech Translation finora sviluppati. Verranno inoltre delineate sia le sfide future che i ricercatori attivi in questo campo dovranno affrontare, sia le prospettive di sviluppo di una tecnologia nata in tempi relativamente recenti. Infine, come per la traduzione automatica, verrà analizzato l’aspetto relativo alla valutazione della qualità.

Nell’ultima parte dell’elaborato verrà poi presentato un test di valutazione condotto su Google Translate e Microsoft Translator, due applicazioni per la traduzione automatica disponibili gratuitamente per smartphone e tablet. Tali applicazioni verranno valutate in merito al servizio di Speech-to-Speech Translation da esse offerto, secondo alcune fra le tante modalità di valutazione della qualità descritte nei capitoli precedenti. L’obiettivo del test presentato in questo studio è quello di verificare il grado di affidabilità e accuratezza della traduzione automatica del parlato offerta dalle applicazioni valutate, nonché quello di fornire un’indicazione sull’effettiva utilità di questi strumenti per un utente che se ne serva per interagire con interlocutori multilingue, pur non parlando la loro lingua.

Sulla scorta di due esempi concreti, i risultati del test di valutazione, esposti negli ultimi capitoli dell’elaborato, porteranno alla scoperta delle potenzialità e dei limiti di una tecnologia in evoluzione, che, al pari della traduzione automatica online, potrebbe arrivare a diffondersi fino a coprire piccole, ma importanti fette di mercato, soddisfacendo bisogni ad oggi non soddisfatti da interpreti e traduttori umani.

I

La traduzione automatica o Machine Translation

1.1 Introduzione

Il primo capitolo di questo elaborato è incentrato sulla traduzione automatica, che verrà analizzata sotto diversi punti di vista. Prima però di procedere a questa analisi, occorre definire l'oggetto di ricerca.

La traduzione automatica è qualcosa che ormai ci circonda, pur passando spesso inosservata. Basti pensare a quante volte, navigando su internet da casa o in ufficio, si ricorre alla traduzione automatica per le pagine web in lingua straniera: grazie a servizi come Google Translate o Bing Translator, con un solo click si ha accesso quasi istantaneo a una traduzione dei contenuti della pagina nella propria lingua madre. Come è facile intuire, non si tratta di una traduzione “tradizionale”, frutto del lavoro di ore e ore di un traduttore professionista. Al contrario, è il prodotto di uno speciale software, per l'appunto, di traduzione automatica. Intuitivamente, è dunque facile distinguere la traduzione automatica da quella classica: si tratta di una traduzione effettuata da un software, da una macchina, invece che da un essere umano.

Ai fini di questa indagine è tuttavia necessario definire l'oggetto di ricerca in modo oggettivo e scientifico. Una definizione abbastanza dettagliata di traduzione automatica (in inglese *Machine Translation*), intesa come disciplina di studio, viene fornita nella *Routledge Encyclopedia of Translation Technology* a cura di Sin-Wai Chan. Secondo tale definizione:

Machine Translation (MT) is a sub-field of computational linguistics (CL) or natural language processing (NLP) that investigates the use of software to translate text or speech from one natural language to another. The core of

MT itself is the automation of the full translation process, which is different with the related terms such as machine-aided human translation (MAHT), human-aided machine translation (HAMT) and computer-aided translation (CAT).

(Liu, Zhang, 2015: 105)

In questa definizione, gli autori distinguono fra traduzione automatica pura, ovvero quella che prevede una completa automazione del processo di traduzione, e altri tipi di traduzione che prevedono invece l'interazione fra esseri umani e macchine, come la MAHT, la HAMT e la CAT.

La definizione fornita da Hutchins & Somers (in Gaspari, 2011: 14) si discosta invece da quella di Liu e Zhang per la mancanza di una divisione così netta. Hutchins & Somers, infatti, includono nella definizione di MT tutti i “sistemi computerizzati responsabili della produzione di traduzioni da una lingua naturale all'altra, con o senza assistenza da parte di persone” (*ibid.*).

Entrambe le definizioni citate pongono l'accento sul fatto che i sistemi di traduzione automatica si occupano della traduzione di lingue naturali, da intendersi in contrapposizione ai linguaggi artificiali o di programmazione (*ibid.*: 14). Inoltre, se la definizione fornita da Hutchins & Somers non specifica che la traduzione automatica debba limitarsi esclusivamente a testi scritti, la definizione fornita più recentemente da Liu e Zhang lo rende esplicito, includendo nella MT sia la traduzione di testi scritti, sia quella di testi orali. Infatti, come si vedrà nel capitolo II, da alcuni anni la traduzione automatica può essere applicata anche alla lingua parlata.

Nel presente elaborato, per motivi di chiarezza, le espressioni “traduzione automatica” e “MT” verranno utilizzate per indicare la disciplina o la traduzione automatica di testi scritti, in contrapposizione alla traduzione automatica del linguaggio parlato, a cui si farà riferimento come “SST” (dall'inglese *Speech-to-Speech Translation*) o come “traduzione automatica del parlato”.

1.2 Cenni di storia della traduzione automatica

Di seguito verranno illustrate le principali tappe della storia della traduzione automatica, una storia “fatta di alti e bassi, di grandi aspirazioni e di fortissime disillusioni” (Chiari, 2011: 31).

L'idea di servirsi di macchine per tradurre il linguaggio naturale è molto antica: risale addirittura al XVII secolo (*ibid.*), quando si diffusero i concetti di lingua universale e di dizionario meccanico (Hutchins, 2010). Tuttavia, è solo nel XX secolo che ebbe inizio la vera e propria storia della traduzione automatica, con i brevetti del francese George Artsrouni e del russo Petr Smirnov-Trojanskij, depositati nel 1933. Il brevetto di Artsrouni riguardava una sorta di dizionario meccanico multilingue, mentre quello di Trojanskij si spingeva oltre, proponendo un dispositivo per la traduzione multilingue che sfruttava un metodo per la codifica e la decodifica delle funzioni grammaticali basato sull'esperanto (Hutchins, 2010). Trojanskij fu un autentico precursore della traduzione automatica, anche se la sua proposta non ebbe molta risonanza al di fuori dei confini russi (Hutchins, 1995).

Del tutto all'oscuro dei brevetti di Artsrouni e Trojanskij, nel 1946 e nel 1947, il cristallografo inglese Andrew Booth e il matematico statunitense Warren Weaver discussero della possibilità di utilizzare il computer, da poco inventato, per la traduzione dei linguaggi naturali (Hutchins 1995; Hutchins, 2010). Fu però solo nel luglio 1949 che, con la pubblicazione del memorandum di Warren Weaver, la traduzione automatica catturò l'attenzione di diversi studiosi e la ricerca negli Stati Uniti iniziò a farsi più consistente. Nel suo memorandum Weaver, basandosi sulle sue conoscenze in materia di crittografia, statistica, teoria dell'informazione, logica e universali linguistici, avanzava proposte specifiche per risolvere il problema dell'ambiguità del linguaggio (Hutchins, 2010).

Nel 1952, presso il Massachusetts Institute of Technology (MIT), si tenne la prima conferenza sulla traduzione automatica, che vide la partecipazione di quasi tutti i ricercatori attivi in questo campo. Già in questa sede gli esperti del settore convennero che la FAHQT (Fully Automatic High Quality Translation) rappresentava un obiettivo irraggiungibile, dal momento che

l'intervento umano, nella fase precedente alla traduzione (pre-editing) o in quella successiva (post-editing), sembrava essere inevitabile, anche se alcuni speravano che si trattasse di una necessità temporanea (*ibid.*).

Un grande impulso alla ricerca fu dato dal cosiddetto “Georgetown experiment”, condotto il 7 gennaio 1954 presso la Georgetown University (*ibid.*), quando “con un vocabolario di sole 250 parole e sole 6 regole grammaticali fu tradotto in inglese un campione selezionato di frasi russe” (Riediger, Galati, 2012: 7). Questa dimostrazione ebbe un grande impatto sull'opinione pubblica, tanto da stimolare l'ingente finanziamento della ricerca sulla MT negli USA e l'inizio di progetti simili anche in altri paesi, fra cui l'URSS (Hutchins, 2010).

Nel 1954 fu anche fondata la prima rivista del settore (*Mechanical Translation*) e venne pubblicata la prima tesi di dottorato sulla traduzione automatica (redatta da Anthony G. Oettinger). Gli anni successivi videro la nascita di diversi gruppi di ricerca sulla MT in diverse parti del mondo, fra cui l'Italia (il gruppo milanese diretto da Silvio Ceccato; Hutchins, 2010).

Nel primo decennio di ricerca sulla MT vennero sviluppati tre principali approcci:

- il modello di traduzione diretta, che prevedeva una serie di regole per la traduzione da una determinata lingua di partenza (SL, ovvero *Source Language*) verso una determinata lingua di arrivo (TL, ovvero *Target Language*), con una quantità minima di analisi e di riorganizzazione a livello sintattico (*ibid.*);
- il modello interlingua, basato su rappresentazioni astratte indipendenti sia dalla SL che dalla TL. La traduzione, in questo caso, avveniva in due fasi distinte: dalla SL all'interlingua e dall'interlingua alla TL (Hutchins, 2010);
- l'approccio transfer, basato su un modulo di analisi della SL, un modulo di sintesi della TL e su “moduli intermedi (denominati appunto moduli di trasferimento o transfer)” (Naldi, 2014: 60), che realizzavano il passaggio da una lingua all'altra.

Negli Stati Uniti le principali lingue di ricerca nel campo della MT erano l'inglese e il russo, per evidenti ragioni politiche e militari, e solo in una

seconda fase la ricerca si estese anche ad altre combinazioni linguistiche. Centri di ricerca erano diffusi in tutto il paese, sia presso università e istituzioni pubbliche, sia presso multinazionali (come IBM). L'Unione Sovietica aveva accolto la sfida della ricerca sulla MT con altrettanto entusiasmo, ma qui i progressi furono talvolta penalizzati dall'inadeguatezza dei computer e degli strumenti a disposizione (Hutchins, 2010).

Tutti gli anni Cinquanta furono caratterizzati da un forte ottimismo nei confronti della ricerca sulla MT e a metà anni Sessanta si contavano gruppi di ricerca in diversi paesi, oltre a quelli già nominati, come Cina, Messico, Giappone e la maggior parte delle nazioni europee. Come riportato da Hutchins, alla luce degli sviluppi nell'elaborazione dati e nella linguistica formale, per molti ricercatori dell'epoca la creazione di sistemi di traduzione completamente automatici sembrava essere a portata di mano (*ibid.*). Tuttavia, ben presto la complessità delle problematiche linguistiche che ci si trovava ad affrontare divenne sempre più evidente, e il sentimento di entusiasmo sfociò in una crescente delusione (*ibid.*).

Nel 1964 gli enti statunitensi che fino ad allora avevano sponsorizzato la ricerca sulla MT (per lo più agenzie militari e di intelligence) si rivolsero alla National Science Foundation affinché venisse creata una commissione per l'analisi dello stato della ricerca. Venne così creato l'ALPAC (Automatic Language Processing Advisory Committee), che nel 1966 produsse l'omonima relazione "ALPAC Report". In questa relazione il comitato ALPAC descriveva la traduzione automatica come più lenta, più costosa e meno accurata di quella umana; proseguiva suggerendo di concentrare la ricerca sullo sviluppo di dispositivi di supporto ai traduttori e di interrompere il finanziamento della ricerca sulla MT, che non sembrava essere una tecnologia promettente (Hutchins, 2010). Il rapporto dell'ALPAC ebbe un impatto considerevole sul mondo accademico e causò un forte rallentamento della ricerca sulla MT per oltre un decennio sia negli Stati Uniti che altrove.

La ricerca non si fermò del tutto, ma sicuramente ci fu un cambio di rotta: mentre la ricerca pre-ALPAC si era concentrata prevalentemente sull'approccio diretto, la ricerca post-ALPAC fu dominata da modelli indiretti, basati su approccio interlingua o transfer. Fra gli esperimenti di

successo del decennio post-ALPAC spicca il sistema canadese Météo. Questo sistema, in uso a partire dal 1976, fu introdotto per tradurre i bollettini meteorologici fra l'inglese e il francese e venne progettato specificatamente per trattare il vocabolario e la sintassi tipiche di questo dominio (*ibid.*).

La seconda parte degli anni Settanta è anche caratterizzata dalla diffusione di sistemi operativi e commerciali, tra cui si ricordano Systran, Logos e METAL. Il sistema Systran, in particolare, venne acquistato dalla Commissione delle Comunità Europee nel 1976 e venne installato anche presso numerosi enti intergovernativi, inclusi la NATO e l'Agenzia internazionale per l'energia atomica, e aziende, come General Motors. Benché Systran, Logos e METAL fossero pensati per un uso generale, negli anni Settanta e Ottanta vennero sviluppati anche sistemi per settori o domini specifici (Hutchins, 2010).

Gli anni Ottanta furono caratterizzati dalla predominanza dell'approccio transfer, seguito dal modello interlingua, e da un cospicuo aumento della ricerca sulla MT in Giappone, dove moltissime aziende del settore informatico iniziarono a investire anche su questa tecnologia, e in altri paesi dell'Asia, che fino ad allora non si erano interessati alla traduzione automatica (*ibid.*).

A partire dal 1989 la ricerca si concentrò su nuovi metodi di traduzione automatica basati su corpora, ovvero ampie raccolte di testi in formato elettronico. Fra questi metodi si distinguono in particolare la traduzione automatica basata su esempi (EBMT), fondata sull'idea secondo cui tradurre spesso implica un processo di ricerca di esempi analoghi a quello che ci si appresta a tradurre, per verificare come questi siano stati tradotti in precedenza, e la traduzione automatica statistica (SMT), caratterizzata dall'impiego di metodi statistici di analisi e sintesi e dall'assenza di regole linguistiche. La ricerca proseguì anche nella direzione dei sistemi basati su regole, sia con approccio transfer che interlingua, presso diversi gruppi di ricercatori (Hutchins: 2010).

Gli anni Ottanta e Novanta furono anche caratterizzati da un crescente interesse per la traduzione automatica del parlato (si veda capitolo II) e dalla diffusione delle cosiddette *translator's workbench*, le stazioni di lavoro per i traduttori, messe in commercio nel 1991 (Zanettin, 2001: 27). Queste

comprendono diversi strumenti utili al traduttore, come un programma di videoscrittura, software di gestione terminologica, strumenti per la ricerca di concordanze e programmi per la gestione di memorie di traduzione (Hutchins, 2010).

A partire dagli anni Ottanta, inoltre, al concetto di traduzione automatica si è affiancato quello di traduzione assistita (in inglese *Machine Aided Translation*) ed è nata la distinzione fra “traduzione umana assistita dalle macchine (MAHT, *Machine Aided Human Translation*)” e “traduzione automatica assistita dagli umani (HAMT, *Human Aided Machine Translation*)” (Zanettin, 2001: 24).

Negli anni Novanta ha avuto inizio anche la ricerca nel campo della valutazione della MT, inizialmente basata esclusivamente sulla valutazione soggettiva di diversi fattori, fra i quali accuratezza, scorrevolezza e comprensibilità, ma a partire dal 2000 effettuata anche sulla base di metodi automatici o semiautomatici (ad esempio il metodo BLEU introdotto dal gruppo IBM). In questo decennio è aumentato anche l'utilizzo di sistemi di traduzione automatica, soprattutto da parte di multinazionali e enti governativi, anche grazie alla graduale diffusione dei computer (Hutchins, 2010).

Negli anni Novanta il mercato della traduzione automatica è stato ulteriormente influenzato dalla nascita di internet, che, con pagine web e posta elettronica, ha portato con sé una nuova domanda di traduzione. Hanno dunque fatto la prima comparsa servizi di traduzione automatica online, fra cui si ricorda Babel Fish, lanciato nel dicembre 1997 dal motore di ricerca AltaVista. Questo servizio, basato su un sistema Systran, ha regalato alla traduzione automatica una visibilità mondiale, rendendola facilmente accessibile a tutti gli utenti della rete. Babel Fish, pur non essendo il primo servizio di MT reso disponibile online, si distingueva dai tentativi precedenti per il fatto di essere aperto a tutti, senza necessità di alcun tipo di abbonamento (Gaspari, Hutchins, 2007: 200). Il lancio di Babel Fish è stato poi seguito da quello di molti altri servizi di traduzione automatica online, per lo più gratuiti.

All'epoca la qualità delle traduzioni offerte da questo tipo di piattaforma online era piuttosto bassa, ma il vantaggio consisteva nell'incredibile

rapidità del servizio; inoltre, poiché gli utenti di internet si accontentavano spesso di una traduzione approssimativa del senso generale del messaggio, i sistemi di MT online si sono mostrati in grado di soddisfare una domanda di traduzione che altrimenti, con ogni probabilità, sarebbe rimasta insoddisfatta. D'altro canto però, la traduzione automatica online, all'epoca piuttosto scadente, ha influenzato in termini negativi la percezione della traduzione automatica in generale (Hutchins, 2010).

Per quanto riguarda invece le architetture dei sistemi di traduzione automatica, a partire dalla fine degli anni Novanta si è fatta strada l'idea di un sistema ibrido, che combini cioè i metodi statistici della SMT o della EBMT con approcci basati su regole linguistiche (*ibid.*). È in questa direzione che si muove la ricerca al momento della stesura del presente elaborato.

1.3 Architetture dei sistemi di traduzione automatica

I sistemi di traduzione automatica attualmente disponibili dispongono di diversi tipi di architetture, che variano per approccio, complessità ed efficacia. Semplificando, è possibile classificare i sistemi di traduzione automatica in due principali categorie; a seconda dell'architettura su cui questi si basano, si può distinguere fra (Chiari, 2011: 31):

- sistemi di traduzione automatica basati su regole (in inglese *Rule Based Machine Translation*);
- sistemi statistici di traduzione automatica (*Statistical Machine Translation*).

I sistemi che adottano la **traduzione automatica basata su regole** furono i primi ad essere creati negli anni Cinquanta (Gaspari, 2011: 23) e si fondano sulla seguente idea:

uno strumento, per poter svolgere un compito di traduzione, deve possedere un dizionario e una grammatica delle lingue da trattare i più ampi possibile

e una serie di regole di conversione che permettano di passare da una frase in una lingua di partenza [...] a una frase nella lingua di arrivo.

(Chiari, 2011: 31)

I sistemi basati su regole si differenziano ulteriormente in (Gaspari, 2011: 23):

- sistemi con approccio diretto;
- sistemi con approccio transfer;
- sistemi con approccio basato su interlingua.

L'approccio diretto, impiegato dai primissimi sistemi di MT, quelli di prima generazione (Zanettin, 2001: 22), è il più semplice. Si basa sulla “sostituzione delle parole della lingua di partenza con (potenziali) equivalenti traduttivi in quella di arrivo” (Gaspari, 2011: 23). Il contesto non viene considerato da questo tipo di sistema, e, in presenza di parole polisemiche o omografe nel testo di partenza, o di diverse possibilità di traduzione nella lingua di arrivo, il sistema seleziona una soluzione o accezione, in genere la più frequente. Anche a livello sintattico, questo tipo di architettura non prevede una rielaborazione approfondita, e la qualità dell'output così prodotto non è particolarmente elevata (*ibid.*).

L'approccio transfer è più sofisticato rispetto a quello diretto. Come nei sistemi diretti, si parte “dal trasferimento lessicale nella lingua di arrivo delle parole dell'input” (*ibid.*: 24), ovvero del testo di partenza, ma, una volta terminata questa fase, l'output viene adattato in base alle regole morfosintattiche della lingua di arrivo. Il risultato è un testo più scorrevole e naturale rispetto a quello prodotto da un sistema con approccio diretto (*ibid.*).

Il terzo approccio impiegato dai sistemi basati su regole è quello fondato sull'interlingua. “L'interlingua consiste in un linguaggio artificiale di rappresentazione semantica, indipendente tanto dalla lingua di partenza quanto da quella di arrivo” (*ibid.*). In questo tipo di sistema la traduzione avviene in due fasi distinte: il testo di partenza viene convertito nell'interlingua tramite alcune regole, e a partire dall'interlingua viene

generato il testo di arrivo tramite altre regole. Il vantaggio principale offerto da questo tipo di sistema è che, non essendo legato a una specifica combinazione linguistica, è applicabile virtualmente a ogni lingua. Tuttavia, “l'esperienza ha dimostrato che mettere a punto un'interlingua efficace è estremamente difficile” (Gaspari, 2011: 24), perché questa dovrebbe mantenere una certa “equidistanza rispetto alle varie lingue che devono essere elaborate” (*ibid.*).

Come sottolineato da Zanettin, rispetto all'approccio diretto, gli approcci transfer e interlingua (i cosiddetti sistemi di seconda generazione) prevedono “la possibilità di arrivare alla realizzazione di sistemi in grado di gestire più coppie di lingue e direzioni di traduzione” (Zanettin, 2001: 22), ma sono anche accompagnati da una “maggiore complessità gestionale” (*ibid.*).

Un forte limite dell'architettura basata su regole è che le regole linguistiche impiegate sono unidirezionali. Ciò significa che:

un sistema transfer che traduce dall'inglese all'italiano non è reversibile, e per consentire la traduzione nella direzione opposta (dall'italiano all'inglese) è necessario [...] creare le nuove regole che consentano la traduzione nell'altro senso.

(Gaspari, 2011: 25)

Questo implica chiaramente un considerevole investimento ogni qualvolta si intende ampliare la rosa di lingue gestita dal sistema di MT.

La seconda tipologia di MT è rappresentata dai sistemi che adottano una **traduzione automatica statistica**. Questo tipo di sistema, nella sua forma più “pura”

non presuppone alcuna conoscenza linguistica ma funziona solamente in base a metodi di appaiamento statistico, limitandosi a cercare corrispondenze su un corpus di traduzioni esistenti.

(Zanettin, 2001: 23)

I sistemi statistici funzionano grazie a “banche date testuali bilingui” (*ibid.*: 22), ovvero grandi raccolte di testi in formato elettronico e delle rispettive versioni tradotte (corpora paralleli), che vengono allineate a livello di frase.

Le corrispondenze traduttive del testo di partenza vengono individuate grazie a un'analisi probabilistica del corpus nella lingua di partenza e della sua versione tradotta (Gaspari, 2011: 25). La MT statistica si basa cioè su algoritmi che stimano “le probabilità che una o più parole della lingua di partenza vengano tradotte con una o più parole corrispondenti nella lingua di arrivo” (*ibid.*). I candidati traduttivi così individuati vengono poi filtrati e vengono selezionati “quelli effettivamente più plausibili nella lingua di arrivo” (*ibid.*).

La traduzione automatica statistica, sviluppatasi negli anni Novanta e oggi molto diffusa, offre diversi vantaggi. Innanzitutto la realizzazione di un sistema di MT statistico richiede un investimento minore rispetto a quella di un sistema basato su regole. È infatti sufficiente disporre di una quantità adeguata di corpora paralleli allineati, la cui compilazione oggi è resa più facile dalla presenza della rete (anche se occorre precisare che il reperimento di questo tipo di risorsa per lingue poco diffuse o di scarso interesse commerciale può essere difficoltoso; Gaspari, 2011: 26). In secondo luogo, l'output prodotto da un sistema di MT statistico non è altro che “una ricombinazione di traduzioni già esistenti ed effettuate di norma da traduttori competenti” (*ibid.*): risulterà quindi più idiomatico e scorrevole, ma solo laddove esista già un'ampia quantità di testi tradotti in modo tradizionale.

Ovviamente, anche la traduzione automatica statistica, come quella basata su regole, presenta dei limiti. “In particolare va evidenziata la settorialità” (*ibid.*: 26) dei sistemi statistici. Questi sistemi vengono infatti addestrati su corpora paralleli, che tendono però a essere rappresentativi di un numero limitato di tipologie testuali: i corpora per l'addestramento, in genere, si concentrano su determinati settori (come quello giuridico-legale, scientifico, ecc.) e vengono costruiti a partire da tipologie testuali facilmente reperibili in formato elettronico e in più lingue (come leggi dell'Unione europea, manualistica, ecc.). Di conseguenza, una volta addestrato su corpora di questo tipo, il sistema di MT statistico estrarrà traduzioni legate a quella particolare tipologia di testi. La performance del sistema sarà dunque soddisfacente solamente nel caso in cui il testo da tradurre sia analogo alla tipologia testuale contenuta nei corpora usati in fase di

addestramento (Gaspari, 2011: 26).

Per superare queste problematiche, a partire dagli anni Novanta la ricerca è andata concentrandosi sempre di più sui cosiddetti **sistemi ibridi**, basati fondamentalmente su un approccio statistico, che viene però corretto tramite l'applicazione di “regole linguistiche e traduttive esplicite” (*ibid.*) normalmente assenti nei sistemi statistici puri.

1.4 Utilità e impieghi della traduzione automatica

Dopo avere elencato vantaggi e limiti dei diversi sistemi di MT esistenti, è importante sottolineare che non esiste ancora, e probabilmente non esisterà mai, un sistema di traduzione completamente automatico, in grado di produrre traduzioni di alta qualità per qualsiasi tipo di testo (Gaspari, 2011: 15). Allo stato presente, occorre sacrificare almeno uno dei seguenti requisiti:

- alta qualità del testo d'arrivo;
- completa automazione del sistema di MT;
- applicabilità a qualsiasi tipo di testo.

La compresenza di questi tre parametri nell'ambito della traduzione automatica, nonostante i considerevoli progressi tecnologici, è infatti tutt'ora impossibile (*ibid.*).

Alla luce di tutte queste considerazioni, sorge spontaneo chiedersi: la traduzione automatica funziona? E ancora: la traduzione automatica è davvero utile? Non esiste una risposta univoca a queste domande, poiché sono i diversi i fattori che entrano in gioco. Infatti, non solo occorre distinguere fra le diverse tipologie di traduzione automatica, ma anche fra le esigenze dell'utente, fra le tipologie di testo da tradurre, fra le situazioni d'uso, e così via (Zanettin, 2001: 21).

Una volta accettata l'impossibilità di ottenere da un software una traduzione del tutto automatica, di alta qualità e per qualsiasi tipo di testo, la traduzione automatica offre comunque vantaggi significativi, che la rendono utile in

particolari contesti e per determinate categorie di utenti. Fra i principali vantaggi compaiono senza alcun dubbio l'elevata velocità di esecuzione, "improponibile per un essere umano" (Gaspari, 2011: 18). L'uso della traduzione automatica riduce fortemente i tempi di consegna e questo, talvolta, si rivela un fattore decisivo nel mondo degli affari e del commercio, dove spesso si ha bisogno di un accesso rapido alle informazioni contenute in un testo in lingua straniera, tanto che le tempistiche necessarie per una traduzione "tradizionale" possono ridurre il valore delle informazioni stesse o renderle addirittura inutili (Dillinger e Lommel, 2004: 12). Inoltre i sistemi di traduzione automatica, al contrario di un traduttore umano, sono sempre disponibili e accessibili, in qualsiasi momento della giornata (*ibid.*).

Un altro vantaggio importante è rappresentato dall'ampio ventaglio di lingue di partenza e arrivo offerte, anch'esso al di fuori della portata di un traduttore umano (Gaspari, 2011: 18). Sommando questi due importanti aspetti alla gratuità della maggior parte dei sistemi di traduzione automatica disponibili online, o ai vantaggi economici in termini di tempo risparmiato e facilità d'accesso offerti dai software offline, è dunque impossibile non riconoscere l'utilità della traduzione automatica.

Infatti, in proporzione ai notevoli vantaggi economici offerti, i costi per la gestione di un sistema di MT sono relativamente bassi e riguardano principalmente la personalizzazione del sistema di MT, nel caso in cui gli utenti intendano farne un uso sistematico (Dillinger e Lommel, 2004: 15).

Per quanto riguarda invece le imperfezioni e i limiti che ancora caratterizzano la performance dei sistemi disponibili oggi, come ricorda Hutchins, la traduzione è un mezzo che ha lo scopo di facilitare la comunicazione interlinguistica. In molti casi per la MT l'importanza di un risultato perfetto è relativa: molto più importante è che il testo prodotto sia comprensibile e fruibile (Hutchins, 2005). La valutazione dell'utilità di un sistema di MT dovrebbe dunque tenere in considerazione lo scopo primario di questo tipo di strumento.

Il grado di utilità di un sistema di MT può però variare in base a numerosi fattori. In generale, la traduzione automatica può essere utile quando occorre tradurre ingenti moli di testo in tempi ristretti, quando non si conosce affatto la lingua del testo che si sta leggendo e si vuole avere un'idea generale dei

contenuti, quando non si è disposti a pagare un traduttore professionista per il servizio, quando è sufficiente una traduzione grezza, ma rapida, del testo, e infine per tradurre testi che presentano espressioni ricorrenti e che non risultano particolarmente ambigui, come testi tecnico-informatici, scientifici, ecc.. Rimangono naturalmente esclusi testi letterari o retorici.

In modo schematico è possibile suddividere gli usi della traduzione automatica come segue:

- traduzione automatica come *dissemination tool*, ovvero come base di partenza per una traduzione finita. “La traduzione automatica è in questo senso solo una fase di un più ampio processo” (Monti, 2004: 65), che può comprendere anche il pre-editing, il post-editing (o entrambi). Come *dissemination tool*, la traduzione automatica viene utilizzata principalmente da traduttori professionisti per testi altamente ripetitivi e semplici dal punto di vista sintattico, ma densi dal punto di vista terminologico. Il vantaggio offerto dalla traduzione automatica in questo caso, oltre a quelli già menzionati, è l'elevata coerenza terminologica, essenziale nella traduzione scientifica, legale, tecnica, ecc. (*ibid.*);
- traduzione automatica come *information access tool*, ovvero come strumento di accesso alle informazioni. In questo caso la traduzione automatica viene utilizzata all'interno di sistemi di *information retrieval*, su database testuali (Monti, 2004: 66);
- traduzione automatica come *assimilation tool*, ovvero come strumento per la rapida comprensione di un testo in lingua straniera (Monti, 2004: 66). Questo uso della MT non prevede, di norma, post-editing, dal momento che non è necessario che la qualità del testo tradotto sia alta, ma è sufficiente che esso risulti fruibile. L'uso della MT come *assimilation tool* riguarda principalmente i sistemi di MT online ed è il pubblico generale a farvi per lo più ricorso, sebbene, in forma minore, la MT a scopi informativi venga utilizzata anche da agenzie di intelligence e organizzazioni internazionali (Hutchins, 2005). Benché questo uso della traduzione automatica sia sempre stato considerato secondario rispetto alla funzione di *dissemination tool*, è interessante notare come invece sia diventato sempre più rilevante a partire dalla

diffusione dei primi sistemi di MT online, contribuendo a far conoscere la MT al grande pubblico (Monti, 2004: 66);

- traduzione automatica come *interchange tool*, ovvero per il “rapido scambio di informazioni” (Monti, 2004: 66). Questo uso della MT riguarda principalmente la traduzione di testi di carattere informativo, destinati ad uso interno in grandi aziende e istituzioni. In questo caso, per ovvie ragioni connesse a tempistiche e costi, si preferisce la traduzione automatica a quella umana. Nella funzione di *interchange tool*, la MT viene utilizzata anche per la traduzione di corrispondenza ed email, anche se originariamente non era stata progettata per assolvere a questo compito (Hutchins, 2005);
- traduzione automatica come *entertainment tool*, ovvero come strumento di intrattenimento. Questo uso della MT riguarda i servizi di MT online ed è considerato un uso improprio della traduzione automatica (Monti, 2004: 66). In questo caso, l'utente effettua traduzioni *back-and-forth* (Gaspari, Hutchins, 2007: 200), ovvero da una lingua A a una lingua B e poi nuovamente dalla lingua B alla lingua A, oppure sottopone al sistema di MT proverbi, testi letterari o volutamente ambigui, per verificare le soluzioni traduttive proposte, che risulteranno necessariamente ricche di errori e “stranezze”;
- traduzione automatica come *learning tool*, ovvero come strumento per l'apprendimento di una lingua straniera (Monti, 2004: 67). In questo caso, i sistemi di MT online vengono utilizzati come supporto da utenti che studiano una particolare lingua straniera.

In seguito verranno descritte le principali categorie di utenti della MT e l'uso che essi fanno dei software di traduzione automatica.

1.5 I principali utenti dei servizi di traduzione automatica

1.5.1 Introduzione

Come si evince dalla sezione precedente, i servizi di traduzione automatica rispondono alle esigenze di diverse categorie di utenti, che possono essere raggruppati come segue:

- istituzioni, organizzazioni internazionali e agenzie governative (Dillinger e Lommel, 2004: 4);
- grandi aziende o multinazionali che operano in mercati multilingue e hanno contatti con utenti multilingue (*ibid.*), ma anche agenzie di traduzione;
- traduttori professionisti, che utilizzano la MT per ottenere una prima bozza di traduzione, su cui intervenire in un secondo momento;
- traduttori occasionali e utenti del pubblico generale.

Di seguito verranno analizzate nel dettaglio le categorie di utenti menzionate, le esigenze di ciascuna di queste e i diversi utilizzi della MT in base all'utente.

1.5.2 Istituzioni, organizzazioni internazionali e agenzie governative

Fra i principali utenti dei servizi di traduzione automatica compaiono ormai da decenni istituzioni e organizzazioni internazionali. Sin dagli esordi della traduzione automatica, istituzioni ed enti governativi hanno mostrato grande interesse per le potenzialità di questa tecnologia, finanziando la ricerca e lo sviluppo di sistemi di MT e più tardi di software per il riconoscimento vocale o per la traduzione automatica del parlato, come si vedrà nel capitolo II. Alla base di questo interesse si trova ovviamente la necessità di mantenere contatti internazionali, anche se questa si declina secondo modalità che variano da paese a paese e da istituzione a istituzione.

Negli Stati Uniti, nazione che ha dato un impulso considerevole

all'avanzamento della ricerca nel campo della traduzione automatica, l'interesse per questo tipo di tecnologia era (e resta) dettato principalmente da motivi di sicurezza nazionale. Il mercato statunitense è infatti essenzialmente monolingue, mentre a livello internazionale l'inglese viene usato come lingua franca nei rapporti commerciali e, per questo, il motivo economico-commerciale non è mai stato rilevante per il finanziamento della ricerca in questo campo (AA.VV., 2006: 33). Sin dagli anni Cinquanta, sono state principalmente agenzie militari e di intelligence a finanziare la ricerca sulla MT negli USA (come si legge in 1.2). In particolare, il DARPA (Defense Advanced Research Projects Agency), l'agenzia di ricerca del Dipartimento della Difesa degli USA, ha finanziato diversi progetti nell'ambito della HLT (Human Language Technology), inclusi alcuni nel campo della traduzione automatica. La necessità di tradurre testi scritti o orali da qualsiasi lingua verso l'inglese si è fatta ancora più stringente dopo gli attacchi terroristici dell'11 settembre 2001 (AA. VV., 2006: 33), in seguito ai quali il DARPA ha lanciato diversi progetti finalizzati alla “raccolta di informazioni in lingue diverse e da fonti multiple” (*ibid.*). Fra i progetti finanziati dal DARPA nell'ambito delle HLT compaiono:

- EARS (Effective Affordable Reusable Speech-to-Text), dal 2002 al 2004 (Chen et al., 2006: 1596);
- TIDES, “un programma per la traduzione, la ricerca, l'estrazione e la sintesi di informazione” (AA.VV., 2006: 35);
- GALE (Global Autonomous Language Exploitation), lanciato nel 2005 e incentrato “sull'estrazione di informazioni da testo multilingue e da documenti audio all'interno di un dominio illimitato” (*ibid.*: 34). Questo progetto affronta anche la traduzione automatica del parlato;
- TRANSTAC (TRANSLation system for TACTical use; Lee, 2015: 621);
- BOLT (Broad Operational Language Translation; *ibid.*).

In Europa, fra le principali istituzioni a investire nella ricerca sulla MT, come è facile intuire, compare l'Unione Europea. L'importanza della traduzione (e dell'interpretazione) a livello europeo si spiega con il fatto che l' “Unione Europea è nata multilingue” (AA.VV., 2006: 16). Infatti, contrariamente agli Stati Uniti, l'UE è formata da 28 stati membri

caratterizzati da lingue e culture diverse. Per rispettare questa componente multilingue, nel 1958 la CEE e l'Euratom hanno adottato un regolamento, tutt'ora in vigore, che sancisce che “le lingue ufficiali degli Stati membri sono lingue ufficiali dell'Unione europea” (Ranucci-Fisher, 2004: 79). Ne consegue che

tutti i regolamenti e gli altri testi di portata generale delle istituzioni comunitarie devono essere redatti e pubblicati in tutte le lingue ufficiali e che i testi diretti dalle istituzioni ad uno Stato membro, o ad una persona appartenente alla giurisdizione di uno Stato membro, devono essere redatti nella lingua di tale Stato.

(ibid.)

Per tenere fede a questi obblighi, la Commissione Europea si è dotata di un servizio di traduzione interno e di un servizio di interpretazione, garantiti rispettivamente dal Directorate General for Translation (DG Translation)¹ e dal Directorate General for Interpretation (DG Interpretation o SCIC).² Nel corso degli anni la domanda interna di traduzione è aumentata in modo costante, soprattutto in seguito all'allargamento dell'Unione Europea, tanto che oggi la Commissione Europea vanta uno dei servizi di traduzione più grandi al mondo³ non solo in termini di pagine tradotte (oltre 2 milioni nel 2014), ma anche di traduttori (oltre 2.000) e di lingue di lavoro (24 lingue ufficiali e altre quando necessario).⁴ Risulta evidente che la traduzione fra le varie lingue ufficiali, che oggi formano oltre 500 coppie linguistiche, richiede uno sforzo notevole: “il mantenimento di molte lingue ha un costo, sia in termini economici [...] sia in termini di impegno” (AA.VV., 2006: 16). Per questo motivo la Commissione Europea “è impegnata nella ricerca sulle HLT e ha finanziato nel tempo alcuni progetti di ricerca sulla traduzione automatica” (AA. VV., 2006: 32).

Già nel 1976 “la Commissione ha cominciato a mettere a punto un sistema di

¹ Commissione Europea, The European Commission's in-house translation service, marzo 2016. Visitato: 12/03/2016. Link: http://ec.europa.eu/dgs/translation/whoweare/index_en.htm

² Commissione Europea, About Dg Interpretation, aprile 2015. Visitato: 12/03/2016. Link: http://ec.europa.eu/dgs/scic/about-dg-interpretation/index_en.htm

³ Commissione Europea, The European Commission's in-house translation service, marzo 2016. Visitato: 12/03/2016. Link: http://ec.europa.eu/dgs/translation/whoweare/index_en.htm

⁴ *Ibid.* Link: http://ec.europa.eu/dgs/translation/whoweare/index_en.htm

traduzione automatica noto come **EC Systran**” (Ranucci-Fisher, 2004: 81), basato sulla tecnologia Systran. Come riporta Elisa Ranucci-Fisher della DG Traduzione, nel 2004 questo sistema permetteva di tradurre fino a 2000 pagine all'ora, potendo contare su una serie di dizionari specializzati divisi per settore, su un database terminologico e su una base documentaria (*ibid.*: 82). EC Systran veniva utilizzato da un lato dai traduttori, per produrre una prima bozza di traduzione da sottoporre poi a post-editing e revisione (Hutchins, 2005); dall'altro, in forma maggiore, dagli amministratori delle varie direzioni generali della Commissione Europea, sia per la traduzione che per la redazione di testi in lingua straniera (Ranucci-Fisher, 2004: 82). L'utilizzo di questo sistema di traduzione automatica e il ricorso sempre più sistematico a strumenti informatici hanno aumentato in modo considerevole la produttività in termini di traduzione (*ibid.*: 80), riducendo al contempo i costi e accorciando le tempistiche, fattore, questo, di assoluto rilievo, soprattutto in relazione a testi urgenti e a diffusione limitata o non destinati alla pubblicazione.

Da giugno 2013 è attivo il servizio MT@EC, che sta per Machine Translation at European Commission.⁵ Questo servizio è stato sviluppato dalla DG Traduzione, in collaborazione con la DG Informatica (DIGIT), nell'ambito del programma ISA (Interoperability Solutions for Public Administrations; AA.VV., 2014: 8). MT@EC si basa sulla traduzione automatica statistica di Moses e traduce da e verso tutte le 24 lingue ufficiali dell'Unione Europea, per un totale di 552 coppie linguistiche (*ibid.*: 3-4). Questo sistema di MT statistico è addestrato principalmente su documenti ufficiali dell'UE, il che lo rende particolarmente adatto per la traduzione di documenti relativi a politiche europee. Attualmente viene utilizzato dalla Commissione Europea, da numerose istituzioni europee e anche da diversi uffici di pubblica amministrazione degli stati membri dell'UE.⁶

MT@EC è progettato per rimuovere le barriere linguistiche ancora esistenti fra gli Stati Europei e per facilitare la circolazione di informazioni e know-

⁵ Commissione Europea, Isa Actions: Machine Translation Service, dicembre 2015. Visitato: 12/03/2016. Link: http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action_en.htm

⁶ Commissione Europea, Isa News: The advantages of using machine translation in Public administration, novembre 2015. Visitato: 12/03/2016. Link: http://ec.europa.eu/isa/news/2015/the-advantages-of-using-machine-translation-in-public-administration_en.htm

how presso le pubbliche amministrazioni. La peculiarità di questo servizio è infatti quella di offrire agli impiegati della pubblica amministrazione dei vari paesi europei la possibilità di tradurre documenti riservati e settoriali in modo rapido e sicuro (AA.VV., 2014: 2), con risultati soddisfacenti anche dal punto di vista qualitativo.

Si può quindi concludere che, a differenza degli Stati Uniti, in Europa la traduzione automatica rappresenta un'esigenza dettata dalla componente multilingue della società e del mercato: ancora oggi, solo il 39% dei cittadini europei utilizza una lingua diversa dalla propria lingua madre per comunicare online, mentre il 46% non è ancora in grado di leggere contenuti in lingua straniera (*ibid.*). Lo sviluppo e la diffusione di sistemi di traduzione automatica e di traduzione automatica del parlato contribuiranno a far superare le barriere linguistiche che ancora oggi dividono i popoli europei, permettendo nel contempo di mantenere intatti multilinguismo e multiculturalismo, caratteristiche che si trovano a fondamento dell'Europa.

1.5.3 Grandi aziende, multinazionali e agenzie di traduzione

Fra i maggiori utenti dei servizi di traduzione automatica compaiono anche le aziende: principalmente multinazionali e grandi aziende, ma anche agenzie di traduzione e aziende di piccole-medie dimensioni. In questo caso la MT è utile a diversi scopi: ottenere traduzioni in tempi rapidi in circostanze in cui non è essenziale un'alta qualità dell'output, mantenere un'elevata coerenza terminologica, abbattere i costi traduttivi, gestire grandi volumi di testi (Hutchins, 2005).

Nel caso specifico delle agenzie di traduzione, la traduzione automatica viene utilizzata per fornire ai traduttori una prima bozza di traduzione da revisionare (Dillinger e Lommel, 2004: 4). A livello aziendale, invece, la traduzione automatica viene utilizzata come *dissemination tool*, *assimilation tool* e infine come *interchange tool*.

Nella funzione di *dissemination tool*, la MT permette alle aziende di risparmiare tempo e denaro; tuttavia, perché il risultato della traduzione

automatica risulti soddisfacente, sono necessarie alcune condizioni (Hutchins, 2005). Innanzitutto è consigliabile utilizzare il software di MT per testi tecnici appartenenti a un dominio limitato, come manuali operativi, la cui traduzione richiede molto tempo (*ibid.*). In secondo luogo, l'azienda deve avere a disposizione risorse adeguate, non solo in termini di personale, ma anche di database terminologici, di corpora tradotti per l'addestramento del software e di linee guida per il pre-editing e il post-editing (*ibid.*).

Anche il software di traduzione automatica deve possedere certi requisiti per poter soddisfare le esigenze dell'azienda. In particolare deve essere un sistema *client-server*, deve essere personalizzabile in base alle esigenze specifiche dell'azienda e deve possedere un grande dizionario generale e dizionari tecnici relativi al settore in cui opera l'azienda. Infine deve essere compatibile con il sistema operativo installato sui computer dell'azienda e deve essere dotato di strumenti semplici per la gestione e l'aggiornamento dei dizionari (Hutchins, 2005).

La traduzione automatica viene utilizzata dalle aziende anche come *assimilation tool*: in questo caso può essere già sufficiente la traduzione grezza prodotta dal sistema di MT, oppure può essere necessario un rapido post-editing (Hutchins, 2005).

Un'altra applicazione della MT presso le aziende è quella di *interchange tool*. In questo caso la MT viene spesso usata per tradurre email di lavoro o pagine internet. La traduzione automatica di email dà spesso risultati peggiori rispetto a quella di testi ben scritti, dal momento che i sistemi di MT non sono stati progettati per tradurre il linguaggio idiomatico e sgrammaticato ricorrente in questo tipo di corrispondenza. Per lo stesso motivo, e a maggior ragione, anche l'uso dei sistemi di MT per la traduzione di lettere o email verso una lingua straniera totalmente sconosciuta è sconsigliabile (Hutchins, 2005). Nel 2005 Hutchins ipotizzava, in futuro, il probabile sviluppo di software progettati in modo specifico per la traduzione di corrispondenza commerciale e per la redazione di lettere in lingua straniera (*ibid.*).

A livello aziendale, la traduzione automatica è sempre stata presente per lo più sotto forma di software offline, che venivano (e vengono tuttora) acquistati presso case produttrici di sistemi di MT e adattati alle esigenze

specifiche dell'azienda tramite personalizzazione, oppure sviluppati dall'azienda stessa per uso interno (sistemi di MT in-house; Hutchins, 2005). Da alcuni anni, tuttavia, dietro pagamento, le aziende hanno anche la possibilità di accedere a sistemi di MT online personalizzabili. Un servizio di questo tipo è offerto da Microsoft, tramite il Microsoft Translator Hub, un'estensione del servizio Microsoft Translator che consente alle imprese di creare un sistema di traduzione automatica personalizzato, il quale, adattandosi alla terminologia usata dall'impresa stessa, è in grado di fornire output di qualità superiore rispetto a un comune sistema di MT online.⁷

Alcune aziende fanno ricorso anche ai servizi gratuiti di traduzione automatica online, soprattutto per la traduzione del sito web aziendale in più lingue. In questo caso, invece di commissionare la localizzazione del sito web a un traduttore professionista, l'azienda opta per una soluzione più economica, riportando sul proprio sito web il link a un servizio di MT online gratuito. Questo indica da un lato la consapevolezza di molte aziende della necessità di superare le barriere linguistiche per potere essere competitivi sul mercato globale, ma dall'altro l'incapacità di distinguere fra traduzione a scopo di *dissemination* e di *assimilation*, nonché una scarsa conoscenza dei sistemi di MT online e dei loro limiti. Infatti, affidando la traduzione del proprio sito web a un sistema gratuito di traduzione automatica, l'azienda, pur ottenendo un risparmio considerevole nell'immediato, rischia un sicuro danno d'immagine nel lungo periodo (Hutchins, 2005).

Risulta quindi evidente la necessità di sensibilizzare gli utenti della MT in merito alle aspettative che realisticamente si possono avere da questo tipo di tecnologia, oltre che in merito ai suoi vantaggi e ai suoi limiti. Inoltre è compito delle aziende produttrici di sistemi di MT fornire agli acquirenti, siano essi aziende o agenzie di traduzione, delle linee guida su come utilizzare il software di MT nel migliore dei modi.

⁷ Microsoft, Microsoft Translator: Microsoft Translator Hub. Visitato: 21/04/2016. Link: <https://www.microsoft.com/en-us/translator/hub.aspx>

1.5.4 Traduttori professionisti

Per molti decenni i traduttori professionisti si sono dimostrati ostili nei confronti dei nuovi strumenti tecnologici per la traduzione. Per quanto riguarda nello specifico la traduzione automatica, questa nuova tecnologia veniva vista come una vera e propria minaccia alla professione del traduttore (Hutchins, 2009: 14).

Negli anni Ottanta e Novanta, l'introduzione e la diffusione di strumenti informatici come database terminologici, accesso a internet e soprattutto memorie di traduzione, ha portato la maggior parte dei traduttori professionisti ad avvicinarsi alla tecnologia. Oggi, tutti i traduttori sono consapevoli dell'importanza di questo tipo di aiuto per il proprio lavoro (*ibid.*) e quasi tutti utilizzano le apposite stazioni di lavoro per tradurre.

Nel corso degli anni, anche la percezione della MT è cambiata nel mondo della traduzione professionale. Questo è dovuto innanzitutto a un miglioramento nella qualità delle prestazioni dei sistemi di MT, e, in secondo luogo, alla consapevolezza da parte dei traduttori dei limiti di questa tecnologia. Il traduttore moderno non teme più che la traduzione automatica possa sostituirsi alla sua professione, perché sa che (almeno) per alcune tipologie di testi questo non è possibile: per testi linguisticamente sofisticati e non ripetitivi, come testi letterari o giuridici, il traduttore umano è, e resterà, senza rivali (Hutchins, 2002). Per i testi che invece possono essere sottoposti a traduzione automatica, è comunque necessario l'intervento umano nella fase antecedente alla traduzione o in quella successiva. Inoltre, riconoscendo l'utilità dei software di MT, che per un traduttore risiede soprattutto nella rapidità, sono in molti oggi a chiedere software sempre più automatici e performanti, che vadano oltre il semplice *matching* di parole o frasi (Hutchins, 2009: 14).

Secondo l'AMTA (Association for Machine Translation in the Americas), la traduzione automatica sta ormai diventando uno strumento utilizzato comunemente dai traduttori professionisti,⁸ che talvolta ricorrono ad essa per

⁸ Amta, Amtaweb: MT for Translators, 2014. Visitato: 19/11/2015. Link: <http://amtaweb.org/mt-for-translators/>

velocizzare la traduzione di documentazione tecnica o manualistica. Si tratta di testi ripetitivi, lunghi centinaia o migliaia di pagine, che richiedono molto tempo per essere tradotti “manualmente” e possono risultare piuttosto “noiosi” per un traduttore professionista (Hutchins, 2005); ma anche di contratti, brevetti, comunicazioni aziendali interne,⁹ ecc. In questi casi, l'uso di un software di MT consente un notevole aumento della produttività.

Il traduttore resta inoltre la figura più qualificata per occuparsi del post-editing di un testo già tradotto in modo automatico (e in misura minore anche del pre-editing). Infatti, come si è capito ben presto, il revisore di una traduzione automatica deve conoscere sia la lingua d'arrivo che quella di partenza per potere correggere il testo nel modo migliore dal punto di vista linguistico (Hutchins, 2005). In questo senso, per il traduttore professionista, la traduzione automatica rappresenta una fonte di guadagno parallela a quella tradizionale della traduzione manuale.¹⁰

La diffusione della traduzione automatica ha portato anche allo sviluppo di vere e proprie linee guida per il post-editing, che si divide oggi in post-editing leggero e post-editing per la pubblicazione, in base all'entità dell'intervento da parte del revisore. Inoltre, riconoscendo la necessità di creare figure qualificate per il post-editing, diverse società di servizi linguistici offrono da qualche anno corsi specifici.¹¹

Infine si può dire che la figura del traduttore abbia una certa importanza per lo sviluppo della traduzione automatica, poiché il dialogo fra i traduttori professionisti che si occupano di post-editing e i programmatori di software di MT permette a questi ultimi di migliorare le performance del software, adattandolo alle esigenze dei traduttori stessi.¹²

È dunque ormai evidente che traduzione automatica e umana possano coesistere in relativa armonia; anzi, la traduzione automatica sembra aprire nuovi scenari e nuove possibilità di guadagno per i traduttori professionisti (Hutchins, 2002).

⁹ *Ibid.* Link: <http://amtaweb.org/mt-for-translators/>

¹⁰ *Ibid.* Link: <http://amtaweb.org/mt-for-translators/>

¹¹ *Ibid.* Link: <http://amtaweb.org/mt-for-translators/>

¹² *Ibid.* Link: <http://amtaweb.org/mt-for-translators/>

1.5.5 Traduttori occasionali e utenti del pubblico generale

Oggi traduttori occasionali e pubblico generale rappresentano una grossa fetta degli utenti dei servizi di traduzione automatica, se non addirittura la categoria principale di utenti (Gaspari, Hutchins, 2007: 203). In questo caso, la traduzione automatica viene usata principalmente come *assimilation tool* (Hutchins, 2005) anche se sono frequenti anche altri usi (*dissemination tool, learning tool, entertainment tool*).

Come facilmente intuibile, il tipo di traduzione automatica principalmente (se non esclusivamente) utilizzato dal pubblico generale è quello della MT online. Dalla sua comparsa negli anni Ottanta, con il servizio fornito da Systran per la rete Minitel (Hutchins, 2010), la MT online ha fatto enormi passi avanti in termini di diffusione, tanto che oggi la maggior parte dei portali internet offre servizi gratuiti di traduzione automatica (Ping, 2009: 165), per moltissime coppie di lingue. La MT è disponibile in rete sia sotto forma di servizi appositamente dedicati alla traduzione (come Google Translate, Bing Translator, ecc.), sia come possibilità offerta direttamente da siti internet il cui contenuto compare in una lingua diversa dalla lingua madre di chi naviga (come il servizio di traduzione istantanea di messaggi offerto da Facebook e altri social network). Da alcuni anni, inoltre, la traduzione automatica è disponibile anche sotto forma di app per dispositivi mobili, come smartphone o tablet, e ciò ha contribuito ulteriormente alla diffusione della MT presso gli utenti del pubblico generale. I servizi di MT online, come Google Translate, si sono evoluti non solo in termini di accessibilità e disponibilità, ma anche di prestazioni, offrendo agli utenti nuove possibilità, come ad esempio quella di ottenere la traduzione automatica di un testo fotografato tramite la fotocamera dello smartphone o di un testo pronunciato oralmente, e infine quella di accedere a traduzioni automatiche da dispositivo mobile anche in assenza di una connessione dati.¹³

Per quanto riguarda nello specifico gli utenti della MT online, è piuttosto

¹³ Google, Google Translate: All Languages. Visitato: 09/03/2016. Link: http://translate.google.com/about/intl/en_ALL/languages/

difficile reperire informazioni dettagliate sulle tipologie di testi tradotti, sulle aspettative o sugli scopi, come evidenziavano Gaspari e Hutchins già nel 2007 (Gaspari, Hutchins, 2007: 203). È però noto che questa categoria di utenti si avvale di servizi come quelli offerti da Google Translate non solo per la traduzione di porzioni di siti web in lingua straniera o di corrispondenza (da e verso lingue straniere), ma in generale per qualsiasi tipologia di testo, di qualsiasi provenienza. Oltre agli usi già elencati nei paragrafi precedenti, molti utenti sembrano inoltre utilizzare i servizi di MT online per la traduzione di poche parole, presumibilmente in lingue di cui hanno già una conoscenza parziale, utilizzando dunque la MT come una sorta di dizionario elettronico (Hutchins, 2009: 16).

Come già accennato, non di rado gli utenti della MT online ne fanno un uso scorretto, sottoponendo al sistema di MT testi del tutto inadatti, contenenti espressioni idiomatiche, ambigue, proverbi o strutture complesse, oppure testando l'efficacia del servizio tramite la *back translation* (*ibid.*: 15).

Alla luce dei recenti sviluppi, il ricorso alla traduzione automatica online e il numero di utenti occasionali sembrano destinati ad aumentare. Anche in questo caso, tuttavia, sarebbe auspicabile una maggiore sensibilizzazione rispetto al corretto funzionamento e ai limiti della MT, e alle aspettative realistiche che si possono avere da questa tecnologia.

1.6 I limiti della traduzione automatica

1.6.1 Premessa

Prima di analizzare i limiti specifici dei sistemi di traduzione automatica, occorre partire dal presupposto che inevitabilmente sia i traduttori professionisti che i sistemi di MT compiono degli errori (Dillinger e Lommel, 2004: 20). Le ragioni sono diverse e possono essere attribuibili a mancanze da parte del traduttore, a un'incompletezza del sistema di MT, o a un testo di partenza contenente errori o passaggi poco chiari. Occorre sempre considerare, inoltre, che la conoscenza che l'autore ha del testo redatto, che

poi verrà sottoposto a traduzione, è in genere superiore rispetto a quella del traduttore, sia in termini linguistici che di contenuto. Nel caso di un testo tecnico, ad esempio, l'autore del testo sarà presumibilmente un madrelingua con un'ottima conoscenza dell'argomento del testo e del lessico specifico. Il traduttore, al contrario, traduce in genere da una lingua straniera verso la propria lingua madre, e raramente ha la stessa conoscenza approfondita di un "addetto ai lavori" rispetto all'argomento del testo da tradurre. Questa disparità di conoscenza può portare ad errori nell'output della traduzione ed è tanto più evidente nel caso in cui il traduttore non sia un traduttore in carne ed ossa, bensì una macchina, la cui conoscenza sarà ancora più limitata, in quanto derivante esclusivamente dalle risorse messe a punto dai programmatori del software di traduzione stesso (ad esempio, dizionario, regole grammaticali, corpora paralleli, ecc.; Dillinger e Lommel, 2004: 20-21).

Ciò premesso, sistemi di traduzione automatica diversi presentano limiti diversi, che emergono in diversa misura, come si evince dai paragrafi precedenti. Tutti i sistemi di MT fino ad ora progettati sono però accomunati da un unico, grande limite, che li distingue dal traduttore umano: essi "sono privi di conoscenza del mondo e pertanto non sono in grado di fare uso del buon senso" (Gaspari e Zanchetta, 2011: 63). In termini pratici ciò significa che i sistemi di MT sono spesso incapaci di disambiguare espressioni che per un traduttore umano risulterebbero di comprensione immediata, e questo porta agli errori di traduzione o alle traduzioni "assurde" ben note agli utenti della MT.

Le principali difficoltà incontrate da un sistema di traduzione automatica derivano dalle proprietà stesse del linguaggio umano e si presentano dunque per qualsiasi coppia linguistica. Basti pensare a come, in qualsiasi lingua, esistano parole che possono assumere diversi significati in base al contesto (come, in italiano, *pésca* e *pèsca*). Se per un traduttore umano, dopo un'analisi del contesto, è piuttosto semplice effettuare una scelta traduttiva sensata, per il sistema di traduzione automatica questa operazione può rivelarsi molto complessa, o addirittura impossibile (Gaspari e Zanchetta, 2011: 63-64).

Un esempio è rappresentato dalle ambiguità lessicali, ovvero da quelle

parole che possono appartenere a diverse categorie grammaticali e assumere significato diverso in base a questa. Questo tipo di ambiguità è particolarmente diffuso in lingue morfologicamente povere, come l'inglese (*ibid.*: 65), in cui termini comuni come “control”, “ship”, “love” possono essere sia sostantivi che verbi (dando luogo a traduzioni differenti, come “controllo” vs. “controllare”, “nave” vs. “inviare”, ecc.).

Un altro tipo di ambiguità è rappresentato dai nomi propri (soprattutto di marchi o aziende), nel caso in cui questi siano “omografi di parole del lessico comune” (Gaspari e Zanchetta, 2011: 68). Basti pensare ad aziende come Windows o Apple, nomi propri che un traduttore lascerebbe invariati, ma che un software di MT potrebbe scambiare rispettivamente per “finestre” e “mela”. Per la traduzione automatica scritta, una potenziale disambiguazione in questo senso può essere fornita dall'iniziale maiuscola che contraddistingue i nomi propri, ma questo principio non si applica a lingue come il tedesco, dove tutti i sostantivi hanno l'iniziale maiuscola (*ibid.*: 69), né vale per i software di traduzione automatica del parlato, che partono da un input orale.

Ulteriori difficoltà per la MT sono rappresentate da traduzioni che dipendono dal contesto extra-linguistico, anafore pronominali, ambiguità nell'ampiezza dei riferimenti, ambiguità strutturale dei complementi, espressioni idiomatiche e infine collocazioni (Gaspari e Zanchetta, 2011: 66-70). Queste ultime risultano di difficile resa soprattutto per i sistemi basati su regole, mentre i sistemi statistici, grazie a corpora di testi tradotti da traduttori umani, danno in genere luogo a traduzioni più idiomatiche (*ibid.*: 70).

Per superare queste difficoltà e migliorare la performance dei sistemi di MT esistono diversi metodi. I principali sono pre-editing, post-editing e redazione del testo in linguaggio controllato.

1.6.2 Come affrontare i limiti della traduzione automatica: pre-editing, post-editing e linguaggio controllato

Come già accennato nella sezione precedente, pre-editing, post-editing e redazione di testi in linguaggio controllato permettono di migliorare la qualità di una traduzione automatica. In particolare, il post-editing agisce sul testo già tradotto e rappresenta una sorta di revisione, effettuata in genere da un traduttore. Il pre-editing e il linguaggio controllato sono invece due metodi applicabili al testo prima che questo venga tradotto. Il post-editing è già stato descritto nella sezione 1.5.4 di questo elaborato, perciò ora ci si concentrerà su pre-editing e linguaggio controllato.

Il pre-editing “prende le mosse da un testo di partenza esistente e lo adatta correggendone parti più o meno estese” (Gaspari e Zanchetta, 2011: 72). Esso ha lo scopo di correggere o semplificare il testo di partenza, così da migliorare la qualità della traduzione automatica e da ridurre la quantità di post-editing necessaria (Aston, 2011: 33). Il pre-editing può consistere nella semplice correzione di errori di punteggiatura, spaziatura, spelling e dell'uso di maiuscole/minuscole, ma può arrivare anche ad includere forme più articolate di intervento sul testo, come l'esplicitazione di elementi impliciti (soggetti, preposizioni, ecc.) e la semplificazione della struttura sintattica (come la semplificazione di gruppi nominali complessi o delle forme verbali; *ibid.*: 37-41). Per compiere un pre-editing efficace, è utile conoscere i punti di forza e i limiti, nonché gli errori ricorrenti dello specifico sistema di traduzione automatica che si utilizzerà.

La scrittura controllata si distingue dal pre-editing in quanto prevede l'adozione di un linguaggio controllato per la redazione del testo stesso da sottoporre a traduzione (Gaspari e Zanchetta, 2011: 72). Questo metodo viene adottato nei casi in cui il testo venga scritto appositamente per essere tradotto in modo automatico, “come può avvenire, ad esempio, nel caso della redazione di manualistica” (*ibid.*: 72) per un pubblico multilingue.

La scrittura controllata si basa sull'applicazione di regole che permettono di redigere un input che risulti il meno problematico possibile per la traduzione automatica. In genere vengono applicate alla redazione del testo sia regole

generali (ovvero valide per ogni lingua), sia regole specifiche per la lingua *source*. Fra le principali regole generali compaiono di solito le seguenti: evitare le forme verbali passive, scrivere frasi brevi (non oltre le 20 parole), evitare le espressioni idiomatiche, le forme impersonali e le nominalizzazioni, utilizzare termini di uso comune, esplicitare tutti i riferimenti, evitando i pronomi e ricorrendo eventualmente a ripetizioni, ecc. (Gaspari e Zanchetta, 2011: 72-76). Le regole riferite a una particolare lingua *source* possono invece essere molto più precise e riferirsi ad esempi concreti. Considerando un input in italiano, alcune regole per la scrittura controllata potrebbero includere i seguenti suggerimenti: utilizzare “anche se” al posto di “benché” per introdurre subordinate concessive (evitando così l’uso del congiuntivo, che potrebbe risultare problematico da tradurre); non utilizzare la parola “solo” in funzione avverbiale, ma preferire “solamente” (per evitare confusione con “solo” utilizzato in funzione aggettivale; *ibid.*: 73); utilizzare il termine “fiasco”, dotato di più significati, solo nell’accezione primaria (recipiente di vetro utilizzato per il vino), evitando espressioni come “ho fatto fiasco”, ecc.

La redazione di testi in linguaggio controllato, purché effettuata sulla base di regole chiare e facilmente applicabili, può dare origine a traduzioni di qualità adeguata, anche se non perfetta, riducendo così il post-editing necessario, ed è diffusa nella scrittura tecnica (Gaspari e Zanchetta, 2011: 74). Occorre però considerare quanto sia complesso mettere a punto regole di scrittura controllata efficaci, facilmente comprensibili e memorizzabili dai redattori dell'input (*ibid.*: 78). In generale, è sempre bene assicurarsi che il dispendio di energia e risorse necessario a compiere pre-editing e/o post-editing o a redigere il testo in linguaggio controllato per la MT non superi quello che sarebbe necessario per tradurre il testo manualmente.

1.7 La qualità nella traduzione automatica

1.7.1 Premessa

Valutare la qualità di una traduzione, sia essa “tradizionale” o automatica, può essere un compito piuttosto arduo, dal momento che non esiste un metodo di valutazione standardizzato (Dillinger e Lommel, 2004: 20), né un'unica traduzione “corretta”, se si considera che ogni testo può essere tradotto in diversi modi (Kit e Wong, 2008: 307). Ogni valutazione in questo campo è altamente soggettiva e variabile: ad esempio, utenti finali, addetti ai lavori/committenti e traduttori hanno punti di vista e aspettative diverse e inevitabilmente valuteranno lo stesso testo tradotto in modi diversi. Inoltre, i parlanti di lingue diverse hanno differenti livelli di tolleranza verso eventuali errori di traduzione: in genere più è elevata la loro conoscenza della lingua d'arrivo, minore sarà il loro bisogno di una traduzione e minore sarà la loro tolleranza (Dillinger e Lommel, 2004: 20). Questi fattori entrano in gioco anche nella valutazione di una traduzione automatica, intesa come prodotto.

Prima di analizzare le diverse modalità di valutazione della qualità nella traduzione automatica, occorre da un lato definire il concetto stesso di qualità, e dall'altro comprendere perché può essere importante valutare questo aspetto nella traduzione automatica.

Il concetto di qualità può essere definito in vari modi, in base al contesto a cui si fa riferimento. Per l'ambito di ricerca di questo elaborato, le seguenti definizioni di qualità risultano particolarmente funzionali:

1. “il complesso delle caratteristiche che rendono un oggetto adatto all'uso e alla funzione cui è destinato” (*Vocabolario della Lingua Italiana* in Viezzi, 1999: 142);
2. “la presenza nei prodotti di caratteristiche che vanno incontro alle esigenze dei clienti, garantendone la soddisfazione” (Juran e Godfrey in Osimo, 2004: 1).

In traduzione (e anche nell'ambito della traduzione automatica) la qualità si

declina, e può quindi essere misurata, in base a diversi parametri, quali fedeltà al testo di partenza, completezza delle informazioni, scorrevolezza del testo, correttezza grammaticale, stile, ecc., ma occorre sempre ricordare che la traduzione può essere vista come un prodotto o un servizio che ha lo scopo di soddisfare un'esigenza (quella della comunicazione interlinguistica e interculturale). Infatti, come evidenziano le definizioni di “qualità” riportate sopra, aspetti non trascurabili della qualità riguardano le esigenze degli utenti finali del testo tradotto e la funzione della traduzione stessa. Questi due aspetti si rivelano ancora più fondamentali quando si valuta l'efficacia di un sistema di MT (si veda 1.7.3).

Nell'ambito della traduzione automatica la valutazione della qualità assume un significato particolarmente importante, perché permette di aggiornare e migliorare continuamente il sistema di MT alla luce degli errori e dei problemi individuati in fase di valutazione della qualità degli output prodotti. In questo modo il software di traduzione automatica può fornire traduzioni via via migliori, soddisfacendo al meglio i bisogni degli utenti. Proprio per questo motivo, diverse piattaforme di MT online offrono agli utenti la possibilità di valutare la qualità delle traduzioni. Google Translate, ad esempio, possiede una sezione “Feedback”, che consente agli utenti di esprimere critiche o commenti sulla traduzione fornita dal sistema o suggerimenti su come migliorare il servizio, mentre il sito MyMemoryTranslated, che offre traduzioni allineate in un'enorme memoria di traduzione multilingue,¹⁴ dà agli utenti la possibilità di votare le traduzioni o eliminare eventuali allineamenti errati.

1.7.2 Metodi per la valutazione della qualità di una traduzione automatica

Esistono due tipologie principali di valutazione della qualità nell'ambito della traduzione automatica: la prima si basa sul giudizio umano (dunque

¹⁴ MyMemoryTranslated, MyMemory: tecnologia di nuova generazione per le memorie di traduzione. Visitato: 21/04/2016. Link: <https://mymemory.translated.net/doc/it/>

soggettivo), la seconda invece su meccanismi automatici.

Nel primo caso gli utenti sono chiamati a esprimere un giudizio sulla qualità della traduzione prodotta automaticamente, sulla base di diversi parametri (come scorrevolezza, adeguatezza, fedeltà al testo *source*, completezza delle informazioni, ecc.) a cui può essere associata una scala di valori (Euromatrix, 2007: 43-45). Altri metodi per valutare “manualmente” la qualità di una traduzione automatica includono:

- il tempo impiegato per la lettura del testo tradotto; questo criterio di valutazione si fonda sull'idea secondo cui più il valore delle parole al minuto (WPM, *Words Per Minute*) si avvicina a quello tipico del linguaggio naturale, migliore è la qualità della traduzione (*ibid.*: 45);
- il tempo impiegato per il post-editing della traduzione (*ibid.*: 45-46);
- test di comprensione del testo tradotto, per verificare la completezza e la chiarezza delle informazioni contenute nell'output (*ibid.*: 46-47).

Un ulteriore metodo di valutazione diffuso fra molti utenti della MT, ma considerato scorretto da studiosi e ricercatori è quello della *back-and-forth translation*. Questo metodo si basa sul presupposto che traducendo un testo dalla lingua A alla lingua B e poi ritraducendo l'output così ottenuto dalla lingua B alla lingua A, il risultato dovrebbe essere uguale al testo iniziale in lingua A (Kit e Wong, 2008: 308). In realtà come spiegato sulla pagina FAQ di SDL Free Translation¹⁵

[w]hen you translate the text into another language, some words can be translated in a number of ways. [...] Consequently when the text is translated back into the original language these words may be translated differently.

In generale si può dire che la valutazione soggettiva della qualità presenta alcuni vantaggi, dal momento che un giudice in carne ed ossa può valutare contemporaneamente diversi aspetti della traduzione, inclusi il senso generale e sfumature difficilmente percepibili da un sistema di valutazione automatico. Questo tipo di valutazione presenta però anche degli svantaggi,

¹⁵ SDL, SDL Freetranslation.com: Help and FAQ. Visitato: 12/03/2016. Link: https://www.freetranslation.com/help/?_sp=11aef9e8d44c59f2.1457539980151#sense

dal momento che persone diverse, anche sulla base degli stessi parametri di valutazione, potrebbero dare giudizi diversi alla stessa traduzione (Euromatrix, 2007: 49). La stessa persona, chiamata a valutare lo stesso testo più volte, potrebbe valutarlo in modo diverso, sulla scorta delle informazioni acquisite durante la prima valutazione. Inoltre sono molti i fattori che possono influenzare il giudizio umano, quali l'umore, la stanchezza, la personale tolleranza verso gli errori, e così via (Kit e Wong, 2008: 307).

Per questo motivo negli ultimi decenni sono stati sviluppati metodi automatici di valutazione della qualità nella MT. Questi sistemi di valutazione, generalmente basati su algoritmi, presentano numerosi vantaggi: sono oggettivi e coerenti, possono essere utilizzati per diversi tipi di sistemi di MT e di testo, in quanto generali, sono rapidi ed economici, permettono di ridurre al minimo l'intervento umano e infine possono essere usati sistematicamente durante la fase di sviluppo di un sistema di MT per testarne le modifiche. Il principale svantaggio deriva invece dal fatto che non sono dotati della capacità di giudizio tipica degli esseri umani e non sono in grado di cogliere sottili differenze fra un sistema di MT e l'altro (Euromatrix, 2007: 49).

Fra questi metodi, il più noto e più usato è BLEU (BiLingual Evaluation Understudy; Kit e Wong, 2008: 309), sviluppato nel 2001 da Kishore Papineni et al. per IBM. Questo algoritmo, presentato come una alternativa rapida alla valutazione umana, si basa sull'idea secondo cui “the closer the machine translation is to a professional human translation, the better it is” (Euromatrix, 2007: 51). BLEU si basa cioè su un confronto di n-grammi allineati fra la traduzione prodotta automaticamente e una o più traduzioni di riferimento, frutto del lavoro di traduttori: maggiore è il numero di n-grammi che corrispondono, migliore sarà la qualità della traduzione. In generale, BLEU sembra corrispondere alla valutazione umana di accuratezza e scorrevolezza del testo, anche se, come ogni altro metodo di valutazione, presenta dei punti deboli (*ibid.*: 51-54). In particolare, secondo uno studio condotto da Koehn e Monz nel 2006, BLEU sottostimerebbe sistematicamente la qualità delle traduzioni prodotte da sistemi basati su regole (2007: 51). Inoltre, BLEU non percepisce i cambiamenti sintattici, ha prestazioni inferiori se usato per valutare testi brevi, e infine, basando la sua

valutazione unicamente sul confronto di n-grammi, può dare luogo a falsi risultati, nel caso ad esempio di una traduzione che in ogni frase utilizzi molti degli n-grammi presenti nel testo di riferimento, ricombinandoli tuttavia secondo un ordine sbagliato (Euromatrix, 2007: 51-53). Per ovviare a questi problemi, spesso BLEU viene utilizzato in combinazione con altri metodi automatici di valutazione, quali ad esempio NIST, come nello studio condotto da Chanyu Kit e Tak Ming Wong (Kit e Wong, 2008: 310). NIST è un parametro di valutazione sviluppato dal National Institute of Standards and Technology degli Stati Uniti sulla base di BLEU, ma si differenzia da questo per essere più incentrato sull'accuratezza lessicale dei singoli n-grammi (*ibid.*).

1.7.3 Valutare la qualità di un sistema di traduzione automatica

La valutazione della qualità del testo tradotto da un sistema di traduzione automatica è una componente fondamentale della valutazione del sistema stesso, ma non coincide con essa. Ci sono infatti altri fattori che entrano in gioco nella valutazione della qualità, ovvero dell'efficacia, di un sistema di traduzione automatica.

In primo luogo, come ribadito più volte, occorre prendere in considerazione le esigenze dell'utente: nel caso di un traduttore occasionale o di un utente della MT online, un sistema “di qualità” sarà un sistema in grado di fornire traduzioni abbastanza accurate e comprensibili di testi appartenenti a vari domini (si pensi ai siti web, alle email, ecc.); in questo caso è sufficiente che l'utente testi il sistema di MT tramite la traduzione di un testo random (Dillinger e Lommel, 2004: 23). Nel caso di utenti che fanno un uso più sistematico della MT (ad esempio grandi aziende), la valutazione del sistema di traduzione automatica non può basarsi sulla traduzione di un testo casuale e sulla conseguente valutazione della qualità dell'output, ma deve prendere in considerazione molti altri fattori.

Delle interessanti linee guida su come valutare e scegliere un sistema di MT sono fornite in “LISA Best Practice Guide”. Qui gli autori suggeriscono di

prendere in considerazione i seguenti aspetti: la “conoscenza” che il sistema di MT ha del dominio a cui appartengono i testi che verranno poi tradotti, le lingue supportate dal sistema (incluse le direzioni di traduzione), il formato dei file supportati, le dimensioni del dizionario, il grado di personalizzazione del sistema, e infine la possibilità di aggiornare il dizionario e le altre risorse del sistema (Dillinger e Lommel, 2004: 23-27).

Si può concludere che la valutazione della qualità nella MT non si applichi unicamente al prodotto della traduzione automatica (ovvero al testo tradotto), ma anche al sistema di traduzione automatica stesso. Quest’ultimo tipo di valutazione ha come scopo da un lato quello di permettere la messa a punto e il miglioramento delle tecnologie di traduzione automatica e dall’altro quello di consentire agli utenti di scegliere il sistema più adatto alle proprie esigenze.

1.8 Dalla traduzione automatica alla traduzione automatica del parlato

Nel prossimo capitolo si vedrà come, partendo dalla traduzione automatica di testi scritti, il mondo della ricerca si sia spinto oltre, iniziando a interessarsi alla possibilità di tradurre automaticamente anche la lingua parlata. Nel giro di alcuni anni, da questa idea sono nati i primi sistemi automatici di *Speech-to-Speech Translation*, in grado di utilizzare input orali per ricavare traduzioni, poi riprodotte con voce artificiale da un modulo di sintesi vocale. Il sogno di creare uno strumento in grado di tradurre conversazioni fra interlocutori multilingue sembrava diventare per la prima volta una possibilità concreta e per i ricercatori attivi nel campo della traduzione automatica si annunciavano nuove opportunità, ma anche nuove sfide.

II

La traduzione automatica del parlato o Speech-to-Speech Translation

2.1 Introduzione

Il pesce Babele è piccolo, giallo, ricorda una sanguisuga ed è probabilmente la cosa più strana dell'Universo. [...] [S]e vi ficcate un pesce Babele in un orecchio, immediatamente capirete qualsiasi cosa vi si dica in qualsivoglia lingua.

(Adams, 2012: 45)

Il pesce Babele (in inglese Babelfish) nasce nel 1979 dalle pagine del romanzo fantascientifico *Guida galattica per gli autostoppisti* di Douglas Adams. Questo curioso animale, il cui nome evoca l'immagine biblica della Torre di Babele, viene descritto da Adams come un pesce che, una volta inserito nell'orecchio, consente di comunicare con qualsiasi persona, in qualunque lingua del mondo: in breve, una sorta di interprete simultaneo universale.

Il pesce Babele, tuttavia, non è l'unico “traduttore universale” a comparire in un'opera letteraria e/o cinematografica; un altro esempio è quello dello *universal translator* della celebre serie televisiva *Star Trek*, ma ne esistono molti altri, diffusi soprattutto nel genere *science fiction*.

Questi traduttori universali descritti da scrittori e sceneggiatori incarnano un sogno che l'uomo ha da sempre: quello di creare uno strumento in grado di tradurre simultaneamente qualsiasi conversazione, in qualsiasi lingua, consentendo così di superare le barriere linguistiche e culturali che separano i vari popoli del pianeta.

Pur essendo lontani dal raggiungimento di questo obiettivo, oggi la

traduzione automatica del parlato non è più confinata alla fantascienza. Di fatto, i grandi progressi compiuti negli ultimi decenni “nelle tecnologie della microelettronica e dell'information technology [...], nelle ricerche dell'elaborazione del linguaggio naturale e del riconoscimento del parlato” (Lazzari, 2002: 211) hanno portato alla realizzazione di sistemi che permettono di tradurre in modo automatico testi pronunciati oralmente. Questi sistemi, seppur lontani dall'essere infallibili, sono ormai accessibili anche al grande pubblico: basti pensare al servizio di traduzione automatica del parlato offerto da Google Translate e da numerose applicazioni per smartphone, o a Skype Translator, in grado di effettuare la traduzione automatica delle telefonate multilingue effettuate sulla piattaforma Skype. Questo capitolo è dedicato proprio alla traduzione automatica del parlato, una nuova tecnologia di cui verranno analizzate storia, funzionamento, limiti, prospettive e sfide future e infine aspetti legati alla qualità.

2.2 La traduzione automatica del parlato o Speech-to-Speech Translation

La traduzione automatica del parlato, o traduzione da parlato a parlato, nota anche come Speech-to-Speech Translation (SST o S2ST) in inglese, può essere definita come

a process that takes the conversational speech phrase in one language as an input and translated speech phrases in another language as the output

(Dureja, Gautam, 2015: 28)

In altre parole, grazie a questa tecnologia,

[i]l parlato nella lingua di partenza viene tradotto in parlato nella lingua di arrivo, in modo che le persone possano parlare tra di loro in diverse lingue, usando il computer come interprete.

(AA. VV., 2006: 55)

Negli ultimi anni la SST è diventata via via più rilevante, almeno nel panorama scientifico delle Human Language Technologies, tanto che, come ricorda Satoshi Nakamura, la traduzione automatica del parlato compare fra le dieci tecnologie emergenti che cambieranno il mondo (Nakamura, 2009: 35) nell'articolo “10 Emerging Technologies That Will Change Your World”, pubblicato sulla *MIT Technology Review* nel febbraio del 2004.

Il carattere fortemente innovativo della SST viene sottolineato anche nel report “Tecnologie del linguaggio per l'Europa”, dove la traduzione automatica del parlato viene descritta come innovazione di rottura (dall'inglese *disruptive innovation*; AA. VV., 2006: 31). Questa espressione, usata in contrapposizione al concetto di innovazione di sostegno (*sustained innovation*), fa riferimento a tutte quelle innovazioni tecnologiche che non migliorano le prestazioni di un prodotto già esistente (come invece fanno le innovazioni di sostegno), ma al contrario le peggiorano, almeno nel breve periodo, offrendo prestazioni più scadenti rispetto a quelle dei prodotti già in commercio da tempo. Le innovazioni di rottura possiedono però “altre caratteristiche che hanno valore per nuovi clienti” (AA.VV., 2006: 30), prendono piede in mercati diversi da quelli esistenti ed estendono la propria quota di mercato nel tempo, oppure portano quella che era una nicchia di mercato ad assumere dimensioni rilevanti (*ibid.*: 30-31). Secondo questa definizione, la traduzione automatica del parlato si configura chiaramente come innovazione di rottura, dal momento che

non è facile vedere come questa tecnologia possa supportare gli interpreti umani, ed inoltre se confrontata con la traduzione umana, le sue prestazioni, sebbene stiano migliorando, rimangono tuttavia molto inferiori.

(AA.VV., 2006: 31)

È dunque presumibile che questa tecnologia possa prendere piede in nuovi mercati, non coperti al momento da traduttori o interpreti umani, soddisfacendo nuovi bisogni ed estendendo la sua quota di mercato nel tempo (*ibid.*: 31), proprio come è già accaduto per altre innovazioni di rottura, quali la telefonia mobile e la fotografia digitale. Accostando la Speech-to-Speech Translation a questi esempi ben più noti di innovazione

di rottura, non risulta difficile comprendere perché essa sia stata inserita fra le tecnologie emergenti più rivoluzionarie.

2.3 Le principali tappe storiche della traduzione automatica del parlato

Dal punto di vista storico, la traduzione automatica del linguaggio parlato “rappresenta un'area di ricerca piuttosto recente nell'ambito delle tecnologie sul linguaggio umano” (Lazzari, 2004: 72). I primi segni di interesse nella ricerca in questo campo si registrano a partire dalla seconda metà degli anni Ottanta, quando, sull'onda degli sviluppi dei sistemi per la traduzione automatica di testi scritti, alcuni ricercatori iniziarono a interessarsi anche alla possibilità di tradurre la lingua parlata. Già all'epoca, questa si configurava come una sfida ancora più ardua della traduzione automatica, a causa delle difficoltà intrinseche al parlato stesso.

La SST attirò per la prima volta l'attenzione di esperti e ricercatori del settore delle telecomunicazioni durante il forum ITU Telecom World, tenutosi a Ginevra nel 1983 (Nakamura, 2009: 36) e organizzato dalla International Communication Union (ITU), una delle agenzie specializzate delle Nazioni Unite.¹⁶ In quell'occasione la multinazionale giapponese NEC Corporation effettuò una dimostrazione di un prototipo di sistema di SST (*ibid.*: 36).

Un'altra tappa storica fondamentale è il 1986, anno in cui venne fondato in Giappone l'Advanced Telecommunications Research Institute International (ATR) e contestualmente venne lanciato un progetto di ricerca sulla traduzione automatica del parlato che vide coinvolti ricercatori provenienti da diversi istituti di ricerca attivi sia in Giappone che in altri paesi. A quello di ATR fecero ben presto seguito, in tutto il mondo, altri progetti di ricerca (Nakamura, 2009: 36): la SST si apprestava così a diventare un settore di ricerca di interesse internazionale.

¹⁶ ITU, ITU: About ITU. Visitato: 13/03/2016. Link: <http://www.itu.int/en/about/Pages/default.aspx>

Il 1991 segna invece la nascita del consorzio C-STAR (l'acronimo di Consortium for Speech Translation Advanced Research), un progetto iniziato con una collaborazione bilaterale e informale fra laboratori di ricerca impegnati nella SST (Chan, 2004: 27). A questo consorzio si deve il merito di avere dimostrato per primo “la fattibilità delle tecnologie SLT attraverso due principali dimostrazioni su scala planetaria nel 1995 e nel 1999” (Lazzari, 2004: 72) e di avere incoraggiato la cooperazione fra il mondo accademico e industriale nell'ambito della ricerca sulla SST (Grazina, 2010: 47).

Nel 1993 fu il governo tedesco a dare un importante impulso alla ricerca, lanciando Verbmobil (Lazzari, 2004: 72), un progetto su larga scala che fu finanziato dal Ministero Federale Tedesco dell'Istruzione e della Ricerca e da diversi partner del mondo dell'industria e che vide coinvolti centinaia di ricercatori (Lee, 2015: 620). Verbmobil era un sistema di Speech-to-Speech Translation per la traduzione automatica di interazioni spontanee in negoziazioni faccia a faccia (Bub et al., 1997) e poteva essere applicato a tre diversi domini, ovvero la programmazione di appuntamenti, la pianificazione di viaggi e l'assistenza remota per PC (Lee, 2015: 620).

Così come la fondazione del Consorzio C-STAR, anche Verbmobil svolse un ruolo fondamentale per lo sviluppo dei sistemi di SST: con Verbmobil non solo si propose “lo scenario applicativo della comunicazione multilingue faccia a faccia” (Lazzari, 2004: 72), ma si iniziarono anche ad esplorare diversi approcci per risolvere il problema della traduzione, quali il trasferimento semantico e l'interlingua, la traduzione automatica basata su esempi e quella statistica (*ibid.*: 72-73).

A partire dal 2000 la ricerca nell'ambito della SST ha compiuto passi avanti, iniziando a testare i sistemi di SST all'interno di scenari reali, esplorando la traduzione di testi appartenenti a nuovi domini, e prendendo in considerazione condizioni acustiche e modi di parlare più realistici (Lee, 2015: 620). Questo è stato possibile grazie al graduale miglioramento delle tecnologie di riconoscimento del parlato e di traduzione automatica, senza le quali la Speech-to-Speech Translation non sarebbe possibile.¹⁷ A Verbmobil

¹⁷ Microsoft, Microsoft Research: How technology can bridge language gaps. Visitato: 30/03/2016. Link: <http://research.microsoft.com/en-us/research/stories/speech-to-speech.aspx>

hanno dunque fatto seguito, nel XXI secolo, molti altri progetti di traduzione automatica del linguaggio parlato. A questo proposito, vale la pena ricordare NESPOLE! (Negotiation through SPOken Language in E-Commerce), lanciato nel 2000. Il progetto, finanziato dalla Commissione Europea e dalla National Science Foundation negli Stati Uniti, è stato frutto di una collaborazione fra diversi partner internazionali: tre laboratori di ricerca europei (IRST di Trento, ISL dell'Università di Karlsruhe e CLIPS dell'Università Joseph Fourier di Grenoble), un gruppo di ricercatori statunitensi (ISL della Carnegie Mellon University di Pittsburgh), e due partner del mondo dei servizi e dell'industria (APT Trento, un ente di promozione turistica, e Aethra, una società di telecomunicazioni con sede ad Ancona; Metze et al., 2002). NESPOLE! assisteva utenti non esperti di SST, di madrelingua inglese, francese o tedesca, nella richiesta di informazioni sul turismo invernale in Trentino. Attraverso una sorta di videoconferenza, l'utente veniva messo in contatto con l'operatore italiano dell'APT, il cui compito era di fornire le informazioni che poi venivano tradotte vocalmente dal sistema NESPOLE! (Metze et al., 2002).

Sempre in ambito europeo si colloca TC-STAR, un progetto a lungo termine finanziato dalla Commissione Europea a partire dal 2004 e incentrato sulla traduzione del parlato fra lo spagnolo, l'inglese e il cinese all'interno di un dominio illimitato (Lee, 2015: 621).

Negli Stati Uniti, il DARPA ha lanciato diversi programmi per l'avanzamento tecnologico nella traduzione del parlato (*ibid.*), fra cui si ricordano GALE (Global Autonomous Language Exploitation), TRANSTAC (TRANSlation system for TACTical use) e BOLT (Broad Operational Language Translation). Questi programmi, progettati per l'impiego nell'ambito di operazioni militari e di sicurezza nazionale, erano perlopiù incentrati sulla traduzione fra l'arabo e l'inglese. Altri sistemi di SST con scopi simili erano MASTOR (Multilingual Automatic Speech-to-Speech Translator), sviluppato da IBM, e IraqComm, sviluppato dallo Stanford Research Institute (SRI). Entrambi i software funzionavano su laptop e sono stati utilizzati dalle truppe americane nel corso di operazioni militari in Iraq (Lee, 2015: 621).

Una tappa storica più recente nello sviluppo dei sistemi di traduzione automatica del parlato risale al 2008, l'anno in cui si sono svolti i Giochi

Olimpici di Pechino. Per questa occasione è stato lanciato un progetto di Speech-to-Speech Translation con lo scopo di assistere i turisti stranieri nell'organizzazione del loro viaggio (*ibid.*). Il prototipo sviluppato, nato dalla collaborazione di NLPR-CAS (National Laboratory of Pattern Recognition - Chinese Academy of Science), Università di Karlsruhe e Carnegie Mellon University, supportava tre lingue (cinese, inglese e spagnolo) e funzionava su laptop e palmari dotati di connessione wireless (Stüker et al., 2006 in Lee, 2015: 621).

Nel frattempo, nel 2006 veniva fondato in Asia A-STAR (Asian Speech Translation Advanced Research), un nuovo consorzio per la ricerca sulle tecnologie SST che riuniva numerosi gruppi di ricerca attivi in Asia. In collaborazione con ASTAP (Asia-Pacific Telecommunity Standardization Program), A-STAR ha iniziato a promuovere la standardizzazione dei protocolli internazionali di comunicazione nell'area della Speech-to-Speech Translation¹⁸ (Lee, 2015: 621) e nel 2009 ha lanciato il primo sistema di SST basato su network di tutta l'Asia. Questo sistema consentiva la comunicazione in tempo reale fra più utenti di lingue asiatiche diverse e ha dimostrato l'attuabilità dei protocolli di Speech-to-Speech Translation basati su network.¹⁹ Nel 2010 A-STAR si è espanso, diventando così un consorzio globale, che ha preso il nome di U-STAR e che oggi conta 26 membri (*ibid.*). Negli ultimi anni, la ricerca nell'ambito della traduzione automatica del parlato ha compiuto progressi tali da rendere disponibili al pubblico applicazioni e software di Speech-to-Speech Translation. Uno degli esempi più noti è sicuramente Skype Translator, presentato da Microsoft durante la Microsoft's Worldwide Partner Conference del 2014 attraverso una dimostrazione durante la quale Skype Translator è stato utilizzato per tradurre una videochiamata fra un interlocutore inglese e un'interlocutrice tedesca. Da dicembre 2015 la funzionalità è disponibile sotto forma di applicazione per pc ed è accessibile a tutti gli utenti gratuitamente e senza

¹⁸ Ustar, Ustar Consortium: About us. Visitato: 13/03/2016.

Link: <http://www.ustar-consortium.com/aboutus.html#background>

¹⁹ Ustar, Ustar Consortium: What is U-STAR?. Visitato: 13/03/2016.

Link: <http://www.ustar-consortium.com/>

obbligo di registrazione.²⁰ Le lingue attualmente supportate per la traduzione di chiamate audio e video sono limitate (cinese, inglese, francese, tedesco, italiano e spagnolo),²¹ ma l'obiettivo di Microsoft è quello di introdurre nuove combinazioni linguistiche, migliorando così il servizio offerto agli oltre 300 milioni di utenti di Skype.²²

Poco dopo il lancio di Skype Translator, anche Google ha annunciato l'introduzione della funzione di traduzione automatica del parlato per l'applicazione Google Translate per smartphone. Attualmente Google Translate supporta oltre 100 lingue, anche se la funzione “Conversazione” (ovvero quella che consente la traduzione automatica del parlato) è disponibile solo per alcune di queste, come si vedrà in seguito.²³ Oltre agli esempi menzionati, sono ormai diffuse molte applicazioni di traduzione automatica e traduzione automatica del parlato per smartphone e tablet. Poiché il numero di queste app è in costante crescita, all'interno di un panorama molto vario (esistono app più o meno accurate) e in rapido mutamento (applicazioni disponibili per il download in un dato momento potrebbero non esserlo più pochi mesi dopo, perché ritirate dal commercio), in questo elaborato si è scelto di trattare le app di traduzione automatica del parlato limitatamente ai loro possibili impieghi (si veda sezione 2.6) e al loro funzionamento, che verrà descritto nel capitolo III sulla scorta di due esempi specifici di applicazione. Nella prossima sezione, invece, verrà analizzato in generale il funzionamento di un sistema di Speech-to-Speech Translation.

²⁰ Skype, Blogs Skype: Skype Translator Preview Access Just Got Easier!. Visitato: 06/04/2016. Link: <http://blogs.skype.com/2015/05/12/skype-translator-preview-access-just-got-easier/>

²¹ Skype, Support Skype: Quali lingue sono disponibili in Skype Translator?. Visitato: 06/04/2016. Link: <https://support.skype.com/it/faq/FA34543/quali-lingue-sono-disponibili-in-skype-translator>

²² Skype, Blogs Skype: Skype Translator Preview Access Just Got Easier!. Visitato: 06/04/2016. Link: <http://blogs.skype.com/2015/05/12/skype-translator-preview-access-just-got-easier/>

²³ Google, Google Translate: Lingue. Visitato: 06/04/2016. Link: http://translate.google.it/about/intl/it_ALL/languages/

2.4 Funzionamento dei sistemi di traduzione automatica del parlato

Come già accennato nelle sezioni precedenti, i sistemi di traduzione automatica del parlato consentono di tradurre in modo automatico e quasi istantaneo un testo orale, pronunciato da un parlante, da una lingua A a una lingua B. La traduzione, fornita nel giro di pochi secondi, può avere forma scritta, nel caso della *Spoken Language Translation* (SLT), oppure forma orale, nel caso dei veri e propri sistemi di Speech-to-Speech Translation (AA.VV., 2006: 54-55).

Gli strumenti SLT, simili ai sistemi di SST, ma meno ambiziosi, partono da un input orale per produrre un output (tradotto) in forma scritta; possono quindi essere utilizzati come supporto alla comunicazione, pur non consentendo un dialogo vero e proprio. Altri strumenti con scopo simile sono i cosiddetti “voice activated phrase lookup systems” (Zhang, 2003), frasari elettronici attivati tramite input vocale. Un prodotto di questo genere è Phraselator, un dispositivo portatile creato da VoxTec per la traduzione vocale unidirezionale, che consente all'utente di scegliere fra migliaia di parole o frasi predefinite mostrate sul display e di ottenere la traduzione corrispondente in diverse lingue. In questo caso, la “pronuncia” dell'output non avviene grazie a un modulo di sintesi vocale, ma grazie a traduzioni pre-registrate, pronunciate da parlanti nativi delle lingue *target* offerte da Phraselator (*ibid.*). Anche questo strumento non rappresenta un vero e proprio sistema di SST e non consente una comunicazione bidirezionale, ma è stato utilizzato con successo dall'esercito americano durante operazioni in Croazia, Iraq, Indonesia, ecc. (Hutchins, 2009: 17).

I veri e propri sistemi di Speech-to-Speech Translation, invece, presentano un maggiore grado di complessità. Infatti la tecnologia SST consente, almeno sulla carta, una sorta di dialogo vero e proprio fra parlante della lingua A e parlante della lingua B, in cui il sistema di SST assume il ruolo di “mediatore linguistico”.

Per consentire questo tipo di comunicazione un sistema di SST si avvale di tre tecnologie distinte (Nakamura, 2009: 35):

- riconoscimento vocale, o *Automatic Speech Recognition (ASR)*, che permette alla macchina di riconoscere l'input vocale di chi parla; questa è una tecnologia ampiamente diffusa oggi, come dimostrano l'utilizzo di sistemi di dettatura e il fatto che smartphone, automobili e altri dispositivi elettronici siano spesso dotati di funzioni attivabili tramite comandi vocali o di “assistenti virtuali”, a cui è possibile chiedere informazioni semplicemente parlando (si pensi a Siri sui dispositivi Apple o alla app Ask Ziggy su Windows Phone).²⁴ Nella traduzione automatica del parlato il riconoscimento vocale avviene in genere in pochi secondi: deve essere più rapido rispetto alla digitazione manuale, altrimenti, a parità di tempo impiegato, l'alternativa più vantaggiosa per gli utenti risulterebbe essere quest'ultima, indubbiamente più precisa rispetto alla tecnologia ASR (è l'utente stesso a digitare il testo da tradurre e dunque ad averne il controllo);
- traduzione automatica (MT o SLT), che consente di tradurre in modo automatico le parole riconosciute tramite l'ASR; il modulo di traduzione automatica può presentare diversi tipi di architettura, come spiegato nella sezione 1.3 (ad esclusione degli approcci transfer e diretto, che non vengono adottati nella SST);
- sintesi vocale, o *Text to Speech Synthesis (TTS)*, che sintetizza il testo tradotto nella lingua di arrivo, riproducendo artificialmente la voce umana, idealmente preservando l'identità del parlante della SL (*Source Language*) e le caratteristiche prosodiche del testo *source*.

In termini pratici, il funzionamento di un sistema di SST può essere riassunto schematicamente come segue

[1] la voce di chi parla viene immessa nel computer tramite un microfono ed elaborata da un programma di riconoscimento vocale che trasforma i suoni

²⁴ Microsoft, Microsoft Research: How technology can bridge language gaps. Speech-to-speech translation promises to help connect our world. Visitato: 30/03/2016. Link: <http://research.microsoft.com/en-us/research/stories/speech-to-speech.aspx>

in lettere e parole. Una volta che un testo orale è stato immesso e trasformato in testo scritto, esso viene tradotto e infine elaborato da un programma di sintesi vocale che trasforma le parole scritte in suoni emessi da un altoparlante.

(Zanettin, 2001: 40)

Un'ulteriore rappresentazione del funzionamento di un sistema di Speech-to-Speech Translation è fornita in Figura 1 (adattata da: AA.VV., 2006: 54):

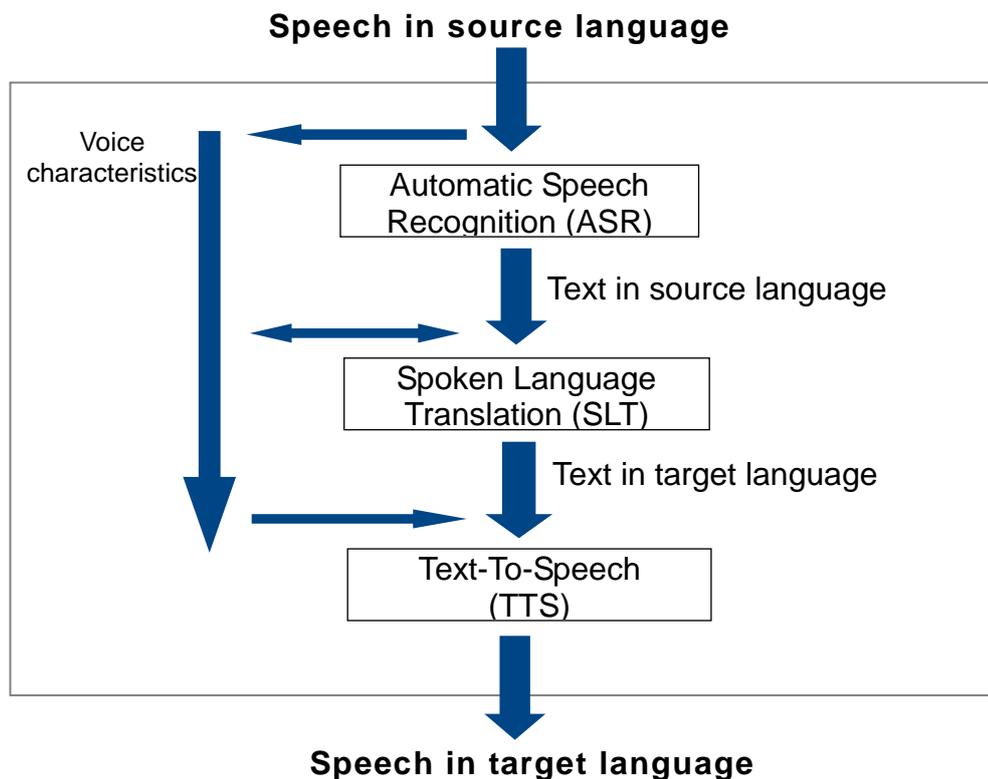


Figura 1

Funzionamento di un sistema di traduzione automatica del parlato.

Come si vede in figura, il parlato della lingua *source* viene convertito in testo dal modulo di riconoscimento vocale automatico (ASR). L'output prodotto dal modulo ASR passa poi al modulo di traduzione automatica (MT o SLT), che analizza il testo (includendo eventuali false partenze, esitazioni, frasi sgrammaticate e altri elementi tipici del parlato) e lo traduce automaticamente. L'output prodotto da questo modulo, ovvero il testo nella lingua *target*, costituisce l'input per il modulo di sintesi vocale (TTS), che sintetizza il parlato nella lingua *target*, sfruttando anche le informazioni

sulla lingua *source* (ad esempio, dati relativi alla prosodia) fornite dal modulo ASR (AA.VV., 2006: 54).

Per quanto riguarda il funzionamento specifico del modulo MT, sono due gli approcci adottati (Grazina, 2010: 6-7): l'approccio interlingua, il primo ad essere applicato ai sistemi di SST, adottato ad esempio nel sistema NESPOLE! (si veda sezione 2.3); l'approccio statistico, introdotto più tardi e oggi prevalente. Il funzionamento dei moduli ASR e TTS, invece, si basa spesso su database di conversazioni registrate; in questo caso si parla di “corpus-based speech recognition and synthesis” (Nakamura, 2009: 39).

La Figura 1 mostra quanto un sistema di SST possa essere complesso, dal momento che si basa su diverse tecnologie interdipendenti, introdotte in tempi relativamente recenti e ancora da perfezionare.

Esiste tuttavia un'alternativa ancora più ambiziosa della normale SST: la SST simultanea, una delle sfide più ardue nel campo della traduzione della lingua parlata (Grazina, 2010: 9). Idealmente, un sistema di SST simultaneo offrirebbe la traduzione simultanea di qualsiasi tipo di discorso e potrebbe essere utilizzato per tradurre ad esempio il contenuto di lezioni, seminari e presentazioni (*ibid.*). È facile intuire come in questo caso le sfide da affrontare si moltiplichino: alle “tradizionali” difficoltà della tecnologia SST si aggiungono quelle legate alla traduzione simultanea, una tecnica in cui il fattore tempo, la segmentazione della frase e la riformulazione dei contenuti rivestono un ruolo fondamentale.

Nonostante le difficoltà, la traduzione automatica simultanea del parlato, ormai da alcuni anni oggetto di studio di alcuni gruppi di ricercatori internazionali, produce già i primi risultati: un esempio è il Lecture Translator, sviluppato dall'università tedesca di Karlsruhe. Questo sofisticato sistema, finora testato durante alcune lezioni tenutesi presso l'università stessa, si avvale delle tecnologie ASR e MT per fornire la traduzione automatica e simultanea delle lezioni dal tedesco all'inglese, in forma scritta (è dunque assente il modulo TTS); ogni traduzione è consultabile da PC o smartphone,²⁵ e viene poi pubblicata online su un sito accessibile agli studenti dell'università.

²⁵ Karlsruhe Institute of Technology, Interactive System Labs: Lecture Translation. Visitato: 12/04/2016. Link: <http://isl.anthropomatik.kit.edu/english/1520.php/>

2.5 Problematiche e sfide della traduzione automatica del parlato

2.5.1 Analisi delle problematiche e delle sfide della traduzione automatica del parlato

Se la FAHQMT rimane un sogno nell'ambito della traduzione automatica di testi scritti, questo vale a maggior ragione per la traduzione automatica del parlato (Zong e Seligman, 2005). Questo tipo di tecnologia presenta infatti le stesse problematiche della MT tradizionale, dal momento che si avvale di un modulo di traduzione automatica, e numerose problematiche aggiuntive, che nel corso degli ultimi anni sono state superate solo parzialmente.

Un primo grande ostacolo è rappresentato dal linguaggio parlato stesso. Occorre considerare, infatti, che un sistema di SST, contrariamente a un sistema di MT, traduce testi orali, idealmente dialoghi. Questo rappresenta un compito molto più complesso rispetto alla traduzione di testi scritti (Nakamura, 2009: 36), soprattutto se si considera che tradizionalmente la MT si è sempre concentrata su testi “ben formulati” e caratterizzati da un linguaggio tecnico-scientifico (Hutchins, 2009: 16). La lingua parlata, con cui si deve confrontare la Speech-to-Speech Translation, presenta caratteristiche molto diverse rispetto a quella scritta, nonché difficoltà aggiuntive per una macchina: basti pensare a false partenze, riformulazioni e ripetizioni, all'uso di termini colloquiali o di espressioni idiomatiche, alla sintassi tipica del parlato, e così via. Questi elementi possono mettere in difficoltà non solo il modulo di riconoscimento vocale, ma anche quello di traduzione automatica.

Il parlato, inoltre, non è visualizzabile graficamente; ciò significa, ad esempio, che quello che in un testo scritto è facilmente riconoscibile come nome proprio grazie all'iniziale maiuscola, non è altrettanto riconoscibile all'interno di una conversazione, soprattutto se questa viene analizzata da una macchina che non ha conoscenza del mondo. Allo stesso modo, non è possibile visualizzare la punteggiatura in un testo pronunciato (Nakamura,

2009: 36). Il sistema di SST deve quindi essere progettato in modo da riuscire a scomporre la frase enunciata negli elementi che la compongono, percepire pause e cambi di intonazione. Non di rado però insorgono problemi derivanti da una scorretta interpretazione della punteggiatura della frase da parte del sistema, ad esempio nel caso di frasi interrogative scambiate dal sistema per frasi affermative (si pensi all'italiano, dove spesso la domanda viene segnalata esclusivamente dall'intonazione della frase).

Occorre inoltre ricordare che la comunicazione, anche in una situazione in cui sono due parlanti della stessa lingua ad interagire *vis à vis*, “può essere soggetta a molte incomprensioni per la mancanza di un senso comune condiviso” (Lazzari, 2002: 213); tali incomprensioni non sono di certo assenti in una conversazione mediata da un sistema di SST. La comunicazione, come ricorda Lazzari, è spesso orientata al raggiungimento di un obiettivo (*ibid.*) e si basa su meccanismi sottili, difficilmente percepibili da una macchina (come modestia, eufemismo, ironia, ecc.).

Ostacoli inerenti il parlato che interessano in modo specifico la performance del modulo di riconoscimento vocale riguardano la velocità di eloquio, le espressioni non verbali (come risate, colpi di tosse, ecc.; Lazzari, 2002: 213), e la variabilità dell'input vocale (Hutchins, 2009: 16), che si traduce in diversi accenti e pronunce, che possono variare anche fra i parlanti della stessa lingua. Per quanto riguarda quest'ultimo aspetto, occorre considerare che solamente i sistemi di SST *speaker-dependent* possono essere “addestrati” a riconoscere il modo di parlare di un determinato utente; i sistemi *speaker-independent*, invece, pur avendo il pregio di essere utilizzabili da tutti, necessariamente non potranno essere altrettanto precisi nel riconoscimento vocale (Zong e Seligman, 2005).

Ulteriori difficoltà per il modulo ASR hanno origine dal rumore ambientale (Lazzari, 2002: 213), che può interferire con il riconoscimento vocale del vero e proprio “testo” da tradurre. In particolare, nel caso delle app per la traduzione automatica del parlato, che possono essere utilizzate ovunque, compresi i luoghi più affollati, l'input vocale viene trasmesso attraverso il microfono interno del cellulare, e, nel caso non avvenga una corretta soppressione del rumore di sottofondo, questo rischia di pregiudicare la performance del sistema di SST.

Sebbene negli ultimi anni si sia verificato un forte miglioramento nelle prestazioni dei sistemi di riconoscimento vocale, esistono ancora circostanze in cui il livello di accuratezza non è sufficiente. In un sistema di SST, in cui un corretto riconoscimento vocale è condizione fondamentale per ottenere un output tradotto di qualità, eventuali errori del modulo ASR si riflettono anche sulla performance del modulo di traduzione automatica, che tenderà a propagarli, generando traduzioni errate o insensate (Nakamura, 2009: 36).

Anche escludendo errori da parte del modulo ASR, la traduzione automatica della lingua parlata risulta comunque particolarmente difficoltosa. Ciò, fra gli altri motivi, è dovuto al fatto che la comunicazione dialogica è interpersonale e in genere dipendente dal contesto (Hutchins, 2009: 16); questo significa che molti elementi possono essere compresi correttamente solo nel contesto immediato della conversazione, oppure vengono dati per scontati dagli interlocutori e non vengono espressi esplicitamente.

Anche la struttura sintattica e le caratteristiche della frase possono variare da lingua scritta a lingua parlata, rendendo più arduo il compito del modulo di traduzione automatica del sistema di SST; nel parlato alcuni elementi della frase vengono talvolta sottintesi (ellissi), la frase può rimanere incompleta (Hutchins, 2005) e, in italiano, sono frequenti costrutti sintattici come la dislocazione a destra o a sinistra, che modificano l'ordine tradizionale della frase SVO e possono facilmente generare errori di traduzione.

Infine, per quanto riguarda il modulo di sintesi vocale, più che di problematiche vere e proprie, è corretto parlare di sfide; queste riguardano ad esempio la capacità del modulo TTS di riprodurre le emozioni espresse dal parlante e la sua prosodia, di dare un'intonazione corretta alle domande, ecc.

Nella sezione successiva si vedrà che, come per la traduzione automatica, anche nel caso della traduzione automatica del parlato esistono alcuni accorgimenti che consentono di ridurre le probabilità di errore e di sopperire ai limiti della tecnologia SST.

2.5.2 La cooperazione dell'utente come strumento per sopperire ai limiti dei sistemi di traduzione automatica del parlato

Come già sottolineato nella sezione 1.4 in riferimento alla traduzione automatica, anche per la traduzione automatica del linguaggio parlato è attualmente impossibile creare un sistema che sia in grado di tradurre in modo del tutto automatico qualsiasi tipologia di discorso, appartenente a qualsiasi dominio, ottenendo allo stesso tempo un output di alta qualità.

In genere l'approccio dei ricercatori alla traduzione automatica del parlato è sempre stato caratterizzato da una forte consapevolezza di questi limiti. I primi sistemi di SST sviluppati, ad esempio, erano molto limitati in termini linguistici: l'utente doveva attenersi a regole ben precise e utilizzare unicamente le frasi che potevano essere riconosciute dal sistema (Grazina, 2010: 5). Successivamente, sempre per ridurre le difficoltà connesse alla traduzione della lingua parlata, sono stati introdotti sistemi di SST che prevedevano un uso limitato a un solo, o pochi argomenti di conversazione (Hutchins, 2009: 16), come NESPOLE!, incentrato principalmente sull'ambito turistico (Metze et al., 2002), o IraqComm, un sistema di SST sviluppato nell'ambito del progetto TRANSTAC del DARPA e in grado di tradurre fra l'inglese e l'arabo conversazioni limitate all'ambito medico, militare o di primo soccorso (Frandsen et al., 2008: 527). Questo restringeva l'ampiezza del vocabolario utilizzato dagli interlocutori e rendeva la conversazione più prevedibile.

Alcuni sistemi, invece, non prevedevano una traduzione bidirezionale, bensì unidirezionale (come Phaselator), mentre altri ancora erano *speaker-dependent*, oppure erano dotati di funzioni che permettevano all'utente di interagire con la macchina, consentendogli ad esempio di verificare la correttezza della traduzione o di segnalare errori e incomprensioni.

Negli ultimi anni la situazione è molto cambiata: gli evidenti miglioramenti delle tecnologie su cui si basa la Speech-to-Speech Translation hanno reso più ambiziosi i ricercatori attivi in questo campo. Questo ha permesso il lancio di applicazioni o software commerciali per la traduzione automatica del parlato che, come tali, spesso non si limitano a coprire un solo dominio

(o argomento) di conversazione e non sono *speaker-dependent*, ma possono essere utilizzati da tutti.

Questi sistemi di SST, tuttavia, presentano ancora delle imperfezioni, e, sebbene sia sorprendente la loro accuratezza nel tradurre alcune conversazioni o frasi, non di rado forniscono ancora traduzioni errate, che si rivelano più o meno comprensibili per l'utente. Proprio il fatto di essere *speaker-independent* e applicabili a qualsiasi tipo di discorso e dominio rende i sistemi di SST di ultima generazione più esposti agli errori.

È dunque ragionevole supporre che anche strumenti come Skype Translator, un'applicazione in grado di tradurre (almeno teoricamente) qualsiasi tipo di conversazione telefonica, trarrebbero beneficio da una cooperazione da parte dell'utente. Infatti, come già suggerito nel 2005 da Zong e Seligman, a maggiori livelli di *linguistic coverage* (ovvero la capacità del sistema di SST di gestire diverse strutture linguistiche e un ampio vocabolario) e di *topical coverage* (la capacità del sistema di SST di coprire diversi argomenti di conversazione), corrisponde una maggiore probabilità di errore ed è quindi necessaria una maggiore cooperazione da parte dell'utente (Zong e Seligman, 2005) per sopperire alle eventuali mancanze della macchina.

Occorre tenere a mente che nel 2005, anno di pubblicazione del saggio di Zong e Seligman, lo stato dell'arte era molto diverso rispetto al momento della stesura di questo elaborato: la tecnologia SST era confinata ai laboratori e certamente aveva ancora molta strada da fare per poter essere davvero utilizzabile, tanto che un sistema di traduzione automatica del parlato, applicabile a scenari di vita quotidiana e che non richiedesse una cooperazione da parte dell'utente, veniva etichettato come troppo ambizioso per l'epoca. Gli autori avevano come obiettivo quello di incoraggiare un'applicazione sul campo di questa tecnologia, che avrebbe richiesto, per l'appunto, un certo livello di *user cooperation* (*ibid.*).

Nonostante queste evidenti differenze, la teoria di questi due autori sembra conservare la propria validità anche oggi: innanzitutto, come nel 2005, non esiste ancora un sistema di SST “ideale”, che consenta all'utente di ignorare il dispositivo di traduzione e di parlare come se si stesse rivolgendo a un interlocutore della sua stessa lingua, passando da un argomento all'altro liberamente, utilizzando un'ampia gamma di vocaboli, strutture grammaticali

ed espressioni idiomatiche e inserendo nel discorso false partenze ed esitazioni (Zong e Seligman, 2005); in secondo luogo, come per la MT, ora che la traduzione automatica del parlato è finalmente uscita dal laboratorio ed è accessibile al grande pubblico, è necessario sensibilizzare gli utenti rispetto a cosa è realistico aspettarsi da questa tecnologia, ai suoi limiti, ai modi per arginarli e all'uso più corretto di un sistema di SST. Occorre tenere a mente che anche gli interpreti umani possono incorrere in malintesi o errori di traduzione, e spesso, per risolvere i propri dubbi, si trovano a dover chiedere all'interlocutore di ripetere o riformulare la frase “problematica” (*ibid.*): sarebbe dunque irrealistico aspettarsi che le prestazioni di una macchina superino quelle di un interprete.

Dal momento che nella traduzione automatica del parlato alcuni errori di traduzione derivano in realtà da errori di riconoscimento vocale, e considerando che è difficile per l'utente intervenire sull'output fornito dal modulo di ASR se questa possibilità non è prevista dal sistema di SST stesso, una prima e immediata forma di cooperazione da parte dell'utente può concentrarsi sul rendere il riconoscimento vocale il più semplice possibile. Escludendo i casi in cui il sistema di SST fornisce all'utente la possibilità di correggere manualmente o tramite input vocale eventuali errori di riconoscimento (come succede, ad esempio, in alcuni sistemi automatici di dettatura o app per la SST), la cooperazione da parte dell'utente può avvenire a monte del riconoscimento vocale. Secondo le linee guida fornite da Zong e Seligman, per facilitare il riconoscimento vocale e quindi la conversazione, gli utenti della SST dovrebbero (Zong e Seligman, 2005):

- evitare di utilizzare il sistema di SST in ambienti troppo rumorosi;
- parlare in modo chiaro e comprensibile, scandendo bene le parole, senza però usare un tono di voce troppo alto, che causerebbe una distorsione del suono;
- evitare parole rare, forme dialettali o espressioni idiomatiche (che, anche in caso di corretto riconoscimento, creerebbero problemi per il modulo di traduzione automatica);
- fornire possibilmente input ben formati, dunque senza false partenze o riformulazioni.

Altre forme di cooperazione richieste all'utente potrebbero prevedere: la correzione di errori di riconoscimento vocale tramite input vocale o altri mezzi, l'addestramento di modelli acustici *speaker-dependent*, lo spelling di parole che non vengono riconosciute immediatamente, la risoluzione di ambiguità lessicali o strutturali, e così via (Zong e Seligman, 2005). Anche un feedback sulla traduzione potrebbe rivelarsi utile (Seligman, 2000) ai fini della conversazione e per un futuro miglioramento del modulo di traduzione automatica. Come già evidenziato, però, mentre i primi accorgimenti elencati in questa sezione sono applicabili a qualsiasi sistema di SST, questi ultimi possono essere messi in atto solamente se il sistema stesso lo prevede.

Ammettendo che una forma di cooperazione più o meno pronunciata da parte dell'utente porterebbe a un miglioramento della qualità della traduzione del parlato, occorre tuttavia chiedersi se gli utenti della SST siano effettivamente disposti a fornire questo tipo di contributo, e se sì, in quale misura. Secondo Seligman la risposta a questa domanda, e quindi la disponibilità degli utenti della SST a mettere in campo forme di cooperazione e correzione, dipende in larga parte dall'applicazione pratica della traduzione automatica del parlato. In circostanze in cui si renda necessaria una traduzione la più accurata possibile, come in una conversazione d'affari, in caso di emergenza o in campo militare, la disponibilità degli utenti a interagire con il sistema di SST dovrebbe aumentare (*ibid.*). Con il miglioramento della tecnologia SST, anche il livello di cooperazione richiesto ai suoi utenti dovrebbe diminuire gradualmente (Zong e Seligman, 2005).

2.6 Applicazioni pratiche della tecnologia SST

Da circa un decennio la tecnologia SST è finalmente uscita dal laboratorio per farsi strada nel mondo reale. Questo sviluppo è dovuto da un lato al miglioramento delle diverse tecnologie alla base della Speech-to-Speech Translation, e dall'altro alla nascita del mobile computing e al miglioramento dei sistemi informatici in termini di velocità e potenza di calcolo, fattori che,

uniti alla diffusione delle reti, hanno reso possibile l'installazione di sistemi di SST anche su dispositivi mobili (Nakamura, 2009: 42; Van der Meer, 2014). Oggi la tecnologia SST è accessibile a un pubblico vasto, sia da dispositivi mobili dedicati alla traduzione automatica, sia da computer, e infine da smartphone o tablet, previo download di apposite app.

Le possibili applicazioni pratiche della tecnologia SST sono varie e si inscrivono principalmente nei seguenti ambiti:

- ambito turistico: una delle applicazioni più ovvie della Speech-to-Speech Translation è quella della comunicazione in ambito turistico (Hutchins, 2009: 17). Nel corso degli anni diversi gruppi di ricercatori hanno sviluppato sistemi di SST progettati appositamente per l'uso da parte di turisti durante soggiorni all'estero (si veda NESPOLE! in 2.3). Oltre a questi, oggi i turisti possono disporre anche di numerose app per la traduzione del parlato, scaricabili su smartphone o tablet e utilizzabili in presenza o assenza di una connessione dati (a seconda dell'app). Se da un lato è vero che ormai l'inglese è diffuso a livello internazionale e facilita le comunicazioni anche in ambito turistico, è altrettanto vero che sono ancora in molti a non parlarlo: in questo contesto le app di traduzione automatica del parlato consentono ai turisti all'estero di avere sempre a portata di mano uno strumento in grado di assisterli nella richiesta di indicazioni o informazioni, nella prenotazione di voli o alloggi, e in generale nella comunicazione con la popolazione locale;
- campo militare: in questo caso dispositivi mobili di traduzione automatica del parlato vengono utilizzati da militari e forze dell'ordine nel corso di operazioni all'estero, per comunicare con la popolazione locale. In campo militare la tecnologia SST viene impiegata ormai da tempo dall'esercito degli Stati Uniti, che nel corso degli anni ha adottato numerosi sistemi di SST, fra cui DIPLOMAT (un sistema sviluppato dalla Carnegie Mellon University appositamente per l'utilizzo in campo militare; Hutchins, 2009: 17), IraqComm (si veda sezione 2.5.2) e il frasario elettronico Phraselator (si veda sezione 2.4). La traduzione automatica del parlato si è rivelata utile anche in situazioni di emergenza, come durante la fase di

aiuti umanitari seguita allo tsunami che nel 2004 devastò le coste di Indonesia, Thailandia, Sri Lanka e altri paesi asiatici (Hutchins, 2009: 17);

- ambito commerciale: la traduzione automatica del parlato è stata applicata in passato anche a trattative commerciali, come nel progetto Verbmobil (si veda sezione 2.3). In uno dei suoi possibili scenari di impiego, Verbmobil consentiva a interlocutori di lingua giapponese e tedesca, coinvolti in una trattativa commerciale in inglese, di ricorrere all'aiuto della traduzione automatica del parlato (da giapponese a inglese e da tedesco a inglese) nel caso in cui la loro conoscenza attiva dell'inglese si dimostrasse insufficiente a proseguire la conversazione (Bub et al., 1997). Benché l'ambito commerciale rappresenti un potenziale scenario di impiego della SST, è ancora preferibile ricorrere ad un interprete professionista, in considerazione del fatto che la tecnologia SST presenta ancora imperfezioni e che la buona riuscita della trattativa commerciale, fortemente dipendente dal corretto svolgimento della comunicazione, è in genere di fondamentale importanza per le parti coinvolte;
- campo medico: la traduzione automatica del parlato potrebbe essere utilizzata anche per mediare la comunicazione fra medico e paziente, soprattutto nel caso in cui quest'ultimo appartenga a una minoranza culturale e linguistica, o nel caso in cui non si disponga di un interprete professionista. Già nel 2009 Hutchins delineava questo possibile scenario di impiego per la SST, pur ammettendo che nella comunicazione fra medico e paziente straniero entrano spesso in gioco anche fattori culturali o ambientali, che un sistema di SST da solo non è in grado di gestire (Hutchins, 2009: 17).

Potenziati applicazioni pratiche della tecnologia SST che potrebbero delinearsi in futuro riguardano poi il settore pubblico (nei casi in cui sia prevista, ad esempio, un'interazione fra le autorità pubbliche e immigrati), l'ambito mediatico e le interazioni B2C (Business to Customer, come nel caso di call center o servizio clienti; Van der Meer, 2014).

2.7 Prospettive e sfide future per la Speech-to-Speech Translation

Come evidenziato nelle sezioni precedenti, la traduzione automatica del parlato ha già fatto enormi passi avanti rispetto ai primi sistemi sviluppati negli anni Ottanta e Novanta; tuttavia, le sfide da affrontare nei prossimi anni sono numerose. Innanzitutto i sistemi di SST, oggi in grado di tradurre fra poche coppie di lingue, dovranno supportare un numero sempre maggiore di combinazioni linguistiche (Nakamura, 2009: 44) per potere essere davvero accessibili a migliaia di utenti in diversi paesi. Questo richiederà un grande sforzo, dal momento che per ogni lingua aggiunta al sistema di SST sono necessari grandi corpora di dati relativi alla lingua parlata e alla lingua scritta, monolingue e bilingue (*ibid.*). Secondo Alex Waibel, professore di informatica presso la Carnegie Mellon University di Pittsburgh e il Karlsruher Institut für Technologie, l'aggiunta di nuove combinazioni linguistiche richiederà nuovi metodi per la progettazione dei sistemi di SST, un processo in cui dovranno essere coinvolte le varie comunità linguistiche. Solo così, infatti, si potrebbe adattare la tecnologia SST a nuove lingue senza dover ricorrere all'intervento di esperti con competenze sia informatiche che linguistiche (Van der Meer, 2014).

Una seconda sfida per la SST consisterà nel miglioramento del riconoscimento vocale in ambienti molto rumorosi, come ristoranti, aeroporti, stazioni ferroviarie, e così via (*ibid.*). La sfida principale, tuttavia, riguarderà sempre la lingua parlata. Attualmente molti sistemi di SST vengono utilizzati per la traduzione di singole frasi, spesso brevi (come nel caso delle app per smartphone o tablet). La traduzione di lunghi interventi frammentati e “disordinati” dal punto di vista sintattico, invece, resta ancora un compito molto complesso, che dà spesso luogo a output di bassa qualità (*ibid.*). Per migliorare le prestazioni della SST anche in quest'ultimo caso, dovranno essere elaborati nuovi metodi di “content extraction, summarization, and semantic and discourse modeling” (Van der Meer, 2014). Un'altra questione fondamentale riguarda l'affidabilità della traduzione automatica del parlato. Come sottolineato da Mark Seligman, ricercatore

attivo nel campo della SST e della MT, “[u]sers won't trust automatic translation for serious uses unless effective verification and correction capability is provided” (Van der Meer, 2014). La possibilità di verificare la correttezza della traduzione fornita, ed eventualmente di correggerla, non solo aumenterebbe la fiducia degli utenti nei confronti della tecnologia SST, ma consentirebbe anche agli utenti di fornire un feedback utile al miglioramento della tecnologia stessa (*ibid.*).

Se la Speech-to-Speech Translation riuscirà ad affrontare con successo tutte queste sfide, la sua diffusione a livello commerciale, ora limitata, potrà verosimilmente aumentare, come è già successo in passato per altre innovazioni di rottura (si veda 2.2). Infatti, se da un lato si può ipotizzare che l'inglese si confermerà come lingua franca utilizzata a livello mondiale (accompagnata forse da poche altre lingue, come lo spagnolo e il cinese mandarino), dall'altro è difficile pensare che tutti saranno in grado di parlare correntemente queste lingue, e che le circa 6000-7000 lingue esistenti sul pianeta si possano ridurre a poche decine (Van der Meer, 2014). In questo senso, in futuro la tecnologia SST potrebbe rappresentare un valido sostegno alla comunicazione qualora non esista una lingua condivisa tra gli interlocutori e qualora sia impossibile o economicamente svantaggioso ricorrere a un interprete professionista (basti pensare all'ambito turistico).

2.8 La valutazione della qualità nella traduzione automatica del parlato

Come sottolineato nella sezione 1.7.1 di questo elaborato, valutare la qualità di una traduzione non è un compito affatto semplice, principalmente perché “quality of a translation is a subjective measure” (Tirronen, 2011: 26). Se si considera la valutazione della qualità di un testo orale, la situazione si complica ulteriormente. Nel saggio “Aspetti della qualità in interpretazione”, Viezzi sottolinea “l'intrinseca difficoltà di analisi di un prodotto evanescente come il testo orale” (Viezzi, 1999: 142). Pur non spingendosi fino a paragonare l'output di un sistema di traduzione automatica del parlato a

quello prodotto da un interprete umano, una difficoltà simile a quella descritta da Viezzi sembra riscontrabile anche nella valutazione della traduzione automatica del parlato.

Benché sia “extremely difficult theoretically to evaluate the accuracy of speech translation” (Nakamura, 2009: 39), come per altre *Human Language Technologies*, anche per la traduzione automatica del parlato la valutazione della qualità è fondamentale, non solo per gli utenti stessi della SST, ma anche, e soprattutto, ai fini del progresso scientifico e tecnologico in questo campo.²⁶ A questo scopo ogni anno vengono organizzati alcuni *International evaluation workshops*, workshop internazionali durante i quali gli istituti di ricerca partecipanti sono chiamati a creare sistemi di SST partendo da un “common dataset” (Nakamura, 2009: 41), in quella che diventa una vera e propria competizione. La valutazione dei vari sistemi così creati, dei loro pregi e punti deboli, rappresenta un importante motore per lo sviluppo della tecnologia SST (*ibid.*).

Attualmente, come per la MT, la comunità scientifica adotta diversi metodi per valutare la qualità nella traduzione automatica del parlato (per una definizione di qualità, si veda 1.7.1). Innanzitutto è possibile distinguere le varie tipologie di valutazione in base all'oggetto della valutazione stessa. Nella SST, infatti, un primo tipo di valutazione si concentra sulla qualità audio dell'output, ovvero sulla performance del modulo di sintesi vocale;²⁷ un secondo tipo di valutazione è invece incentrato sul “meaning preservation”,²⁸ quindi sul contenuto dell'output tradotto e sulla “equivalence between the source and the target language” (Tirronen, 2011: 26).

Per quest'ultimo tipo di valutazione, Nakamura descrive un metodo simile a quello adottato per la valutazione della qualità nella traduzione automatica (Nakamura, 2009: 39): escludendo la valutazione della performance del modulo di sintesi vocale, è sufficiente sottoporre al sistema SST una serie di frasi (pronunciate oralmente) da tradurre, e valutare l'output prodotto,

²⁶ ELRA (European Language Resources Association), Human Language Technologies Evaluation: Why is Evaluation important?. Visitato: 11/04/2016. Link: <http://www.hlt-evaluation.org/spip.php?article139>

²⁷ ELRA (European Language Resources Association), Human Language Technologies Evaluation: Speech-to-Speech Translation, Measures. Visitato: 11/04/2016. Link: <http://www.hlt-evaluation.org/spip.php?article146>

²⁸ *Ibid.* Link: <http://www.hlt-evaluation.org/spip.php?article146>

secondo parametri oggettivi o soggettivi.

L'adozione di parametri di valutazione oggettivi, come BLEU, NIST (si veda sezione 1.7.2) e WER (Word Error Rate), è indubbiamente vantaggiosa: da un lato, producendo come risultato valori numerici, permette di confrontare le prestazioni di sistemi diversi in modo oggettivo (Nakamura, 2009: 39); dall'altro consente ai ricercatori di effettuare in modo rapido ed economico una grande quantità di valutazioni (Grazina, 2010), processo fondamentale durante le fasi di sviluppo di un nuovo sistema di SST.

Un'altra opzione è quella della valutazione soggettiva, basata cioè sul giudizio di traduttori e interpreti professionisti, linguisti, parlanti bilingue o altri esperti (Grazina, 2010: 11). In questo caso la valutazione può essere effettuata tramite il confronto fra l'output prodotto dal sistema di SST e una o più traduzioni di riferimento preparate in precedenza (Nakamura, 2009: 39), oppure fra testo *source* e testo *target*. A seconda dei casi, la valutazione soggettiva dell'output può basarsi anche su test di comprensione o questionari di vario tipo (compilabili anche da utenti monolingue), ma in genere, anche in questi casi, alla base c'è comunque un confronto fra l'output automatico e una traduzione di riferimento (effettuata idealmente da un interprete umano).²⁹

Pur presentando alcuni svantaggi pratici, quali i costi elevati, e altri svantaggi relativi all'efficacia (si veda sezione 1.7.2) la valutazione soggettiva resta molto importante: “because of the linguistic knowledge humans possess, subjective evaluation reflects the actual, true quality of the translation” (Tirronen, 2011: 26).

È inoltre possibile sia valutare il sistema SST nel suo complesso, sia valutare separatamente i moduli che lo compongono, in modo da identificare i singoli moduli responsabili di eventuali errori contenuti nell'output (si veda Rossato et al., 2002). Un metodo per compiere questo tipo di valutazione è quello di verificare dapprima la correttezza dell'output del modulo ASR nella lingua *source* (quando possibile), e solo in un secondo luogo la correttezza della traduzione effettuata dal modulo MT (Tirronen, 2011: 27).

Occorre infine considerare che lo scenario di impiego dei sistemi di SST è

²⁹ *Ibid.* Link: <http://www.hlt-evaluation.org/spip.php?article146>

quello di una conversazione fra interlocutori che parlano due lingue diverse. Ecco perché è importante che questi sistemi vengano testati anche tramite *field experiments*, ovvero test condotti nel mondo reale, o, in generale, in situazioni in cui l'interazione fra gli interlocutori dipende dalla “mediazione” garantita dal sistema di traduzione automatica del parlato. L'importanza di questo tipo di test valutativo viene sottolineata anche da Nakamura, che, a proposito dei metodi oggettivi di valutazione della qualità, spiega come questi, pur avendo molti lati positivi, non possano fornire un'indicazione sulla performance del sistema di SST nel mondo reale (Nakamura, 2009: 39). Un *field experiment* prevede in genere l'interazione fra interlocutori che parlano lingue diverse e può essere condotto, ad esempio, in uffici di informazioni turistiche o in altri contesti in cui abitualmente avviene il contatto fra persone di lingue e culture diverse. Ai partecipanti può essere chiesto di portare a termine incarichi più o meno complessi; in un test condotto in ambito turistico tali incarichi potrebbero includere la richiesta di indicazioni o informazioni, la prenotazione di un alloggio, e così via. Infine viene valutata l'efficacia della comunicazione mediata dal sistema SST, analizzando ad esempio il numero di frasi che è stato necessario ripetere, il numero di frasi tradotte correttamente, il numero di incarichi portati a termine dai partecipanti, ecc., a seconda dei parametri scelti dai ricercatori. Nel prossimo capitolo verrà presentato un test di valutazione di due applicazioni per la traduzione automatica del parlato, condotto alla luce delle precedenti considerazioni sulla valutazione della qualità in SST. Di tale test verranno descritte in particolare la metodologia, gli obiettivi, lo svolgimento e le modalità di valutazione adottate.

III

Google Translate e Microsoft Translator: Valutazione di due applicazioni per la traduzione automatica del parlato

3.1 Introduzione

In questo capitolo verrà presentato un test di valutazione condotto su due applicazioni per smartphone e tablet create da due multinazionali concorrenti che operano nel settore informatico e da anni si occupano di traduzione automatica. Le applicazioni oggetto di valutazione sono Google Translate (sviluppata da Google Inc.) e Microsoft Translator (sviluppata da Microsoft Corporation), che verranno analizzate e valutate unicamente in merito al servizio di traduzione automatica del parlato.

È importante sottolineare che con il test descritto in questo capitolo non si intende dare un giudizio complessivo della performance delle due applicazioni. La valutazione è da considerarsi limitata innanzitutto al servizio di traduzione automatica del parlato e in secondo luogo agli output prodotti dalle due applicazioni nell'ambito del test condotto.

In questo capitolo verrà presentata la metodologia adottata per lo svolgimento del test. Verranno descritti dapprima gli obiettivi del test, le caratteristiche delle applicazioni d'oggetto d'analisi, il materiale selezionato per la traduzione automatica del parlato e i partecipanti; infine verrà illustrato lo svolgimento del test di valutazione, nonché i criteri adottati per valutare i risultati, che verranno presentati nel capitolo IV.

3.2 Obiettivi del test

L'obiettivo del test di valutazione descritto in questo capitolo è quello di valutare, analizzare e confrontare gli output prodotti da due applicazioni di traduzione automatica del parlato, utilizzate per la traduzione di singole frasi appartenenti a un ambito specifico di conversazione, nelle combinazioni linguistiche italiano-inglese e italiano-tedesco. La scelta di questo tipo di test è supportata dalla tesi di Nakamura (si veda anche 2.8), secondo il quale per la Speech-to-Speech Translation

[i]f the evaluation of the speech synthesis module is not included, evaluation is made by feeding a number of test sentences into the system, and evaluating the quality of the output.

(Nakamura, 2009: 39)

Come già sottolineato, l'obiettivo del test descritto non è quello di dare un giudizio complessivo ed esauriente della performance delle due applicazioni oggetto di analisi. Si è infatti consapevoli del fatto che la portata del test condotto non è sufficiente per una valutazione di questo tipo. La valutazione effettuata sarà dunque da considerarsi limitata in primo luogo al servizio di traduzione automatica del parlato e in secondo luogo agli output prodotti dalle due applicazioni per le combinazioni linguistiche scelte e nell'ambito del test condotto.

In questo test si è deciso di prendere in considerazione unicamente gli output prodotti dai moduli ASR e MT, che verranno valutati separatamente, come spiegato nella sezione 3.7.1 del presente capitolo. Per quanto riguarda la valutazione dell'output del modulo ASR, l'obiettivo principale è quello di verificare in quali e quanti casi un riconoscimento vocale scorretto sia responsabile di errori di traduzione nell'output tradotto. Nella sezione 2.5.1 si è infatti visto come spesso gli errori del modulo ASR si riflettano sulla performance del modulo MT. In aggiunta a questo tipo di valutazione, in generale verranno presi in considerazione anche dati relativi all'accuratezza del riconoscimento vocale e all'usabilità delle due applicazioni.

Per quanto riguarda il modulo MT, invece, l'obiettivo del test di valutazione

è quello di verificare se e in quanti casi gli output prodotti risultino usufruibili e comprensibili, valutandone altresì la qualità dal punto di vista traduttivo.

Come già accennato, si è scelto di escludere il modulo TTS dalla valutazione, in considerazione del fatto che generalmente non è responsabile di errori gravi. Benché infatti alcuni sistemi di SST dispongano di moduli di sintesi vocale di qualità superiore rispetto ad altri (ovvero in grado di produrre output più naturali), in genere la qualità audio dell'output sintetizzato non ostacola la comprensione da parte dell'interlocutore. Inoltre quando separata dalla valutazione della MT, la valutazione del modulo TTS si basa in genere su parametri relativi alle caratteristiche audio dell'output sintetizzato ed è finalizzata a migliorare le prestazioni del modulo stesso, cosicché questo possa produrre output più naturali, meno artificiali e sempre più vicini al modo di parlare degli utenti. In questa sede, invece, il focus principale dell'analisi è sulla traduzione e sulla capacità delle applicazioni valutate di “mediare” la comunicazione fra interlocutori di lingue diverse. La valutazione del modulo TTS non sembrava dunque rientrare nell'obiettivo dell'analisi condotta in questo studio.

Si è infine considerato che nelle applicazioni oggetto di analisi, la qualità audio dell'output prodotto dal modulo TTS non pregiudicherebbe in alcun modo la riuscita di un'ipotetica conversazione multilingue, dal momento che gli utenti non dipendono dalla sintesi vocale, ma possono leggere l'output tradotto dal modulo MT anche sul display dello smartphone, come si vedrà in seguito.

3.3 Applicazioni oggetto di valutazione

3.3.1 Applicazioni scelte

Le applicazioni oggetto di valutazione nell'ambito del test condotto sono Google Translate e Microsoft Translator, sviluppate rispettivamente da Google Inc. e Microsoft Corporation. La scelta di queste due applicazioni è stata dettata da diverse ragioni. Innanzitutto entrambe sono scaricabili gratuitamente da smartphone e tablet e facilmente accessibili ad un vasto pubblico. Entrambe, inoltre, offrono un servizio di Speech-to-Speech Translation *speaker-independent* e non limitato a un solo dominio o argomento di conversazione. Ciò significa che possono essere utilizzate da qualsiasi utente che parli una delle lingue supportate, per tradurre potenzialmente qualsiasi frase verso una seconda lingua fra quelle disponibili. Si tratta infine di applicazioni che presentano funzioni e servizi molto simili, dunque paragonabili, ma essendo state sviluppate da aziende concorrenti, si basano su sistemi di SST senza dubbio differenti.

Pur offrendo servizi di traduzione automatica molto diversificati, come si vedrà di seguito, nell'ambito del test condotto le applicazioni scelte sono state valutate solo limitatamente al servizio di traduzione automatica del parlato. Nelle prossime sezioni, verranno descritte nel dettaglio le caratteristiche e il funzionamento di Google Translate e Microsoft Translator.

3.3.2 Google Translate

La prima applicazione oggetto di valutazione è Google Translate, sviluppata da Google Inc., disponibile gratuitamente per i sistemi operativi Android e iOS e scaricabile da smartphone e tablet (in questo caso scaricata da smartphone). L'applicazione si basa sul servizio di MT online offerto dalla piattaforma Google Translate e produce quindi traduzioni automatiche fondate sull'approccio statistico (per una definizione di traduzione

automatica statistica, si veda sezione 1.3).

L'app Google Translate (a cui da qui si farà riferimento anche come GT) traduce in circa 90 lingue (rispetto alle oltre 100 supportate dalla piattaforma online) e, oltre alla classica traduzione automatica di testo digitato, offre servizi studiati appositamente per l'utilizzo da dispositivo mobile.³⁰ Questi includono ad esempio:

- le modalità “Scatti” e “Visualizzazione”,³¹ che consentono di utilizzare la fotocamera per sottoporre all'app il testo da tradurre desiderato;³² grazie a queste funzioni è sufficiente inquadrare (modalità “Visualizzazione”) o fotografare (modalità “Scatti”) il testo da tradurre per ottenerne la traduzione istantanea;
- la modalità “Scrittura”, che consente di tracciare il testo da tradurre con il dito, utilizzando il touchscreen del dispositivo;³³ questa funzione si rivela utile soprattutto per lingue come il cinese e il giapponese, che non utilizzano l'alfabeto latino, bensì caratteri, pittogrammi e ideogrammi (generalmente non disponibili nella tastiera touch dei dispositivi mobili venduti in Europa);
- la modalità “Conversazione”, che consente di intrattenere conversazioni bilingue;³⁴ questa funzione, disponibile per circa 40 lingue,³⁵ si basa sulla tecnologia Speech-to-Speech Translation.

Alcune delle funzioni elencate sono disponibili anche offline (ovvero in assenza di connessione internet) previo download dei pacchetti dati relativi alle lingue desiderate (opzione disponibile solo per dispositivi Android o per dispositivi Apple limitatamente alla modalità “Visualizzazione”). Occorre ricordare, inoltre, che alcune funzioni sono disponibili solo in un numero

³⁰ Google, Google Play Apps: Google Traduttore. Visitato: 20/04/2016. Link: <https://play.google.com/store/apps/details?id=com.google.android.apps.translate>

³¹ Google, Google Translate: Lingue. Visitato: 20/04/2016. Link: http://translate.google.it/about/intl/it_ALL/languages/

³² Google, Google Play Apps: Google Traduttore. Visitato: 20/04/2016. Link: <https://play.google.com/store/apps/details?id=com.google.android.apps.translate>

³³ Google, Google Translate: Lingue. Visitato: 20/04/2016. Link: http://translate.google.it/about/intl/it_ALL/languages/

³⁴ *Ibid.* . Link: http://translate.google.it/about/intl/it_ALL/languages/

³⁵ Google, Google Play Apps: Google Traduttore. Visitato: 20/04/2016. Link: <https://play.google.com/store/apps/details?id=com.google.android.apps.translate>

ridotto di lingue, rispetto alle circa 90 offerte dall'applicazione.³⁶

La valutazione di Google Translate nell'ambito del test descritto in questo capitolo verte unicamente sulla traduzione automatica del parlato, ovvero sulla modalità “Conversazione” dell'app. Per accedere a questa funzione, dalla schermata principale è sufficiente selezionare le lingue desiderate e successivamente l'icona del microfono, come mostrato in Figura 2.

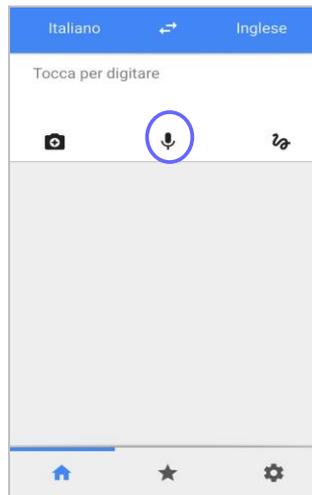


Figura 2
Schermata principale dell'app Google Translate.

Una volta selezionata l'icona del microfono, si aprirà una nuova schermata (Figura 3).



Figura 3
Schermata della modalità “Conversazione” (GT).

³⁶ *Ibid.* Link: <https://play.google.com/store/apps/details?id=com.google.android.apps.translate>

All'utente sarà sufficiente pronunciare la frase da tradurre nella lingua selezionata (visualizzata in rosso) come *source* (input vocale nella lingua *source*); nella parte superiore della schermata comparirà poi la trascrizione della frase pronunciata così come riconosciuta dal modulo ASR. A partire da questo primo output verrà prodotta una traduzione automatica nella lingua *target*, che comparirà immediatamente nella parte inferiore dello schermo (output tradotto in forma scritta). Qualora non si avvii automaticamente la sintesi vocale della traduzione, sarà possibile ascoltarla selezionando nuovamente l'icona del microfono; in questo modo si interromperà il processo di riconoscimento vocale e la traduzione verrà riprodotta dal modulo di sintesi vocale (output vocale nella lingua *target*). Per passare alla direzione opposta di traduzione, è sufficiente selezionare sul display la nuova lingua che si desidera utilizzare come *source*. Toccando invece il microfono al centro della schermata (Figura 3) verrà attivata la funzione “Ascolto in entrambe le lingue”, che permette all'applicazione di rilevare automaticamente la lingua dell'input vocale immesso.

La funzione “Conversazione” può essere particolarmente utile sia nel caso in cui l'utente che si avvale della traduzione automatica non sia in grado di pronunciare l'output prodotto, sia nel caso in cui due utenti di lingue diverse vogliano intrattenere una conversazione mediata dall'app. Di fatto, la modalità “Conversazione” offre una vera e propria traduzione automatica del parlato pensata per un contesto interattivo: si parte da un input orale, che viene riconosciuto dal modulo ASR quasi istantaneamente, per poi passare al modulo di traduzione automatica, che lo traduce altrettanto velocemente (avvalendosi in questo caso di un approccio statistico); infine l'output tradotto viene riprodotto dal modulo di sintesi vocale.

Durante questi passaggi è possibile dapprima verificare l'accuratezza del riconoscimento vocale, il che consente all'utente di ripetere o digitare manualmente la frase da tradurre nel caso in cui il riconoscimento non sia andato a buon fine, e in un secondo momento visualizzare il testo tradotto in forma scritta. Non esiste tuttavia alcuna possibilità di intervento sulla traduzione fornita dall'applicazione, né la possibilità di fornire un feedback (opzione prevista solo per la piattaforma online di Google Translate). Se l'utente lo desidera, è possibile invece salvare le traduzioni in un elenco di

preferiti, selezionando l'apposita icona a forma di stella dalla schermata principale dell'applicazione.

Al momento della stesura di questo elaborato, Google Translate offre la traduzione automatica del parlato per circa 30³⁷ delle 90 lingue supportate dall'applicazione.

3.3.3 Microsoft Translator

La seconda app oggetto di valutazione è Microsoft Translator, sviluppata da Microsoft Corporation e basata sulle API di Microsoft Translator, che possono essere integrate in qualsiasi sito web, applicazione, dispositivo elettronico e sistema operativo dietro pagamento di una quota mensile.³⁸ L'applicazione (a cui da qui si farà riferimento anche come MsT) è compatibile con i sistemi operativi Android, iOS e con Windows Phone ed è scaricabile gratuitamente da smartphone, tablet e smartwatch (in questo caso è stata scaricata da smartphone).

L'app Microsoft Translator, basata sulla traduzione automatica statistica, consente di tradurre testi scritti o pronunciati oralmente e offre le seguenti funzioni:

- traduzione automatica scritta di testo digitato;
- traduzione automatica tramite fotocamera: avviando questa funzione, è sufficiente accedere alle foto già salvate sul dispositivo o scattare una foto al testo desiderato per ottenere la traduzione automatica. Questa funzione è particolarmente utile durante i soggiorni all'estero per tradurre automaticamente cartelli, insegne, menù o volantini,³⁹ alla pari delle corrispondenti modalità “Scatti” e “Visualizzazione” offerte da GT;
- funzione di traduzione automatica del parlato: questa funzione,

³⁷ Google, Google Translate: Lingue. Visitato: 20/04/2016. Link:

http://translate.google.it/about/intl/it_ALL/languages/

³⁸ Microsoft, Microsoft Translator: Getting started using the Translator API. Visitato: 20/04/2016. Link:

<https://www.microsoft.com/en-us/translator/getstarted.aspx>

³⁹ Microsoft, Microsoft Translator: Microsoft Translator for iPhone. Visitato: 20/04/2016. Link:

<https://www.microsoft.com/en-us/translator/iPhone.aspx>

disponibile per circa 50 lingue, consente all'utente di pronunciare brevi frasi che vengono poi tradotte e riprodotte dall'applicazione.⁴⁰ Come nel caso della corrispondente modalità “Conversazione” di GT, anche questa funzione si basa sulla tecnologia Speech-to-Speech Translation;

- funzione di traduzione automatica di conversazioni: questa funzione, disponibile per una decina di lingue, consente di tradurre automaticamente un input vocale, ma in questo caso l'output tradotto non viene riprodotto dal modulo di sintesi vocale, bensì mostrato in forma scritta sul display del dispositivo, all'interno di una schermata bilingue (Figura 4). Questo accorgimento è pensato per consentire la conversazione anche in ambienti rumorosi.⁴¹ Inoltre, la funzione di traduzione automatica di conversazioni consente di collegare lo smartphone a uno smartwatch (su entrambi i dispositivi deve essere installata l'applicazione Microsoft Translator e selezionata la funzione apposita): questo consente a due interlocutori di utilizzare due dispositivi diversi per intrattenere una conversazione a breve distanza.⁴² La traduzione automatica del parlato avviene in questo caso grazie al modulo ASR e al modulo MT.

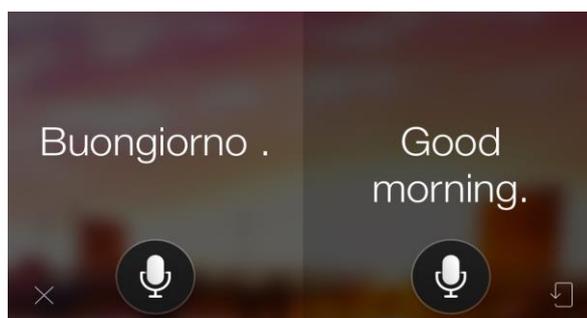


Figura 4
Schermata bilingue della funzione conversazione (MsT).

⁴⁰ *Ibid.* Link: <https://www.microsoft.com/en-us/translator/iPhone.aspx>

⁴¹ Apple, iTunes Apple Store: Microsoft Translator. Visitato: 20/04/2016. Link: <https://itunes.apple.com/it/app/microsoft-translator/id1018949559?mt=8>

⁴² Microsoft, Microsoft Translator: Microsoft Translator for iPhone. Visitato: 20/04/2016. Link: <https://www.microsoft.com/en-us/translator/iPhone.aspx>

Anche per Microsoft Translator, alcune delle funzioni descritte sono disponibili offline: in assenza di connessione internet è comunque possibile tradurre testi o immagini previo download dei pacchetti dati relativi alle lingue desiderate. La modalità offline non è invece prevista per la traduzione del parlato. Come nell'app Google Translate, anche in Microsoft Translator è poi possibile salvare le traduzioni per recuperarle in un secondo momento. Le lingue supportate dall'applicazione sono circa 50, tuttavia alcune funzioni sono disponibili per un numero più limitato di lingue (ad esempio, la funzione di conversazione), anche se Microsoft dichiara di porsi come obiettivo quello di migliorare il servizio ampliando la rosa di lingue disponibili per la traduzione.⁴³

Anche il test di valutazione condotto su MsT verte unicamente sulla funzione di traduzione automatica del parlato. Benché esista una funzione pensata appositamente per la traduzione di conversazioni bilingue (funzione di traduzione automatica di conversazioni), la funzione di traduzione automatica del parlato è l'unica a comprendere anche la sintesi del parlato, che tuttavia non verrà valutata in questa sede in quanto poco soggetta a errori (si ricorda però che un sistema di SST completo si basa su tre tecnologie: ASR, MT e TTS). Questa funzione risulta inoltre quasi del tutto equivalente alla modalità “Conversazione” di GT scelta per il test, sia in termini di funzionamento che in termini di “grafica” delle schermate.

Per accedere alla funzione di traduzione automatica del parlato di Microsoft Translator, occorre selezionare l'icona del microfono dalla schermata principale dell'applicazione. Si aprirà quindi una schermata da cui sarà possibile selezionare le lingue desiderate e, una volta effettuata la scelta, sarà sufficiente selezionare l'icona del microfono per avviare il riconoscimento vocale (si veda Figura 5). Quando il microfono diventa verde, si può iniziare a parlare.

⁴³ *Ibid.* Link: <https://www.microsoft.com/en-us/translator/iPhone.aspx>

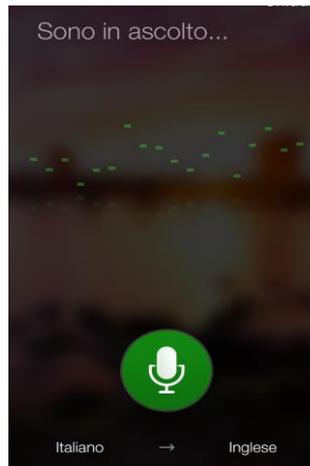


Figura 5
Schermata della funzione di traduzione automatica del parlato (MsT).

Dopo avere pronunciato la frase da tradurre, occorreranno alcuni secondi perché essa appaia sullo schermo così come riconosciuta dal modulo ASR. Il testo riconosciuto dal modulo di ASR viene tradotto quasi istantaneamente dal modulo di MT, che produce un output in lingua *target* immediatamente visibile sul display del dispositivo. Occorrerà attendere ancora qualche secondo perché il modulo TTS effettui la sintesi vocale della traduzione.

Utilizzando questa funzione, per invertire la direzione di traduzione occorre selezionare entrambe le lingue e modificarle una ad una manualmente. Per continuare invece a tradurre nella direzione selezionata, è sufficiente toccare nuovamente l'icona del microfono.

Come GT, anche MsT consente all'utente di visualizzare l'output del modulo ASR, permettendogli così di individuare immediatamente eventuali errori di riconoscimento vocale e di correggerli tramite la digitazione manuale o la ripetizione della frase. Non è invece possibile intervenire sull'output tradotto dal modulo MT, né fornire un feedback sulla traduzione, esattamente come nel caso di GT. Gli utenti, tuttavia, possono avere accesso a una serie di forum online suddivisi per categoria, creati da Microsoft allo scopo di raccogliere idee e suggerimenti per il miglioramento della tecnologia e dei prodotti di Microsoft Translator.⁴⁴

⁴⁴ Microsoft, Microsoft Translator User Voice: Microsoft Translator Apps – Feedback and Support. Visitato: 20/04/2016. Link: <https://microsofttranslator.uservoice.com/>

3.4 Materiale utilizzato per il test

L'ambito per cui si è scelto di testare le performance delle due app è quello turistico. Come si è visto nella sezione 2.6, i sistemi di Speech-to-Speech Translation installati su dispositivi mobili possono infatti essere utilizzati da utenti che soggiornano all'estero, ma non parlano la lingua locale, oppure possiedono una competenza linguistica limitata. In questo caso i sistemi di SST offrono supporto nella richiesta di informazioni o indicazioni, e in generale nell'interazione con interlocutori stranieri o con i residenti del luogo.

La scelta dell'ambito turistico per questo test è stata dettata inoltre dalle seguenti considerazioni:

- allo stato dell'arte attuale, per la traduzione in ambiti come quello medico e commerciale in genere si preferisce ancora ricorrere a un interprete professionista, piuttosto che utilizzare strumenti di Speech-to-Speech Translation *speaker-independent* e applicabili a un dominio illimitato, come quelli analizzati in questo capitolo; in ambito turistico, invece, è raro che venga richiesto l'intervento di un interprete (escluse situazioni particolari, come nel caso di viaggi di gruppo, visite guidate, viaggi d'affari, ecc.). In questo scenario le applicazioni di traduzione automatica del parlato possono rappresentare un'alternativa tecnologica ai tradizionali frasari per turisti;
- le app di traduzione automatica sono perlopiù pensate per un uso generale e non rappresentano lo strumento più adatto per la traduzione di testi ricchi di terminologia settoriale; per la traduzione del parlato in ambito “tecnico” (come potrebbe essere quello medico o commerciale), nel corso degli anni sono stati progettati sistemi specifici e più idonei all'adempimento di questo compito;
- il fatto che le app scelte per il test siano progettate per l'utilizzo su dispositivi mobili facilmente trasportabili, quali smartphone e tablet, le rende particolarmente adatte all'uso da parte di turisti.

Per tutti questi motivi, l'ambito turistico è parso essere lo scenario più

realistico fra quelli disponibili, nonché il più vicino all'utilizzo reale che si suppone venga fatto delle applicazioni oggetto d'analisi.

Per effettuare il test sono state selezionate ventuno frasi in italiano, tratte e adattate dalle versioni italiano-inglese e italiano-tedesco di una nota collana di frasari per turisti. Questo tipo di frasario, contenente glossari e frasi utili tradotte, copre in genere diversi temi legati al viaggio, inclusi la prenotazione dell'alloggio, la richiesta di informazioni e indicazioni, le ordinazioni al ristorante, gli spostamenti, gli acquisti, ecc. Le frasi selezionate sono in genere abbastanza brevi e variano per grado di complessità. Ai fini della valutazione sono state modificate parzialmente o integrate, ad esempio tramite l'inserimento, ove necessario, di toponimi specifici.

Tre frasi sono presenti in due varianti, che si differenziano proprio per la scelta del toponimo: si tratta di frasi contenenti richieste di informazioni, che sono state completate con il nome di una città della Gran Bretagna nella prima variante e con il nome di una località tedesca nella seconda variante. In questi casi, ovviamente, la prima variante è stata sottoposta esclusivamente a traduzione verso l'inglese, mentre la seconda esclusivamente a traduzione verso il tedesco. Considerando anche le frasi che presentano doppia variante, il numero totale delle frasi utilizzate per il test sale a 24.

Leggendo le frasi selezionate si noterà come la maggior parte di esse presenti una forma interrogativa. Si è infatti pensato che l'uso di frasi interrogative sarebbe stato più realistico, considerato l'ambito scelto per il test: si presuppone infatti che un turista che si trova all'estero, ma non parla la lingua locale, limiti le interazioni con i residenti allo stretto necessario, ad esempio alla richiesta di informazioni.

Un elenco completo delle frasi utilizzate per il test con il rispettivo output riconosciuto dal modulo ASR e output tradotto dal modulo MT, suddiviso per combinazione linguistica e per partecipante, è consultabile in Appendice 1.

3.5 Partecipanti

Per lo svolgimento del test sono stati selezionati due soggetti di età compresa fra i 21 e i 25 anni, di nazionalità e madrelingua italiana, ma di diversa provenienza geografica. Nello specifico, il partecipante 1 è originario di una regione dell'Italia centrale (Toscana), mentre il partecipante 2 di una regione dell'Italia settentrionale (Lombardia); entrambi i partecipanti presentano un accento tipico dell'area di provenienza.

I due partecipanti avevano il compito di leggere ad alta voce le frasi selezionate per la traduzione, pronunciandole nel microfono dello smartphone e fornendo così l'input al modulo di riconoscimento vocale dell'applicazione testata, come spiegato nella sezione 3.6.

Selezionando due partecipanti di due regioni diverse d'Italia si è voluto verificare se e in che misura fattori quali l'accento e la pronuncia diversa di particolari fonemi influissero sul riconoscimento vocale del modulo ASR. È infatti noto che l'italiano regionale settentrionale (ovvero l'italiano diffuso nelle regioni settentrionali) si differenzia dalla varietà toscana dell'italiano regionale centrale per diversi aspetti, inclusi quelli fonetici (Sobrero, Miglietta, 2010). Le differenze possono riguardare ad esempio la pronuncia della vocale “e” o quella delle affricate palatali [tʃ] e [dʒ], e la caratteristica presenza della gorgia nella variante toscana (*ibid.*).

La provenienza geografica dei due partecipanti è stata dunque tenuta in considerazione nella valutazione dell'output prodotto dal modulo ASR (si vedano i capitoli IV e V). Tale valutazione non ha come obiettivo quello di identificare la variante di italiano regionale più facilmente comprensibile nella fase di riconoscimento vocale, bensì quello di verificare se in un sistema di SST *speaker-independent* esistano o meno differenze nel riconoscimento del parlato di italiani regionali diversi.

3.6 Svolgimento del test

Il test è stato condotto secondo le medesime modalità dapprima sull'applicazione Google Translate, e in un secondo momento su Microsoft Translator.

Come visto nella sezione 3.4, per il test sono state selezionate in totale 24 frasi, per le quali si disponeva di una traduzione di riferimento contenuta nei frasari dai quali le frasi stesse sono state tratte. Si è valutato il riconoscimento vocale e la traduzione automatica per entrambe le applicazioni nelle seguenti combinazioni linguistiche: italiano-inglese e italiano-tedesco. Per il test in ciascuna di queste combinazioni linguistiche sono state utilizzate 21 frasi, dal momento che 3 frasi erano presenti in una versione doppia (come spiegato nella sezione 3.4).

Ai partecipanti è stato chiesto di leggere le frasi in modo chiaro e udibile, evitando se possibile false partenze ed esitazioni, come suggerito nella sezione 2.5.2 sulla cooperazione dell'utente, ma senza per questo suonare innaturali o modificare la propria pronuncia. Ai partecipanti è inoltre stato spiegato lo scopo del test, ovvero quello di valutare gli output prodotti dalle due applicazioni, immaginandone l'uso da parte di un turista italiano all'estero.

Ogni partecipante ha pronunciato tutte le 21 frasi selezionate per i test italiano-inglese e italiano-tedesco e per entrambe le applicazioni. Il test è stato dunque condotto in quattro diverse fasi per ogni partecipante (una fase per combinazione linguistica e per applicazione); fra una fase e l'altra è stata effettuata una pausa di qualche minuto. Le frasi da leggere sono state fornite ai partecipanti in forma scritta e sono state lette nello stesso ordine, con pause di qualche secondo fra una frase e l'altra in modo da consentire l'annotazione dei risultati del riconoscimento vocale e della traduzione automatica.

Il test è stato condotto in ambienti poco rumorosi o con isolamento acustico, in modo da evitare interferenze con il riconoscimento vocale. Per ogni frase è stata annotata dapprima la trascrizione prodotta dal modulo ASR così come appariva sulla schermata dell'applicazione (inclusi i segni di punteggiatura) e

poi la traduzione effettuata dal modulo MT, così come presentata sulla schermata dell'applicazione (inclusi i segni di punteggiatura). In entrambe le app infatti, nella funzione di traduzione automatica del parlato l'output tradotto viene fornito sia in forma scritta sul display del dispositivo, sia in forma orale grazie alla sintesi vocale del modulo TTS (si vedano 3.3.2 e 3.3.3). In questo test si è scelto di utilizzare le traduzioni presentate in forma scritta sul display come punto di partenza per la valutazione dell'output del modulo MT. In caso contrario sarebbe stato necessario effettuare la trascrizione dell'output prodotto dal modulo TTS. Tuttavia, si è considerato il fatto che la percezione dell'intonazione di frasi riprodotte dal modulo di sintesi vocale è soggettiva; trascrivendo manualmente queste frasi, il rischio sarebbe stato quello di alterarle in qualche modo, aggiungendo ad esempio un punto interrogativo nelle frasi originariamente interrogative, pur essendo questo assente dalla traduzione presentata in forma scritta sulla schermata dell'app.

Per quanto riguarda invece la possibilità di valutare la qualità della traduzione effettuata dal modulo MT sottoponendo ai valutatori l'output tradotto sintetizzato, sono state individuate difficoltà pratiche. Non potendo convocare i valutatori in sede di test, sarebbe infatti stato necessario registrare gli output sintetizzati e riprodotti durante il test per poi distribuirli ai valutatori stessi, con il rischio di generare file dalla qualità audio non ottimale che avrebbero ostacolato la valutazione della MT. Si ritiene inoltre che sia preferibile valutare una traduzione “nero su bianco”, ovvero in forma scritta.

Durante la fase di riconoscimento vocale sono stati annotati eventuali errori di riconoscimento,⁴⁵ nonché il numero di volte per cui è stato necessario ripetere la frase prima che avvenisse un corretto riconoscimento vocale. Infatti, quando il riconoscimento vocale non avveniva correttamente al primo tentativo, al partecipante veniva chiesto di ripetere la frase, cercando eventualmente di scandire meglio le parole non riconosciute. Se neanche al secondo tentativo la frase veniva riconosciuta correttamente dal modulo ASR, il partecipante passava a quella successiva. Si è supposto infatti che in

⁴⁵ Ogni qualvolta non avveniva un corretto riconoscimento vocale per una o più parole della frase, si registrava un errore di riconoscimento vocale.

uno scenario reale, qualora l'utente si accorga di un errore nel riconoscimento vocale (grazie alla trascrizione della frase riconosciuta mostrata sul display), ripeta la frase una seconda volta; è invece probabile che qualora questa non venga riconosciuta correttamente neppure al secondo tentativo, l'utente opti per una riformulazione.

Nel caso in cui durante il test, a fronte di un riconoscimento vocale andato a buon fine, fosse invece l'output tradotto a contenere degli errori, al partecipante non veniva chiesto di ripetere la frase. Occorre infatti considerare che in genere l'utente che usufruisce del servizio di SST delle applicazioni esaminate non conosce la lingua *target* e quindi difficilmente sarà in grado di individuare eventuali errori presenti nella traduzione.

3.7 Valutazione dei risultati

3.7.1 Premessa

Come già sottolineato, in questo test si è scelto di valutare esclusivamente gli output prodotti dai moduli ASR e MT, escludendo dunque il modulo TTS. Le performance di questi due moduli sono state valutate separatamente, in due fasi e con modalità distinte, che verranno descritte in seguito. La valutazione si è basata su:

- la trascrizione dell'output prodotto dal modulo ASR, così come appariva sul display del dispositivo; questa trascrizione viene fornita direttamente dalle applicazioni oggetto d'analisi, in modo da consentire all'utente di verificare l'accuratezza del riconoscimento vocale. L'output prodotto dal modulo ASR costituisce l'input per il modulo MT (come spiegato nella sezione 2.4) e, nel caso contenga errori, può portare a traduzioni errate o fuorvianti. La fase di riconoscimento vocale è dunque molto delicata e un test di un'applicazione di traduzione automatica del parlato non può prescindere dalla valutazione della qualità dell'output del modulo ASR;

- l'output prodotto dal modulo MT, presentato in forma scritta sul display del dispositivo. Nelle applicazioni analizzate la traduzione prodotta dal modulo MT viene sia pronunciata oralmente grazie alla sintesi vocale, sia mostrata sul display in forma scritta (si vedano 3.3.2 e 3.3.3). La valutazione si è basata sulle traduzioni scritte così come fornite dall'app (inclusi i segni di punteggiatura), dal momento che, come già spiegato, in questa sede non si intende valutare l'output del modulo TTS, e “if the evaluation of the TTS is not desired, the textual input to the TTS can be evaluated” (Dorr et al., 2011: 750).

3.7.2 Valutazione dell'output del modulo ASR

Per quanto riguarda la performance del modulo ASR, il test di valutazione ha diversi obiettivi: in primo luogo quello di verificare la correttezza e l'accuratezza del riconoscimento vocale di frasi tipiche della lingua parlata e frequentemente utilizzate in ambito turistico; in secondo luogo quello di verificare l'accuratezza del riconoscimento vocale di toponimi (in questo caso indicanti città non italiane); in terzo luogo quello di testare la capacità del modulo ASR di essere *speaker-independent* e di riconoscere le frasi pronunciate da parlanti con accenti differenti; e infine quello di verificare l'incidenza degli errori di riconoscimento vocale sulla traduzione automatica. Una volta annotati e raccolti tutti i dati relativi al riconoscimento vocale, ai fini della valutazione, per ogni partecipante, applicazione e combinazione linguistica di traduzione è stato contato il numero di frasi riconosciute correttamente al primo tentativo, il numero di quelle riconosciute correttamente al secondo tentativo e infine il numero di frasi non riconosciute correttamente neppure al secondo tentativo. Occorre sottolineare che, benché si sia scelto di non valutare gli output tradotti relativi al primo tentativo di riconoscimento vocale in tutti i quei casi in cui vi sono stati due tentativi (come si vedrà nella sezione 3.7.3.2), la distinzione fra frasi riconosciute al primo tentativo e frasi riconosciute al secondo tentativo resta rilevante nella valutazione del modulo ASR per

motivi legati all'usabilità dell'app stessa. Si è infatti considerato che, dal punto di vista dell'utente, un'app in cui il riconoscimento vocale avvenga correttamente al primo tentativo nella maggior parte dei casi sia preferibile rispetto a un'app in cui il riconoscimento vocale richiede più tentativi.

A partire dai dati relativi ai parametri sopra elencati, si è calcolato il numero di errori di riconoscimento vocale che hanno condotto a errori di traduzione nell'output del modulo MT. Considerando lo scopo del servizio di SST offerto dalle applicazioni oggetto di valutazione (ovvero quello di garantire la comunicazione), questo dato sembrava essere il più rilevante.

I risultati del test relativo al modulo ASR e la valutazione del suo output saranno presentati nella sezione 4.3 del capitolo IV.

3.7.3 Valutazione dell'output del modulo MT

Si è scelto di valutare l'output prodotto dal modulo MT (in forma scritta) secondo le modalità in uso per la valutazione della traduzione automatica scritta. Nakamura sostiene infatti che in un test in cui si sottopongono al sistema di SST singole frasi da tradurre e si esclude l'output del modulo TTS dalla valutazione, è possibile valutare la traduzione della lingua parlata attraverso gli stessi metodi impiegati per la valutazione della traduzione automatica scritta (Nakamura, 2009: 39).

If the evaluation of the speech synthesis module is not included, evaluation is made by feeding a number of test sentences into the system, and evaluating the quality of the output. In this sense, the method for evaluating speech translation is essentially the same as that for evaluating automatic text translation. For speech translation, however, the utterances that are evaluated are not strings of text but speech.

(Nakamura, 2009: 39)

Una volta deciso di valutare l'output scritto prodotto dal modulo MT secondo queste modalità, si aprivano diverse possibilità. Come spiegato nelle sezioni 1.7.2 e 2.8, esistono due metodi di valutazione principali nell'ambito della

traduzione automatica. I primi sono metodi soggettivi, ovvero basati sul giudizio umano, i secondi sono metodi oggettivi e automatici, generalmente fondati su algoritmi.

Pur nella consapevolezza dei numerosi vantaggi offerti dai metodi oggettivi, si è optato per una valutazione soggettiva per diversi motivi. Innanzitutto la quantità di dati da analizzare e valutare non era tale da rendere necessario il ricorso a metodi automatici. In genere, infatti, questi metodi vengono applicati alla valutazione di grandi moli di dati per ridurre costi e tempistiche, come accade nelle fasi di sviluppo di un nuovo sistema di MT o SST. In secondo luogo, metriche automatiche come BLEU non sono adatte alla valutazione di frasi brevi quali quelle scelte per il test condotto in questo studio (si veda sezione 1.7.2). Infine si è condiviso il punto di vista secondo il quale

[a]t present, the most reliable way to test a speech-to-speech MT system is subjective evaluation.

(Tirronen, 2011: 26)

Considerando infatti le caratteristiche delle frasi selezionate per il test, si presume che un valutatore umano bilingue, ovvero con conoscenza sia della lingua *source* che della lingua *target*, possa giudicare efficacemente la qualità della traduzione. Lo scopo di un sistema di SST utilizzato in ambito turistico è infatti quello di consentire una comunicazione basilare fra l'utente e il suo interlocutore. Mentre un parametro di valutazione oggettiva difficilmente sarà in grado di cogliere le sfumature nella traduzione, di tenere in considerazione lo scopo del sistema oggetto di valutazione, o distinguere fra errori gravi che impediscono la comprensione (come l'assenza della negazione in una frase originariamente negativa), e errori lievi (Dorr et al., 2011), un valutatore umano bilingue può utilizzare le proprie conoscenze linguistiche e del mondo per esprimere un giudizio che rifletta davvero la qualità della traduzione (Tirronen, 2011: 26). Infatti “human understanding of the real world allows a human judge to assess the practical importance of errors in a translation” (Dorr et al., 2011: 751).

Infine è importante sottolineare che nonostante la diffusione dei metodi per

la valutazione oggettiva, la valutazione soggettiva è stata e viene tuttora utilizzata in numerosi studi, dal momento che la nozione stessa di qualità in ambito traduttivo è in qualche modo “soggettiva”.

The general notion of quality of a translation is [...] somewhat subjective, and the most accepted measure, the Gold Standard, so to speak, is to have truly bilingual human judges compare the source language inputs to the target language outputs and then somehow provide an opinion about the quality of the translation.

(Dorr et al., 2011: 750)

3.7.3.1 Scelta dei valutatori

Per la valutazione dell'output prodotto dal modulo MT sono stati scelti in totale sei soggetti di madrelingua italiana con elevate competenze linguistiche nelle lingue *target* scelte per il test. In considerazione del carattere soggettivo delle valutazioni espresse da giudici umani, era infatti necessario interpellare più di un soggetto, in modo da potere così confrontare le risposte dei diversi valutatori, calcolarne la media, e ottenere un risultato “normalizzato” (Dorr et al., 2011: 755).

Dei sei valutatori selezionati, tre sono stati chiamati ad esprimere il proprio giudizio sull'output prodotto dal modulo MT nell'ambito dei test condotti per la combinazione linguistica italiano-inglese, mentre gli altri tre valutatori hanno fatto lo stesso per la combinazione linguistica italiano-tedesco. Dei sei valutatori scelti, tre sono traduttori (o laureandi in traduzione presso la ex SSLMIT di Forlì) e tre interpreti (o laureandi in interpretazione presso la stessa università). Tutti i valutatori lavorano con le lingue per le quali è stato chiesto loro di esprimere un giudizio nell'ambito del test (italiano e inglese o italiano e tedesco) e hanno avuto esperienze di traduzione o interpretazione sia in direzione passiva (dalla lingua straniera alla lingua madre), sia attiva (dalla lingua madre alla lingua straniera).

Tutti i valutatori sono di età compresa fra i 24 e i 25 anni e sono accomunati da un percorso formativo simile, che include o ha incluso studi universitari e

periodi più o meno lunghi di soggiorno all'estero. Nessun valutatore si distingue per avere maturato un numero di anni di esperienza come traduttore o interprete nettamente superiore rispetto agli altri valutatori. Si suppone dunque che le competenze linguistiche maturate dai valutatori nella lingua straniera possano considerarsi più o meno equivalenti.

In questo caso, per la valutazione della qualità dell'output della MT non si è optato né per “truly bilingual judges” (Dorr et al., 2011: 750), né per soggetti di madrelingua inglese e tedesca. Di fatto i valutatori scelti sono chiamati a valutare traduzioni verso una lingua diversa dalla loro lingua madre, compito che normalmente può risultare difficoltoso. Tuttavia, considerati fattori quali le difficoltà riscontrate nel reperire soggetti di madrelingua straniera o bilingue perfetti, la semplicità delle frasi selezionate per il test e le elevate competenze linguistiche dei valutatori scelti, il ricorso a soggetti di madrelingua italiana non è parso presentare evidenti svantaggi rispetto alla scelta di valutatori di madrelingua straniera o bilingue perfetti.

Un altro elemento a favore della scelta di questi sei valutatori è infine rappresentato dalla loro familiarità con il lessico e la fraseologia ricorrenti nelle frasi selezionate per il test, acquisita presumibilmente durante le esperienze prolungate di soggiorno all'estero. In una valutazione soggettiva è infatti importante che i valutatori conoscano l'argomento della traduzione da valutare; in caso contrario, un giudizio negativo potrebbe essere dovuto non tanto alla bassa qualità della traduzione, quanto piuttosto alla scarsa conoscenza che il giudice stesso ha dell'argomento (Dorr et al., 2011: 754-755).

3.7.3.2 Scelta dei parametri di valutazione e preparazione dei questionari di valutazione

Come si legge in numerosi studi, fra i criteri utilizzati più frequentemente per la valutazione soggettiva della traduzione automatica compaiono la *semantic adequacy* e la *fluency*, misurati in genere su una scala a punti o tramite la classifica (da migliore a peggiore) delle traduzioni oggetto di

valutazione. Altri criteri diffusi e misurati allo stesso modo sono la *understandability* (Dorr et al., 2011), la *clarity* e l'*informativeness* (Euromatrix, 2007). Esistono poi metodi di valutazione soggettiva basati sul tempo impiegato per il post-editing, su test di comprensione, sul tempo impiegato per la lettura della traduzione, ecc. (si veda. 1.7.2).

Per la valutazione delle traduzioni prodotte dalle app GT e MsT si è deciso di optare per una valutazione basata sui criteri di *adequacy* e *fluency*, misurati su una scala standard da 1 a 5 punti. Metodi alternativi, come la misurazione del tempo impiegato per il post-editing o per la lettura del testo tradotto, o test di comprensione non apparivano idonei al test condotto, in considerazione del fatto che l'input somministrato alle applicazioni non era un vero e proprio “testo”, bensì era costituito da singole frasi, molto brevi e slegate fra loro, per la cui valutazione sarebbe poco indicativo misurare il tempo di lettura o post-editing, così come sarebbe insensato utilizzare un test di comprensione visto l'argomento scelto.

I parametri di *adequacy* (ovvero equivalenza semantica o fedeltà) e di *fluency* (scorrevolezza del testo) sono parsi i più adatti alla valutazione. Si tratta infatti di parametri incentrati su aspetti fondamentali in una traduzione:

- l'equivalenza semantica è usata per misurare la fedeltà della traduzione rispetto all'originale a livello di contenuto e significato (Euromatrix, 2007). Viene tipicamente valutata a partire dal confronto fra la frase originale e l'output tradotto, oppure fra questo e altre traduzioni di riferimento e misurata su una scala di cinque punti, dove 5 corrisponde a equivalenza perfetta e 1 alla totale assenza di equivalenza;
- la scorrevolezza o *fluency* misura “the degree to which the translation is well-formed according to the grammar of the target language” (Euromatrix, 2007: 43) e viene tipicamente misurata su una scala da 5 a 1, dove 5 equivale a una formulazione perfetta e 1 equivale a una formulazione incomprensibile; la *fluency* di una traduzione non deve essere necessariamente valutata da soggetti bilingue, dal momento che non è necessario un confronto con l'originale (Dorr et al., 2011: 754).

L'idea alla base della distinzione fra questi due parametri è che una traduzione può contenere errori grammaticali e non essere scorrevole, ma allo stesso tempo riportare tutte le informazioni contenute nell'originale (Euromatrix, 2007). Necessariamente poi, una traduzione scorrevole e fedele risulta anche comprensibile; per questo motivo, se adottati insieme, i parametri equivalenza semantica e *fluency* sembrano coprire anche *understandability* o *clarity*.

Uno svantaggio dell'utilizzo di equivalenza semantica e *fluency* come criteri di valutazione è che talvolta i valutatori hanno difficoltà a distinguere questi due aspetti della traduzione (Euromatrix, 2007) e per entrambi i criteri attribuiscono alla traduzione punteggi simili. Per tentare di superare questo problema, si è deciso di dividere il questionario di valutazione somministrato ai sei valutatori in due sezioni distinte: la prima è dedicata esclusivamente alla valutazione della *fluency* degli output tradotti (in inglese e in tedesco), per i quali non viene mostrata la relativa frase originale; la seconda è invece dedicata alla valutazione dell'equivalenza semantica ed è basata su un confronto fra frase originale e traduzione. In questo modo, nella valutazione del parametro *fluency*, i valutatori non dovrebbero lasciarsi influenzare dal fattore “fedeltà all'originale”.

Per la stesura del questionario ci si è basati su esempi di scala di valutazione tratti da “Machine Translation Evaluation and Optimization” (Dorr et al., 2011) e parzialmente adattati. Per la valutazione della *fluency* o scorrevolezza è stata adottata una scala di 5 punti, dove 5 corrisponde a “inglese perfetto” o “tedesco perfetto” (a seconda del questionario), mentre 1 corrisponde a “inglese incomprensibile” e “tedesco incomprensibile”. I punti da 4 a 2 sono lasciati in bianco, per evitare di associarvi formulazioni ambigue o ritenute tali da confondere il valutatore (ad esempio, formulazioni quali “inglese/tedesco poco fluente” o “inglese/tedesco non madrelingua”).

Per la valutazione dell'equivalenza semantica è stata creata una scala da 5 a 1, dove 5 corrisponde a traduzione “perfettamente equivalente per significato e intento” all'originale e 1 corrisponde a traduzione “per nulla equivalente per significato e intento”. Anche in questo caso i punti compresi fra 4 e 2 sono stati lasciati in bianco.

Prima di ogni sezione del questionario sono state riportate istruzioni

dettagliate per la compilazione del questionario stesso ed è stata fornita una breve spiegazione del significato di scorrevolezza/*fluency* ed equivalenza semantica. Nel corpo del questionario sono state riportate le frasi in lingua *target* prodotte dai moduli MT delle due applicazioni, nonché il relativo input in lingua *source* (presente esclusivamente nella seconda parte del questionario), seguite dalle apposite scale di valutazione a scelta multipla. Per le frasi *source* che il modulo ASR non ha riconosciuto correttamente al primo tentativo e che i partecipanti hanno ripetuto sono stati riportati unicamente gli output tradotti relativi al secondo, nonché ultimo tentativo di riconoscimento vocale, in considerazione del fatto che in uno scenario reale qualora l'utente si accorga di un errore nel riconoscimento vocale, tenderà generalmente a ripetere la frase e si avvarrà della traduzione ottenuta all'ultimo tentativo effettuato. Sarebbe dunque insensato valutare l'output tradotto a partire da una frase ritenuta errata già dopo la fase di riconoscimento vocale e non corrispondente alla frase *source* desiderata. La traduzione ottenuta a partire dall'input ASR relativo al secondo tentativo di riconoscimento vocale, invece, è stata considerata valida indipendentemente dal fatto che il riconoscimento stesso fosse andato a buon fine, in considerazione del fatto che, come già sottolineato, in uno scenario reale, in questo caso l'utente opterebbe verosimilmente per una riformulazione della frase *source*.

Il numero totale di frasi valutate ammonta a 42 per il questionario italiano-inglese e a 44 per il questionario italiano-tedesco. Occorre infatti considerare che, benché le frasi utilizzate per il test fossero 21 per ogni combinazione linguistica, queste hanno dato luogo a molte più traduzioni. Sono infatti stati riscontrati numerosi casi in cui le due applicazioni hanno offerto traduzioni diverse per la stessa frase; in altri casi, per la stessa frase esistevano delle differenze persino fra l'output ottenuto dal partecipante 1 e quello ottenuto dal partecipante 2.

Nel questionario sono stati riportati, e quindi sottoposti a valutazione, tutti gli output che presentavano differenze, anche a livello di punteggiatura. Per quanto riguarda in particolare quest'ultimo aspetto, dopo avere notato la totale assenza di segni di interpunzione nell'output scritto prodotto dal modulo MT di GT, in contrasto con quello prodotto da MsT, si è deciso di

procedere come segue:

- nel caso in cui, per la stessa frase in lingua *source*, l'output tradotto da GT coincidesse con quello prodotto da MsT, ad esclusione del punto fermo finale (generalmente presente nella versione offerta da MsT, ma assente da quella di GT), nel questionario si è riportato l'output in questione senza alcun segno di punteggiatura finale. Si è infatti considerato il fatto che in genere una frase priva di punto finale viene percepita come affermativa/negativa analogamente a una frase che termina con il punto fermo;
- per uniformità, tutti gli output tradotti terminanti con un punto fermo (prodotti da MsT) sono stati riportati nel questionario senza alcun segno di punteggiatura finale;
- nel caso in cui, per la stessa frase interrogativa in lingua *source*, l'output tradotto da GT, privo di punteggiatura, coincidesse con quello prodotto da MsT a parte il punto interrogativo presente solo in quest'ultimo, nel questionario sono stati riportati entrambi gli output inclusa la rispettiva punteggiatura. Il punto interrogativo è infatti un segno di interpunzione la cui omissione viene in genere percepita maggiormente rispetto a quella del punto fermo. Ai valutatori è stato dunque chiesto di valutare entrambi gli output, benché la differenza risiedesse unicamente nella punteggiatura.

Le frasi sono state riportate nel questionario secondo un ordine casuale, per evitare che i valutatori si trovassero a valutare in sequenza, e di conseguenza a confrontare, diverse versioni della stessa frase in lingua *target* o diverse traduzioni della stessa frase *source*. I valutatori, inoltre, erano all'oscuro della provenienza delle frasi da valutare; si è deciso infatti di non specificare né il nome dell'applicazione da cui aveva origine ciascun output valutato, né il fatto che si trattava di frasi tradotte da applicazioni di traduzione automatica per smartphone, per evitare che i valutatori si lasciassero influenzare da eventuali pregiudizi o opinioni personali sull'una o l'altra applicazione, o in generale sulla traduzione automatica.

I questionari adottati per la valutazione della qualità dell'output tradotto sono riportati interamente in Appendice 2 (questionario italiano-inglese) e in

Appendice 3 (questionario italiano-tedesco), mentre esempi delle scale di valutazione utilizzate per *fluency* ed equivalenza semantica sono mostrati in Figura 6 e Figura 7.

<p>Scorrevolezza / Fluency</p> <p>Come giudichi la fluency della traduzione? E'</p> <ul style="list-style-type: none"><input type="radio"/> 5 tedesco perfetto<input type="radio"/> 4<input type="radio"/> 3<input type="radio"/> 2<input type="radio"/> 1 tedesco incomprensibile

Figura 6
Scala di valutazione della *fluency* adottata
nel questionario italiano-tedesco.

<p>Equivalenza semantica</p> <p>Rispetto all'originale la traduzione fornita è</p> <ul style="list-style-type: none"><input type="radio"/> 5 perfettamente equivalente per significato e intento<input type="radio"/> 4<input type="radio"/> 3<input type="radio"/> 2<input type="radio"/> 1 per nulla equivalente per significato e intento

Figura 7
Scala di valutazione dell'equivalenza semantica adottata
in entrambi i questionari.

3.8 Presentazione dei risultati

Nel prossimo capitolo verranno presentati i risultati del test di valutazione descritto nel presente capitolo, nonché le relative modalità di compilazione. In particolare il capitolo IV sarà incentrato sulla valutazione dell'output prodotto dal modulo ASR e sulla valutazione della qualità delle traduzioni prodotte dal modulo MT a partire dall'input fornito dal modulo ASR.

IV

Risultati del test di valutazione

4.1 Introduzione

In questo capitolo verranno presentati i risultati del test di valutazione condotto su due applicazioni di traduzione automatica del parlato, Google Translate (GT) e Microsoft Translator (MsT), secondo le modalità descritte nel capitolo III. Di seguito verranno descritte le modalità di compilazione dei risultati relativi al riconoscimento vocale in GT e MsT e verranno innanzitutto esposti i dati più significativi inerenti alle performance dei moduli ASR delle due applicazioni valutate. Dalla valutazione del modulo ASR si passerà a quella del modulo MT, per la quale verranno esposti i risultati raccolti tramite i questionari somministrati a sei valutatori. La parte dedicata alla MT verrà accompagnata anche da un'analisi dell'accordo fra i valutatori e, parallelamente, delle potenziali cause di disaccordo. Si vedrà infine come, in base a questa analisi, alcune risposte non sono state incluse nel calcolo della media dei punteggi attribuiti dai valutatori all'output tradotto da ciascuna applicazione.

4.2 Compilazione dei risultati relativi al riconoscimento vocale in GT e MsT

La valutazione dell'output del modulo ASR è avvenuta secondo le modalità descritte nella sezione 3.7.2.

Per effettuare la valutazione, per ciascun test effettuato e per ciascuna combinazione linguistica sono stati annotati i seguenti dati:

- numero di frasi riconosciute correttamente dal modulo ASR al primo tentativo;
- numero di frasi non riconosciute correttamente dal modulo ASR al primo tentativo, ovvero frasi in cui, a una prima lettura da parte del partecipante, non sono state riconosciute correttamente una o più parole;
- numero di frasi riconosciute correttamente dal modulo ASR al secondo tentativo;
- numero di frasi non riconosciute correttamente neppure al secondo e ultimo tentativo effettuato;
- numero di frasi contenenti errori di riconoscimento vocale (ovvero per le quali una o più parole non sono state riconosciute correttamente) che hanno condotto a errori di traduzione nella fase successiva di traduzione automatica.

In secondo luogo, per entrambe le app i dati sul riconoscimento vocale ottenuti per il test relativo alla combinazione linguistica italiano-inglese sono stati incrociati con quelli ottenuti per la combinazione italiano-tedesco, al fine di ottenere un quadro generale della performance di riconoscimento vocale di GT e MsT. A questo scopo, per il calcolo di eventuali errori di riconoscimento vocale, per ogni app e ciascun partecipante sono state adottate le seguenti convenzioni:

- alle frasi utilizzate sia nel test italiano-inglese che in quello italiano-tedesco e non riconosciute correttamente in entrambi i casi è stato attribuito il valore di 1. Ad esempio, una frase non riconosciuta correttamente al primo tentativo né nel test italiano-inglese, né in quello italiano-tedesco viene contata con il valore di 1 fra le frasi non riconosciute correttamente al primo tentativo;
- alle frasi utilizzate sia nel test italiano-inglese che in quello italiano-tedesco, riconosciute correttamente solo in uno di questi due casi, è stato attribuito il valore di 0,5. Ad esempio, una frase riconosciuta correttamente nel test italiano-inglese, ma non in quello italiano-tedesco viene contata fra le frasi non riconosciute correttamente al primo tentativo con il valore di 0,5, poiché in un caso su due il

riconoscimento vocale era andato a buon fine;

- analogamente, alle frasi utilizzate nei test in entrambe le combinazioni linguistiche e riconosciute correttamente al secondo tentativo solo in uno dei due casi è stato attribuito un punteggio di 0,5. Ad esempio, una frase riconosciuta correttamente al secondo tentativo nel test italiano-inglese, ma non in quello italiano-tedesco viene contata fra quelle riconosciute correttamente al secondo tentativo con il valore di 0,5, invece che 1;
- alle frasi contenenti toponimi specifici e presenti in due varianti (una per combinazione linguistica) è stato attribuito il valore di 1, dal momento che il calcolo è stato effettuato su un totale di 24 frasi (totale che include le frasi presenti in due varianti, come spiegato nella sezione 3.4).

Come si vedrà in seguito, nelle tabelle e nei grafici relativi alla performance dei moduli ASR di GT e MsT sono stati riportati per completezza tutti i dati sul riconoscimento vocale raccolti durante il test di valutazione. Tuttavia, come spiegato in 3.7.2, il dato più rilevante ai fini dello studio condotto è rappresentato da “Errori ASR che hanno condotto a errori in MT”, ovvero il numero di frasi non riconosciute correttamente che hanno influenzato negativamente la performance del modulo MT, provocando errori di traduzione. Come si potrà notare, talvolta questo valore è inaspettatamente inferiore rispetto a quello indicante il numero di frasi non riconosciute correttamente dal modulo ASR. Ciò è dovuto al fatto che alcuni errori presenti al secondo tentativo di riconoscimento vocale non sono stati reputati tali da provocare sostanziali modifiche nella traduzione e, nel caso questa contenesse errori che ne ostacolavano la comprensione, ciò non poteva essere imputato a un errato riconoscimento vocale.

Si vedano a titolo d’esempio i casi seguenti:

- per entrambi i partecipanti e entrambe le combinazioni linguistiche del test, la frase “Può chiamarmi un taxi?” è stata riconosciuta da MsT come “Puoi chiamarmi un taxi” sia al primo che al secondo tentativo di riconoscimento vocale (si veda Appendice 1). Questo errore può tuttavia essere considerato minore, dal momento che in inglese non

esiste la distinzione fra la seconda persona singolare “tu” e la forma di cortesia “Lei” (entrambe vengono tradotte con “you”), mentre in tedesco per richieste di questo tipo si tende ad usare la forma di cortesia (come confermato dalla traduzione effettuata dal modulo MT di MsT a partire dall'input riconosciuto dal modulo ASR; si veda Appendice 1). Si è dunque considerato che eventuali errori presenti nell'output tradotto a partire da questo input ASR non fossero imputabili all'errore di riconoscimento in esso contenuto;

- per un partecipante, la frase “Avete una cartina della città?” è stata riconosciuta da MsT come “Avete la cartina della città” (si veda Appendice 1) sia al primo che al secondo tentativo di riconoscimento vocale. In questo caso l'errore riguardante la scelta fra articolo indeterminativo e determinativo è stato classificato come minore e si è considerato che eventuali errori contenuti nell'output tradotto non fossero riconducibili all'input ASR.

Accanto all'analisi generale del riconoscimento vocale, è stato analizzato anche il riconoscimento vocale dei toponimi inseriti nelle frasi *source* utilizzate per il test, dal momento che rientrava fra gli obiettivi della valutazione del modulo ASR (si veda 3.7.2). Per effettuare tale analisi, per ogni applicazione testata, ogni combinazione linguistica e ogni partecipante sono stati contati i casi in cui il toponimo inserito nelle frasi è stato riconosciuto correttamente al primo tentativo, quelli in cui il toponimo inserito è stato riconosciuto al secondo tentativo, e infine i casi in cui il toponimo non è stato riconosciuto correttamente, portando necessariamente a errori di traduzione nella fase successiva di traduzione automatica.

I risultati della valutazione del modulo ASR verranno esposti nella prossima sezione, mentre per consultare un elenco completo degli output prodotti dai moduli ASR di GT e MsT, utilizzati poi come input per il modulo MT, si rimanda all'Appendice 1, dove è riportato l'elenco completo delle frasi utilizzate per il test con il rispettivo output ASR (relativo all'ultimo tentativo di riconoscimento effettuato).

4.3 Risultati del riconoscimento vocale in GT e MsT

Di seguito verranno esposti i risultati relativi al riconoscimento vocale effettuato da GT e MsT per entrambi i partecipanti e entrambe le combinazioni linguistiche in cui è stato condotto il test. In particolare, in Figura 8 e Figura 9 sono riportate le tabelle relative ai dati sul riconoscimento vocale raccolti rispettivamente durante i test italiano-inglese (Figura 8) e italiano-tedesco (Figura 9). I dati sono suddivisi per partecipante e applicazione.

I grafici in Figura 10 e Figura 11 rappresentano invece i dati generali relativi al riconoscimento vocale effettuato da entrambe le applicazioni rispettivamente per il partecipante 1 (Figura 10) e per il partecipante 2 (Figura 11), senza alcuna distinzione fra le combinazioni linguistiche.

Infine, la tabella in Figura 12 riporta i dati relativi al riconoscimento vocale dei toponimi da parte dei moduli ASR di GT e MsT. I dati sono suddivisi per applicazione, partecipante e combinazione linguistica.

	Google Translate		Microsoft Translator	
	Partecipante 1	Partecipante 2	Partecipante 1	Partecipante 2
N. frasi riconosciute correttamente al 1° tentativo/totale frasi	17/21	18/21	13/21	17/21
N. frasi non riconosciute correttamente al 1° tentativo/totale frasi	4/21	3/21	8/21	4/21
N. frasi riconosciute correttamente al 2° tentativo/n. frasi non riconosciute correttamente al 1° tentativo	3/4	2/3	4/8	1/4
N. frasi non riconosciute correttamente/totale frasi	1/21	1/21	4/21	3/21
Frase non riconosciute correttamente che hanno provocato errori di traduzione	1	1	3	2

Figura 8
Dati relativi al riconoscimento vocale per il test italiano-inglese.

	Google Translate		Microsoft Translator	
	Partecipante 1	Partecipante 2	Partecipante 1	Partecipante 2
N. frasi riconosciute correttamente al 1° tentativo/totale frasi	17/21	18/21	13/21	16/21
N. frasi non riconosciute correttamente al 1° tentativo/totale frasi	4/21	3/21	8/21	5/21
N. frasi riconosciute correttamente al 2° tentativo/n. frasi non riconosciute correttamente al 1° tentativo	3/4	3/3	5/8	1/5
N. frasi non riconosciute correttamente/totale frasi	1/21	0	3/21	4/21
Frase non riconosciute correttamente che hanno provocato errori di traduzione	1	0	2	2

Figura 9
Dati relativi al riconoscimento vocale per il test italiano-tedesco.

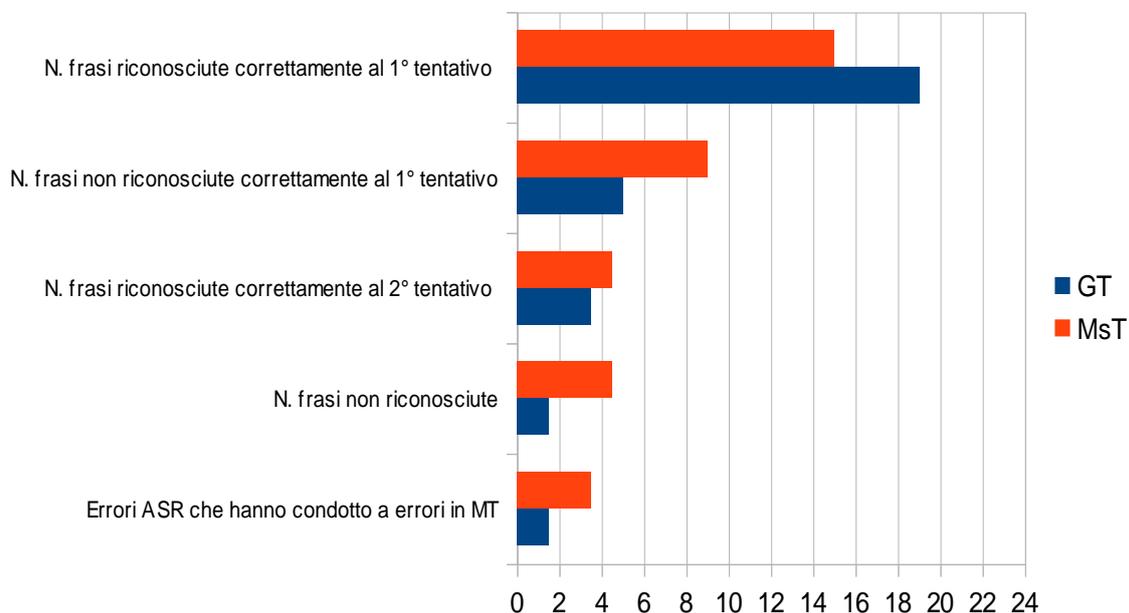


Figura 10
Risultati relativi al riconoscimento vocale per il partecipante 1.

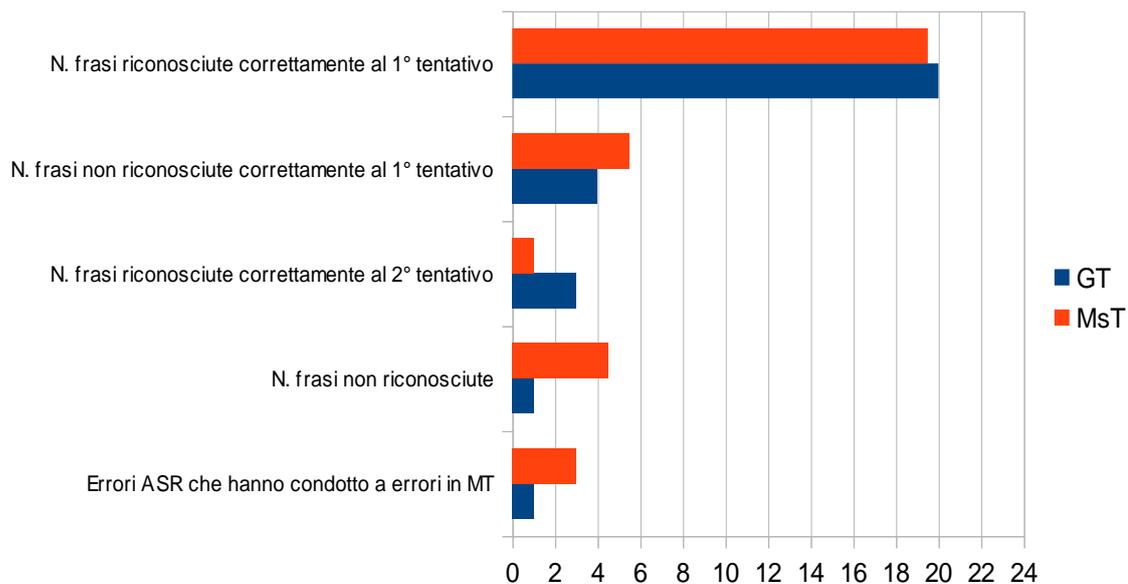


Figura 11
Risultati relativi al riconoscimento vocale per il partecipante 2.

	Google Translate				Microsoft Translator			
	Partecipante 1		Partecipante 2		Partecipante 1		Partecipante 2	
	Test it-en	Test it-de	Test it-en	Test it-de	Test it-en	Test it-de	Test it-en	Test it-de
N. toponimi riconosciuti correttamente al 1° tentativo/ totale toponimi	2/3	2/3	2/3	2/3	2/3	2/3	2/3	2/3
N. toponimi non riconosciuti correttamente al 1° tentativo/ totale toponimi	1/3	1/3	1/3	1/3	1/3	1/3	1/3	1/3
N. toponimi riconosciuti correttamente al 2° tentativo/ totale toponimi	1/3	0/3	0/3	1/3	0/3	0/3	0/3	0/3
N. toponimi non riconosciuti	0	1	1	0	1	1	1	1
N. toponimi riconosciuti/ totale toponimi	3/3	2/3	2/3	3/3	2/3	2/3	2/3	2/3

Figura 12
Risultati relativi al riconoscimento vocale dei toponimi presenti nelle frasi *source*.

In Figura 8 e Figura 9 i risultati del riconoscimento vocale sono calcolati su un totale di 21 frasi, poiché per ciascun test in ogni combinazione linguistica sono state utilizzate 21 frasi *source*, come spiegato nella sezione 3.6. In Figura 10 e Figura 11 invece il totale è rappresentato da 24 frasi e include tutte le frasi utilizzate nell'ambito del test di valutazione.

Analizzando le tabelle e i grafici riportati nelle figure 8,9,10 e 11 emerge che generalmente l'accuratezza del riconoscimento vocale durante il test è stata abbastanza alta per entrambi i partecipanti ed entrambe le app; parallelamente l'incidenza degli errori di riconoscimento vocale sulla traduzione automatica è relativamente bassa, come mostrato nell'ultima riga dei grafici in Figura 10 e Figura 11 (indicante il parametro “Errori ASR che hanno condotto a errori in MT”).

Confrontando le performance delle due app nei test dei due partecipanti emergono tuttavia delle differenze. Le figure 8 e 9 evidenziano infatti come GT abbia registrato risultati migliori per entrambe le combinazioni linguistiche e per entrambi i partecipanti sia per il parametro “N. frasi riconosciute correttamente al 1° tentativo”, sia per “N. frasi riconosciute correttamente al 2° tentativo”, sia infine per “N. frasi non riconosciute correttamente”. Anche il numero di frasi non riconosciute correttamente dal modulo ASR che hanno poi condotto a errori nella traduzione del modulo MT è inferiore per GT.

Questi dati sono confermati dalle figure 10 e 11, rappresentative del quadro generale della performance dei moduli ASR di GT e MST nell'ambito del test condotto. Se per GT la percentuale di frasi riconosciute al primo tentativo si attesta al 79,2% per il partecipante 1 e all' 83,3% per il partecipante 2, per MsT questi valori si attestano rispettivamente al 62,5% e all'81,25%. Parallelamente, per MsT il valore indicante il numero di frasi non riconosciute al primo tentativo è più alto che per GT. GT ha poi riconosciuto al secondo tentativo il 14,6% di frasi sul totale per il partecipante 1 e il 12,5% per il partecipante 2. MsT, invece, al secondo tentativo ha riconosciuto correttamente il 18,75% del totale delle frasi per il partecipante 1 e il 4,2% per il partecipante 2. Questo fa sì che il totale di frasi riconosciute correttamente dal modulo ASR di GT (inclusi dunque il primo e il secondo tentativo di riconoscimento vocale) ammonti al 93,75% per il

partecipante 1 e al 95,8% per il partecipante 2. Per MsT questo parametro si attesta invece all'81,25% per entrambi i partecipanti.

Per quanto riguarda infine il parametro “Errori ASR che hanno condotto a errori in MT”, equivalente al numero di frasi non riconosciute correttamente che hanno provocato errori di traduzione, emergono nuovamente differenze sia fra GT e MsT, sia fra i due partecipanti. Analizzando i grafici riportati in Figura 10 e Figura 11, si può notare che solo in 2,5 casi errori del modulo ASR di GT hanno portato a errori di traduzione (rispettivamente 1,5 casi su 24 per il partecipante 1 e un solo caso su 24 per il partecipante 2). Questo dato si attesta invece a 6,5 per MsT, con 3,5 casi su 24 per il partecipante 1 e 3 casi per il partecipante 2. Calcolando la percentuale di accuratezza del riconoscimento vocale sulla base del numero di frasi non riconosciute correttamente che hanno condotto a errori di traduzione, per GT si ottiene un valore del 93,75% per il partecipante 1 e del 95,8% per il partecipante 2, mentre per MsT gli stessi valori si attestano all'85,4% per il partecipante 1 e all'87,5% per il partecipante 2, con una differenza rispetto a GT di 8,35 punti percentuali nel primo caso e di 8,3 nel secondo caso. Questi dati dovranno essere tenuti in considerazione nella valutazione della qualità degli output prodotti dal modulo MT, alla luce del fatto che alcuni errori di traduzione risulteranno derivanti da errori già presenti nell'output del modulo ASR.

Per quanto riguarda il riconoscimento vocale dei toponimi, la Figura 12 mostra come per ogni partecipante e ogni combinazione linguistica l'inserimento di toponimi nelle frasi *source* sia stato occasionalmente fonte di problemi nella fase di riconoscimento vocale, dando luogo a errori (si vedano le figure 8, 9, 10 e 11). In tutti i casi registrati, un toponimo su tre non è stato riconosciuto correttamente al primo tentativo; talvolta il riconoscimento è avvenuto al secondo tentativo, mentre in altri casi non è mai avvenuto. Analizzando in particolare il parametro “N. toponimi riconosciuti/ totale toponimi” in Figura 12, si nota che in due casi i risultati ottenuti da GT sono migliori rispetto a quelli di MsT; in generale tuttavia, le performance di GT e MsT nel riconoscimento dei toponimi sembrano essere abbastanza equivalenti. I risultati relativi al riconoscimento vocale dei toponimi sono inclusi nei dati generali esposti in precedenza e, come tali, verranno tenuti in considerazione nella fase di valutazione della qualità degli

output prodotti dal modulo MT.

I dati finora esposti verranno discussi più approfonditamente nel capitolo V, dove verranno confrontate e valutate le prestazioni dei moduli ASR di GT e MsT, verranno analizzate e discusse le differenze riscontrabili fra il partecipante 1 e il partecipante 2 in termini di accuratezza del riconoscimento vocale e infine verrà commentato il riconoscimento vocale dei toponimi inseriti nelle frasi utilizzate per il test di valutazione.

4.4 Valutazione dell'output tradotto da GT e MsT

La valutazione dell'output prodotto dai moduli MT di Google Translate e Microsoft Translator è avvenuta secondo le modalità descritte nella sezione 3.7.3. Tale valutazione si è basata su questionari somministrati a un totale di sei valutatori (tre per combinazione linguistica), chiamati a esprimere il proprio giudizio su ogni output tradotto sulla base dei parametri *fluency* ed equivalenza semantica.

Per ogni questionario compilato sono stati annotati tutti i giudizi espressi dai valutatori. Per ogni esempio valutato sono poi stati confrontati i giudizi raccolti ed è stato calcolato il livello di accordo presente fra i valutatori, sulla scorta di precedenti esperienze descritte in “Machine Translation Evaluation and Optimization” (Dorr et al., 2011). Questo ha rappresentato un passaggio fondamentale: si è deciso infatti di “accantonare” le risposte per le quali i valutatori si sono mostrati in disaccordo, in quanto ritenute meno attendibili e rappresentative rispetto a quelle caratterizzate da accordo. Per le risposte escluse dalla valutazione si è poi cercato di capire cosa avesse indotto i valutatori a esprimere giudizi discordanti (si veda 4.5.2). A partire dalle risposte caratterizzate da accordo, invece, per ogni esempio valutato è stata calcolata la media dei punteggi attribuiti dai valutatori per i parametri *fluency* ed equivalenza semantica. A partire da questi valori è poi stato calcolato il punteggio medio ottenuto dagli output di ogni applicazione per i due parametri valutati. Questa operazione è stata ripetuta per entrambe le combinazioni linguistiche, come si vedrà in seguito.

Le prossime sezioni saranno dedicate all'accordo fra i valutatori, all'analisi delle potenziali cause di disaccordo e infine all'esposizione dei risultati derivanti dalle valutazioni prese in considerazione.

4.5 Accordo fra i valutatori

4.5.1 Analisi dell'accordo fra i valutatori

Come spiegato, tutti gli output tradotti dai moduli MT delle applicazioni oggetto di analisi sono stati valutati da tre diversi soggetti, sia in merito alla *fluency* sia in merito all'equivalenza semantica. Questo ha permesso di analizzare l'accordo fra i valutatori sia per il test italiano-inglese sia per quello italiano-tedesco.

Per ogni combinazione linguistica sono state contate le risposte identiche e quelle diverse. Di queste, tuttavia, solo alcune sono state classificate come indicanti disaccordo. In una scala di valutazione da 1 a 5, infatti, è possibile calcolare il livello di accordo fra i valutatori includendo, oltre alle risposte identiche, anche le risposte che presentano un solo punto di differenza (Dorr et al., 2011: 755). Alla luce di tale considerazione, si è deciso di classificare come indicanti accordo sia le risposte identiche, sia quelle che si differenziano per un massimo di un punto, mentre le risposte che si differenziano per due o più punti sono state classificate come indicanti disaccordo fra i valutatori.

Nel questionario relativo al test italiano-inglese i soggetti sono stati chiamati a valutare 42 output tradotti, come spiegato in 3.7.3.2, fornendo ciascuno 42 giudizi sulla *fluency* degli output e 42 sull'equivalenza semantica rispetto alle frasi originali.

Nel questionario italiano-inglese, per quanto riguarda la *fluency*, 10 risposte su 42 sono identiche per tutti e tre i valutatori e 32 sono diverse (almeno un valutatore ha dato una risposta diversa dagli altri). Tuttavia, di queste 32 risposte diverse, 25 si differenziano per un solo punto e indicano dunque un alto livello di accordo. La valutazione della *fluency* degli output in inglese è

dunque caratterizzata da un accordo fra i valutatori pari all'83,3% (si veda Figura 13). Inoltre è importante sottolineare che nei casi di disaccordo le risposte dei valutatori si differenziano in genere per due punti e molto raramente per tre punti. Questo significa che nessun valutatore ha giudicato perfetta dal punto di vista della *fluency* una frase classificata da un altro valutatore come incomprensibile.

Per il parametro equivalenza semantica 12 risposte su 42 sono identiche per tutti e tre i valutatori, mentre 30 sono diverse. Di queste, 24 si differenziano per un punto e indicano un livello alto di accordo. Per questo parametro l'accordo fra i valutatori è dell'85,7% (si veda Figura 13). Anche in questo caso le risposte che evidenziano disaccordo si differenziano generalmente per due punti e raramente per tre.

Nel questionario relativo al test italiano-tedesco i soggetti sono stati chiamati a valutare 44 output tradotti, come spiegato in 3.7.3.2, e ciascuno ha dunque fornito 44 giudizi sulla *fluency* degli output e 44 sull'equivalenza semantica rispetto alle frasi originali.

In questo questionario, per il parametro *fluency*, in 10 casi su 44 i valutatori hanno fornito risposte identiche, in 34 casi risposte diverse. Di queste, 22 risposte si differenziano per un solo punto e possono quindi considerarsi caratterizzate da un alto livello di accordo. Per questo parametro la percentuale di accordo fra i valutatori ammonta dunque al 72,7% (si veda Figura 13).

Per il parametro equivalenza semantica i casi di risposta identica sono 7 su 44; in 37 casi le risposte sono diverse. Tuttavia, in 19 casi queste risposte si differenziano per un solo punto e possono essere incluse nel calcolo delle risposte indicanti accordo. La percentuale di accordo ammonta dunque a 59,1% (si veda Figura 13). Anche per il questionario italiano-tedesco nei casi di disaccordo le risposte dei valutatori presentano in genere due punti di differenza e raramente tre.

Come mostrato in Figura 13, in generale l'accordo fra i valutatori è parso essere piuttosto alto, pur variando da parametro a parametro e da un questionario all'altro. Benché per la combinazione linguistica italiano-inglese l'accordo fra i valutatori si attesti a livelli più alti rispetto alla combinazione italiano-tedesco, per tutti i parametri valutati in entrambe le

combinazioni linguistiche l'accordo è comunque superiore al 50%. Il livello di accordo raggiunto è dunque stato considerato sufficientemente alto da consentire l'utilizzo esclusivo delle risposte identiche o con una differenza massima di un punto ai fini del calcolo del punteggio medio attribuito agli output prodotti dai moduli MT di GT e MsT. In altre parole, il numero di risposte indicanti accordo era tale da consentire l'esclusione delle risposte indicanti disaccordo dal calcolo dei punteggi medi attribuiti a ogni app, senza pregiudicare l'attendibilità dei risultati.

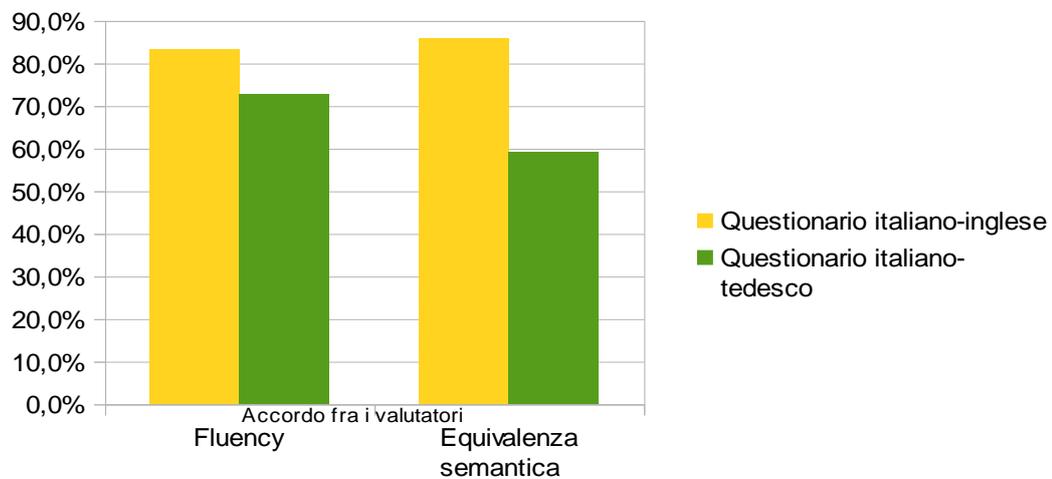


Figura 13
Dati percentuali sull'accordo fra i valutatori.

4.5.2 Discussione delle potenziali cause di disaccordo fra i valutatori e trattamento delle risposte indicanti disaccordo

Come sottolineato nelle sezioni precedenti, le risposte caratterizzate da una differenza superiore a un punto non sono state incluse nel calcolo del punteggio medio attribuito da ogni soggetto agli output di GT e MsT. Per la coppia di lingue italiano-inglese questo ha portato a escludere dalla valutazione finale il 16,7% delle risposte raccolte per il parametro *fluency* e il 14,3% per il parametro equivalenza semantica. Per la coppia di lingue italiano-tedesco è invece stato escluso dalla valutazione finale il 27,3% delle risposte per il parametro *fluency* e il 40,9% per il parametro

equivalenza semantica.

Pur considerando, come già sottolineato, che in tutti i casi il livello di accordo supera il 50% e che le risposte indicanti disaccordo presentano spesso solo due punti di differenza e in pochi casi tre, vale la pena cercare di individuare le cause di disaccordo fra i valutatori e le motivazioni per cui la percentuale di disaccordo non è stabile nelle due direzioni linguistiche.

In generale, sembra che i soggetti coinvolti nella valutazione abbiano attribuito un peso diverso a errori di traduzione derivanti dal mancato riconoscimento dei toponimi inseriti nelle frasi *source*. Questo ha inciso soprattutto sulla valutazione della *fluency*, poiché se alcuni valutatori hanno giudicato la sostituzione del toponimo con un termine di uso comune come un errore che rendeva l'intera frase incomprensibile, altri hanno tenuto in maggiore considerazione la formulazione complessiva della frase, dando meno peso al singolo termine sbagliato (in questo caso il toponimo). Nel valutare la *fluency* sembra inoltre che alcuni soggetti abbiano attribuito punteggi più bassi rispetto ad altri alle frasi che, pur essendo corrette dal punto di vista grammaticale, risultavano poco idiomatiche o forse di difficile interpretazione senza un contesto più ampio. In questo caso è dunque possibile che valutando la *fluency* alcuni valutatori abbiano giudicato esclusivamente la correttezza grammaticale degli esempi da valutare, mentre altri si siano soffermati anche sull'aspetto dell'idiomaticità o abbiano tenuto in considerazione anche il senso della frase.

Per quanto riguarda il parametro equivalenza semantica, è possibile che alcuni valutatori abbiano tenuto in maggiore considerazione la *fluency* della frase rispetto ad altri, attribuendo punteggi più bassi a output che, pur rispecchiando il senso della frase *source*, non erano corretti grammaticalmente. Nel caso invece in cui l'output contenesse tutte le informazioni della frase originale, senza rispecchiarne però l'intento (come per frasi *source* interrogative tradotte in forma affermativa), i valutatori sembrano avere dato diverso peso al fattore intento. Analizzando le risposte raccolte tramite i questionari emerge ad esempio che alcuni valutatori hanno giudicato la trasformazione di una frase interrogativa in affermativa come un errore grave dal punto di vista dell'equivalenza semantica; invece altri, ove la traduzione (affermativa) conteneva tutte le informazioni presenti nella

frase originale (interrogativa), hanno dato minore peso alla discordanza di intento fra input e output. Infine, anche l'aspetto della punteggiatura ha generato valutazioni talvolta discordanti, poiché alcuni valutatori hanno attribuito punteggi più bassi rispetto ad altri alle frasi interrogative che non terminavano con l'adeguato segno di interpunzione.

In generale, occorre sottolineare che ogni individuo ha una diversa tolleranza nei confronti degli errori e soggetti diversi possono percepire lo stesso errore in modi diversi. È dunque inevitabile che tali differenze emergano all'interno di una valutazione soggettiva che vede coinvolti più partecipanti. Proprio per questo motivo nel presente test di valutazione, al fine di ottenere risultati “normalizzati”, si è deciso di calcolare la media dei punteggi attribuiti a ogni app valutata e, a tale scopo, di considerare unicamente le risposte indicanti accordo.

4.6 Risultati della valutazione degli output tradotti

I dati che verranno esposti di seguito derivano da calcoli effettuati esclusivamente sulla base delle risposte caratterizzate da accordo, estratte da quelle raccolte tramite i questionari di valutazione somministrati. Nella tabella in Figura 14 sono riportate le medie dei punteggi ottenuti da GT e MsT nella valutazione di *fluency* ed equivalenza semantica per le coppie di lingue italiano-inglese e italiano-tedesco. Per un elenco completo delle valutazioni attribuite da ciascun valutatore agli output tradotti si rimanda invece all'Appendice 4.

Media valutazioni/ parametro	Google Translate		Microsoft Translator	
	Italiano-inglese	Italiano-tedesco	Italiano-inglese	Italiano-tedesco
Media valutazioni/ <i>fluency</i>	3,18	3,33	3,76	3,81
Media valutazioni/ equivalenza semantica	2,9	2,69	3,23	3,03

Figura 14
Risultati finali delle valutazioni degli output tradotti da GT e MsT.

Come emerge dai dati riportati in Figura 14, entrambe le app valutate hanno registrato prestazioni paragonabili in termini di qualità dell'output tradotto. Tuttavia, si può osservare come, per tutti i parametri valutati ed entrambe le coppie di lingue del test, MsT abbia ottenuto un punteggio medio leggermente superiore a quello di GT. Se infatti per la combinazione italiano-inglese GT registra punteggi medi di 3,18 per la *fluency* e 2,9 per l'equivalenza semantica, MsT registra un punteggio di 3,76 per la *fluency* e di 3,23 per l'equivalenza semantica. Per la coppia di lingue italiano-tedesco ancora una volta MsT ottiene risultati leggermente migliori, con un punteggio medio di 3,81 per il parametro *fluency* e di 3,03 per l'equivalenza semantica, a fronte dei 3,33 e 2,69 punti ottenuti da GT per gli stessi parametri.

Confrontando invece i dati relativi alle coppie di lingue di traduzione, si nota come ancora una volta non sussistano differenze cospicue. Tuttavia, per entrambe le applicazioni l'output tedesco ha ottenuto punteggi leggermente superiori all'output inglese in termini di *fluency*. Per il parametro equivalenza semantica la situazione è invece opposta: è l'output inglese a ottenere un punteggio superiore, seppur di poco, a quello tedesco.

Per una discussione dei risultati della valutazione dell'output MT finora presentati si rimanda alla sezione 5.3 del capitolo V, che sarà interamente dedicato all'analisi delle performance registrate dai moduli ASR e MT di GT e MsT nell'ambito del test condotto.

V

Discussione dei risultati

5.1 Introduzione

In questo capitolo verranno discussi i risultati presentati nel capitolo IV.

La prima parte del capitolo sarà incentrata sul riconoscimento vocale, i cui dati, esposti nella sezione 4.3, verranno discussi alla luce degli obiettivi della valutazione. L'analisi verterà su un confronto fra GT e MsT in termini di accuratezza del riconoscimento vocale, sulla questione della punteggiatura negli output trascritti, sul riconoscimento dei toponimi e infine su un confronto fra i dati raccolti per il partecipante 1 e il partecipante 2.

La seconda parte del capitolo sarà invece dedicata alla discussione dei dati derivanti dalla valutazione dell'output prodotto dai moduli MT, per il quale verranno confrontati i risultati ottenuti da GT e MsT e verrà discussa l'incidenza del riconoscimento vocale sulla traduzione. Infine verranno confrontate le valutazioni ottenute dall'output inglese e da quello tedesco e verranno descritte le principali problematiche e tipologie di errore riscontrate nelle varie traduzioni prodotte da GT e MsT.

5.2 Discussione dei risultati del riconoscimento vocale

5.2.1 Introduzione

Di seguito verranno discussi i risultati derivanti dalla valutazione della performance dei moduli ASR di GT e MsT. L'analisi dei risultati si concentrerà su diversi aspetti, sulla base degli obiettivi esposti nella sezione

3.7.2. In primo luogo, verranno discusse e confrontate le performance di GT e MsT in termini di accuratezza del riconoscimento vocale e di incidenza sulla performance dei moduli MT; successivamente verranno analizzati il riconoscimento vocale dei toponimi presenti nelle frasi *source* e la capacità delle due applicazioni di essere *speaker-independent*, ovvero di riconoscere frasi pronunciate da utenti con accenti e modi di parlare differenti.

5.2.2 Confronto fra le performance dei moduli ASR di GT e MsT

Fra gli obiettivi della valutazione del modulo ASR elencati in 3.7.2 compariva quello di verificare l'accuratezza del riconoscimento vocale delle frasi utilizzate per il test, tratte dall'ambito turistico, per poi analizzare in un secondo momento l'incidenza di eventuali errori di riconoscimento vocale sulla traduzione automatica.

Analizzando i risultati esposti nella sezione 4.3 si evince che in generale, nel test condotto, l'accuratezza del riconoscimento vocale è stata alta sia per GT che per MsT: in tutti i casi registrati la percentuale di accuratezza del riconoscimento vocale supera l'80%, toccando valori superiori al 90% per GT. Parallelamente, sono limitati i casi in cui il riconoscimento vocale non è andato a buon fine, così come i casi in cui gli errori avvenuti in fase di riconoscimento vocale hanno provocato poi errori di traduzione.

Dal punto di vista di un potenziale utente, dunque, entrambe le performance possono essere ritenute soddisfacenti in termini di riconoscimento vocale, benché questa considerazione si limiti all'ambito del test condotto e dipenda dalle caratteristiche della frase input pronunciata (come si noterà nella sezione 5.2.4 relativa al riconoscimento dei toponimi) e benché ci sia margine di miglioramento. Tuttavia, come già sottolineato in 4.3, nella fase di compilazione dei risultati sono state rilevate alcune differenze fra le performance di GT e MsT sia in termini di accuratezza del riconoscimento vocale, sia in termini di usabilità.

Nell'ambito del test condotto, in generale l'output prodotto dal modulo ASR di GT si è dimostrato più accurato rispetto a quello prodotto dal

corrispondente modulo di MsT GT ha riconosciuto un maggior numero di frasi al primo tentativo, caratteristica che attribuisce a questa applicazione una maggiore usabilità in termini di riconoscimento vocale rispetto a MsT. Inoltre anche al secondo tentativo di riconoscimento vocale GT ha registrato risultati migliori. In diversi casi, infatti, per GT il riconoscimento fallito al primo tentativo è andato a buon fine al secondo tentativo. Lo stesso non può essere detto per MsT, per il quale in diverse occasioni anche il secondo tentativo di riconoscimento vocale si è rivelato problematico. Nell'ottica di un utente che decida di utilizzare un'app per la traduzione automatica del parlato come GT o MsT per comunicare con un interlocutore straniero, la necessità di ripetere più volte la frase da tradurre prima che questa venga riconosciuta correttamente rappresenta un elemento di disturbo non trascurabile, che rischia di ostacolare lo svolgimento della conversazione stessa o di indurre l'utente a interrompere l'utilizzo dell'app.

Queste considerazioni iniziali su GT e MsT sono confermate dalle percentuali inerenti al numero di frasi non riconosciute correttamente neppure al secondo tentativo. Come già evidenziato, analizzando i valori associati a questo parametro per entrambe le applicazioni, appare evidente che nel test condotto la performance di GT risulta superiore a quella di MsT in termini di accuratezza del riconoscimento vocale.

Infine occorre considerare i dati relativi al numero di frasi non riconosciute correttamente dal modulo ASR che hanno poi condotto a errori di traduzione. Benché, come spiegato nella sezione 4.2, questo valore risulti talvolta inferiore rispetto a quello del parametro indicante le frasi non riconosciute correttamente e non dipenda obbligatoriamente da esso, anche per questo parametro GT ha ottenuto risultati migliori rispetto a MsT. Questo non comporta necessariamente che l'output tradotto da GT risulterà qualitativamente superiore a quello tradotto da MsT; significa invece che in un minor numero di casi eventuali errori contenuti nell'output tradotto da GT saranno imputabili a errori di riconoscimento vocale.

Alla luce di queste considerazioni, si può concludere che il modulo ASR di MsT non ha prodotto output più accurati rispetto a quelli di GT, peraltro più veloce nel riconoscimento vocale. Al contrario, GT si è dimostrato non solo più accurato nel riconoscimento vocale per tutti i partecipanti e i test svolti,

ma anche caratterizzato da un maggiore grado di usabilità per l'utente, caratteristica legata sia alla velocità istantanea con cui avviene il riconoscimento vocale in GT, tale per cui non esistono di fatto tempi morti, sia alla percentuale di accuratezza riscontrata al primo tentativo di riconoscimento.

Occorre naturalmente considerare che i risultati esposti e le considerazioni che ne derivano hanno validità limitata al test condotto in questo studio e non implicano necessariamente che in ogni condizione il modulo ASR di GT superi le prestazioni di quello di MsT.

5.2.3 La punteggiatura nell'output ASR prodotto da GT e MsT

Nel confronto fra le performance dei moduli ASR di GT e MsT, una menzione particolare va alla questione della punteggiatura.

Come spiegato nella sezione 3.6, durante lo svolgimento del test sono stati annotati gli output prodotti dai moduli ASR delle app testate così come apparivano sul display del dispositivo, inclusi i segni di punteggiatura. A questo proposito, sono emerse alcune differenze fra l'output ASR trascritto prodotto da GT e quello prodotto da MsT, poi rispecchiate anche dalle differenze di punteggiatura rilevate nell'output tradotto dai moduli MT delle due applicazioni (si veda 3.7.3.2). Infatti, mentre l'output ASR prodotto da GT si presenta sempre privo di punteggiatura finale nella sua forma scritta, l'output ASR prodotto da MsT termina con chiari segni di interpunzione. Pur essendo presenti, questi non sempre corrispondono al reale intento della frase *source* pronunciata: sono stati registrati diversi casi in cui frasi originariamente interrogative terminano con il punto fermo nella versione scritta dell'output ASR di MsT.

Confrontando gli output ASR prodotti da entrambe le applicazioni con i rispettivi output tradotti, si è notato che l'assenza della punteggiatura negli output ASR di GT non ha necessariamente pregiudicato l'accuratezza e la correttezza della traduzione prodotta dal rispettivo modulo MT. Sono infatti stati registrati sia casi in cui frasi originariamente interrogative sono state

tradotte correttamente dal punto di vista dell'intento (nonostante l'assenza di punto interrogativo nell'output ASR, ovvero nell'input usato dal modulo MT), sia casi in cui frasi originariamente interrogative sono state tradotte erroneamente in forma affermativa (si veda Appendice 1). In genere le frasi interrogative che denotano chiaramente un intento interrogativo (grazie a pronomi e avverbi interrogativi, ecc.), pur essendo prive di segno di interpunzione finale, hanno dato origine a output tradotti interrogativi; le frasi interrogative prive di elementi quali pronomi e avverbi interrogativi hanno invece dato spesso origine a output tradotti affermativi (ovvero con la sintassi tipica della forma affermativa). Questo, tuttavia, non significa necessariamente che in casi come questo l'errore di traduzione (ovvero l'uso della forma affermativa al posto di quella interrogativa) derivi dall'assenza di punteggiatura nell'output ASR: effettuando una prova, sono stati registrati casi in cui aggiungendo manualmente alla frase un punto interrogativo in fase di riconoscimento vocale, e dunque a fronte di un output ASR con la punteggiatura corretta, la traduzione fornita da modulo MT di GT non ha subito alcuna modifica.

Per quanto riguarda invece MsT, sembra che il segno di interpunzione finale dell'output ASR non abbia influenzato sistematicamente l'output tradotto, dal momento che, se da un lato gli output ASR che terminano con un punto interrogativo hanno generato output tradotti interrogativi, dall'altro gli output ASR relativi a frasi *source* interrogative, che terminano però con un punto fermo, hanno generato talvolta output tradotti con sintassi tipicamente interrogativa (nonostante il punto fermo finale dell'output ASR) e in altri casi output tradotti con sintassi affermativa (in accordo con il segno di punteggiatura dell'output ASR).

Alla luce di queste considerazioni, non sembra che l'assenza o la presenza della punteggiatura nell'output dei moduli ASR abbia condizionato in modo assoluto la qualità della traduzione prodotta dai moduli MT a partire dall'input ASR, dal momento che nel test condotto gli output ASR privi di punteggiatura non hanno sistematicamente dato origine a traduzioni errate; viceversa gli output ASR con punteggiatura inclusa non hanno sistematicamente dato origine a traduzioni corrette.

5.2.4 Il riconoscimento vocale dei toponimi

Per il modulo ASR il secondo obiettivo del test di valutazione era quello di verificare l'accuratezza del riconoscimento vocale dei toponimi inseriti nelle frasi *source*. Si è infatti considerato che il corretto riconoscimento di eventuali toponimi rappresentasse un elemento di fondamentale importanza nell'ambito turistico scelto per il test. Infatti, come sottolineato in precedenza, spesso le interazioni turistiche sono caratterizzate da richieste di informazioni o indicazioni stradali e possono includere toponimi, i quali devono essere riconosciuti correttamente perché l'interazione possa procedere con successo.

Per testare il comportamento dei moduli ASR di GT e MsT in presenza di toponimi, alcune delle frasi tratte dai frasari turistici selezionati sono state modificate parzialmente tramite l'inserimento di toponimi specifici, indicanti città straniere, come spiegato in 3.4. In totale le frasi *source* contenenti toponimi erano sei, di cui tre utilizzate per il test svolto nella combinazione linguistica italiano-inglese e tre per il test italiano-tedesco.

Già durante lo svolgimento del test si è potuto notare come il riconoscimento di alcuni di questi toponimi risultasse particolarmente difficoltoso per GT e MsT, indipendentemente dal partecipante che li pronunciava. Infatti, se per quattro dei sei toponimi scelti (ovvero Londra, Edimburgo, Berlino e Norimberga) il riconoscimento è avvenuto al primo tentativo per entrambi i partecipanti ed entrambe le applicazioni, per Plymouth e Dresda tale processo è stato molto più problematico (si veda Appendice 1). In nessun caso questi due toponimi sono stati riconosciuti al primo tentativo. Al secondo tentativo il toponimo Dresda è stato riconosciuto correttamente in un solo caso, precisamente dal modulo ASR di GT nel test del partecipante 2; lo stesso vale per Plymouth, riconosciuto correttamente al secondo tentativo in un solo caso, precisamente dal modulo ASR di GT nel test del partecipante 1. Il riconoscimento vocale di questi due toponimi si è rivelato problematico (e in alcuni casi impossibile) sia per GT che per MsT, che hanno mostrato performance quasi del tutto equivalenti; tuttavia, in due casi GT ha registrato una performance migliore rispetto a MsT, portando a termine il

riconoscimento vocale al secondo tentativo, che è invece nuovamente fallito in MsT (si veda 4.3).

La difficoltà riscontrata nel riconoscere due dei sei toponimi selezionati può essere dovuta da un lato alla pronuncia del toponimo stesso. In generale questo può essere un fattore decisivo soprattutto nei casi in cui il toponimo indichi una città straniera che non ha un equivalente tradotto nella lingua *source*, per la quale un utente senza conoscenza della lingua straniera avrà molto probabilmente difficoltà di pronuncia. Nell'ambito del test condotto questo è forse il caso di Plymouth, città del Regno Unito priva di traduzione in italiano. Dall'altro lato, il fatto che città straniere ormai note come Londra, Edimburgo, Berlino e Norimberga siano state riconosciute immediatamente dai moduli ASR di GT e MsT fa ipotizzare che durante il riconoscimento vocale in entrambe le app entrino in gioco meccanismi statistici, tali per cui parole usate poco frequentemente o che compaiono raramente nel corpus di addestramento (nel caso in cui i moduli ASR siano basati su corpora) vengono riconosciute meno accuratamente rispetto a parole comuni o che compaiono frequentemente nel corpus di addestramento. Nel test effettuato i toponimi Dresda e Plymouth sono stati ricondotti dai moduli ASR di GT e MsT a parole di uso comune, come ad esempio “brutta” o “30” per Dresda e “primo” o “prima” per Plymouth (si veda Appendice 1). Si tratta di termini sicuramente frequenti nella lingua italiana e verosimilmente anche in un eventuale corpus di addestramento rappresentativo della lingua parlata italiana, dove con ogni probabilità i termini “primo”, “prima”, “brutta” e “30” avrebbero più occorrenze rispetto ai toponimi sopra indicati. Alla luce di queste considerazioni si può supporre che gli errori di riconoscimento vocale dei toponimi menzionati siano riconducibili a un approccio statistico dei moduli ASR. Tuttavia, dal momento che non sono stati reperiti dati precisi sul funzionamento del riconoscimento vocale in GT e MsT, non è possibile affermare con certezza che sia basato su un approccio statistico.

Si suppone infine che il riconoscimento vocale di toponimi da parte di GT e MsT migliori avvicinandosi alla località indicata dai toponimi stessi, poiché è probabile che queste app facciano uso della tecnologia di geolocalizzazione, accedendo così ai dati relativi alla posizione in cui si

trova lo smartphone usato, sul quale è installata l'app stessa. Se questo fosse vero, l'app avrebbe accesso anche ai dati relativi alle località situate nelle vicinanze della posizione registrata e il riconoscimento vocale dei toponimi potrebbe trarre vantaggio da questa possibilità. Tuttavia, occorre considerare che il test di valutazione di questo studio è stato condotto in Italia, quindi a centinaia di chilometri dalle città indicate dai toponimi scelti e necessariamente il servizio di geolocalizzazione non può aver contribuito in alcun modo al riconoscimento vocale di alcuno dei toponimi del test. Anche in questo caso non è stato possibile reperire informazioni precise sul funzionamento dei moduli ASR di GT e MsT a conferma di questa ipotesi, che rimane quindi speculativa.

Per verificare l'accuratezza del riconoscimento vocale delle frasi contenenti toponimi e consultarne la trascrizione si rimanda all'Appendice 1, dove è riportato l'elenco completo delle frasi utilizzate per il test con il rispettivo output ASR relativo all'ultimo tentativo di riconoscimento vocale effettuato.

5.2.5 Il riconoscimento vocale per il partecipante 1 e il partecipante 2

Uno degli obiettivi della valutazione dei moduli ASR di GT e MsT era quello di testare la loro capacità di essere *speaker-independent*, ovvero di riconoscere correttamente le frasi pronunciate da utenti diversi, in modi diversi (si veda 3.7.2). Per questo motivo, per il test sono stati selezionati due partecipanti provenienti da regioni d'Italia diverse, con accento, cadenza e modo di parlare differenti, come spiegato nella sezione 3.5.

Analizzando i dati esposti nella sezione 4.3 si nota come, pur senza addestramento iniziale, le due applicazioni si siano mostrate in grado di riconoscere abbastanza accuratamente le frasi pronunciate da entrambi i partecipanti. Ciononostante, confrontando i risultati ottenuti dai due partecipanti in sede di riconoscimento vocale emergono alcune differenze, più o meno marcate.

La prima differenza riguarda la percentuale di frasi riconosciute al primo

tentativo, che in tutti i casi registrati risulta superiore per il partecipante 2. Il primo tentativo di riconoscimento vocale ha infatti fornito risultati più accurati per il partecipante 2, mentre il partecipante 1 si è trovato a dover ripetere le frasi *source* in diverse occasioni. Questo divario fra partecipante 1 e 2 nei risultati relativi al primo tentativo di riconoscimento vocale risulta particolarmente evidente per MsT, mentre la differenza è minima per GT.

Con il secondo tentativo di riconoscimento vocale le differenze emerse fra il partecipante 1 e il partecipante 2 si sono appianate, conducendo, per il parametro indicante il numero di frasi non riconosciute correttamente neppure al secondo e ultimo tentativo di riconoscimento, a valori comparabili (per GT) o addirittura identici (per MsT). L'accuratezza del riconoscimento vocale si attesta infatti a percentuali perfettamente equivalenti o quasi equivalenti per entrambi i partecipanti (si veda 4.3).

Per quanto riguarda infine il numero di frasi non riconosciute correttamente che hanno causato errori di traduzione, i valori ottenuti dai due partecipanti presentano una leggera differenza a favore del partecipante 2 per entrambe le applicazioni (si veda 4.3).

I dati raccolti fanno emergere che, se a un primo tentativo di riconoscimento vocale la quantità di frasi problematiche è stata superiore per il partecipante 1, dopo il secondo e ultimo tentativo di riconoscimento si è giunti a un'equivalenza quasi perfetta fra il numero di frasi riconosciute per il partecipante 1 e per il partecipante 2. Considerando dunque la quantità di frasi riconosciute immediatamente e di casi in cui è stata necessaria una ripetizione della frase, le applicazioni valutate hanno mostrato un grado di usabilità leggermente superiore (GT) o nettamente superiore (MsT) per il partecipante 2. Tuttavia, in termini di accuratezza del riconoscimento vocale nessuna delle due app ha registrato una performance considerevolmente superiore per uno dei due partecipanti.

Per quanto riguarda invece l'incidenza del riconoscimento vocale su eventuali errori di traduzione, i risultati mostrano come questa sia leggermente superiore per il partecipante 1, ma ancora una volta non vi è una cospicua differenza fra i due partecipanti. La somiglianza dei valori ottenuti dai due partecipanti per questo parametro e per quello indicante l'accuratezza del riconoscimento vocale deriva presumibilmente dal fatto che, in molti casi

le frasi non riconosciute correttamente, che in alcuni hanno provocato anche errori di traduzione, sono più o meno le stesse per il partecipante 1 e il partecipante 2 e includono le frasi contenenti toponimi non riconosciuti (si veda Appendice 1).

Si può dunque concludere che nel test svolto in questo studio entrambe le applicazioni si sono mostrate effettivamente *speaker-independent*, producendo output ASR di qualità equivalente o paragonabile per il partecipante 1 e il partecipante 2. La diversa provenienza geografica dei due partecipanti non sembra dunque avere avuto un ruolo determinante nell'accuratezza finale del riconoscimento vocale. Tuttavia, se si considera l'usabilità dei moduli ASR di GT e MsT nell'ambito del test condotto, il partecipante 2 è stato favorito e ha riscontrato meno problemi in fase di riconoscimento vocale (come indicato dal parametro relativo al numero di frasi riconosciute correttamente al primo tentativo nella sezione 4.3). Questo indica che nel test condotto il modo di parlare del partecipante 2 è risultato di più immediata comprensione per i moduli ASR di GT e MsT, mentre fattori quali il tono di voce più alto usato durante il test o la pronuncia di alcuni fonemi tipica della varietà toscana dell'italiano regionale centrale potrebbero avere penalizzato il partecipante 1 al primo tentativo di riconoscimento vocale. È difficile tuttavia determinare con certezza quali siano stati gli elementi alla base delle difficoltà riscontrate dal partecipante 1 durante il riconoscimento vocale.

Occorre tenere presente, infine, che i risultati discussi finora hanno validità limitata al test condotto e non indicano in alcun modo una maggiore facilità di riconoscimento vocale per l'uno o l'altro italiano regionale in termini assoluti. Individui diversi possono infatti avere pronunce e modi di parlare diversi, pur provenendo dalla stessa area geografica e pur condividendo la varietà regionale di italiano; inoltre, in circostanze diverse lo stesso individuo può pronunciare la stessa frase in modi più o meno differenti.

5.3 Discussione dei risultati relativi alla valutazione degli output tradotti

5.3.1 Introduzione

Per quanto riguarda la valutazione della traduzione automatica del parlato, l'obiettivo individuato nella sezione 3.2 per il test condotto era quello di verificare se e in quanti casi gli output prodotti dai moduli MT di GT e MsT risultassero utilizzabili e comprensibili, valutando la qualità della traduzione a partire dall'output scritto presentato sul display del dispositivo. Immaginando un turista italiano all'estero quale utente ipotetico del servizio di Speech-to-Speech Translation offerto dalle applicazioni valutate, si voleva dunque verificare in quale misura GT e MsT si sarebbero mostrate in grado di “mediare” la comunicazione con un ipotetico interlocutore straniero. Per fare questo si è chiesto ad alcuni valutatori di esprimere un giudizio sulla *fluency* e l'equivalenza semantica delle traduzioni fornite dalle due applicazioni, ritenendo che la combinazione di questi parametri rispecchiasse gli obiettivi della valutazione.

Dal momento che il test di valutazione condotto in questo studio ha visto coinvolte due applicazioni concorrenti per la traduzione automatica, la discussione dei risultati relativi alla valutazione degli output tradotti non potrà prescindere da un confronto fra la performance di GT e quella di MsT effettuato alla luce degli obiettivi esposti. I risultati raccolti tramite i questionari somministrati verranno analizzati sotto diversi aspetti, che includeranno *fluency*, equivalenza semantica, un confronto fra i risultati relativi all'output tedesco e quelli relativi all'output inglese, e per ultimo un'analisi dell'incidenza del riconoscimento vocale sugli errori di traduzione rilevati. Infine verranno analizzate sommariamente le principali categorie di errori riscontrati nelle traduzioni prodotte dalle due applicazioni.

L'analisi degli output tradotti da GT e MsT esposta nelle sezioni seguenti si fonda da un lato sui dati esposti nella sezione 4.6, i quali forniscono un quadro della valutazione delle performance dei moduli MT di GT e MST; dall'altro, si farà riferimento in più punti a output specifici (o parte di essi),

riportati nell'elenco completo degli output prodotti da entrambe le applicazioni (si veda Appendice 1).

5.3.2 *Fluency* ed equivalenza semantica negli output tradotti da GT e MsT

La discussione dei risultati relativi alla valutazione di *fluency* ed equivalenza semantica negli output tradotti da GT e MsT si basa interamente sui dati esposti nella sezione 4.6, dove sono riportati i punteggi medi ottenuti dalle due applicazioni per i due parametri valutati. Per una panoramica più completa dei giudizi attribuiti da ciascun valutatore a ogni output valutato si rimanda invece all'Appendice 4.

Analizzando i dati esposti nella sezione 4.6, si nota che entrambe le applicazioni hanno registrato risultati paragonabili in termini di qualità dell'output tradotto: per ciascun parametro e ciascuna combinazione linguistica, la differenza nei punteggi ottenuti da GT e MsT si riduce a qualche numero decimale. I punteggi medi ottenuti da GT e MsT per i parametri *fluency* ed equivalenza semantica non superano in nessun caso il 4, collocandosi così al centro della scala di valutazione fornita ai valutatori (si veda 3.7.3.2). Tali valori derivano dal calcolo della media dei punteggi attribuiti da ciascun valutatore agli output tradotti da GT e MsT in inglese e tedesco. Si riferiscono dunque a una prestazione “media” dei moduli MT delle due applicazioni: per alcuni output giudicati perfetti o buoni dal punto di vista della *fluency* o dell'equivalenza semantica, ce ne sono stati altri giudicati incomprensibili o per nulla equivalenti.

Benché le prestazioni di GT e MsT si siano mostrate paragonabili in termini di qualità dell'output tradotto, vale la pena effettuare un confronto sulla base dei diversi parametri valutati. In generale, si è notato come per tutti i parametri e per entrambe le combinazioni linguistiche MsT abbia ottenuto risultati migliori, seppur di poco, rispetto a GT, anche a fronte di un riconoscimento vocale contenente più errori (si veda 5.3.3). Gli output tradotti da MsT sono stati giudicati sia più corretti dal punto di vista della

fluency (dunque più corretti dal punto di vista grammaticale e più idiomatici), sia più fedeli al significato delle frasi *source*. Pur non riscontrando una proporzionalità diretta fra i punteggi relativi alla *fluency* e quelli relativi all'equivalenza semantica, questo risultato sembra indicare l'esistenza di un legame fra i due parametri di valutazione utilizzati. Si può infatti supporre che, risultando più scorrevoli e presumibilmente più comprensibili, agli occhi dei valutatori gli output prodotti da MsT rispecchiassero più fedelmente il significato delle frasi originali. Viceversa è probabile che una *fluency* inferiore negli output di GT in alcuni casi abbia in qualche modo ostacolato la comprensione del significato della frase, conducendo a punteggi inferiori anche per il parametro equivalenza semantica. Quindi, benché i questionari somministrati siano stati divisi in due sezioni distinte, dedicate rispettivamente alla valutazione della *fluency* e dell'equivalenza semantica, nel tentativo di aiutare i valutatori a non confondere questi due parametri, si è riscontrato un certo livello di correlazione fra essi.

Da un confronto fra GT e MsT si potrebbe concludere che a livello generale gli output tradotti dalle due applicazioni valutate risultano in media parzialmente corretti dal punto di vista della *fluency* (con valori compresi fra un massimo di 3,81 e un minimo di 3,18) e parzialmente equivalenti all'originale dal punto di vista semantico (con valori compresi fra 3,23 e 2,69), con livelli di *fluency* ed equivalenza semantica leggermente superiori per MsT; in pratica, tuttavia, occorrerebbe distinguere caso per caso, dal momento che per entrambe le applicazioni sono stati riscontrati sia output classificati come perfetti, sia output giudicati incomprensibili e/o per nulla equivalenti alle frasi *source*.

Partendo dal presupposto che *fluency* ed equivalenza semantica rappresentino criteri fondamentali nel determinare la qualità di una traduzione e, conseguentemente, la sua capacità di soddisfare il bisogno di comunicazione fra interlocutori multilingue, si può concludere che nell'ambito del test condotto GT e MsT si sono mostrati solo parzialmente in grado di soddisfare questo bisogno. Ciò significa che in uno scenario reale simile a quello pensato per il test di valutazione del presente studio, solo in alcuni casi GT e MsT riuscirebbero a “mediare” con efficacia la comunicazione fra un turista

italiano e un interlocutore straniero, mentre in altri la comunicazione fra i due interlocutori risulterebbe problematica o fallirebbe. In particolare, sulla base dei risultati raccolti, solo in una percentuale limitata di casi la qualità della traduzione garantirebbe la piena comprensione fra gli interlocutori (tale percentuale è superiore per MsT). Nella maggior parte dei casi la comunicazione risulterebbe parzialmente compromessa da alcuni errori contenuti nella traduzione, che potrebbero ostacolare la comprensione da parte dell'interlocutore straniero. Tuttavia nel caso specifico di problematiche relative alla *fluency* della traduzione, considerando uno scenario in cui un turista italiano utilizza il servizio di SST offerto da GT e MsT per interagire con un interlocutore madrelingua (in questo caso di madrelingua inglese o tedesca), è legittimo ipotizzare che la comprensione dell'output tradotto avverrebbe correttamente nonostante la presenza di alcuni errori grammaticali, a fronte della capacità di analisi e disambiguazione e della piena padronanza della lingua tipiche di un madrelingua. Infine, solo in un numero limitato di casi la comprensione fra gli interlocutori sarebbe gravemente compromessa, risultando impossibile o conducendo a malintesi.

5.3.3 Incidenza del riconoscimento vocale sulla qualità dell'output tradotto

Fra gli output tradotti giudicati più problematici dal punto di vista della *fluency* e/o dell'equivalenza semantica compaiono senza dubbio le frasi per le quali il riconoscimento vocale non è andato a buon fine. Fra queste si distinguono in particolar modo le richieste di indicazioni contenenti toponimi non riconosciuti correttamente dal modulo ASR, che rappresentano il caso principale in cui un errore di riconoscimento vocale ha provocato un errore di traduzione. Per tali frasi, il modulo ASR si è reso responsabile di errori presenti nell'output tradotto, influenzando negativamente il punteggio che i valutatori hanno attribuito alla traduzione.

Ad esempio sono stati riscontrati casi in cui, non riconoscendo correttamente

il toponimo presente nella frase *source*, il modulo ASR lo ha sostituito con termini di uso comune, che sono poi stati tradotti coerentemente dal modulo MT. Questo ha prodotto output giudicati sia difficilmente comprensibili o incomprensibili dal punto di vista della *fluency*, sia non equivalenti all'originale dal punto di vista semantico. Questo caso è più frequente per MsT, dal momento che GT si è dimostrato leggermente più accurato nel riconoscimento vocale dei toponimi inseriti nelle frasi *source* e conseguentemente anche nella traduzione di tali toponimi. Questo non implica però che le traduzioni delle frasi contenenti toponimi siano necessariamente migliori in GT: se da un lato è vero che GT è riuscito in più casi a riconoscere e tradurre correttamente i toponimi inseriti nelle frasi *source*, dall'altro occorre considerare che talvolta gli output in questione contengono altre tipologie di errore, che non derivavano da errori di riconoscimento vocale, ma hanno comunque abbassato la qualità della traduzione.

A questo proposito, è interessante notare come MsT, pur avendo registrato performance peggiori rispetto a GT in termini di riconoscimento vocale generale e di riconoscimento vocale dei toponimi, abbia comunque ottenuto giudizi leggermente superiori a GT nella valutazione della qualità degli output tradotti per entrambi i parametri misurati. Occorre tenere presente che, sia per GT che per MsT, alcuni degli output contenenti toponimi non riconosciuti o errori di traduzione derivanti da errori di riconoscimento vocale sono stati esclusi dal calcolo della media dei risultati, in quanto caratterizzati da valutazioni indicanti disaccordo per il parametro *fluency* e/o per il parametro equivalenza semantica (si veda Appendice 4). Tuttavia, anche includendo tutte le risposte dei valutatori nel calcolo dei punteggi medi ottenuti dalle due app per i parametri *fluency* ed equivalenza semantica, il risultato resta invariato: l'output tradotto da MsT ottiene punteggi medi leggermente superiori a quello tradotto da GT.

In conclusione, nonostante un riconoscimento vocale meno accurato e dunque una maggiore incidenza degli errori di riconoscimento vocale sugli errori compiuti dal modulo MT, MsT sembra avere prodotto traduzioni qualitativamente superiori rispetto a quelle prodotte da GT, seppur di poco. Dal canto suo, GT sembra più affidabile nella traduzione di toponimi, dal

momento che il loro riconoscimento vocale pare essere più accurato in questa applicazione.

5.3.4 *Fluency* ed equivalenza semantica nell'output inglese e nell'output tedesco

Analizzando i dati esposti nella sezione 4.6 per le coppie di lingue italiano-inglese e italiano-tedesco, emergono leggere differenze nei punteggi medi ottenuti dall'output inglese e dall'output tedesco per i parametri valutati. In particolare, come già sottolineato in 4.6, per entrambe le applicazioni gli output in tedesco risultano leggermente più scorrevoli rispetto a quelli in inglese. La situazione si inverte invece per il parametro equivalenza semantica: sembra infatti che nonostante la maggiore *fluency*, se confrontati con quelli in inglese, gli output in tedesco rispecchino meno fedelmente significato e intento delle frasi *source*, seppur di poco.

Considerando che in un sistema statistico di traduzione automatica (come quello alla base di GT e MsT) generalmente la qualità della traduzione è tanto maggiore quanto maggiori sono le dimensioni del corpus o dei corpora utilizzati per l'addestramento del sistema stesso, ci si poteva aspettare che l'output inglese risultasse qualitativamente superiore all'output tedesco, in considerazione del fatto che generalmente i corpora in inglese hanno dimensioni vastissime, di solito maggiori rispetto a quelle dei corpora utilizzati per l'addestramento nelle altre lingue.

Questa aspettativa è confermata dai risultati relativi all'equivalenza semantica, che indicano che le traduzioni verso l'inglese risultano più fedeli alle frasi originali, ma non dai punteggi medi relativi al parametro *fluency*. Tuttavia occorre ricordare che per il calcolo di tali punteggi è stata effettuata una selezione delle risposte fornite dai valutatori, tale per cui, come sottolineato più volte, alcune risposte sono state escluse in quanto caratterizzate da disaccordo. Per la combinazione italiano-tedesco il numero di risposte escluse è stato superiore rispetto alla combinazione italiano-inglese e questo potrebbe avere in qualche modo influito sul calcolo della

media dei punteggi ottenuti dai diversi output per *fluency* ed equivalenza semantica. A questo proposito, si è notato, ad esempio, che includendo tutte le valutazioni sulla *fluency* nel calcolo della media ottenuta da ogni app per ogni combinazione linguistica, i valori ottenuti dall'output inglese e da quello tedesco si equivalgono (attestandosi intorno a un punteggio medio di circa 3,2 per GT e di circa 3,6 per MsT). Al di là di questo, è importante sottolineare che la valutazione degli output in inglese e in tedesco è fondata su giudizi soggettivi forniti da valutatori diversi. I soggetti chiamati a valutare gli output in tedesco non sono gli stessi chiamati a valutare l'inglese: i valutatori di una combinazione linguistica potrebbero avere dato diverso peso a eventuali errori riscontrati nelle traduzioni rispetto a quelli dell'altra combinazione, oppure potrebbero avere una sensibilità diversa nei confronti degli errori.

Un altro aspetto interessante che emerge analizzando gli output tradotti in tedesco e inglese da GT è la presenza di un termine inglese (“excuse”) all'interno di un output tedesco (“Excuse spricht Italienisch”, traduzione della frase source “Scusi, parla italiano?”; si veda Appendice 1). Una possibile spiegazione di questa “anomalia” è stata individuata in “Everything you ever wanted to know about Google Translate, and finally got the chance to ask”, un articolo pubblicato sul portale TAUS, dove viene illustrato il funzionamento di Google Translate. Secondo l'autore dell'articolo, l'ampia gamma di combinazioni linguistiche di traduzione offerte da Google Translate è resa possibile dal fatto che per molte di queste combinazioni linguistiche l'inglese viene usato come “lingua ponte” nel processo traduttivo. Ciò significa ad esempio che per effettuare una traduzione dal finlandese allo zulu Google Translate tradurrà dapprima il testo *source* dal finlandese all'inglese e in un secondo momento tradurrà l'inglese in zulu (Greene, 2016). Questa strategia ha permesso a Google Translate di ampliare la sua rosa di lingue e combinazioni linguistiche di traduzione, includendo anche combinazioni per le quali non esistono ancora corpora paralleli allineati di dimensioni sufficienti all'addestramento. Tuttavia questo stratagemma aumenta le probabilità che insorgano errori nella traduzione (*ibid.*). L'articolo si riferisce al servizio di traduzione automatica offerto dalla piattaforma online Google Translate, ma appare probabile che lo stesso

meccanismo entri in atto nell'applicazione.

Detto ciò, la coppia linguistica italiano-tedesco dovrebbe disporre di un numero sufficiente di testi tradotti allineati da consentire la traduzione diretta dall'italiano al tedesco di parole di uso comune come “Scusi”. Tuttavia è possibile che nell'applicazione GT, ideata per la traduzione di testi di dimensioni ridotte o di conversazioni, la traduzione italiano-tedesco si basi almeno in parte sull'inglese come “lingua ponte”, forse a causa di un numero insufficiente di testi allineati relativi alla lingua parlata (si ricorda che originariamente Google Translate è nato come servizio per la traduzione di testi scritti). Questo spiegherebbe sia la presenza di un termine inglese all'interno dell'output tedesco tradotto da GT, sia la somiglianza riscontrata fra la struttura sintattica di alcuni output in inglese e delle corrispondenti versioni in tedesco, sia infine la specularità di alcuni errori di traduzione presenti in entrambi gli output. Per quanto riguarda questi ultimi aspetti, un esempio che sembra confermare questa teoria è rappresentato dalle frasi “What time served breakfast” e “Zu welcher Zeit das Frühstück serviert”, traduzioni della frase *source* “A che ora servite la colazione?” (si veda Appendice 1). Si può notare come in entrambi gli output tradotti il verbo servire sia coniugato al participio passato (rispettivamente “served” e “serviert”); questo errore può derivare dal fatto che l'italiano “servite” indica sia la seconda persona plurale dell'indicativo presente del verbo servire, sia il participio passato riferito a un soggetto femminile plurale. In inglese il verbo è stato tradotto erroneamente proprio come participio passato e lo stesso errore si ripresenta in tedesco. L'ipotesi è dunque che l'errore in tedesco sia dovuto a un errore già presente nell'output inglese che ha fatto da “ponte” nella traduzione fra italiano e tedesco.

Le stesse considerazioni sembrano valere per il funzionamento di MsT: anche per questa app alcuni degli output in tedesco sembrano rispecchiare la sintassi dei corrispondenti output in inglese, proprio come per GT. Ciò fa supporre che anche per MsT vi sia l'utilizzo dell'inglese come “lingua ponte”, ma anche in questo caso non sono state reperite informazioni certe a conferma di questa ipotesi.

Se le considerazioni finora espresse fossero corrette, confermerebbero l'idea secondo la quale, sia per GT che per MsT, l'output inglese dovrebbe risultare

leggermente migliore o almeno equivalente in termini di qualità rispetto a quello tedesco, che potrebbe contenere errori derivanti dal doppio passaggio linguistico in sede di traduzione.

5.3.5 Analisi delle principali problematiche riscontrate nelle traduzioni

Alcuni degli errori di traduzione rilevati sono ricorrenti, ovvero si ripresentano in forma simile in più di un output tradotto. Tali errori sono spesso stati riscontrati sia negli output prodotti da GT, sia negli output prodotti da MsT. Questi talvolta sono identici (si veda Appendice 1); in altri casi sono diversi, ma presentano errori simili, riconducibili ad una o più tipologie.

Questa sezione sarà dedicata proprio alla discussione delle problematiche e degli errori di traduzione più ricorrenti negli output valutati. Con questa analisi, tuttavia, non si ha la pretesa di fornire un quadro esaustivo degli errori o delle particolarità delle traduzioni prodotte dalle app valutate: non si procederà cioè all'analisi di ciascun output, né si descriveranno nel dettaglio tutti gli errori presenti nelle traduzioni. Lo scopo di questa breve sezione è invece quello di fornire una panoramica generale delle principali problematiche riscontrate, la quale potrebbe sia rivelarsi utile per i potenziali utenti del servizio di Speech-to-Speech Translation offerto da GT e MsT, sia fungere da base per studi futuri incentrati su questo aspetto o finalizzati ad apportare migliorie ai sistemi di traduzione di GT e MsT.

Gli errori ricorrenti riscontrati negli output tradotti da GT e MsT possono essere suddivisi schematicamente nelle seguenti categorie (parzialmente adattate da Grazina, 2010):

- errori di riconoscimento vocale: alcuni degli errori contenuti nelle traduzioni prodotte dalle due app valutate derivano da errori di riconoscimento vocale (si veda 5.3.3). È il caso di toponimi o nomi propri non riconosciuti correttamente dal modulo ASR e sostituiti con termini di uso comune (si veda Appendice 1);

- errori prodotti dal modulo MT, i quali si dividono ulteriormente in:
 - parole mancanti: alcuni output sono caratterizzati da parole mancanti, talvolta “parole piene”, ovvero dotate di un significato preciso, come sostantivi, aggettivi, verbi, ecc., talvolta “parole vuote” o *function words*, prive di significato, ma dotate di una precisa funzione grammaticale, come l'inglese “do” in funzione di verbo ausiliare. Tali output non solo risultano incompleti, ma talvolta anche difficili da comprendere. A titolo di esempio si riportano i seguenti casi, tratti dall'Appendice 1: “From which platform the train to London” (verbo e ausiliare mancanti), “How much?” (verbo e soggetto mancanti), “Könnte sprechen langsamer” (soggetto mancante);
 - ordine delle parole: alcuni output presentano errori a livello di ordine o posizione delle parole nella frase, il quale, soprattutto in tedesco, segue regole ben precise. Per questo tipo di errore si riportano alcuni output in tedesco caratterizzati da *Satzstellung* errata (si veda Appendice 1): “Wie viel es kostet” (ordine invertito fra soggetto e verbo), “Welche Zeit Sie Frühstück servieren” (ordine invertito fra soggetto e verbo);
 - termini errati o errori di senso: in alcuni casi gli output tradotti presentano errori derivanti da una scorretta disambiguazione della frase *source* o da una scelta lessicale sbagliata. Errori che rientrano in questa categoria compaiono nelle seguenti frasi (si veda Appendice 1): “Where is the nearest distributor”, dove “distributor” non è la traduzione corretta del termine *source* “distributore”, che, se usato in contesto turistico, indica chiaramente una stazione di rifornimento, in inglese “petrol station” o “gas station”. In questo caso, il modulo MT, privo di conoscenza del mondo, non ha compiuto la scelta lessicale più adeguata al contesto; “Leider spricht Italienisch”, dove dall'analisi della frase *source* “Scusi parla italiano?” il termine “scusi” è stato erroneamente interpretato dal modulo MT come il soggetto di “parla”. In questo caso ci si trova in presenza di una scorretta disambiguazione della frase *source*, poiché il modulo MT non ha

riconosciuto la forma di cortesia usata per il verbo “parlare”, riconducendo la forma coniugata “parla” alla terza persona singolare dell'indicativo presente;

- frasi interrogative: diverse frasi originariamente interrogative sono state tradotte in inglese e tedesco con sintassi tipicamente affermativa; in questo caso è difficile stabilire se l'errore sia dovuto al riconoscimento vocale o a un errore del modulo di traduzione automatica. Un esempio: “La colazione è inclusa?”, che ha dato origine, fra gli altri, agli output “Breakfast is included” e “Frühstück ist inklusive”, entrambi affermativi.

Nella classificazione delle tipologie di errori ricorrenti in una traduzione automatica Grazina (2010) include anche gli errori di punteggiatura. Tuttavia, nell'ambito dello studio condotto, considerando da un lato lo scopo della traduzione automatica del parlato offerta dalle due applicazioni e lo scenario d'uso ipotizzato, e dall'altro il fatto che le traduzioni prodotte vengono pronunciate dal modulo TTS, si è concluso che eventuali errori di punteggiatura non rappresentano un ostacolo evidente alla comprensione degli output tradotti, benché, come si è visto in 4.5.2, alcuni dei valutatori coinvolti abbiano dato maggior peso a questo aspetto rispetto ad altri.

VI

Conclusioni

6.1 Contenuti principali dell'elaborato

Il presente elaborato è partito da una panoramica generale sulla traduzione automatica, anche nota come Machine Translation (MT), per approdare all'analisi della Speech-to-Speech Translation (SST), una tecnologia che utilizza il riconoscimento vocale, la traduzione automatica e la sintesi vocale per tradurre automaticamente input vocali da una lingua A a una lingua B, producendo output tradotti che vengono riprodotti acusticamente.

Il primo capitolo, incentrato sulla traduzione automatica di testi scritti, è partito da un excursus storico sulla nascita e l'evoluzione di questa tecnologia per poi analizzare il funzionamento e la struttura dei sistemi di MT, i possibili impieghi e le diverse categorie di utenti di tali strumenti. Oltre ai numerosi vantaggi offerti dalla traduzione automatica, sono stati delineati anche i limiti e le problematiche inerenti a questa tecnologia, che rendono di fatto impossibile allo stato attuale della conoscenza il raggiungimento della Fully Automatic High Quality Translation (FAHQT). Sono stati poi analizzati i diversi metodi utilizzati per arginare i limiti della traduzione automatica, migliorando la qualità del prodotto finale (pre-editing, post-editing e redazione in linguaggio controllato). Da qui si è passati alla definizione di qualità nell'ambito della traduzione automatica, per analizzarne infine i possibili metodi di valutazione.

Dalla traduzione automatica di testi scritti il focus è passato alla traduzione automatica di input orali nel secondo capitolo, incentrato sulla traduzione automatica del parlato. In un percorso simile a quello che ha scandito il primo capitolo, sono state analizzate le principali tappe storiche che hanno

portato alla nascita e alla diffusione di questa tecnologia, il funzionamento e la struttura dei sistemi di SST, nonché le problematiche e le sfide che accompagnano la traduzione automatica della lingua parlata. Anche in questo caso sono stati proposti alcuni metodi per sopperire ai limiti dei sistemi di SST attualmente disponibili, basati ad esempio sulla cooperazione dell'utente sia nella fase di riconoscimento vocale, sia, in misura minore, nella fase successiva alla traduzione. È seguita una descrizione delle diverse applicazioni pratiche della tecnologia SST, delle sfide che i ricercatori attivi in questo campo dovranno affrontare nei prossimi anni e delle prospettive future di questa tecnologia. Infine, come per la Machine Translation, anche per la Speech-to-Speech Translation è stata affrontata la questione della valutazione della qualità, un tema di fondamentale importanza per lo studio descritto negli ultimi capitoli dell'elaborato.

Tale studio si è incentrato sulla valutazione di Google Translate e Microsoft Translator, due applicazioni per la traduzione automatica che per alcune coppie di lingue offrono, fra gli altri, anche un servizio di Speech-to-Speech Translation. Nel capitolo III sono stati presentati gli aspetti metodologici del test di valutazione condotto, a partire dagli obiettivi del test e dalle app oggetto di valutazione, passando per i partecipanti, fino allo svolgimento del test e alle modalità di raccolta dei dati e di redazione dei questionari utilizzati per la valutazione della qualità degli output tradotti.

I risultati della valutazione dei moduli di riconoscimento vocale e di traduzione automatica sono stati esposti nel capitolo IV e discussi nel capitolo V, dove sono state confrontate le performance di Google Translate e Microsoft Translator sia in termini di accuratezza del riconoscimento vocale, sia in termini di correttezza della traduzione fornita, e sono stati analizzati tutti i dati raccolti alla luce degli obiettivi esposti nel capitolo III.

6.2 Conclusioni raggiunte

Da questa analisi è emerso che il servizio di traduzione automatica del parlato offerto gratuitamente da queste due applicazioni presenta ancora alcune imperfezioni legate sia al riconoscimento vocale, che non sempre viene portato a termine correttamente al primo tentativo, sia alla qualità e alla fruibilità delle traduzioni fornite dal modulo di traduzione automatica. Benché la possibilità di tradurre automaticamente e gratuitamente conversazioni o interazioni orali da una lingua all'altra sia di fatto straordinaria, dallo studio condotto sono emersi i limiti del servizio offerto da applicazioni gratuite quali Google Translate e Microsoft Translator.

Nel valutare l'utilità e l'usabilità di tali applicazioni, occorre tuttavia prendere in considerazione diversi fattori. Innanzitutto è importante sottolineare il fatto che si tratta, in questo caso specifico, di applicazioni gratuite e scaricabili su dispositivi mobili, progettate per l'utilizzo da parte di un pubblico ampio. Nel panorama dei sistemi di SST, tali applicazioni hanno scopi e caratteristiche simili a quelle dei sistemi di MT online nel panorama degli strumenti di traduzione automatica. Nel caso specifico del servizio di Speech-to-Speech Translation, ciò significa che tali applicazioni non sono certamente progettate per un uso intensivo e sistematico in contesti altamente tecnici, come ad esempio l'ambito medico-ospedaliero o quello commerciale, nei quali dalla comunicazione dipendono rispettivamente le sorti di un paziente e l'esito di una trattativa commerciale (benché in generale tali ambiti rappresentino potenziali scenari di impiego della tecnologia SST). Al contrario, proprio come i sistemi di MT online, queste applicazioni sono progettate per l'utilizzo da parte di utenti non esperti, in contesti quotidiani di interazione. Alla luce di ciò e in considerazione del potenziale uso che un utente medio può fare del servizio di traduzione automatica del parlato offerto da un'applicazione come Google Translate o Microsoft Translator, si può concludere che l'obiettivo della tecnologia SST in questo caso non è tanto quello di garantire traduzioni perfette degli input orali formulati, bensì quello di produrre traduzioni fruibili, che, pur presentando talvolta errori o imprecisioni, consentano la comunicazione e la

comprensione reciproca fra utenti di lingue diverse. La qualità della traduzione richiesta in questo caso non è la stessa normalmente richiesta, ad esempio, a un interprete professionista nella mediazione di una trattativa commerciale. E ciò deve necessariamente essere tenuto in considerazione nella valutazione del servizio di Speech-to-Speech Translation offerto da questo tipo di applicazione.

In secondo luogo, è probabile che il servizio di SST offerto da queste e altre applicazioni simili migliori nel tempo, proprio come è successo per i sistemi di traduzione automatica online. Nel caso in cui le applicazioni adottino sistemi di SST basati sulla traduzione automatica statistica (come Google Translate e Microsoft Translator), il miglioramento in termini di qualità degli output tradotti potrà derivare sia da un aumento delle dimensioni dei corpora allineati utilizzati per l'addestramento, sia dai feedback sul servizio forniti dagli utenti. Anche le tecnologie di riconoscimento e di sintesi vocale sono in continua evoluzione e potranno contribuire in futuro al miglioramento delle prestazioni delle applicazioni per la SST, o più in generale dei sistemi di SST. Nel presente elaborato si è visto come la tecnologia MT ha compiuto passi da gigante dalla sua nascita; è dunque ragionevole aspettarsi che, nonostante le numerose sfide da affrontare, anche la traduzione automatica del parlato vada incontro a simili progressi nei prossimi decenni.

Infine, occorre considerare la difficoltà intrinseca alla traduzione della lingua parlata, che, come si è visto nel presente elaborato, presenta caratteristiche ben diverse rispetto alla lingua scritta. Se è vero che anche un interprete professionista può trovarsi in difficoltà nel corso del lavoro ed essere soggetto a errori o incomprensioni, a maggior ragione non ci si può aspettare una prestazione perfetta da parte di una macchina priva di conoscenza del mondo e delle competenze specifiche che contraddistinguono un interprete.

Considerando tutti questi aspetti e sulla base della valutazione effettuata nel presente studio, si può concludere che, nonostante i limiti riscontrati, le applicazioni per la traduzione automatica del parlato rappresentano uno strumento utile per tutti quegli utenti che, per necessità, si trovano a dovere interagire con interlocutori stranieri con i quali non condividono alcuna lingua, in contesti non tecnici, come ad esempio l'ambito turistico. In

contesti altamente tecnici invece, come l'ambito commerciale o quello medico, è ancora probabilmente sconsigliabile fare affidamento sul servizio di traduzione automatica del parlato offerto da applicazioni gratuite come quelle analizzate.

Allo stato attuale, per ottenere prestazioni soddisfacenti dalle app per la SST, gli utenti devono però essere consapevoli dei limiti di questa tecnologia e delle problematiche connesse alla traduzione della lingua parlata. Sarebbe dunque auspicabile sensibilizzare i potenziali utenti rispetto all'uso corretto delle applicazioni per la SST e alle aspettative che realisticamente si possono avere sulla qualità delle traduzioni. Attenendosi a semplici indicazioni, come quelle esposte nel capitolo III di questo elaborato, gli utenti potrebbero ottenere un servizio di traduzione mediamente soddisfacente, che certamente migliorerà nei prossimi decenni.

Per quanto riguarda nello specifico le due applicazioni valutate nel presente studio, basandomi sull'esperienza acquisita, la conclusione raggiunta è che Microsoft Translator rappresenta un valido supporto alla comunicazione informale e non specializzata e può sostituire il classico frasario cartaceo per turisti e viaggiatori, fornendo traduzioni generalmente più idiomatiche e accurate rispetto a quelle fornite da Google Translate. Occorre tuttavia prestare molta attenzione al riconoscimento vocale, verificandone costantemente la correttezza e correggendo manualmente eventuali errori. Si è infatti visto come l'applicazione Microsoft Translator sia meno precisa di Google Translate in fase di riconoscimento vocale. È inoltre consigliabile utilizzare l'applicazione in luoghi non troppo rumorosi, proprio per non ostacolare il riconoscimento vocale delle frasi pronunciate, oppure, qualora ciò sia impossibile, utilizzare funzioni diverse da quella di traduzione automatica del parlato. Infine, occorre armarsi di pazienza, poiché il servizio di traduzione del parlato offerto da Microsoft Translator impiega qualche secondo per portare a termine il riconoscimento vocale e la traduzione, senza considerare il fatto che in alcuni casi potrebbero essere necessari più tentativi prima di ottenere una traduzione corretta o quantomeno comprensibile della frase pronunciata.

6.3 Prospettive future

Per quanto riguarda infine lo studio condotto e presentato in questo elaborato, sono emersi sia alcuni limiti che è opportuno mettere in evidenza, sia alcuni aspetti che potrebbero essere approfonditi in studi futuri.

Dal punto di vista strettamente metodologico, lo studio condotto ha fatto emergere ad esempio alcuni aspetti che potrebbero essere migliorati o corretti in studi futuri, soprattutto per quanto riguarda la raccolta delle valutazioni soggettive sulla qualità della traduzione automatica. Infatti nel presente studio, analizzando i giudizi sulla qualità degli output tradotti espressi dai valutatori e raccolti tramite gli appositi questionari, è stata registrata una percentuale di risposte per le quali i soggetti coinvolti nella valutazione si sono trovati in disaccordo. Il livello di disaccordo riscontrato non ha impedito di ottenere risultati attendibili per la valutazione della qualità delle traduzioni prodotte dalle applicazioni analizzate; tuttavia ha reso necessaria l'esclusione di alcune risposte (indicanti disaccordo) dal calcolo della media dei giudizi forniti dai valutatori. Benché un certo livello di disaccordo fra i valutatori sia “fisiologico” all'interno di una valutazione soggettiva che raccoglie i giudizi di più soggetti, studi successivi potrebbero limitarne l'entità intervenendo ad esempio sulla parte introduttiva dei questionari di valutazione distribuiti. I questionari adottati nel presente studio presentavano infatti una breve introduzione dove veniva fornita una spiegazione dei parametri di valutazione adottati; tuttavia valutatori diversi sembrano avere interpretato in modo differente gli aspetti da valutare negli output tradotti, dando diverso peso a eventuali errori di traduzione rilevati. Per ricerche future di questo tipo potrebbe essere utile includere nella parte introduttiva del questionario alcuni esempi di valutazione già compilati, corredati da una spiegazione del motivo per cui all'output tradotto in quell'esempio è stato attribuito un determinato punteggio per quello specifico parametro. Potrebbe anche essere utile, ove possibile, riunire i soggetti coinvolti nella valutazione per fornire delucidazioni sugli aspetti da valutare nella traduzione e sui parametri adottati, promuovendo altresì uno scambio di opinioni fra i valutatori stessi.

Ricerche future potrebbero inoltre coinvolgere valutatori bilingue perfetti per la valutazione di *fluency* ed equivalenza semantica, oppure valutatori madrelingua per la valutazione della *fluency* e valutatori bilingue per la valutazione dell'equivalenza semantica. Potrebbe anche essere interessante confrontare i risultati ottenuti tramite l'impiego di metodi soggettivi e oggettivi di valutazione, per verificare la correlazione fra giudizi soggettivi e parametri oggettivi di valutazione e identificare il metodo oggettivo migliore per la valutazione di singole frasi tradotte da un sistema di Speech-to-Speech Translation.

Studi futuri potrebbero inoltre trattare aspetti che non sono stati presi in considerazione in questa ricerca, ad esempio includendo nella valutazione il modulo di sintesi vocale, escluso nel presente studio. Oppure potrebbero analizzare la performance di Google Translate e Microsoft Translator (o applicazioni simili) in contesti diversi da quello turistico. Infine, sarebbe interessante condurre un vero e proprio *field test* sulle applicazioni valutate in questo studio, riproducendo una situazione il più verosimile possibile e analizzando la loro efficacia sul campo, ovvero la loro capacità di “mediare” l'interazione fra un partecipante di lingua A e un partecipante di lingua B che non hanno modo di comprendersi se non attraverso il servizio di Speech-to-Speech Translation.

Bibliografia

- AA.VV. (2006). “Tecnologie del linguaggio per l'Europa”, report sulle tecnologie del linguaggio finanziato dall'Unione Europea. Online: http://www.tcstar.org/pubblicazioni/ITC_ita.pdf [visitato: 11/11/2015].
- AA.VV. (2014). “MT@EC. Secure Machine Translation for the European Union”, depliant informativo sul servizio MT@EC della Commissione Europea. Online: http://ec.europa.eu/isa/documents/publications/brochure-mt@ec-a5-v3_en.pdf [visitato: 11/11/2015].
- Adams, D. (2012). *Guida galattica per gli autostoppisti*. Traduzione di Laura Serra. Milano: Arnoldo Mondadori Editore.
- Aston, G. (2011). “Tecniche per migliorare la traduzione automatica: post-editing e pre-editing”. In G. Bersani Berselli (2011). 33-45.
- Bersani Berselli, G. a cura di (2011). *Usare la traduzione automatica*. Bologna: CLUEB.
- Bub, T., W. Wahlster, A. Waibel (1997). “Verbmobil: the combination of deep and shallow processing for spontaneous speech translation”. *1997 IEEE International Conference on Acoustics, Speech, And Signal Processing*, vol.1, 71-74.
- Chan, S.W. (2004). *A dictionary of translation technology*. Hong Kong: The Chinese University Press.
- Chan, S.W. a cura di (2015). *Routledge Encyclopedia of Translation Technology*. New York: Routledge.
- Chen, S. F., B. Kingsbury, L. Mangu, D. Povey, G. Saon, H. Soltau e G. Zweig (2006). “Advances in Speech Transcription at IBM Under the DARPA EARS Program”. *IEEE Transactions on audio, speech, and language processing*, vol. 14, n. 5, settembre, 1596-1608.
- Chiari, I. (2011). “Traduzione automatica”. *X la Tangente*, aprile, 31-33.
- Dillinger, M. e A. Lommel (2004). “LISA Best Practice Guide. Implementing Machine Translation”, depliant informativo sulla MT della Localization Industry Standards Association. Online: http://www.mikedillinger.com/papers/MT_BPG2004.pdf [visitato: 16/11/2015].
- Dorr, B., J. Olive, J. McCary, C. Christianson (2011) “Machine Translation Evaluation and Optimization”. In Olive, J., C. Christianson, J. McCary, a cura di, *Handbook of Natural Language Processing and Machine Translation*. New

York: Springer. 745-843.

Dureja, M. e S. Gautam (2015). "Speech-to-Speech Translation: A Review". *International Journal of Computer Applications*, vol. 129, n.13, novembre 2015, 28-30.

Euromatrix (2007). "Survey of Machine Translation Evaluation", progetto di ricerca sulla traduzione automatica finanziato dall'UE. Online: http://www.euromatrix.net/deliverables/Euromatrix_D1.3_Revised.pdf [visitato: 09/03/2016].

Frandsen, M.W., S. Riehemann, K. Precoda (2008). "IraqComm and FlexTrans: A Speech Translation System and Flexible Framework". *SCSS*, vol. 1, 527-532.

Gaspari, F. e J. Hutchins (2007). "Online and Free! Ten Years of Online Machine Translation: Origins, Developments, Current Use and Future Prospects". *Proceedings of Machine Translation Summit XI*, Copenhagen Business School, Copenhagen (Danimarca), 10-14 settembre 2007, 199-206.

Gaspari, F. (2011). "Introduzione alla traduzione automatica". In G. Bersani Berselli (2011). 13-31.

Gaspari, F. e E. Zanchetta (2011). "Scrittura controllata per la traduzione automatica". In G. Bersani Berselli (2011). 63-79.

Grazina, N.M.C. (2010). "Automatic Speech Translation". Tesi di laurea magistrale non pubblicata, Universidade Técnica de Lisboa (UTL), Instituto Superior Técnico. Online: <https://fenix.tecnico.ulisboa.pt/downloadFile/395142104990/dissertacao.pdf> [visitato: 12/04/2016].

Greene, L. (2016). "Everything you ever wanted to know about Google Translate, and finally got the chance to ask". Online: <https://www.taus.net/think-tank/articles/everything-you-ever-wanted-to-know-about-google-translate-and-finally-got-the-chance-to-ask> [visitato: 01/06/2016]

Hutchins, J. (1995). "Machine translation: a brief history". In *Concise history of the language sciences: from the Sumerians to the cognitivists*, E.F.K. Koerner e R.E. Asher, a cura di. Oxford: Pergamon Press. 431-445. Online: <http://www.hutchinsweb.me.uk/history.htm> [visitato: 27/10/2015].

Hutchins, J. (2002). "The state of machine translation in Europe and future prospects". *HLT Central*, gennaio 2002. Online: <http://www.hutchinsweb.me.uk/HLT-2002.pdf> [visitato: 20/06/2016].

Hutchins, J. (2005). "Current commercial machine translation systems and computer-based translation tools: system types and their uses". *International Journal of Translation*, vol. 17, n. 1-2, dicembre, 5-38.

- Hutchins, J. (2009). "Multiple Uses of Machine Translation and Computerised Translation Tools". *International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages*, Besançon, 13-20.
- Hutchins, J. (2010). "Machine translation: a concise history". *Journal of Translation Studies*, vol. 13. Online: <http://www.hutchinsweb.me.uk/history.htm> [visitato: 27/10/2015].
- Kit, C. e T. M. Wong (2008). "Comparative Evaluation of Online Machine Translation Systems with Legal Texts". *Law Library Journal*, vol. 100, 299-321.
- Lazzari, G. (2002). "Verso i traduttori universali del parlato. Progetti in corso e prospettive". *Conferenza TIPI (Tecnologie Informatiche nella Promozione della lingua Italiana)*, 25-26 giugno 2002, 211-215.
- Lazzari, G. (2004). "Comunicare con le tecnologie del linguaggio nell'era globale". *Media Duemila – Mensile di Cultura Informatica e ICT*, anno XXII, n.219, settembre, 71-78.
- Lee, T. (2015). "Speech Translation". In S.W. Chan (2015). 619-631.
- Liu, Q. e X. Zhang (2015). "Machine Translation. General". In S.W. Chan (2015). 105-119.
- Metze, F., J. McDonough, H. Soltau, A. Waibel, A. Lavie, S. Burger, C. Langley, L. Levin, T. Schultz, F. Piansesi, et al. (2002). "The NESPOLE! Speech to Speech Translation System". *Proceedings of the Human Language Technologies 2002*, marzo 2002, San Diego-California, 378-383.
- Monti, J. (2004). "Dal sogno meccanico alla e-translation: la traduzione automatica è realtà?". *Media Duemila – Mensile di Cultura Informatica e ICT*, n. 219, anno XXII, settembre, 60-67.
- Naldi, M. (2014). *Traduzione automatica e traduzione assistita*. Bologna: Esculapio.
- Nakamura, S. (2009). "Overcoming the Language Barrier with Speech Translation Technology". *Science & Technology Trends - Quarterly Review*, n. 31, aprile 2009, 36-48.
- Osimo, B. (2004). *Traduzione e qualità. La valutazione in ambito accademico e professionale*. Milano: Ulrico Hoepli Editore.
- Ping, K. (2009). "Machine Translation". *The Routledge Encyclopedia of Translation Studies*, Routledge, Abingdon (seconda edizione), 162-169.

Ranucci-Fisher, E. (2004). "L'informatica a sostegno del multilinguismo: il servizio di traduzione della Commissione europea". *Media Duemila – Mensile di Cultura Informatica e ICT*, anno XXII, n. 219, settembre, 79-84.

Riediger, H. e G. Galati (2012). "Tecnologie per la traduzione". Online: <http://www.fondazionemilano.eu/moodle/course/view.php?id=278> [visitato: 27/10/2015].

Rossato, S., H. Blanchon, L. Besacier (2002). "Speech-to-speech translation system evaluation: results for French for the Nespole! project first showcase". *ICSLP-2002*, 1905-1908.

Seligman, Mark (2000). "Nine Issues in Speech Translation". *Machine Translation*, vol. 15, n.1/2, 149-185.

Sobrero, A.A. e A. Miglietta (2010). *Introduzione alla linguistica italiana*. Bari: Editori Laterza.

Tirronen, Silja (2011). "Automated Testing of Speech-to-Speech Machine Translation in Telecom Networks". Tesi di laurea magistrale non pubblicata, Università degli Studi di Aalto, School of Electrical Engineering, Department of Signal Processing and Acoustics. Online: <http://lib.tkk.fi/Dipl/2011/urn100412.pdf> [visitato: 16/04/2016].

Van der Meer, J. (2014). "Speech Translation Technology: A conversation between Alexander Waibel and Mark Seligman. Online: <https://www.taus.net/think-tank/articles/translate-articles/speech-translation-technology-a-conversation-between-alexander-waibel-and-mark-seligman> [visitato: 17/04/2016]

Viezzi, M. (1999). "Aspetti della qualità nell'interpretazione". In *Interpretazione simultanea e consecutiva*, C. Falbo, M. Russo, F. Straniero Sergio, a cura di. Milano: Ulrico Hoepli Editore, 140-151.

Zanettin, F. (2001). "Informatica e Traduzione". In *Traduzione, revisione e localizzazione nel terzo millennio: da e verso l'inglese*, C. Monacelli, a cura di. Milano: Franco Angeli editore. 19-45.

Zhang, Y. (2003) "Survey of Current Speech Translation Research". *Multilingual Speech-to-Speech Translation Seminar*, Carnegie Mellon University, Pittsburgh, 2003. Online: https://www.researchgate.net/profile/Xuan_Cheng/publication/228693496_Survey_of_current_speech_translation_research/links/54892e6e0cf289302e30c166.pdf

Zong, C. e M. Seligman (2005). "Toward Practical Spoken Language Translation". *Machine Translation*, vol.19, n.2, giugno 2005, 113-137.

This work focuses on Machine Translation (MT) and Speech-to-Speech Translation, two emerging technologies that allow users to automatically translate written and spoken texts.

The first part of this work provides a theoretical framework for the evaluation of Google Translate and Microsoft Translator, which is at the core of this study. Chapter one focuses on Machine Translation, providing a definition of this technology and glimpses of its history. In this chapter we will also learn how MT works, who uses it, for what purpose, what its pros and cons are, and how machine translation quality can be defined and assessed. Chapter two deals with Speech-to-Speech Translation by focusing on its history, characteristics and operation, potential uses and limits deriving from the intrinsic difficulty of translating spoken language. After describing the future prospects for SST, the final part of this chapter focuses on the quality assessment of Speech-to-Speech Translation applications.

The last part of this dissertation describes the evaluation test carried out on Google Translate and Microsoft Translator, two mobile translation apps also providing a Speech-to-Speech Translation service. Chapter three illustrates the objectives, the research questions, the participants, the methodology and the elaboration of the questionnaires used to collect data. The collected data and the results of the evaluation of the automatic speech recognition subsystem and the language translation subsystem are presented in chapter four and finally analysed and compared in chapter five, which provides a general description of the performance of the evaluated apps and possible explanations for each set of results. In the final part of this work suggestions are made for future research and reflections on the usability and usefulness of the evaluated translation apps are provided.

Schwerpunkte dieser Arbeit sind die maschinelle Übersetzung (MÜ oder MT aus dem Englischen Machine Translation) und die maschinelle Sprachübersetzung, zwei entstehende Technologien, die das automatische Übersetzen von geschriebener und gesprochener Sprache ermöglichen.

Die ersten zwei Kapitel dienen als theoretischer Rahmen für die Studie, die in den letzten Kapiteln präsentiert wird. Genauer gesagt werden im ersten Kapitel die Geschichte der maschinellen Übersetzung von geschriebenem Text, ihre Funktionsweise, die verschiedenen Verwendungsmöglichkeiten und Gruppen von Benutzern sowie die Vor- und Nachteile dieser Technologie und die Qualitätsbewertung abgehandelt.

Im zweiten Kapitel wird dagegen das Thema maschinelle Sprachübersetzung (SST) behandelt. Dieses Kapitel konzentriert sich auf die Geschichte dieser Technologie, die sich aus der traditionellen Maschinenübersetzung entwickelt hat, auf ihre Funktionsweise und Merkmale, auf ihre Grenzen und Schwierigkeiten und noch weiter auf ihre Nutzungsmöglichkeiten und Zukunftsaussichten und im letzten Teil des Kapitels wird das Thema Qualität und Qualitätsbewertung analysiert.

Daraufhin wird ein Test für die Bewertung von zwei Übersetzung-Apps, nämlich Google Translate und Microsoft Translator, erläutert. Im dritten Kapitel wird diese Studie vom Zweck, über die Planung bis zur Methodik für die Erhebung von Daten ausführlich erläutert. Im vierten Kapitel werden die Testergebnisse vorgestellt, die aus der Bewertung des maschinelle Spracherkennung-Systems und des MT-Systems hervorgehen. Im fünften Kapitel werden diese Daten in Bezug auf ihre möglichen Ursachen analysiert und verglichen.

Im letzten Teil dieses Werks werden Forschungsempfehlungen formuliert sowie Schlussfolgerungen über die Nützlichkeit und die Verwendbarkeit der in der Studie bewerteten Apps gezogen.

Appendice 1

Partecipante 1 – Test italiano-inglese

Frase in lingua <i>source</i>	Ultimo riconoscimento vocale in GT / input per traduzione*	Output tradotto GT	Ultimo riconoscimento vocale in MsT / input per traduzione*	Output tradotto MsT
1. Vorrei un biglietto andata e ritorno per Edimburgo.	Vorrei un biglietto andata e ritorno per Edimburgo	I'd like a return ticket to Edinburgh	Vorrei un biglietto andata e ritorno per Edimburgo.	I'd like a return ticket to Edinburgh.
2. Mi scusi, dove si trova la stazione dei treni?	Mi scusi dove si trova la stazione dei treni	Excuse me, where is the train station	Mi scusi dove si trova la stazione dei treni?	Excuse me, where is the train station?
3. Da quale binario parte il treno per Londra?	Da quale binario parte il treno per Londra	From which platform the train to London	Da quale binario parte il treno per Londra?	From which platform does the train leave for London?
4. Quanto costa un biglietto solo andata per Plymouth?	Quanto costa un biglietto solo andata per Plymouth (2° tentativo)	How much does a one-way ticket to Plymouth	Quanto costa un biglietto solo andata per prima. (2° tentativo)	How much is a one-way ticket to first.
5. Dov'è il distributore più vicino?	Dov'è il distributore più vicino	Where is the nearest distributor	Dov'è il distributore più vicino.	Where's the nearest distributor.
6. Si può parcheggiare qui?	Si può parcheggiare qui	You can park here	Si può parcheggiare qui.	You can park here.
7. Esistono abbonamenti settimanali?	Esistono abbonamenti settimanali	There are weekly tickets	Esistono abbonamenti settimanali.	There are weekly tickets.
8. Quanto costa il noleggio per una settimana?	Quanto costa il noleggio di una settimana (2° tentativo)	How much is the rental of a week	Quanto costa il noleggio per una settimana?	How much is the rent for a week?
9. La colazione è inclusa?	La colazione è inclusa (2° tentativo)	Breakfast is included	La colazione è inclusa. (2° tentativo)	Breakfast is included.

10. A che ora servite la colazione?	A che ora servite la colazione	What time served breakfast	A che ora servite la colazione? (2° tentativo)	What time you serve breakfast?
11. Può chiamarmi un taxi?	Può chiamarmi un taxi	You can call a taxi	Puoi chiamarmi un taxi. (2° tentativo)	You can call me a cab.
12. Scusi, dov'è la toilette?	Scusi dov'è la toilette (2° tentativo)	Excuse me where is the toilet	Scusi dov'è la toilette.	Excuse me where's the toilet.
13. Avete una cartina della città?	Avete una cartina della città	You have a map of the city	Avete una cartina della città.	Do you have a map of the city.
14. Scusi, parla italiano?	Scusi parla italiano	Excuse speaks Italian	Scusi parla italiano. (2° tentativo)	Sorry speaks Italian.
15. Potrebbe parlare più lentamente?	Potrebbe parlare più lentamente	Could speak more slowly	Potrebbe parlare più lentamente.	Could you speak more slowly.
16. Dove si acquista lo skipass?	Dove si acquista lo skipass	Where you buy your ski pass	Dove si acquista lo skipass.	Where do you buy your ski pass.
17. Dove si trova l'ambasciata italiana?	Dove si trova l'ambasciata italiana	Where the Italian embassy	Dove si trova l'ambasciata italiana.	Where is the Italian embassy.
18. Buongiorno, ho prenotato una camera a nome di Rossi.	Buongiorno ho prenotato una camera a nome di Rossi	Good morning I booked a room in the name of Rossi	Buongiorno ho prenotato una camera il nome di Rossi. (2° tentativo)	Good morning I made reservation the name of Rossi.
19. Quanto costa?	Quanto costa	How much	Quanto costa?	How much is it?
20. La mia valigia non è arrivata.	La mia valigia non è arrivata	My suitcase did not arrive	Ravioli sono arrivata. (2° tentativo)	Ravioli are coming.
21. Manca l'acqua calda.	Manca l'acqua calda	No hot water	Manca l'acqua calda. (2° tentativo)	Missing hot water.

* Gli output ASR qui riportati sono esclusivamente relativi all'ultimo tentativo di riconoscimento vocale effettuato dal partecipante. Ove non diversamente indicato, il riconoscimento vocale è avvenuto al primo tentativo. Per tutti i casi in cui vi sono stati più tentativi di riconoscimento vocale, viene riportato unicamente l'output ASR derivante dal secondo e ultimo tentativo di riconoscimento vocale (indicato esplicitamente nella tabella tramite la dicitura "2° tentativo"), in considerazione sia del fatto che in tali casi il primo tentativo di riconoscimento vocale era stato ritenuto scorretto già dal partecipante stesso in sede di svolgimento del test di valutazione, sia del fatto che in questi casi l'output tradotto derivante dal primo tentativo di riconoscimento vocale non è rilevante ai fini di questo studio in quanto escluso dalla valutazione della MT.

Partecipante 2 – Test italiano-inglese

Frase in lingua <i>source</i>	Ultimo riconoscimento vocale GT / input per traduzione*	Output tradotto GT	Ultimo riconoscimento vocale MsT / input per traduzione*	Output tradotto MsT
1. Vorrei un biglietto andata e ritorno per Edimburgo.	Vorrei un biglietto andata e ritorno per Edimburgo	I'd like a return ticket to Edinburgh	Vorrei un biglietto andata e ritorno per Edimburgo.	I'd like a return ticket to Edinburgh.
2. Mi scusi, dove si trova la stazione dei treni?	Mi scusi dove si trova la stazione dei treni	Excuse me, where is the train station	Mi scusi dove si trova la stazione dei treni?	Excuse me, where is the train station?
3. Da quale binario parte il treno per Londra?	Da quale binario parte il treno per Londra	From which platform the train to London	Da quale binario parte il treno per Londra?	From which platform does the train leave for London?
4. Quanto costa un biglietto solo andata per Plymouth?	Quanto costa un biglietto solo andata per prima (2° tentativo)	How much it costs only one-way ticket to first	Quanto costa un biglietto solo andata per primo. (2° tentativo)	How much is a ticket one-way first.
5. Dov'è il distributore più vicino?	Dov'è il distributore più vicino	Where is the nearest distributor	Dov'è il distributore più vicino.	Where's the nearest distributor.
6. Si può parcheggiare qui?	Si può parcheggiare qui	You can park here	Si può parcheggiare qui.	You can park here.
7. Esistono abbonamenti settimanali?	Esistono abbonamenti settimanali	There are weekly tickets	Esistono abbonamenti settimanali.	There are weekly tickets.
8. Quanto costa il noleggio per una settimana?	Quanto costa il noleggio per una settimana	How much is the rental for a week	Quanto costa il noleggio per una settimana?	How much is the rent for a week?
9. La colazione è inclusa?	La colazione è inclusa (2° tentativo)	Breakfast is included	La colazione è inclusa. (2° tentativo)	Breakfast is included.
10. A che ora servite la colazione?	A che ora servite la colazione	What time served breakfast	A che ora servite la colazione?	What time you serve breakfast?
11. Può chiamarmi un taxi?	Può chiamarmi un taxi	You can call a taxi	Puoi chiamarmi un taxi. (2° tentativo)	You can call me a cab.

12. Scusi, dov'è la toilette?	Scusi dov'è la toilette (2° tentativo)	Excuse me where is the toilet	Scusi dov'è la toilette.	Excuse me where's the toilet.
13. Avete una cartina della città?	Avete una cartina della città	You have a map of the city	Avete una cartina della città.	Do you have a map of the city.
14. Scusi, parla italiano?	Scusi parla italiano	Excuse speaks Italian	Scusi parla italiano.	Sorry speaks Italian.
15. Potrebbe parlare più lentamente?	Potrebbe parlare più lentamente	Could speak more slowly	Potrebbe parlare più lentamente.	Could you speak more slowly.
16. Dove si acquista lo skipass?	Dove si acquista lo skipass	Where you buy your ski pass	Dove si acquista lo skipass.	Where do you buy your ski pass.
17. Dove si trova l'ambasciata italiana?	Dove si trova l'ambasciata italiana	Where the Italian embassy	Dove si trova l'ambasciata italiana.	Where is the Italian embassy.
18. Buongiorno, ho prenotato una camera a nome di Rossi.	Buongiorno ho prenotato una camera a nome di Rossi	Good morning I booked a room in the name of Rossi	Buongiorno ho prenotato una camera non mi di Rossi. (2° tentativo)	Good morning I booked a room for me.
19. Quanto costa?	Quanto costa	How much	Quanto costa?	How much is it?
20. La mia valigia non è arrivata.	La mia valigia non è arrivata	My suitcase did not arrive	La mia valigia non è arrivata.	My suitcase did not arrive.
21. Manca l'acqua calda.	Manca l'acqua calda	No hot water	Manca l'acqua calda.	Missing hot water.

* Gli output ASR qui riportati sono esclusivamente relativi all'ultimo tentativo di riconoscimento vocale effettuato dal partecipante. Ove non diversamente indicato, il riconoscimento vocale è avvenuto al primo tentativo. Per tutti i casi in cui vi sono stati più tentativi di riconoscimento vocale, viene riportato unicamente l'output ASR derivante dal secondo e ultimo tentativo di riconoscimento vocale (indicato esplicitamente nella tabella tramite la dicitura "2° tentativo"), in considerazione sia del fatto che in tali casi il primo tentativo di riconoscimento vocale era stato ritenuto scorretto già dal partecipante stesso in sede di svolgimento del test di valutazione, sia del fatto che in questi casi l'output tradotto derivante dal primo tentativo di riconoscimento vocale non è rilevante ai fini di questo studio in quanto escluso dalla valutazione della MT.

Partecipante 1 – Test italiano-tedesco

Frasi in lingua <i>source</i>	Ultimo riconoscimento vocale GT / input per traduzione*	Output tradotto GT	Ultimo riconoscimento vocale MsT / input per traduzione*	Output tradotto MsT
1. Vorrei un biglietto andata e ritorno per Norimberga.	Vorrei un biglietto andata e ritorno per Norimberga	Ich würde ein Rückflugticket nach Nürnberg gerne	Vorrei un biglietto andata e ritorno per Norimberga.	Ich möchte eine Rückfahrkarte nach Nürnberg gerne.
2. Mi scusi, dove si trova la stazione dei treni?	Mi scusi dove si trova la stazione dei treni	Entschuldigung, wo ist der Bahnhof	Mi scusi, dove si trova la stazione dei treni.	Entschuldigen Sie mich, wo der Bahnhof ist.
3. Da quale binario parte il treno per Berlino?	Da quale binario parte il treno per Berlino	Von welcher Plattform der Zug nach Berlin	Da quale binario parte il treno per Berlino?	Von welchem Bahnsteig der Zug nach Berlin?
4. Quanto costa un biglietto solo andata per Dresda?	Quanto costa un biglietto solo andata per 30 (2° tentativo)	Wie viel es kostet nur One-Way Ticket zu 30	Quanto costa un biglietto solo andata per dress? (2° tentativo)	Wie viel ein Ticket Einzelfahrt zu kleiden?
5. Dov'è il distributore più vicino?	Dov'è il distributore più vicino	Wo ist die nächste Verteiler	Dov'è il distributore più vicino.	Wo ist der nächste Distributor.
6. Si può parcheggiare qui?	Si può parcheggiare qui	Sie können hier parken	Si può parcheggiare qui.	Hier können Sie parken.
7. Esistono abbonamenti settimanali?	Esistono abbonamenti settimanali	Es gibt Wochenkarten	Esistono abbonamenti settimanali.	Es gibt Wochenkarten.
8. Quanto costa il noleggio per una settimana?	Quanto costa il noleggio per una settimana	Wie hoch ist die Miete für eine Woche	Quanto costa il noleggio per una settimana?	Wie hoch ist die Miete für eine Woche?
9. La colazione è inclusa?	La colazione è inclusa (2° tentativo)	Ist das Frühstück eingeschlossen	La colazione è inclusa. (2° tentativo)	Frühstück ist inklusive.
10. A che ora servite la colazione?	A che ora servite la colazione	Zu welcher Zeit das Frühstück serviert	A che ora servite la colazione? (2° tentativo)	Welche Zeit Sie Frühstück servieren?

11. Può chiamarmi un taxi?	Può chiamarmi un taxi (2° tentativo)	Sie können ein Taxi rufen	Puoi chiamarmi un taxi. (2° tentativo)	Sie können mir ein Taxi rufen.
12. Scusi, dov'è la toilette?	Scusi dov'è la toilette (2° tentativo)	Entschuldigen Sie, wo die Toilette ist	Scusi dov'è la toilette.	Entschuldigen Sie mich, wo ist die Toilette.
13. Avete una cartina della città?	Avete una cartina della città	Sie haben eine Karte der Stadt	Avete una cartina della città. (2° tentativo)	Sie haben eine Karte der Stadt.
14. Scusi, parla italiano?	Scusi, parla italiano	Excuse spricht Italienisch	Scusi parla italiano. (2° tentativo)	Leider spricht Italienisch.
15. Potrebbe parlare più lentamente?	Potrebbe parlare più lentamente	Könnte sprechen langsamer	Potrebbe parlare più lentamente.	Sie könnten etwas langsamer sprechen.
16. Dove si acquista lo skipass?	Dove si acquista lo skipass	Wo Sie kaufen Ihre Skipass	Dove si acquista lo skipass.	Wo kaufen Sie Ihren Skipass.
17. Dove si trova l'ambasciata italiana?	Dove si trova l'ambasciata italiana	Wo die italienische Botschaft	Dove si trova l'ambasciata italiana.	Wo ist die italienische Botschaft.
18. Buongiorno, ho prenotato una camera a nome di Rossi.	Buongiorno ho prenotato una camera a nome di Rossi	Ich guten Morgen buchte ein Zimmer im Namen von Rossi	Buongiorno ho prenotato una camera nome di Rossi. (2° tentativo)	Guten Morgen mietete ich ein Raumname von Rossi.
19. Quanto costa?	Quanto costa	Wie viel es kostet	Quanto costa?	Wie viel kostet es?
20. La mia valigia non è arrivata.	La mia valigia non è arrivata	Mein Koffer nicht angekommen	La mia valigia non è arrivata. (2° tentativo)	Mein Koffer ist nicht angekommen.
21. Manca l'acqua calda.	Manca l'acqua calda	Kein warmes Wasser	Manca l'acqua calda.	Fehlende Warmwasser.

* Gli output ASR qui riportati sono esclusivamente relativi all'ultimo tentativo di riconoscimento vocale effettuato dal partecipante. Ove non diversamente indicato, il riconoscimento vocale è avvenuto al primo tentativo. Per tutti i casi in cui vi sono stati più tentativi di riconoscimento vocale, viene riportato unicamente l'output ASR derivante dal secondo e ultimo tentativo di riconoscimento vocale (indicato esplicitamente nella tabella tramite la dicitura "2° tentativo"), in considerazione sia del fatto che in tali casi il primo tentativo di riconoscimento vocale era stato ritenuto scorretto già dal partecipante stesso in sede di svolgimento del test di valutazione, sia del fatto che in questi casi l'output tradotto derivante dal primo tentativo di riconoscimento vocale non è rilevante ai fini di questo studio in quanto escluso dalla valutazione della MT.

Partecipante 2 – Test italiano-tedesco

Frases in lingua <i>source</i>	Ultimo riconoscimento vocale GT / input per traduzione*	Output tradotto GT	Ultimo riconoscimento vocale MsT / input per traduzione*	Output tradotto MsT
1. Vorrei un biglietto andata e ritorno per Norimberga.	Vorrei un biglietto andata e ritorno per Norimberga	Ich würde ein Rückflugticket nach Nürnberg gerne	Vorrei un biglietto andata e ritorno per Norimberga.	Ich möchte eine Rückfahrkarte nach Nürnberg gerne.
2. Mi scusi, dove si trova la stazione dei treni?	Mi scusi dove si trova la stazione dei treni	Entschuldigung, wo ist der Bahnhof	Mi scusi, dove si trova la stazione dei treni.	Entschuldigen Sie mich, wo der Bahnhof ist.
3. Da quale binario parte il treno per Berlino?	Da quale binario parte il treno per Berlino	Von welcher Plattform der Zug nach Berlin	Da quale binario parte il treno per Berlino?	Von welchem Bahnsteig der Zug nach Berlin?
4. Quanto costa un biglietto solo andata per Dresda?	Quanto costa un biglietto solo andata per Dresda (2° tentativo)	Wie viel kostet ein Ticket nach Dresden	Quanto costa un biglietto solo andata per brutta? (2° tentativo)	Wie viel kostet ein Ticket Einzelfahrt zu schlechte?
5. Dov'è il distributore più vicino?	Dov'è il distributore più vicino	Wo ist die nächste Verteiler	Dov'è il distributore più vicino.	Wo ist der nächste Distributor.
6. Si può parcheggiare qui?	Si può parcheggiare qui	Sie können hier parken	Si può parcheggiare qui.	Hier können Sie parken.
7. Esistono abbonamenti settimanali?	Esistono abbonamenti settimanali	Es gibt Wochenkarten	Esistono abbonamenti settimanali.	Es gibt Wochenkarten.
8. Quanto costa il noleggio per una settimana?	Quanto costa il noleggio per una settimana	Wie hoch ist die Miete für eine Woche	Quanto costa il noleggio per una settimana?	Wie hoch ist die Miete für eine Woche?
9. La colazione è inclusa?	La colazione è inclusa (2° tentativo)	Ist das Frühstück eingeschlossen	La colazione è inclusa. (2° tentativo)	Frühstück ist inklusive.
10. A che ora servite la colazione?	A che ora servite la colazione	Zu welcher Zeit das Frühstück serviert	A che ora servite la colazione?	Welche Zeit Sie Frühstück servieren?

11. Può chiamarmi un taxi?	Può chiamarmi un taxi	Sie können ein Taxi rufen	Puoi chiamarmi un taxi. (2° tentativo)	Sie können mir ein Taxi rufen.
12. Scusi, dov'è la toilette?	Scusi dov'è la toilette (2° tentativo)	Entschuldigen Sie, wo die Toilette ist	Scusi dov'è la toilette.	Entschuldigen Sie mich, wo ist die Toilette.
13. Avete una cartina della città?	Avete una cartina della città	Sie haben eine Karte der Stadt	Avete la cartina della città. (2° tentativo)	Sie haben die Karte der Stadt.
14. Scusi, parla italiano?	Scusi, parla italiano	Excuse spricht Italienisch	Scusi parla italiano.	Leider spricht Italienisch.
15. Potrebbe parlare più lentamente?	Potrebbe parlare più lentamente	Könnte sprechen langsamer	Potrebbe parlare più lentamente.	Sie könnten etwas langsamer sprechen.
16. Dove si acquista lo skipass?	Dove si acquista lo skipass	Wo Sie kaufen Ihre Skipass	Dove si acquista lo skipass.	Wo kaufen Sie Ihren Skipass.
17. Dove si trova l'ambasciata italiana?	Dove si trova l'ambasciata italiana	Wo die italienische Botschaft	Dove si trova l'ambasciata italiana.	Wo ist die italienische Botschaft.
18. Buongiorno, ho prenotato una camera a nome di Rossi.	Buongiorno ho prenotato una camera a nome di Rossi	Ich guten Morgen buchte ein Zimmer im Namen von Rossi	Buongiorno ho prenotato una camera non mi di Rossi. (2° tentativo)	Guten Morgen, ich buchte ein Zimmer für mich.
19. Quanto costa?	Quanto costa	Wie viel es kostet	Quanto costa?	Wie viel kostet es?
20. La mia valigia non è arrivata.	La mia valigia non è arrivata	Mein Koffer nicht angekommen	La mia valigia non è arrivata.	Mein Koffer ist nicht angekommen.
21. Manca l'acqua calda.	Manca l'acqua calda	Kein warmes Wasser	Manca l'acqua calda.	Fehlende Warmwasser.

* Gli output ASR qui riportati sono esclusivamente relativi all'ultimo tentativo di riconoscimento vocale effettuato dal partecipante. Ove non diversamente indicato, il riconoscimento vocale è avvenuto al primo tentativo. Per tutti i casi in cui vi sono stati più tentativi di riconoscimento vocale, viene riportato unicamente l'output ASR derivante dal secondo e ultimo tentativo di riconoscimento vocale (indicato esplicitamente nella tabella tramite la dicitura "2° tentativo"), in considerazione sia del fatto che in tali casi il primo tentativo di riconoscimento vocale era stato ritenuto scorretto già dal partecipante stesso in sede di svolgimento del test di valutazione, sia del fatto che in questi casi l'output tradotto derivante dal primo tentativo di riconoscimento vocale non è rilevante ai fini di questo studio in quanto escluso dalla valutazione della MT.

Appendice 2

QUESTIONARIO DI VALUTAZIONE ITALIANO-INGLESE

Parte 1

Nella parte 1 del questionario ti si chiede di esprimere un giudizio sulla *fluency* o scorrevolezza di 42 frasi in inglese. Dovrai esprimere il tuo giudizio valutando *la fluency* della frase proposta su una scala da 5 a 1, dove 5 equivale a una frase in un inglese perfetto, senza errori di grammatica o in generale errori che nessun madrelingua farebbe mai, e 1 equivalente a una frase in un inglese incomprensibile, con molti errori di grammatica o in generale errori che nessun madrelingua farebbe mai.

Per indicare la risposta puoi evidenziarla (con un colore o utilizzando il grassetto) oppure apporvi una x.

Scorrevolezza / Fluency

(La frase tradotta contiene errori di grammatica o in generale errori che nessun madrelingua farebbe mai?)

Questionario di valutazione – Parte 1

1. I'd like a return ticket to Edinburgh

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

2. From which platform the train to London

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

3. There are weekly tickets

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

4. Could you speak more slowly

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

5. Excuse me, where is the train station

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

6. How much is the rental of a week

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

7. Where is the nearest distributor

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

8. Sorry speaks Italian

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

9. Where you buy your ski pass

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

10. My suitcase did not arrive

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

11. How much is a one-way ticket to first

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

12. You have a map of the city

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

13. What time you serve breakfast?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

14. Excuse me where's the toilet

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

15. Where is the Italian embassy

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

16. Good morning I booked a room for me

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

17. How much is it?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

18. How much it costs only one-way ticket to first

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

19. You can call me a cab

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

20. No hot water

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

21. Excuse me, where is the train station?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

22. Could speak more slowly

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

23. Where's the nearest distributor

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

24. How much does a one-way ticket to Plymouth

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

25. Breakfast is included

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

26. How much is the rent for a week?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

27. Do you have a map of the city

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

28. From which platform does the train leave for London?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

29. Good morning I made reservation the name of Rossi

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

30. Where the Italian embassy

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

31. You can park here

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

32. What time served breakfast

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

33. Excuse me where is the toilet

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

34. Where do you buy your ski pass

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

35. Missing hot water

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

36. How much

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

37. You can call a taxi

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

38. Ravioli are coming

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

39. How much is a ticket one-way first

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

40. Good morning I booked a room in the name of Rossi

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

41. How much is the rental for a week

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

42. Excuse speaks Italian

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 inglese perfetto
 - 4
 - 3
 - 2
 - 1 inglese incomprensibile
-

Parte 2

Nella parte 2 del questionario ti si chiede di esprimere un giudizio sulla traduzione delle precedenti 42 frasi dall'italiano verso l'inglese. Per ogni esempio da valutare troverai: frase originale in lingua *source* (italiano) e output tradotto in lingua *target* (inglese), in modo da potere confrontare la versione originale con quella tradotta. La versione tradotta è riportata in grassetto.

Dovrai esprimere il tuo giudizio valutando la traduzione in merito a equivalenza semantica e di intento, utilizzando una scala da 5 a 1, dove 5 corrisponde a una equivalenza perfetta di significato e intento fra frase originale e traduzione e 1 corrisponde a una mancanza totale di equivalenza di significato e intento fra frase originale e traduzione.

Per indicare la risposta puoi evidenziarla (con un colore o utilizzando il grassetto) oppure apporvi una x.

Equivalenza semantica

(La frase tradotta rispecchia il significato e l'intento della frase originale?)

Questionario di valutazione – Parte 2

1. Mi scusi, dove si trova la stazione dei treni?

Excuse me, where is the train station

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

2. Quanto costa un biglietto solo andata per Plymouth?

How much does a one-way ticket to Plymouth

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

3. Quanto costa il noleggio per una settimana?

How much is the rental of a week

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

4. Dov'è il distributore più vicino?
Where's the nearest distributor

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

5. La colazione è inclusa?
Breakfast is included

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

6. Scusi, parla italiano?
Excuse speaks Italian

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

7. Potrebbe parlare più lentamente?
Could you speak more slowly

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

8. Buongiorno, ho prenotato una camera a nome di Rossi.

Good morning I made reservation the name of Rossi

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

9. Dove si acquista lo skipass?

Where you buy your ski pass

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

10. Avete una cartina della città?

Do you have a map of the city

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

11. Quanto costa?

How much

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

12. Si può parcheggiare qui?

You can park here

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

13. Può chiamarmi un taxi?

You can call me a cab

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

14. A che ora servite la colazione?

What time you serve breakfast?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

15. Dove si trova l'ambasciata italiana?

Where is the Italian embassy

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

16. Scusi, dov'è la toilette?

Excuse me where's the toilet

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

17. Manca l'acqua calda.

Missing hot water

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

18. Quanto costa il noleggio per una settimana?

How much is the rent for a week?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

19. Vorrei un biglietto andata e ritorno per Edimburgo.

I'd like a return ticket to Edinburgh

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

20. Quanto costa un biglietto solo andata per Plymouth?

How much is a one-way ticket to first

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

21. Da quale binario parte il treno per Londra?

From which platform does the train leave for London?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

22. La mia valigia non è arrivata.

Ravioli are coming

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

23. A che ora servite la colazione?

What time served breakfast

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

24. Quanto costa un biglietto solo andata per Plymouth?

How much is a ticket one-way first

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

25. Esistono abbonamenti settimanali?

There are weekly tickets

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

26. Scusi, parla italiano?

Sorry speaks Italian

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

27. Buongiorno, ho prenotato una camera a nome di Rossi.

Good morning I booked a room for me

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

28. Può chiamarmi un taxi?

You can call a taxi

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

29. Dove si acquista lo skipass?

Where do you buy your ski pass

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

30. Potrebbe parlare più lentamente?

Could speak more slowly

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

31. Da quale binario parte il treno per Londra?

From which platform the train to London

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

32. Avete una cartina della città?

You have a map of the city

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

33. Quanto costa il noleggio per una settimana?

How much is the rental for a week

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

34. Scusi, dov'è la toilette?

Excuse me where is the toilet

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

35. La mia valigia non è arrivata.

My suitcase did not arrive

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

36. Dove si trova l'ambasciata italiana?

Where the Italian embassy

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

37. Quanto costa?

How much is it?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

38. Manca l'acqua calda.

No hot water

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

39. Mi scusi, dove si trova la stazione dei treni?

Excuse me, where is the train station?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

40. Quanto costa un biglietto solo andata per Plymouth?

How much it costs only one-way ticket to first

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

41. Buongiorno, ho prenotato una camera a nome di Rossi.

Good morning I booked a room in the name of Rossi

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

42. Dov'è il distributore più vicino?

Where is the nearest distributor

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

Appendice 3

QUESTIONARIO DI VALUTAZIONE ITALIANO-TEDESCO

Parte 1

Nella parte 1 del questionario ti si chiede di esprimere un giudizio sulla *fluency* o scorrevolezza di 44 frasi in tedesco. Dovrai esprimere il tuo giudizio valutando *la fluency* della frase proposta su una scala da 5 a 1, dove 5 equivale a una frase in un tedesco perfetto, senza errori di grammatica o in generale errori che nessun madrelingua farebbe mai, e 1 equivalente a una frase in un tedesco incomprensibile, con molti errori di grammatica o in generale errori che nessun madrelingua farebbe mai.

Per indicare la risposta puoi evidenziarla (con un colore o utilizzando il grassetto) oppure apporvi una x.

Scorrevolezza / Fluency

(La frase tradotta contiene errori di grammatica o in generale errori che nessun madrelingua farebbe mai?)

Questionario di valutazione – Parte 1

1. Wo ist die italienische Botschaft

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

2. Wie viel kostet ein Ticket nach Dresden

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

3. Ich möchte eine Rückfahrkarte nach Nürnberg gerne

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

4. Mein Koffer nicht angekommen

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

5. Sie haben eine Karte der Stadt

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

6. Entschuldigen Sie mich, wo ist die Toilette

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

7. Kein warmes Wasser

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

8. Von welchem Bahnsteig der Zug nach Berlin?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

9. Guten Morgen mietete ich ein Raumname von Rossi

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

10. Wie viel es kostet

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

11. Wo Sie kaufen Ihre Skipass

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

12. Leider spricht Italienisch

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

13. Sie können ein Taxi rufen

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

14. Sie könnten etwas langsamer sprechen

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

15. Sie können mir ein Taxi rufen

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

16. Zu welcher Zeit das Frühstück serviert

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

17. Wie hoch ist die Miete für eine Woche

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

18. Wie viel kostet ein Ticket Einzelfahrt zu schlechte?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

19. Ist das Frühstück eingeschlossen

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

20. Wo ist die nächste Verteiler

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

21. Sie können hier parken

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

22. Entschuldigung, wo ist der Bahnhof

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

23. Es gibt Wochenkarten

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

24. Excuse spricht Italienisch

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

25. Ich würde ein Rückflugticket nach Nürnberg gerne

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

26. Mein Koffer ist nicht angekommen

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

27. Wie viel ein Ticket Einzelfahrt zu kleiden?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

28. Von welcher Plattform der Zug nach Berlin

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

29. Frühstück ist inklusive

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

30. Guten Morgen, ich buchte ein Zimmer für mich

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

31. Wo ist der nächste Distributor

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

32. Könnte sprechen langsamer

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

33. Fehlende Warmwasser

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

34. Sie haben die Karte der Stadt

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

35. Entschuldigen Sie, wo die Toilette ist

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

36. Wo die italienische Botschaft

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

37. Wie hoch ist die Miete für eine Woche?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

38. Entschuldigen Sie mich, wo der Bahnhof ist

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

39. Welche Zeit Sie Frühstück servieren?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

40. Wie viel kostet es?

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

41. Ich guten Morgen buchte ein Zimmer im Namen von Rossi

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

42. Hier können Sie parken

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

43. Wie viel es kostet nur One-Way Ticket zu 30

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

44. Wo kaufen Sie Ihren Skipass

Scorrevolezza / Fluency

Come giudichi la fluency della traduzione? E'

- 5 tedesco perfetto
 - 4
 - 3
 - 2
 - 1 tedesco incomprensibile
-

Parte 2

Nella parte 2 del questionario ti si chiede di esprimere un giudizio sulla traduzione delle precedenti 44 frasi dall'italiano verso il tedesco. Per ogni esempio da valutare troverai: frase originale in lingua *source* (italiano) e output tradotto in lingua *target* (tedesco), in modo da potere confrontare la versione originale con quella tradotta. La versione tradotta è riportata in grassetto.

Dovrai esprimere il tuo giudizio valutando la traduzione in merito a equivalenza semantica e di intento, utilizzando una scala da 5 a 1, dove 5 corrisponde a una equivalenza perfetta di significato e intento fra frase originale e traduzione e 1 corrisponde a una mancanza totale di equivalenza di significato e intento fra frase originale e traduzione.

Per indicare la risposta puoi evidenziarla (con un colore o utilizzando il grassetto) oppure apporvi una x.

Equivalenza semantica

(La frase tradotta rispecchia il significato e l'intento della frase originale?)

Questionario di valutazione – Parte 2

1. Mi scusi, dove si trova la stazione dei treni?

Entschuldigen Sie mich, wo der Bahnhof ist

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

2. Quanto costa un biglietto solo andata per Dresda?

Wie viel kostet ein Ticket nach Dresden

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

3. Quanto costa il noleggio per una settimana?

Wie hoch ist die Miete für eine Woche?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

4. Dov'è il distributore più vicino?

Wo ist der nächste Distributor

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

5. La colazione è inclusa?

Frühstück ist inklusive

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

6. Scusi, parla italiano?

Leider spricht Italienisch

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

7. Potrebbe parlare più lentamente?

Sie könnten etwas langsamer sprechen

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

8. Buongiorno, ho prenotato una camera a nome di Rossi.

Guten Morgen mietete ich ein Raumname von Rossi

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

9. Dove si acquista lo skipass?

Wo kaufen Sie Ihren Skipass

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

10. Avete una cartina della città?

Sie haben die Karte der Stadt

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

11. Quanto costa?

Wie viel kostet es?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

12. Si può parcheggiare qui?

Hier können Sie parken

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

13. Può chiamarmi un taxi?

Sie können ein Taxi rufen

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

14. A che ora servite la colazione?

Zu welcher Zeit das Frühstück serviert

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

15. Dove si trova l'ambasciata italiana?

Wo ist die italienische Botschaft

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

16. Scusi, dov'è la toilette?

Entschuldigen Sie, wo die Toilette ist

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

17. Manca l'acqua calda.

Kein warmes Wasser

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

18. Vorrei un biglietto andata e ritorno per Norimberga.

Ich möchte eine Rückfahrkarte nach Nürnberg gerne

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

19. Quanto costa un biglietto solo andata per Dresda?

Wie viel kostet ein Ticket Einzelfahrt zu schlechte?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

20. Da quale binario parte il treno per Berlino?
Von welchem Bahnsteig der Zug nach Berlin?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

21. La mia valigia non è arrivata.
Mein Koffer nicht angekommen

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

22. A che ora servite la colazione?
Welche Zeit Sie Frühstück servieren?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

23. Quanto costa un biglietto solo andata per Dresda?
Wie viel es kostet nur One-Way Ticket zu 30

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

24. Esistono abbonamenti settimanali?

Es gibt Wochenkarten

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

25. Scusi, parla italiano?

Excuse spricht Italienisch

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

26. Buongiorno, ho prenotato una camera a nome di Rossi.

Guten Morgen, ich buchte ein Zimmer für mich

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

27. Può chiamarmi un taxi?

Sie können mir ein Taxi rufen

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

28. Dove si acquista lo skipass?

Wo Sie kaufen Ihre Skipass

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

29. Potrebbe parlare più lentamente?

Könnte sprechen langsamer

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

30. Da quale binario parte il treno per Berlino?

Von welcher Plattform der Zug nach Berlin

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

31. Avete una cartina della città?

Sie haben eine Karte der Stadt

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

32. Quanto costa il noleggio per una settimana?

Wie hoch ist die Miete für eine Woche

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

33. Scusi, dov'è la toilette?

Entschuldigen Sie mich, wo ist die Toilette

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

34. La mia valigia non è arrivata.

Mein Koffer ist nicht angekommen

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

35. Dove si trova l'ambasciata italiana?

Wo die italienische Botschaft

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

36. Quanto costa?

Wie viel es kostet

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

37. Manca l'acqua calda.

Fehlende Warmwasser

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

38. Mi scusi, dove si trova la stazione dei treni?

Entschuldigung, wo ist der Bahnhof

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

39. Quanto costa un biglietto solo andata per Dresda?

Wie viel ein Ticket Einzelfahrt zu kleiden?

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

40. Buongiorno, ho prenotato una camera a nome di Rossi.

Ich guten Morgen buchte ein Zimmer im Namen von Rossi

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

41. Dov'è il distributore più vicino?

Wo ist die nächste Verteiler

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

42. Vorrei un biglietto andata e ritorno per Norimberga.

Ich würde ein Rückflugticket nach Nürnberg gerne

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

43. La colazione è inclusa?

Ist das Frühstück eingeschlossen

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

44. Si può parcheggiare qui?

Sie können hier parken

Equivalenza semantica

Rispetto all'originale la traduzione fornita è

- 5 perfettamente equivalente per significato e intento
 - 4
 - 3
 - 2
 - 1 per nulla equivalente per significato e intento
-

Appendice 4

Test italiano-inglese

Frase <i>source</i>	Output tradotto valutato	App di provenienza	Valutazioni scorrevolezza / <i>fluency</i>			Valutazioni equivalenza semantica		
Vorrei un biglietto andata e ritorno per Edimburgo.	I'd like a return ticket to Edinburgh	GT – MsT	5	5	4	5	5	5
Da quale binario parte il treno per Londra?	From which platform the train to London	GT	2	2	2	2	3	3
Esistono abbonamenti settimanali?	There are weekly tickets	GT – MsT	5	4	2	2	3	1
Potrebbe parlare più lentamente?	Could you speak more slowly	MsT	4	3	3	3	3	5
Mi scusi, dove si trova la stazione dei treni?	Excuse me, where is the train station	GT	4	5	5	4	5	5
Quanto costa il noleggio per una settimana?	How much is the rental of a week	GT	3	2	4	3	1	4

Dov'è il distributore più vicino?	Where is the nearest distributor	GT	3	3	4	2	3	2
Scusi, parla italiano?	Sorry speaks Italian	MsT	1	1	1	2	2	1
Dove si acquista lo skipass?	Where you buy your ski pass	GT	3	2	2	3	3	3
La mia valigia non è arrivata.	My suitcase did not arrive	GT – MsT	4	4	5	5	4	5
Quanto costa un biglietto solo andata per Plymouth?	How much is a one-way ticket to first	MsT	2	2	2	1	3	1
Avete una cartina della città?	You have a map of the city	GT	5	3	3	3	4	2
A che ora servite la colazione?	What time you serve breakfast?	MsT	3	3	3	3	3	4
Scusi, dov'è la toilette?	Excuse me where's the toilet	MsT	4	5	5	4	5	5
Dove si trova l'ambasciata italiana?	Where is the Italian embassy	MsT	4	5	5	4	5	5
Buongiorno, ho prenotato una camera a nome di Rossi.	Good morning I booked a room for me	MsT	5	4	4	1	2	1
Quanto costa?	How much is it?	MsT	5	5	5	5	5	5

Quanto costa un biglietto solo andata per Plymouth?	How much it costs only one-way ticket to first	GT	2	1	1	2	1	1
Può chiamarmi un taxi?	You can call me a cab	MsT	5	3	3	2	2	3
Manca l'acqua calda.	No hot water	GT	3	3	2	4	4	4
Mi scusi, dove si trova la stazione dei treni?	Excuse me, where is the train station?	MsT	5	5	5	5	5	5
Potrebbe parlare più lentamente?	Could speak more slowly	GT	3	3	3	4	3	4
Dov'è il distributore più vicino?	Where's the nearest distributor	MsT	3	3	4	2	2	2
Quanto costa un biglietto solo andata per Plymouth?	How much does a one-way ticket to Plymouth	GT	3	2	3	2	2	3
La colazione è inclusa?	Breakfast is included	GT – MsT	5	4	5	2	2	2
Quanto costa il noleggio per una settimana?	How much is the rent for a week?	MsT	5	4	5	4	4	5
Avete una cartina della città?	Do you have a map of the city	MsT	4	5	5	3	5	5
Da quale binario parte il treno per Londra?	From which platform does the train leave for London?	MsT	4	4	4	3	3	4

Buongiorno, ho prenotato una camera a nome di Rossi.	Good morning I made reservation the name of Rossi	MsT	3	3	3	3	2	3
Dove si trova l'ambasciata italiana?	Where the Italian embassy	GT	2	2	3	3	3	3
Si può parcheggiare qui?	You can park here	GT	5	5	5	2	2	2
A che ora servite la colazione?	What time served breakfast	GT	3	2	2	3	3	2
Scusi, dov'è la toilette?	Excuse me where is the toilet	GT	4	4	5	4	5	5
Dove si acquista lo skipass?	Where do you buy your ski pass	MsT	4	4	3	4	5	5
Manca l'acqua calda.	Missing hot water	MsT	3	1	2	3	2	2
Quanto costa?	How much	GT	1	2	2	2	2	2
Può chiamarmi un taxi?	You can call a taxi	GT	5	2	3	2	3	3
La mia valigia non è arrivata.	Ravioli are coming	MsT	4	3	1	1	1	1
Quanto costa un biglietto solo andata per Plymouth?	How much is a ticket one-way first	MsT	1	2	1	1	2	1

Buongiorno, ho prenotato una camera a nome di Rossi.	Good morning I booked a room in the name of Rossi	GT	4	3	3	4	3	3
Quanto costa il noleggio per una settimana?	How much is the rental for a week	GT	4	4	5	4	5	4
Scusi, parla italiano?	Excuse speaks Italian	GT	2	1	1	1	1	1

Test italiano-tedesco

Frasesource	Output tradotto valutato	App di provenienza	Valutazioni scorrevolezza / fluency			Valutazioni equivalenza semantica		
			Valutatore 1	Valutatore 2	Valutatore 3	Valutatore 1	Valutatore 2	Valutatore 3
Dove si trova l'ambasciata italiana?	Wo ist die italienische Botschaft	MsT	5	4	5	5	4	4
Quanto costa un biglietto solo andata per Dresda?	Wie viel kostet ein Ticket nach Dresden	GT	5	4	5	4	3	5
Vorrei un biglietto andata e ritorno per Norimberga.	Ich möchte eine Rückfahrkarte nach Nürnberg gerne	MsT	2	3	4	3	2	1
La mia valigia non è arrivata.	Mein Koffer nicht angekommen	GT	3	3	3	3	3	4
Avete una cartina della città?	Sie haben eine Karte der Stadt	GT – MsT	2	2	4	2	4	3
Scusi, dov'è la toilette?	Entschuldigen Sie mich, wo ist die Toilette	MsT	5	4	5	5	4	5
Manca l'acqua calda.	Kein warmes Wasser	GT	3	5	4	2	4	4

Da quale binario parte il treno per Berlino?	Von welchem Bahnsteig der Zug nach Berlin?	MsT	2	3	2	2	3	1
Buongiorno, ho prenotato una camera a nome di Rossi.	Guten Morgen mietete ich ein Raumname von Rossi	MsT	1	2	2	1	1	2
Quanto costa?	Wie viel es kostet	GT	4	3	3	3	4	4
Dove si acquista lo skipass?	Wo Sie kaufen Ihre Skipass	GT	4	2	3	3	3	4
Scusi, parla italiano?	Leider spricht Italienisch	MsT	1	2	1	1	1	1
Può chiamarmi un taxi?	Sie können ein Taxi rufen	GT	4	5	4	3	3	1
Potrebbe parlare più lentamente?	Sie könnten etwas langsamer sprechen	MsT	5	4	5	3	4	3
Può chiamarmi un taxi?	Sie können mir ein Taxi rufen	MsT	4	4	4	3	4	1
A che ora servite la colazione?	Zu welcher Zeit das Frühstück serviert	GT	2	3	3	2	2	1
Quanto costa il noleggio per una settimana?	Wie hoch ist die Miete für eine Woche	GT	3	5	5	4	4	5

Quanto costa un biglietto solo andata per Dresda?	Wie viel kostet ein Ticket Einzelfahrt zu schlechte?	MsT	1	3	3	1	1	1
La colazione è inclusa?	Ist das Frühstück eingeschlossen	GT	2	3	5	3	3	5
Dov'è il distributore più vicino?	Wo ist die nächste Verteiler	GT	3	2	4	1	1	2
Si può parcheggiare qui?	Sie können hier parken	GT	5	5	5	1	1	1
Mi scusi, dove si trova la stazione dei treni?	Entschuldigung, wo ist der Bahnhof	GT	5	5	5	5	4	5
Esistono abbonamenti settimanali?	Es gibt Wochenkarten	GT – MsT	5	4	5	3	3	2
Scusi, parla italiano?	Excuse spricht Italienisch	GT	2	1	2	2	1	1
Vorrei un biglietto andata e ritorno per Norimberga.	Ich würde ein Rückflugticket nach Nürnberg gerne	GT	3	4	2	3	1	1
La mia valigia non è arrivata.	Mein Koffer ist nicht angekommen	MsT	5	5	5	5	5	5

Quanto costa un biglietto solo andata per Dresda?	Wie viel ein Ticket Einzelfahrt zu kleiden?	MsT	1	3	2	1	1	1
Da quale binario parte il treno per Berlino?	Von welcher Plattform der Zug nach Berlin	GT	3	2	2	3	2	2
La colazione è inclusa?	Frühstück ist inklusive	MsT	3	3	4	2	4	3
Buongiorno, ho prenotato una camera a nome di Rossi.	Guten Morgen, ich buchte ein Zimmer für mich	MsT	4	5	4	4	1	1
Dov'è il distributore più vicino?	Wo ist der nächste Distributor	MsT	4	3	3	1	1	4
Potrebbe parlare più lentamente?	Könnte sprechen langsamer	GT	2	3	2	2	3	2
Manca l'acqua calda.	Fehlende Warmwasser	MsT	3	2	3	3	2	4
Avete una cartina della città?	Sie haben die Karte der Stadt	MsT	5	3	4	3	4	2
Scusi, dov'è la toilette?	Entschuldigen Sie, wo die Toilette ist	GT	4	3	4	3	4	4
Dove si trova l'ambasciata italiana?	Wo die italienische Botschaft	GT	3	3	3	3	4	4

Quanto costa il noleggio per una settimana?	Wie hoch ist die Miete für eine Woche?	MsT	4	5	5	4	4	5
Mi scusi, dove si trova la stazione dei treni?	Entschuldigen Sie mich, wo der Bahnhof ist	MsT	3	3	4	3	3	5
A che ora servite la colazione?	Welche Zeit Sie Frühstück servieren?	MsT	2	3	3	3	2	3
Quanto costa?	Wie viel kostet es?	MsT	5	5	5	5	5	5
Buongiorno, ho prenotato una camera a nome di Rossi.	Ich guten Morgen buchte ein Zimmer im Namen von Rossi	GT	3	1	1	3	1	1
Si può parcheggiare qui?	Hier können Sie parken	MsT	5	5	5	1	3	1
Dove si acquista lo skipass?	Wo kaufen Sie Ihren Skipass	MsT	5	5	5	5	3	4
Quanto costa un biglietto solo andata per Dresda?	Wie viel es kostet nur One-Way Ticket zu 30	GT	1	1	1	1	1	1

Ringraziamenti

Un sentito ringraziamento va alla professoressa Bernardini per avere accolto la proposta di realizzare uno studio sulla traduzione automatica del parlato e per avermi guidato con i suoi preziosi consigli nella stesura del presente elaborato, dimostrandosi sempre disponibile e al contempo sinceramente interessata alla tematica scelta e allo sviluppo dello studio condotto.

Un caloroso ringraziamento anche alla professoressa Lecci per il Suo interesse, la collaborazione e la disponibilità mostrata nel corso di questi mesi.

Ringrazio anche la professoressa Mack per avermi aiutato nella scelta del relatore con i suoi consigli, rivelatisi davvero preziosi.

Per ultimo, ma certamente non per importanza, ringrazio calorosamente tutti coloro che hanno partecipato al presente studio. Fra questi, un primo ringraziamento va a Paola e Greta per avere accettato di “prestare la propria voce” al mio progetto e avere preso parte alla prima fase del test di valutazione. Un ringraziamento altrettanto caloroso va anche a Beatrice, Alessandro, Silvia, Ilaria, Luca e Andrea, per il prezioso contributo fornito tramite la compilazione dei questionari di valutazione. Ringrazio tutti i partecipanti per la grande disponibilità dimostrata, nonostante gli impegni lavorativi, accademici e personali: la loro collaborazione è stata davvero fondamentale per la realizzazione di questo progetto.